10

Dilute Bismide Alloys

Christopher A. Broderick

Igor P. Marko

Eoin P. O'Reilly

and

Stephen J. Sweeney

10.1 Introduction

Dilute bismide alloys are III–V semiconductor alloys containing small fractions of substitutional bismuth (Bi) atoms. Similar to the incorporation of nitrogen (N) to form the dilute nitride alloys described in Chapter 9, dilute bismide alloys are characterized by the fact that Bi acts as an isovalent impurity when incorporated into, e.g. (In)GaAs, to form the (In)$GaBi_xAs_{1-x}$ alloy. Similar to the case of dilute nitride alloys, it is the large differences in size (covalent radius) and chemical properties (electronegativity) between Bi atoms and the group-V atoms they replace that brings about this impurity-like behavior, which in practice means that incorporating Bi at dilute concentrations has a significant impact on the material properties. Dilute bismide alloys, like dilute nitrides, are considered to be highly mismatched semiconductor materials. However, while N incorporation primarily affects the conduction band (CB) structure of the material into which it is incorporated, Bi, being significantly larger and more electropositive than N, primarily affects the valence band (VB). Since several general aspects of the physics of dilute bismide alloys are qualitatively similar to those of dilute nitrides, dilute bismide alloys can be considered a naturally complementary material system to the dilute nitrides.

The dilute bismide alloy $GaBi_xAs_{1-x}$ has several novel electronic properties. For example, substitution of As by Bi causes a rapid reduction in the band gap (E_g) with increasing Bi composition x, by up to 90 meV per % Bi at low x [1–6]. This strong Bi-induced decrease in E_g is accompanied by a strong increase in the VB spin-orbit-splitting energy (Δ_{SO}) [5–7], and the changes in both E_g and Δ_{SO} are characterized by strong, composition-dependent bowing, much like the band gap reduction in GaN_xAs_{1-x} [5,8]. These unusual material properties—among others, which we discuss in Section 10.2—have prompted significant interest in the development of dilute bismide alloys for a range of practical applications, including in semiconductor lasers [9–15] and photodiodes [16–19], as well as photovoltaics [20–23], spintronics [7,24,25], and thermoelectrics [26].

The theoretical modeling of dilute bismide alloys is significantly complicated by the impurity-like behavior of the Bi atoms in the alloy, and the resultant strong perturbation of the electronic structure of the host matrix semiconductor. The presence of alloy disorder due to the formation of pairs and larger clusters of Bi atoms in $GaBi_xAs_{1-x}$ has a particularly significant impact on the electronic structure, meaning that any serious attempt to quantify the consequences of Bi incorporation on the material properties must encompass an atomistic viewpoint. In Section 10.2, we provide an overview of the theory of the unusual properties of dilute bismide alloys, as revealed by experimental measurements, and discuss the various atomistic approaches that have been developed to understand the impact of Bi incorporation on the electronic structure. In particular, we focus on an atomistic tight-binding (TB) model we have developed to study the electronic structure of dilute bismide alloys of (In)GaAs(P) and compare the results of calculations undertaken using this approach to a series of experimental measurements of the electronic [5], optical [6,27,28], and spin properties [29] of $GaBi_xAs_{1-x}$ alloys.

While atomistic models provide significant insight into the properties of Bi-containing alloys, they are too computationally expensive to apply directly on the length scales required to model the properties of optoelectronic devices. This motivates the development of continuum models which are suited to calculating the properties of realistically sized Bi-containing device heterostructures. Through a series of detailed atomistic calculations, we review that the main features of the band structure of $GaBi_xAs_{1-x}$ and related alloys can be understood in terms of a band-anticrossing interaction, much like that considered for the dilute nitrides in Chapter 9. However, for $GaBi_xAs_{1-x}$ alloys the situation is the converse of that in GaN_xAs_{1-x}. In GaN_xAs_{1-x}, the band-anticrossing model describes the N composition-dependent coupling between the extended states of the host matrix (GaAs) CB edge and highly localized N-related impurity states which are resonant with the CB. In $GaBi_xAs_{1-x}$, atomistic calculations reveal that a Bi composition-dependent band-anticrossing interaction occurs between the extended states of the host matrix VB edge and highly localized Bi-related impurity states which are resonant with the VB.

Taking into account the p-like nature of the GaAs VB edge eigenstates, as well as the presence of strong spin-orbit-coupling in the VB, it is clear that even the simplest description of the $GaBi_xAs_{1-x}$ VB structure will be more complicated than that of the GaN_xAs_{1-x} CB structure. It is to this issue that we turn our attention in Section 10.3, where we review how atomistic calculations can be used to directly derive and parameterize valence band-anticrossing (VBAC) models for dilute bismide alloys, thereby removing the parametric ambiguity with which phenomenological models of this type are typically associated. Following this approach, we review the derivation of a 12-band $\mathbf{k}\cdot\mathbf{p}$ Hamiltonian for the band structure of (In)$GaBi_xAs_{1-x}$ alloys directly from atomistic supercell calculations, and present a general method for calculating the band offsets in (001)-oriented pseudomorphically strained epilayers and quantum wells (QWs).

The practical application which has to date seen the largest research effort and most significant progress has been the development of $GaBi_xAs_{1-x}$ alloys for applications in GaAs-based semiconductor QW lasers [10–12,15,30]. A critical issue with existing InP-based 1.3- and 1.55-μm QW lasers is that their threshold currents and optical (cavity) losses tend to increase strongly with increasing temperature, due largely to a combination of two intrinsic loss mechanisms: non-radiative Auger recombination [31,32] and inter-valence band absorption (IVBA) [33]. $GaBi_xAs_{1-x}$ alloys having Bi compositions $x > 10\%$ have been demonstrated to have a band structure in which $\Delta_{SO} > E_g$, while also possessing a band gap $E_g \approx 0.8$ eV

(1.55 μm) [5,6]. It has therefore been proposed that this large VB spin-orbit-splitting energy could lead to the suppression of the dominant "conduction-hole-spin-hole" (CHSH) Auger recombination process in 1.55-μm semiconductor lasers [9,10,12] in which the energy and crystal momentum of an electron–hole pair recombining across the band gap excite a hole from the VB edge to the spin-split-off (SO) band, as depicted schematically in Figure 10.1.

The CHSH Auger process is a loss mechanism which accounts for the majority of the current at threshold in conventional InP-based devices, and which also strongly degrades the stability and efficiency of the device operation above room temperature [31]. Suppression of the CHSH Auger recombination pathway is therefore expected to bring about highly efficient, temperature-stable laser operation, opening the route to uncooled GaAs-based telecom lasers with significantly reduced power consumption and providing large energy savings compared to the InP-based technologies currently deployed in optical communication networks [9–12]. Using dilute bismide alloys as a route to realizing long-wavelength semiconductor lasers on GaAs substrates should also enable the growth of monolithic telecom-wavelength vertical-cavity surface-emitting lasers (VCSELs), thereby bringing the benefits of (Al)GaAs-based distributed Bragg reflectors (DBRs) to 1.55 μm. This is appealing from a technological perspective, because of the possibility of taking advantage of the enhanced carrier and optical confinement offered by (Al)GaAs-based heterostructures, as well as the potential to monolithically integrate long-wavelength semiconductor lasers with high-speed GaAs-based microelectronics.

Considerable progress has been made in developing epitaxial growth of $GaBi_xAs_{1-x}$ alloys, leading to the demonstration of an optically pumped bulk-like laser by Tominaga et al. in 2010 [34]. Room temperature electrically pumped lasing from a $GaBi_xAs_{1-x}$ QW laser was first demonstrated at $x = 2.1\%$ by Ludewig et al. in 2013 [30], and the Bi composition in such laser structures has since extended up to

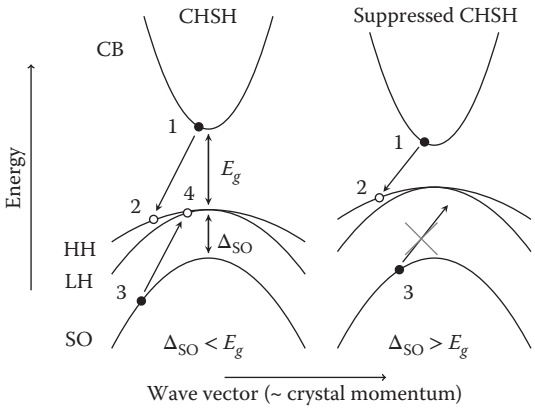

FIGURE 10.1 Left: Schematic illustration of the hot hole-producing CHSH Auger recombination process, in which a CB edge electron (1) recombines with a VB edge hole (2), with the energy released then exciting a VB edge hole (4) to the SO band (3), such that energy and crystal momentum ($\hbar\mathbf{k}$) are conserved. The hot hole generated by this process dissipates its excess energy and crystal momentum primarily via scattering within the lattice, which generates phonons, increasing the lattice temperature and generating significant waste heat. The presence of strong, temperature-dependent CHSH Auger recombination in the materials forming the active region of InP-based 1.3- and 1.55-μm semiconductor lasers means that external cooling equipment is required to extract Auger-generated waste heat, in order to maintain operational stability, but at the cost of significantly increasing power consumption. Right: Schematic illustration of suppressed CHSH Auger recombination, achieved for a band structure in which the spin-orbit-splitting energy SO exceeds the band gap: $\Delta_{SO} > E_g$. In this case, the CHSH process is forbidden by conservation of energy, since the energy made available by an electron–hole pair recombining across the band gap is insufficient to excite a VB edge hole to the SO band. The suppression of IVBA proceeds in a similar manner: for $\Delta_{SO} > E_g$ the energy of a photon emitted by an electron–hole pair recombining across the band gap is insufficient to promote (by optical absorption) an electron from the SO band to a hole state at the VB edge, thereby suppressing photon reabsorption in the active region.

$x = 6.5\%$ by Butkutė et al. [35,36]. However, Bi compositions $x > 10\%$ are required to obtain a band structure having $\Delta_{SO} > E_g$ (and hence suppression of Auger recombination and IVBA processes involving the SO band) in $GaBi_xAs_{1-x}$ alloys. Achieving high-quality $GaBi_xAs_{1-x}$ layers with $x > 10\%$ is challenging due to the difficulty associated with epitaxial growth of Bi-containing alloys (see, e.g., Refs. [37,38] and references therein). While significant progress has been made in developing $GaBi_xAs_{1-x}$ materials and devices over the last 5 years, efforts are ongoing to grow heterostructures with sufficiently high Bi compositions to demonstrate lasing at 1.55 μm, as well as suppression of Auger recombination and IVBA, in this new class of semiconductor alloys.

Given the significant potential of dilute bismide alloys for applications in next-generation semiconductor lasers, our focus throughout much of this chapter is on reviewing the theory of dilute bismide lasers. In Section 10.4 we provide an overview of the theory of GaAs-based dilute bismide QW lasers, using a theoretical model based on the 12-band **k·p** Hamiltonian derived in Section 10.3. Using this approach, we review (1) the impact of Bi on the band structure, density of states (DOS), and optical gain, (2) pathways toward device optimization at low Bi compositions, (3) trends in the expected device performance as a function of Bi composition, and (4) the potential of $GaBi_xAs_{1-x}$ alloys for the development of highly efficient and temperature-stable 1.55-μm lasers. We also review direct comparisons of theoretical calculations of the (1) threshold characteristics, (2) spontaneous emission (SE), and (3) optical gain, to experimental measurements performed on first-generation $GaBi_xAs_{1-x}$ QW laser devices [15,39].

While the majority of research on dilute bismides has to date centered on $GaBi_xAs_{1-x}$ alloys and laser applications, the potential of alternative Bi-containing alloys for a range of practical applications has begun to drive an increasing diversity of research into related dilute bismide semiconductors. It is to these emerging directions in dilute bismide research that we turn our attention in Section 10.5. We discuss the significant potential offered by the quaternary alloys $In_yGa_{1-y}Bi_xAs_{1-x}$ and $GaBi_xN_yAs_{1-x-y}$—for respective applications in InP-based mid-infrared photonics and multijunction solar cells—due to the broad flexibility of their band structures, associated with the ability to engineer the band gap, spin-orbit-splitting energy, strain, and band offsets over wide ranges [12,13,40–45]. We also review the recent development of $GaBi_xAs_{1-x}/GaN_yAs_{1-y}$ type-II QWs, and demonstrate that this novel class of GaAs-based, strain-balanced heterostructures has great potential to deliver emission/absorption across a broad range of near- and mid-infrared wavelengths on, making them of interest for applications in photonics and photovoltaics [46]. Finally, we provide a brief overview of emerging interest in narrow-gap Bi-containing alloys, which are generating strong interest for applications in mid-infrared photonics, as well as in spintronics.

Overall, this chapter should equip the reader with a comprehensive overview of the theory and simulation of dilute bismide alloys and provide an up-to-date overview of ongoing research on this emerging class of semiconductor materials.

10.2 Electronic and Optical Properties of $GaBi_xAs_{1-x}$ Alloys: Atomistic Theory

In this section, we discuss some of the general properties of dilute bismide alloys, the understanding of which has been strongly informed by theoretical interpretations of experimental measurements. Since Bi acts as an isovalent impurity in (In)GaAs(P) and strongly perturbs the electronic structure, alloy disorder, which is inevitably present in real materials, plays an important role in determining the material properties. This, combined with the associated breakdown of the virtual crystal approximation in Bi-containing alloys, means that atomistic calculations are required in order to provide a quantitative understanding of the impact of Bi incorporation on the electronic structure. As such, we present our discussion in this section within the context of atomistic electronic structure calculations. In Section 10.2.1, we trace how such analyses have contributed strongly to developing the current understanding of this unusual class of semiconductor alloys. Next, in Section 10.2.2, we present and compare to experiment the results of atomistic calculations of a range of material properties for realistic, disordered materials. In Section 10.2.3 we

provide a brief outlook for the atomistic theory of dilute bismide alloys, highlighting the tasks and challenges that must be completed and overcome to improve the fundamental understanding of dilute bismides and related emerging classes of highly mismatched semiconductor alloys.

Since $GaBi_xAs_{1-x}$ is the material which has been the subject of most of the research effort on dilute bismide alloys to date, it is to this material that we devote our attention in this section. However, the understanding of the impact of Bi incorporation on the properties of $GaBi_xAs_{1-x}$ alloys has been found to transfer to a range of related materials, so that the theory and material trends we outline in this section can be understood to apply in a general sense to other Bi-containing III–V semiconductor alloys.

10.2.1 Impact of Bi Incorporation on the Electronic Structure: Atomistic Theory

The first atomistic study of the impact of Bi incorporation on the electronic structure of GaAs was undertaken by Janotti et al. [47]. Using a first-principles approach based on density functional theory (DFT), Janotti et al. analyzed the electronic structure of $GaBi_xAs_{1-x}$ alloys, as well as the impact of co-alloying Bi and N in GaAs to form the $GaBi_xN_yAs_{1-x-y}$ quaternary alloy (cf. Section 10.5.2). In Ref. [47] the first detailed calculations for the III–V compound GaBi were also presented. In addition to calculating structural parameters such as the lattice constant and bulk modulus, Janotti et al. presented a calculation of the band structure of bulk GaBi in the zincblende phase. This calculation confirmed the presence of an extremely large spin-orbit-splitting energy in the VB, which first principles calculations indicate to be >2 eV [48]. Furthermore, an unusual (topological) band ordering was identified at the Γ-point in GaBi. As a result of the extremely large spin-orbit coupling, the Γ_{6c} CB edge states were calculated to lie lower in energy than the Γ_{8v} VB edge states, but higher in energy than the Γ_{7v} SO band edge states. As such, GaBi is predicted to be a metallic compound, with its negative band gap $E_g = E(\Gamma_{6c}) - E(\Gamma_{8v}) \simeq -1.5eV$. While the analysis of Janotti et al. correctly suggested that Bi incorporation leads to a reduction of E_g with increasing x in $GaBi_xAs_{1-x}$, the calculated band gap bowing was approximately one-sixth of that calculated for GaN_xAs_{1-x} using the same approach. However, experimental measurements and subsequent theoretical analyses have revealed that replacing 1% of the As atoms in GaAs by Bi (N) causes a reduction of E_g of up to 90 (150) meV, meaning that while the impact of Bi incorporation is not as strong as that of N incorporation, the effects are broadly comparable in magnitude.

The first major advancement in the understanding of the electronic structure of $GaBi_xAs_{1-x}$ came from the work of Zhang et al. [49]. Using first-principles pseudopotential calculations to investigate the electronic structure as a function of Bi composition in a series of ordered $GaBi_xAs_{1-x}$ supercells, Zhang et al. provided a number of key insights into the impact of Bi incorporation on the GaAs electronic structure and also highlighted the differences between Bi and N incorporation in terms of their relative impact on the electronic structure. First, it was calculated that Bi incorporation causes a significant decrease (increase) in E_g (Δ_{SO}). The calculated E_g data for ordered alloy supercells were shown to be in good agreement with experiment for Bi compositions up to $x = 3.125\%$, the largest dilute composition which could be reliably investigated using the small calculational supercells employed. However, no comment on the precise nature of the strong Bi-induced band gap bowing was provided, and no *direct* attempt was made to elucidate its origin. Second, it was demonstrated that a substitutional Bi atom in GaAs forms impurity states which are (1) highly localized about the Bi atomic site, and (2) resonant with the extended states of the GaAs host matrix VB—i.e. lie below the GaAs VB edge in energy. Third, it was demonstrated that application of hydrostatic pressure decreases the energy of the Bi-related resonant states. As such, while the N-related states in the GaN_xAs_{1-x} CB can be pushed into the band gap under the influence of high hydrostatic pressure [50], the Bi-related states in $GaBi_xAs_{1-x}$ are expected to move deeper into the VB with increasing pressure.

The presence of spin-orbit coupling splits the resonant states associated with an isolated Bi impurity into a doublet (lying close in energy to the top of the GaAs VB) and a singlet (lying deep within the VB, due to the large spin-orbit coupling). Zhang et al. focused on the higher-energy Bi-related doublet, $|\psi_{Bi}\rangle$, and

investigated the character and evolution of the these states as a function of Bi composition. Following this approach, it was found that the states $|\psi_{Bi}\rangle$ are formed of a linear combination of host matrix VB states—having wave vectors located along the Δ, Λ, and Σ directions in the Brillouin zone—which are folded back to Γ. This suggests that the Bi-related states, which lie energetically within the VB, form at the expense of a host matrix VB state. This is in contrast to the case of N incorporation in GaAs, in which N is understood to contribute an additional state to the CB. It was shown using a series of ordered $GaBi_xAs_{1-x}$ supercells containing a single substitutional Bi impurity that the calculated energy E_{Bi} of the doublet states $|\psi_{Bi}\rangle$ increases strongly with increasing supercell size (decreasing Bi composition), with the trend stabilizing in large supercells containing >2000 atoms where the impurity limit is approached. A significant number of host matrix states with large wave vectors are then folded back to Γ in these large supercells, and are hence available to contribute to the supercell representation of the states $|\psi_{Bi}\rangle$. As a result, a more accurate representation of the impact of Bi incorporation is obtained in larger supercells. In the impurity limit, Zhang et al. estimated that an isolated substitutional Bi impurity produces resonant doublet states that lie approximately 80 meV below the GaAs VB edge in energy—i.e. lying directly between the GaAs VB and SO band edge states in energy.

While this analysis has played a significant role in informing the current understanding of the fundamental aspects of the $GaBi_xAs_{1-x}$ electronic structure, the charge-patching method used for larger supercells by Zhang et al. is restricted in the size and type of alloys to which it can be applied. For example, such an approach is not well-suited to analyze the properties of randomly disordered alloys at low composition, since it relies on patching the converged charge densities from small supercells into larger supercells: finite-size effects associated with the periodic boundary conditions employed in the underlying small supercell calculation(s) are implicitly retained, which introduces some degree of spurious Bi-Bi interactions in the large supercell calculation(s). In order to overcome these limitations Usman et al. [5] developed a semi-empirical atomistic TB model, based on the sp^3s^* basis set originally introduced by Vogl et al. [51]. Since it employs a localized basis of atomic orbitals, the TB method is well-suited to study the impact of localized impurities on the electronic structure (as we highlighted and reviewed in our discussion of dilute nitride alloys in Chapter 9). Compared to related atomistic approaches, the TB method is computationally cheap and can be used to efficiently study the electronic structure of large alloy supercells containing upward of thousands of atoms. As described in Chapter 9, the TB method has previously been applied with much success to the study of GaN_xAs_{1-x} alloys, where it has provided quantitative predictions of several important material properties in real, disordered alloys (see, e.g., Refs. [50], [52] and [53]).

Using the TB method, in Ref. [5], Usman et al. significantly broadened the scope of the analysis undertaken by Zhang et al. to include detailed investigation of (1) coupling between localized Bi-related states and the extended states of the host matrix VB, (2) the nature of the localized impurity states associated with the formation of pairs and larger clusters of Bi atoms in a realistic, disordered alloy, (3) the impact of alloy disorder on the overall electronic structure, (4) trends in the electronic structure across an extended range of Bi compositions, and (5) the nature of the VB structure of $GaBi_xP_{1-x}$ alloys. The TB method has been highly successful in elucidating the impact of Bi on the electronic and optical properties of $GaBi_xAs_{1-x}$ alloys. For example, the calculated evolution of the band gap, spin-orbit-splitting energy, and electron effective g factor [5,29] are in quantitative agreement with experiment across the full composition range for which experimental data are available (cf. Section 10.2.2). Further investigations based on the TB method have also quantified the origin and consequences of the strong band gap bowing in Bi-containing alloys [42,54], as well as the effect of intrinsic alloy disorder on the optical properties of bulk and QW $GaBi_xAs_{1-x}$ [27,28].

Before describing the TB analysis, it is worthy of note that the prediction by Zhang et al. that Bi forms a resonant state in GaAs, energetically within the VB, is contradictory to the conclusions of Janotti et al. in Ref. [47], whose analysis suggested that a substitutional Bi impurity in GaAs forms a bound state lying approximately 100 meV above the GaAs VB edge in energy. This result of the analysis of Janotti et al. is surprising. It has been known since the 1960s that a substitutional Bi impurity in GaP produces a bound state lying approximately 100 meV above the GaP VB edge in energy [55]. Since the difference in energy

between the 3*p* valence orbitals of P and the 6*p* valence orbitals of Bi is larger than that between the latter and the energy of the 4*p* valence orbitals of As, it is expected that Bi is less likely to form a bound impurity state in GaAs than in GaP, with the energy of the Bi-related impurity states in GaAs expected to be lower than those in GaP (relative to the host matrix VB edge) [5]. That Bi forms a bound state in $GaBi_xAs_{1-x}$ is therefore highly unlikely, given the known energy of the Bi-related states in GaP:Bi and the fact that the corresponding states in GaAs:Bi are expected to be at lower energy based on well-established chemical trends.

It has been proposed [4,54] that a band-anticrossing approach, similar to that for the GaN_xAs_{1-x} CB, can be applied to describe the VB structure in $GaBi_xAs_{1-x}$ alloys. However, this interpretation of the VB structure has generated to a degree of controversy [54,56]. In order to confirm the presence of a Bi-induced VBAC interaction, Usman et al. used the TB method to investigate the validity of this approach for the VB structure of ordered $GaBi_xP_{1-x}$ and $GaBi_xAs_{1-x}$ alloys by directly analyzing the Bi-related doublet states $|\psi_{Bi}\rangle$ [5]. For the case of a substitutional Bi atom in GaP the analysis and interpretation are straightforward, since the Bi-related states lie energetically within the GaP band gap. By explicitly constructing the states $|\psi_{Bi}\rangle$ in a series of ordered $GaBi_xP_{1-x}$ supercells (cf. Section 10.3.1), and by directly computing the strength with which these states couple to the GaP VB edge as a function of Bi composition, it was found that (1) in the impurity limit, an isolated, substitutional Bi impurity in GaP forms bound states which are highly localized about the Bi atomic site and lie approximately 120 meV above the GaP VB edge in energy, and (2) the coupling of the states $|\psi_{Bi}\rangle$ to the GaP VB edge states is precisely of the form $V_{Bi} = \beta\sqrt{x}$, as in GaN_xAs_{1-x} and GaN_xP_{1-x} alloys [5]. The calculated Bi bound state energy for GaP:Bi in the impurity limit is in good agreement with the experimentally measured value of Ref. [55], and the TB analysis explicitly confirms the presence of a VBAC interaction in ordered $GaBi_xP_{1-x}$ alloys [5].

In order to examine the validity of the VBAC model for $GaBi_xAs_{1-x}$, Usman et al. repeated the impurity state analysis described above for $GaBi_xP_{1-x}$ [5,54], and reached broadly similar conclusions to those of Zhang et al. [49]: (1) a substitutional Bi atom in $GaBi_xAs_{1-x}$ forms a set of impurity states that are resonant with the extended states of the GaAs VB, (2) these Bi-related states are highly localized about the Bi atomic site, and (3) the calculated energy E_{Bi} of these Bi-related states varies strongly with supercell size (emphasizing the role of the finite-size effects described above), and converges in the impurity limit, a prediction which is qualitatively the same as, and quantitatively very close to, that of Zhang et al. The difference in the energies E_{Bi} calculated using the first principles pseudopotential and TB methods is likely related in part to the treatment of the supercells: the pseudopotential calculations rely on patching the converged charge density from a 64-atom supercell into larger supercells in order to reach the impurity limit, while the TB method treats all supercells directly. This analysis predicts that Bi forms resonant states in $GaBi_xAs_{1-x}$ which lie approximately 180 meV below the GaAs VB edge in energy in the impurity limit. Usman et al. further found that the coupling between the Bi-related resonant states and the GaAs VB edge states in $GaBi_xAs_{1-x}$ varies with Bi composition as $V_{Bi} = \beta\sqrt{x}$ in ordered alloys, explicitly confirming that the alloy VB edge states can be described in terms of a VBAC interaction. Recent experimental studies have provided direct evidence of the presence of Bi-related resonant states lying energetically within the $GaBi_xAs_{1-x}$ VB [57,58]. These detailed measurements support the general conclusions of the theoretical calculations of Refs. [5], [49] and [54], and provide direct experimental support for this interpretation of the VB structure of $GaBi_xAs_{1-x}$ alloys.

On the basis of DFT calculations, Deng et al. [56] suggested that the VBAC approach breaks down at higher Bi compositions, where alloy disorder effects become important and lead to strong modifications of the VB structure resulting from interactions between neighboring Bi atoms. Detailed analysis of this "band broadening" concept by Virkkala et al. [59] showed that a model Hamiltonian considering Bi–Bi interactions is indeed capable of producing a significant band gap with increasing Bi composition in disordered $GaBi_xAs_{1-x}$ alloys. However, this approach has two significant shortcomings. First, the predicted decrease in the $GaBi_xAs_{1-x}$ band gap is linear in x, meaning that the strong, composition-dependent bowing of the band gap observed in experimental measurements is not captured by the band broadening interpretation of the band structure. Second, the proposed model requires extraction of supercell- and

composition-dependent parameters from DFT calculations in order to define the alloy band gap, which limits predictive power. Furthermore, it has been demonstrated [54,60] that the conclusions of Deng et al. in Ref. [56] were based on a misinterpretation of the VB ordering in the small $GaBi_xAs_{1-x}$ supercells studied. As such, while the strict validity of the VBAC approach must necessarily break down in a disordered alloy, it is noteworthy that the predictions of the model are in quantitative agreement with experimental measurements of E_g and Δ_{SO} in $GaBi_xAs_{1-x}$ across the full range of compositions for which data are available (cf. Section 10.3.2). Detailed investigations suggest that the VBAC model describes well the band edge states in $GaBi_xAs_{1-x}$ alloys, but breaks down for states lying deeper in the VB, particularly in the vicinity of the SO band [29,61]. As we review in Section 10.4, this means that a VBAC approach provides a simple and accurate means by which to describe the properties of semiconductor lasers and related optoelectronic devices, where we are interested only in states lying energetically in the immediate vicinity of the CB and VB edges.

The understanding of the $GaBi_xAs_{1-x}$ electronic structure developed through atomistic calculations and experimental measurements then describes that (1) Bi introduces highly localized impurity states which are resonant with the GaAs VB, (2) the main features of the band edge states, and hence the alloy band structure, can be effectively described in terms of a VBAC interaction between localized Bi-related doublet states (lying between the GaAs VB and SO band edges in energy) and the extended states of the GaAs VB edge, (3) Bi-related singlet states, which lie much deeper (by several eV) within the VB, play a negligible role in determining the details of the electronic structure, and (4) the CB and SO bands in $GaBi_xAs_{1-x}$ are relatively unperturbed by Bi incorporation—their evolution with increasing Bi composition can be described in a conventional manner using, e.g. the virtual crystal approximation. Given the strong impact of alloy disorder on the electronic properties, interpretation of the details of the $GaBi_xAs_{1-x}$ band structure at higher Bi compositions then remains a somewhat open question: detailed and systematic analysis of the relative impact of As–Bi (VBAC-like) and Bi–Bi (band broadening) interactions on the band structure at higher compositions—building upon that already present in the literature—is required to provide definitive insight. However, as we describe throughout this chapter, describing the impact of Bi incorporation via a simple VBAC approach can be used to analyze the VB edge states in (In)$GaBi_xAs_{1-x}$ alloys with a degree of accuracy that is sufficient to describe and predict the optical properties of semiconductor lasers and related devices.

Since our focus in this chapter is on developing theory with a view to the modeling of photonic devices, we have concentrated in this overview on the fundamental features of the dilute bismide electronic structure which are relevant for such applications (primarily the evolution of the band edge energies and character of the band edge eigenstates with Bi composition). We note there have been several additional studies which have either further explored the electronic structure, or which have focused on other aspects of the interesting and unusual properties of dilute bismide alloys. For example, further DFT-based calculations of the electronic structure of $GaBi_xAs_{1-x}$ and related Bi-containing alloys have been carried out by Mbarki et al. [62], and more recently by Polak et al. [63,64]. Calculations of the structural, elastic, and vibrational properties of III-Bi compounds have been undertaken by Ferhat et al. [65,66], and theoretical investigations of the thermo-dynamics and kinetics of the growth of $GaBi_xAs_{1-x}$ alloys have been carried out by Morgan et al. [67,68] and by Punkinnen et al. [69,70]. We direct the reader to the references provided for further information on these and related aspects of the physics and chemistry of Bi-containing compounds and alloys.

10.2.2 Comparison of Atomistic Theory with Experimental Measurements

Having outlined the general features of the electronic structure of $GaBi_xAs_{1-x}$ alloys, we now turn our attention to direct comparisons between theory and experiment for this material system. We reserve detailed discussion of the impact of isolated Bi impurities and quantitative analysis of the VBAC model until Section 10.3.1, and focus here on disordered alloys. The TB method described in Section 10.2.1 provides a particularly useful framework within which to analyze such systems, since it enables direct calculation of

the electronic structure in large supercells containing several thousand atoms, thereby allowing alloy disorder effects *at dilute compositions* to be treated in a realistic manner. Theoretical analysis of the electronic structure of disordered $GaBi_xAs_{1-x}$ alloys and QWs using the TB method—full details of which are beyond the scope of this chapter—can be found in Refs. [5,27–29]. Here, we provide a brief overview of this work, describing some general features of the electronic structure, and comparing the results of TB calculations on disordered alloys directly with experimental measurements of the band gap, spin-orbit-splitting energy, and electron effective g factor.

The electronic structure calculations we review in this section were performed on a series of disordered 4096-atom supercells, which are either free-standing (unstrained) or under compressive pseudomorphic strain. Beginning with a GaAs supercell, a disordered $GaBi_xAs_{1-x}$ supercell was generated by replacing As atoms by Bi atoms at randomly chosen sites on the anion sublattice. In the free-standing supercells, all atomic positions were allowed to relax freely—using a valence force field model based on the Keating potential [71,72]—with no constraints applied to the atoms at the supercell boundaries. For pseudomorphically strained supercells, a macroscopic pseudomorphic strain was imposed by fixing the positions of the atoms on the exterior boundaries and relaxing the positions of the atoms in the interior of the supercell [29]. For the construction of the supercell TB Hamiltonian, the orbital energies at a given atomic site were computed depending on the overall neighbor environment. Local strain effects were taken into account by scaling the interatomic interaction matrix elements based on the relaxed nearest-neighbor bond length d as $\left(\frac{d_0}{d}\right)^{\eta}$, where d_0 is the equilibrium bond length in the equivalent binary compound and η is a dimensionless scaling parameter (the value of which depends on the type and symmetry of the interaction). The bond angle dependence of the interaction matrix elements was represented using the two-center integrals of Slater and Koster [73]. Full details of the TB model are described in Ref. [5].

In order to facilitate the TB analysis of the electronic structure of $GaBi_xAs_{1-x}$ alloys, the fractional Γ character spectrum $G_\Gamma(E)$ serves as a useful tool to probe the nature and origin of the alloy eigenstates. $G_\Gamma(E)$ is calculated in general by projecting a specific choice of host matrix (in this case, GaAs) Γ-point eigenstates onto the full spectrum of zone-center eigenstates for a given alloy supercell. Specifically, using (0) and (1) to respectively denote the unperturbed host matrix and Bi-containing alloy supercell states, $G_\Gamma(E)$ is computed by projecting the unperturbed zone-center eigenstates $\{|\psi_l^{(0)}\rangle\}$ of a $2M$-atom Ga_MAs_M supercell onto the eigenstates $\{|\psi_k^{(1)}\rangle\}$ computed at Γ for a given $Ga_MBi_LAs_{M-L}$ alloy supercell

$$G_\Gamma(E) = \sum_k \sum_{l=1}^{g(E_l)} |\langle \psi_k^{(1)} | \psi_l^{(0)} \rangle|^2 \; T\left(E - E_k\right), \tag{10.1}$$

where $x = \frac{L}{M}$ is the Bi composition, E_k is the energy of the zone-center alloy supercell eigenstate $|\psi_k^{(1)}\rangle$, $g(E_l)$ is the degeneracy of the unperturbed host matrix eigenstate having energy E_l at Γ (so that $g(E_l) = 2$, 4, and 2 for the CB, VB, and SO band edge eigenstates, respectively), and where the "top hat" function $T\left(E - E_k\right)$ is defined so that $G_\Gamma(E_k)$ has a value of unity for a doubly (spin) degenerate host matrix Γ state at energy E_k. As has previously been demonstrated—in Refs. [5], [27] and [29]—computing $G_\Gamma(E)$ for a range of band edge states provides (1) a consistent approach with which to analyze the evolution of the electronic structure in both ordered and disordered alloy supercells, and (2) detailed insight into the nature of the zone-center eigenstates in a Bi-containing alloy, by quantifying their origin in terms of those of the unperturbed host matrix semiconductor.

Figure 10.2 shows the calculated distribution of the Γ character associated with the extended GaAs host matrix heavy-hole (HH), light-hole (LH), and SO band edge states over the full spectrum of zone-center alloy states, for a disordered 4096-atom $Ga_{2048}Bi_{82}As_{1966}$ ($x = 4\%$) supercell under compressive pseudomorphic strain [29]. The top, middle, and bottom panels of Figure 10.2 show, respectively, the $G_\Gamma^{HH}(E_{j,1})$, $G_\Gamma^{LH}(E_{j,1})$ and $G_\Gamma^{SO}(E_{j,1})$ spectra, calculated using the unperturbed GaAs host matrix HH, LH, and SO states

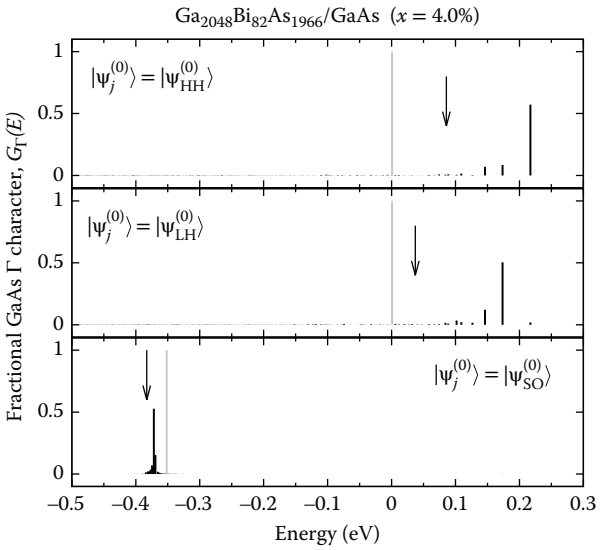

FIGURE 10.2 Black lines: Calculated distribution of the Γ character associated with the GaAs HH (top panel), LH (middle panel), and SO (bottom panel) band edge states over the full spectrum of alloy zone-center valence states in a disordered, 4096-atom $Ga_{2048}Bi_{82}As_{1966}$ ($x = 4\%$) supercell (cf. Equation 10.1). The supercell was placed under compressive pseudomorphic strain corresponding to epitaxial growth on a GaAs substrate, as described in the text. (Adapted from C.A. Broderick et al., *Phys. Rev. B*, 90, 195301, 2014.) Gray lines: Corresponding distribution of the Γ character in a Bi-free $Ga_{2048}As_{2048}$ supercell. The zero of energy is taken at the GaAs VB edge.

$|\psi_{l,0}\rangle = |\psi_{HH,0}\rangle$, $|\psi_{LH,0}\rangle$ and $|\psi_{SO,0}\rangle$ in Equation 10.1. For comparative purposes, the corresponding Γ character of the unperturbed $Ga_{2048}As_{2048}$ host matrix is also shown in Figure 10.2 using gray lines.

Comparing the computed spectra for the host matrix and Bi-containing supercells reveals several important effects of Bi incorporation on the VB structure. First, incorporation of Bi leads to strong hybridization between the GaAs VB edge states and a series of Bi-related states lying at lower energies, within the VB. Second, this hybridization acts to push to the alloy VB edge states upward in energy with increasing Bi composition. In the ordered supercell case, this behavior can be explained directly in terms of a VBAC interaction (cf. Section 10.3.1), and this general behavior is clearly observed in disordered $GaBi_xAs_{1-x}$ alloys across a broad composition range [5]. Third, we note that the Γ character associated with the GaAs HH and LH states is distributed over a large number of VB states in a disordered $GaBi_xAs_{1-x}$ supercell. There are primarily two reasons for this: (1) the localized impurity states associated with substitutional Bi atoms are resonant with a large density of host matrix VB states, which leads to strong energy broadening of the associated quasi-localized resonances (as described for N impurities in Chapter 9) and is reflected by the large number of alloy VB states having small Γ character [54], and (2) the presence of Bi–Bi pairs (in which a single Ga atom has two Bi nearest neighbors) and larger clusters of Bi atoms (having shared Ga nearest neighbors) leads to the formation of localized impurity states which more strongly bind holes—i.e. which lie close to and/or above the GaAs VB edge in energy [5]. This progression of localized Bi-related cluster states leads, in a disordered alloy, to the distribution of a significant fraction of the GaAs Γ character over a range of VB states lying close in energy to the alloy VB edge. This then quantifies quantifying the strong inhomogeneous broadening of the band edge features in $GaBi_xAs_{1-x}$ alloys observed in spectroscopic measurements in terms of the strong impact of Bi-related disorder on the nature of the $GaBi_xAs_{1-x}$ VB [27]. Fourth, in contrast to the character of the alloy VB edge, we note that the feature associated with the Γ character of the GaAs SO band edge states, visible in the bottom panel of Figure 10.2, remains relatively sharp in $GaBi_xAs_{1-x}$. This suggests that the impact of Bi incorporation and alloy disorder on the SO band is less perturbative than it is on the HH- and LH-like states at the VB edge. However, additional

calculations indicate that this distribution of the SO-related Γ character persists as the Bi composition is increased, leading to strong energy broadening and a breakdown of the alloy SO band edge at higher Bi compositions [5]. This suggests a breakdown of a conventional band description of the SO states, which has been confirmed by further analysis [29,61]. Finally, as expected for an epilayer under compressive strain, we see that the highest-energy alloy VB state is predominantly HH-like, having >50% GaAs HH Γ character. We additionally note that both the LH and HH band edges in the alloy are strongly inhomogeneously broadened, which is evidenced by the large number of alloy VB states over which the GaAs VB edge Γ character is distributed [27].

Similar analysis of the conduction states suggests that the impact of Bi incorporation on the CB is generally weak, leading to a downward shift in energy and decrease in the CB-related Γ character that is consistent with alloying effects in conventional semiconductor materials. Overall, these calculated trends confirm that Bi incorporation almost exclusively and strongly perturbs the VB structure, while the CB structure is, by comparison, relatively unaffected [5].

On the basis of these supercell calculations, the variation of the band gap and spin-orbit-splitting energy as a function of Bi composition can be extracted for disordered $GaBi_xAs_{1-x}$ alloys. These values were calculated directly for each supercell on the basis of the calculated $G_\Gamma(E)$ spectra: the band gap was computed as the difference in energy between the (largely GaAs-like) lowest-energy CB state and the alloy VB state having the largest HH or LH GaAs Γ character; the spin–orbit-splitting energy was then computed as the difference in energy between this (typically highest-energy) alloy VB state and a weighted average energy for the alloy SO band (obtained by averaging over the full set of zone-center alloy energies using the value of the SO Γ character as a weight for each contributing eigenstate) [5]. Given the important role played by alloy disorder in determining the details of the VB structure, for each Bi composition considered this analysis was undertaken for five supercells in which the Bi atoms were randomly distributed at different sites on the anion sublattice. The values of E_g and Δ_{SO} were then obtained for that Bi compostion as the average of those calculated using each of the different supercells considered. The results of these calculations are shown in Figure 10.3a, which compares the calculated variation of E_g and Δ_{SO} as a function of Bi composition x in $GaBi_xAs_{1-x}$ to a range of experimental measurements. We note that the theoretical calculations are in excellent agreement with the experimental measurements, across the full composition range for which data are available. In particular, the TB model accurately captures the strong, composition-dependent bowing of E_g and Δ_{SO} demonstrating—in line with the discussion above—that the observed strong decrease (increase) of E_g (Δ_{SO}) with increasing x is primarily attributable to the impact of Bi incorporation on the VB structure, and results largely from the upward shift in energy of the alloy VB edge. The TB analysis confirms that this behavior originates due to hybridization of the GaAs VB edge states with a full distribution of Bi-related localized states in a disordered alloy [5,54].

The Landé effective g factor of conduction electrons (g_e^*) is a key parameter in semiconductor materials, describing the response of the electron spins to externally applied magnetic fields, as well as providing useful information regarding the symmetry of the electron and hole states at the Γ-point in the Brillouin zone. Therefore, in addition to providing information pertinent to evaluating the potential of specific materials for spintronic applications, analysis of the electron spin properties and g factor gives insight into the material band structure, providing valuable data for material characterization and modeling. In particular, g_e^* is extremely sensitive to the separation in energy between zone-center crystal eigenstates, as well as hybridization between zone-center eigenstates caused by, e.g. a reduction of symmetry due to strain [75] or quantum confinement [76]. As a result, calculations of g_e^* provide a stringent test of theoretical models of the material band structure [77].

Optical excitation of a semiconductor material with circularly polarized light is known to generate a spin-polarized population of electrons in the CB. These electrons can then relax by recombining with holes in the VB and emitting photons which must be circularly polarized in order to conserve angular momentum, with the nature of the circular polarization determined by the orientation of the electron spin. As such, Larmor precession of electron spins in the presence of an externally applied magnetic field leads to beating of the circular polarization of the photoluminescence (PL) generated by exciting a material sample with

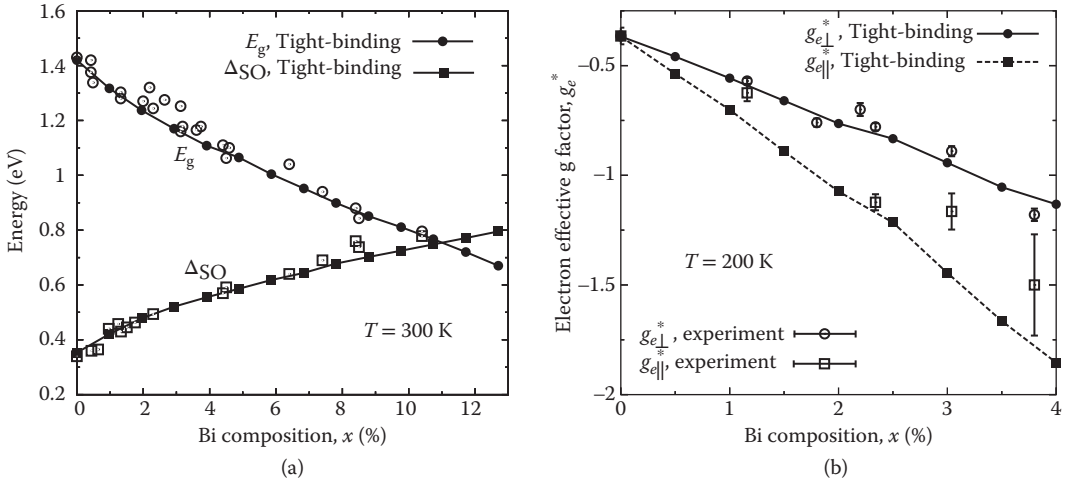

FIGURE 10.3 (a) Comparison between theory and experiment for the variation of the band gap (E_g) and spin–orbit-splitting energy (Δ_{SO}) as a function of Bi composition x in $GaBi_xAs_{1-x}$ at 300 K. Experimental measurements of E_g are depicted by open circles. (The experimental data are taken from S. Tixier et al., *Appl. Phys. Lett*., 82, 2245, 2003; S. Francoeur et al., *Appl. Phys. Lett*., 82, 3874, 2003; W. Huang et al., *J. Appl. Phys*., 98, 053505, 2005; K. Alberi et al., *Phys. Rev. B*, 75, 045203, 2007; Z. Batool et al., *J. Appl. Phys*., 111, 113108, 2012; Yoshida et al., *Jpn. J. Appl. Phys*., 42, 371, 2003.) Experimental measurements of Δ_{SO} are depicted by open squares. (The experimental data are taken from K. Alberi et al., *Phys. Rev. B*, 75, 045203, 2007; Z. Batool et al., *J. Appl. Phys*., 111, 113108, 2012; B. Fluegel et al., *Phys. Rev. Lett*., 97, 067205, 2006.). Closed circles (squares) show the variation of E_g (Δ_{SO}) calculated using the TB model for a series of disordered, 4096-atom supercells. (Adapted from M. Usman et al., *Phys. Rev. B*, 84, 245202, 2011; C.A. Broderick et al., *Semicond. Sci. Technol*., 28, 125025, 2013.). (b) Comparison between theory and experiment for the variation of the transverse and longitudinal components $g^*_{\perp,e}$ and $g^*_{\parallel,e}$ of the electron effective g factor g^*_e as a function of Bi composition x in $GaBi_xAs_{1-x}$ at 200 K. Closed circles (squares) show the variation of $g^*_{\perp,e}$ ($g^*_{\parallel,e}$) calculated using the TB model for a series of disordered, 4096-atom supercells placed under pseudomorphic strain as described in the text. Open circles (squares) show the variation of $g^*_{\perp,e}$ ($g^*_{\parallel,e}$) measured for a series of $GaBi_xAs_{1-x}$/GaAs epitaxial layers using spin quantum beat spectroscopy. (Adapted from Broderick, C.A. et al., *Phys. Rev. B*, 90, 195301, 2014.)

circularly polarized light [78]. Using time-resolved PL spectroscopy, the Larmor frequency $\omega = \frac{g^*_e \mu_B}{\hbar} B$ of these "quantum beats" can be determined, allowing the relevant component of g^*_e to be determined directly for a given relative orientation of the material sample and externally applied magnetic field [76]. As a result of the long coherence times associated with the beating of the PL polarization, this experimental technique, known as spin quantum beat spectroscopy, enables highly accurate determination of the electron Larmor frequency, and hence the effective g factor g^*_e [77,78]. In order to investigate in further detail the impact of Bi incorporation on the electronic structure, as well as to provide insight into the impact of Bi incorporation on the spin properties of electrons, Broderick et al. [29] developed a TB-based model for the calculation of g^*_e in disordered $GaBi_xAs_{1-x}$ alloys. Here, we provide a brief summary of the analysis of g^*_e in $GaBi_xAs_{1-x}$, including comparison to experimental data, full details of which can be found in Ref. [29].

The hybridization between zone-center LH and SO states brought about by the tetragonal deformation and associated reduction of symmetry in pseudomorphically strained layers is known to play a key role in determining the anisotropy of g^*_e in conventional semiconductor alloys, leading to distinct transverse and longitudinal components $g^*_{\perp,e}$ and $g^*_{\parallel,e}$ of the g^*_e tensor [75]. Returning our attention to Figure 10.2, we begin by noting that the calculated distribution of GaAs SO character for this typical pseudomorphically strained $GaBi_xAs_{1-x}$ supercell highlights that the strong inhomogeneous broadening of the LH-like alloy VB states brought about by alloy disorder minimizes this expected strain-induced LH–SO hybridization—i.e. there is no notable projection of the GaAs SO band edge character onto LH-like states close to the alloy

VB edge—providing the first hint that the unusual VB structure of Bi-containing alloys is likely to have a strong impact on g_e^*.

In the absence of strain, g_e^* is expected to be isotropic, with $g_{\perp,e}^* = g_{\parallel,e}^*$ [75]. Comparative calculations we have undertaken on free-standing (unstrained) supercells demonstrate strain-induced mixing between the LH and SO states in ordered supercells, enabling us to conclude that the strong inhibition of such mixing observed in Figure 10.2 arises due to the disorder-induced energy broadening of the GaAs LH character over a significant number of alloy VB states. In order to investigate the impact of Bi incorporation, pseudomorphic strain, and Bi-related alloy disorder on the isotropy of g_e^*, we have therefore first calculated $g_{\perp,e}^*$ and $g_{\parallel,e}^*$ as a function of Bi composition x in a series of free-standing supercells. These calculations show that (1) Bi incorporation leads to a rapid, monotonic increase in the magnitude of g_e^* with increasing x, and (2) the reduced local symmetry present in a disordered $GaBi_xAs_{1-x}$ alloy supercell has a negligible effect on the isotropy of g_e^* [29]. The latter result suggests that the anisotropy of g_e^* observed in experimental measurements performed on $GaBi_xAs_{1-x}$ epitaxial layers arises primarily due to the reduction in symmetry brought about by the impact of biaxial stress on the band structure, much like in conventional semiconductor materials, but that the LH–SO mixing to which this behavior is attributed in conventional semiconductor alloys [75] is not a significant factor determining the anisotropy of g_e^* in $GaBi_xAs_{1-x}$ alloys.

In order to determine the anisotropy of g_e^* in real, pseudomorphically strained epitaxial layers, as well as to compare the results of the theoretical calculations to experiment, Broderick et al. repeated this analysis for disordered $GaBi_xAs_{1-x}$ supercells under the application of a macroscopic strain corresponding to pseudomorphic growth on a GaAs substrate [29] (applied as described above). The results of this analysis are summarized in Figure 10.3b, which shows the calculated variation of $g_{\perp,e}^*$ (closed circles) and $g_{\parallel,e}^*$ (closed squares) with x, as well as experimental measurements of $g_{\perp,e}^*$ (open circles) and $g_{\parallel,e}^*$ (open squares) performed on a series of strained $GaBi_xAs_{1-x}$ epitaxial layers at a temperature of 200 K [29].

In the presence of pseudomorphic strain, we see that g_e^* is strongly anisotropic. As in the unstrained case, we note that the magnitudes of both $g_{\perp,e}^*$ and $g_{\parallel,e}^*$ increase strongly and monotonically with increasing x. We also calculate that (1) the magnitude of $g_{\parallel,e}^*$ exceeds that of $g_{\perp,e}^*$ for non-zero x, in agreement with experiment, and (2) the magnitude of $g_{\parallel,e}^*$ is close to the isotropic value of g_e^* calculated in the absence of strain [29], both of which are expected for compressively strained epitaxial layers [75]. Based on these comparative calculations, we conclude that while the overall symmetry of g_e^* is largely independent of the effects of alloy disorder, the impact of pseudomorphic strain is crucial for determining the experimentally observed anisotropy of g_e^* in $GaBi_xAs_{1-x}$ epitaxial layers. We note that the calculated enhancement of $g_{\perp,e}^*$ with increasing x is in excellent agreement with experiment, confirming not only the accuracy of the detailed picture of the $GaBi_xAs_{1-x}$ VB structure provided by the TB model, but also the strong impact of Bi incorporation on the spin properties of electrons in $GaBi_xAs_{1-x}$ alloys.

The effects of alloy disorder on the calculated variation of $g_{\perp,e}^*$ and $g_{\parallel,e}^*$ with x can be understood in more detail by considering the distribution of host matrix Γ character shown in Figure 10.2. In a disordered $GaBi_xAs_{1-x}$ alloy, the VB edge GaAs Γ character is distributed over a large number of alloy VB states, with the details of the distribution determined by the precise short-range disorder present in the alloy. The details of this distribution of the GaAs Γ character onto VB states lying below the alloy VB edge in energy plays a crucial role in determining the calculated variation of $g_{\perp,e}^*$ and $g_{\parallel,e}^*$ with x, since the contribution from a given alloy VB state to a specific component of g_e^* is weighted by its associated host matrix Γ character [29]. Specifically, mixing of GaAs VB edge Γ character into alloy VB states with greater separation in energy from the alloy CB edge results in contributions to g_e^* which act against the much stronger increase in the magnitudes of $g_{\perp,e}^*$ and $g_{\parallel,e}^*$ than would be expected based solely upon the measured variation of E_g and Δ_{SO} with x [24,29]. We recall from Section 10.2.1 that this breakdown of the simple VBAC picture of the VB structure in the vicinity of the SO band [29] suggests that detailed approaches are typically required to understand material properties involving eigenstates that are energetically remote from the CB and VB edges.

Overall, we conclude that the TB model developed in Ref. [5] provides a detailed and quantitative understanding of the electronic structure of $GaBi_xAs_{1-x}$ and related alloys. The variation of the band gap, spin–orbit-splitting energy, and electron effective g factor as a function of Bi composition calculated using the TB method are in good, quantitative agreement with experiment, and this theoretical approach has delivered significant insight into the impact of Bi incorporation on the $GaBi_xAs_{1-x}$ band structure.

10.2.3 Outlook for Atomistic Theory

Despite the detailed insight atomistic calculations have provided into the impact of Bi incorporation on the electronic and optical properties of $GaBi_xAs_{1-x}$ and related alloys, there is still significant potential to develop and apply atomistic methods and models to further analyze a range of fundamental material properties in this emerging class of III-V semiconductors.

For example, based on the observation that the VBAC picture of the band structure—which very accurately describes the evolution of E_g and Δ_{SO} with x (cf. Section 10.3), as well as the optical properties of dilute bismide semiconductor lasers (cf. Section 10.4)—breaks down for states energetically remote from the VB edge, it is possible to conclude that simple, phenomenological approaches are ill equipped to account for the impact of Bi incorporation and alloy disorder on a number of technologically relevant material properties. This is perhaps unsurprising, since the VBAC model treats only the main features of the band structure and does so in a phenomenological manner. However, this conclusion has significant implications for further theoretical analysis of dilute bismide alloys, implying, e.g. that simple models of the band structure are ill equipped to quantitatively describe key processes such as Auger recombination and IVBA, which involve a broad energy range of alloy VB states including, in particular, states lying close to the SO band edge in energy. As was described for dilute nitride alloys in Chapter 9, detailed and specialized theoretical approaches must therefore be developed in order to obtain a quantitative understanding of complicated processes and properties of this nature.

There have to date been no *direct* theoretical investigations on a fundamental level of the impact of Bi incorporation on a number of important material properties and processes including, e.g. (1) carrier lifetimes, (2) radiative and non-radiative recombination rates, (3) carrier transport, and (4) impact ionization. Detailed understandings of these and related properties must be developed in order to inform the development of Bi-containing alloys for applications in a range of novel semiconductor devices.

Previous theoretical analysis has highlighted the strong influence of Bi-related alloy disorder in determining the details of the electronic structure, meaning that further quantitative theoretical analysis of the properties of Bi-containing alloys is likely to require detailed atomistic treatments. This requirement presents the theoretical community with a series of interesting challenges, centered on developing the capability to describe—on an atomic scale—the impact of short-range disorder on a range of fundamental properties and processes in highly mismatched alloys, where dilute concentrations of impurities strongly perturb the electronic structure in a manner that is not compatible with many commonly employed approximations due to the prominence of alloy disorder effects in determining the material properties.

On this basis, further theoretical research on dilute bismides—perhaps informed by related research on dilute nitride alloys—is likely to deliver not only an improved understanding of this interesting class of semiconductors, but also insights and methods that can be applied to a broad range of emerging highly mismatched materials.

10.3 Modeling Dilute Bismide Band Structure: The k·p Method

Having presented an overview of atomistic electronic structure calculations for $GaBi_xAs_{1-x}$ alloys, in this section, we present a (continuum) $\mathbf{k}\cdot\mathbf{p}$ model of the (In)$GaBi_xAs_{1-x}$ band structure which is well suited to the simulation of the electronic and optical properties of dilute bismide materials and devices.

In Section 10.3.1, we outline how a VBAC approach can be used to provide a simple yet effective interpretation of the band structure of dilute bismide alloys and describe how detailed atomistic calculations have provided a means to deliver quantitative predictive power to this phenomenological picture of the electronic properties.

In Section 10.3.2 we review the use of TB supercell calculations to derive an extended basis set 12-band **k·p** Hamiltonian for (In)GaBi$_x$As$_{1-x}$ alloys. This rigorous analysis verifies that localized impurity states associated with isolated substitutional Bi atoms play the dominant role in determining the main features of the band structure close in energy to the VB edge. The result of this is—despite the fact that more sophisticated atomistic approaches are typically required to understand the detailed properties of these materials—a VBAC-based approach is sufficient to provide a understanding of the electronic and optical properties of dilute bismide devices such as QW lasers in which only quantitative optically active states lying close in energy to the CB and VB edges are of significant importance.

In Section 10.3.3 we present a general method—based on the 12-band **k·p** Hamiltonian—for calculating the band offsets in pseudomorphically strained (In)GaBi$_x$As$_{1-x}$ epitaxial layers and QWs, such as those which are attracting significant attention for the development of highly efficient GaAs- and InP-based semiconductor lasers. The theory outlined in this section lays the foundation for our analysis of dilute bismide QW lasers in Section 10.4.

10.3.1 Valence Band-Anticrossing

As described in Chapter 9, the strong, composition-dependent band gap bowing observed in the dilute nitride alloy (In)GaN$_x$As$_{1-x}$ is well explained in terms of a band-anticrossing interaction between two levels: one at energy E_{CB} associated with the extended CB edge states of the (In)GaAs host matrix, and the second at energy E_{N} associated with the highly localized N-related impurity states. In this simple model, the CB edge energy of the N-containing alloy is given by the lower eigenvalue of the two-band Hamiltonian [8]:

$$H = \begin{pmatrix} E_{\mathrm{N}} & V_{\mathrm{Nc}} \\ V_{\mathrm{Nc}} & E_{\mathrm{CB}} \end{pmatrix}, \tag{10.2}$$

where V_{Nc} is the N composition dependent matrix element describing the interaction between E_{CB} and E_{N}, which varies with N composition x as $V_{\mathrm{Nc}} = \beta\sqrt{x}$.

Bismuth, being the heaviest stable group-V element, is significantly larger and more electropositive than As. It is therefore expected that any Bi-related impurity levels in (In)GaAs should lie close to or below the VB edge in energy and, if present, a VBAC interaction will occur between the Bi-related impurity levels and the VB edge states of the host matrix (cf. Section 10.2.1). The presence of such an interaction has been proposed on the basis of experimental measurements [4], in order to account for the observed composition-dependent bowing of the (In)GaBi$_x$As$_{1-x}$ band gap. While detailed analysis of atomistic supercell calculations [5,54,60] directly confirms the validity of this approach to describing the band structure in the case of ordered alloys, we note that the VBAC model introduced in Ref. [4] is in practice strongly limited by parametric ambiguity which removes its predictive capability. For example, one typically has two parameters (the energy of the Bi-related states E_{Bi} and the VBAC coupling strength β) which must be fitted to a single piece of experimental data (the measured alloy band gap at a given Bi composition), meaning that there is insufficient information available with which to provide an unambiguous fit without resorting to limiting assumptions. For example, initial estimation of the energy of the localized states associated with an isolated substitutional Bi impurity in GaAs produced a value $E_{\mathrm{Bi}} = -0.4$ eV, which lies approximately 60 meV below the SO band edge [4]. However, first principles pseudopotential and large-supercell TB calculations confirm that these states lie above the SO band edge in energy (cf. Section 10.2.1), suggesting that simple estimates based on experimental data produce a qualitatively different description of the VB structure. Usman et al. overcame this limitation by using atomistic TB calculations to directly evaluate the

VBAC parameters, thereby producing a quantitative VBAC approach that provides a predictive capability which has been verified by comparison to a range of experimental data [15,39,54,79]. Here, we outline the TB analysis of the VBAC interaction, as a prelude to the derivation of a suitable **k·p** Hamiltonian with which to model the band structure of (In)$GaBi_xAs_{1-x}$ alloys [54] and heterostructures [15,39,79].

To directly analyze the VBAC interaction, we consider the electronic structure of ordered $GaBi_xAs_{1-x}$ and $InBi_xAs_{1-x}$ alloys by inserting a single substitutional Bi atom in a series of cubic $X_MBi_1As_{M-1}$ (X = Ga, In) supercells containing a total of $2M = 8N^3$ atoms, for $2 \leq N \leq 8$. In the VBAC model the Bi-hybridized (In)$GaBi_xAs_{1-x}$ VB edge states are an admixture of the unperturbed GaAs VB edge states and the T_2 symmetric localized states associated with isolated Bi impurities. Writing the alloy VB edge states as a linear combination in this manner, we obtain an expression for the Bi-related (doublet) states that interact with the (In)GaAs VB edge [5,42,54]

$$|\psi_{\mathrm{Bi},i}\rangle = \frac{|\psi_i^{(1)}\rangle - \sum_n |\psi_n^{(0)}\rangle\langle\psi_n^{(0)}|\psi_i^{(1)}\rangle}{\sqrt{1 - \sum_n |\langle\psi_n^{(0)}|\psi_i^{(1)}\rangle|^2}}, \tag{10.3}$$

where the indices n and i run respectively over the four-fold degenerate states of the (In)GaAs and (In)$GaBi_xAs_{1-x}$ VB edge, respectively. The superscripts (0) and (1) refer respectively to the unperturbed host matrix and alloy VB edge eigenstates.

From Equation 10.3 we have calculated the energy of the Bi-related states, $E_{\mathrm{Bi},i} = \langle\psi_{\mathrm{Bi},i}|\widehat{H}(x)|\psi_{\mathrm{Bi},i}\rangle$, as well as the coupling strength $V_{\mathrm{Bi},i}$ of the VBAC interaction between $|\psi_{\mathrm{Bi},i}\rangle$ and the host matrix VB edge states, $V_{\mathrm{Bi},i} = \langle\psi_{\mathrm{Bi},i}|\widehat{H}(x)|\psi_{\mathrm{VB},i}^{(0)}\rangle$, where $\widehat{H}(x)$ is the full TB Hamiltonian for the Bi-containing supercell and $|\psi_{\mathrm{VB},i}^{(0)}\rangle$ is the unperturbed host matrix VB edge state with which $|\psi_{\mathrm{Bi},i}\rangle$ interacts [5]. Doing this shows directly that (1) a substitutional Bi atom in $Ga_MBi_1As_{M-1}$ or in $In_MBi_1As_{M-1}$ forms a set of four-fold degenerate impurity states (at energy $E_{\mathrm{Bi}} = E_{\mathrm{Bi},i}$), which are highly localized about the Bi site, and (2) the strength of the VBAC interaction between the states $|\psi_{\mathrm{Bi},i}\rangle$ and those of the unperturbed host matrix VB edge varies with Bi composition x as $V_{\mathrm{Bi}} = \beta\sqrt{x}$.

Figure 10.4 shows the calculated values of the difference in energy between the Bi-related states and the unperturbed host matrix VB edge, $\Delta E_{\mathrm{Bi}} = E_{\mathrm{Bi}} - E_{\mathrm{VB}}^{(0)}$, as well as the VBAC coupling parameter β, as a function of supercell size (alloy composition) for a series of ordered $X_MBi_1As_{M-1}$ (X = Ga, In) supercells. We calculate a strong variation of ΔE_{Bi} and β with Bi composition at larger x, with the trends stabilizing as x approaches the impurity limit in the larger (>2000 atom) supercells [5]. In line with the first-principles pseudopotential calculations of Zhang et al. [49], we find that this behavior is attributable to the folding of host matrix VB states back to the Γ point as the supercell size is increased, which makes more zone-center supercell states having large **k** components available to construct $|\psi_{\mathrm{Bi},i}\rangle$ in the larger supercells [54]. Therefore, we conclude that one should refer to large supercell calculations when examining the behavior of a substitutional Bi impurity in (In)GaAs alloys. We also note that $\Delta E_{\mathrm{Bi}} < 0$ in both $GaBi_xAs_{1-x}$ and $InBi_xAs_{1-x}$, indicating that a substitutional Bi atom forms *resonant* states in (In)GaAs which lie below the unperturbed host matrix VB edge in energy. This is distinct from the case of a substitutional Bi impurity in GaP, where Bi forms *bound* states that lie energetically with the host matrix band gap [5,55].

10.3.2 12-Band k·p Hamiltonian for Dilute Bismide Alloys

In order to derive an appropriate **k·p** Hamiltonian for (In)$GaBi_xAs_{1-x}$, we began with an eight-band model for the (In)GaAs host matrix [80], which employs a basis of spin-degenerate zone-center Bloch states for the lowest-energy CB, as well as the HH, LH, and SO VBs. By diagonalizing the (In)GaAs TB Hamiltonian at Γ, we obtain these basis states $|u_1\rangle, \ldots, |u_8\rangle$ directly. Then, we use this set of supercell basis states to construct an eight-band parameter set for (In)GaAs directly from the full TB calculations. This gives a

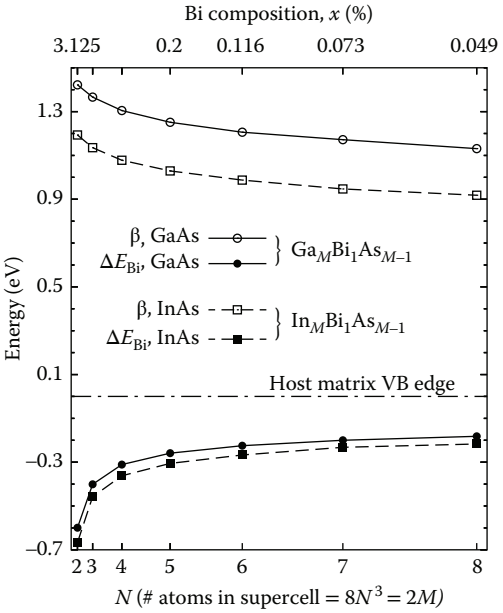

FIGURE 10.4 Calculated difference in energy between the Bi-related localized impurity states and host matrix VB edge, ΔE_{Bi} (closed symbols) and VBAC coupling parameter, β (open symbols) for a substitutional Bi impurity in a series of ordered, cubic $\mathrm{Ga}_M\mathrm{Bi}_1\mathrm{As}_{M-1}$ (circles) and $\mathrm{In}_M\mathrm{Bi}_1\mathrm{As}_{M-1}$ (squares) supercells. (Adapted from C.A. Broderick et al., *Semicond. Sci. Technol.*, 28, 125025, 2013; C.A. Broderick et al., *Phys. Stat. Sol. B*, 250, 773, 2013.) Each supercell contains a total of $2M = 8N^3$ atoms (where N is the number of cubic 8-atom unit cells along each of the x-, y- and z-directions), corresponding to Bi compositions ranging from $x = 0.049\%$ (4096 atoms) to $x = 3.125\%$ (64 atoms). The zero of energy is taken in each case at the host matrix VB edge and is denoted by a dash-dotted line at 0 eV.

parameterization of the eight-band Hamiltonian that reproduces the full TB band structure of an $\mathrm{X}_M\mathrm{As}_M$ (X = Ga, In) supercell with a high degree of accuracy in the vicinity of the zone center [42,54].

Having established this description of the host matrix band structure, we extend the eight-band basis set that describes the host matrix band structure by directly including the four Bi-related localized impurity (doublet) states described by Equation 10.3. The manner in which these Bi-related basis states couple to the host matrix VB edge states is then straightforward to include in the modified **k**·**p** Hamiltonian: the matrix elements between the localized Bi-related states and the extended VB edge states of the host matrix are precisely those evaluated in Section 10.3.1 as $V_{\mathrm{Bi}} = \beta\sqrt{x}$. In this manner, we arrive at a 12-band, extended basis set **k**·**p** Hamiltonian for $(\mathrm{In})\mathrm{GaBi}_x\mathrm{As}_{1-x}$ alloys in which the impact of Bi on the band structure is described primarily via the VBAC interaction. In addition to the VBAC-induced shift in the VB edge energies, there are also conventional (virtual crystal) contributions to each of the band edge energies which depend linearly upon the Bi composition x. We have again used the TB model to directly quantify these effects, by evaluating the matrix elements $\langle u_b|\widehat{H}(x)|u_b\rangle$, where $|u_b\rangle$ are the aforementioned basis states for the host matrix. Following this approach, we find firstly that the Bi-related localized states do not coupled directly to the extended CB edge states of the host matrix. Next, we calculate that the virtual crystal contributions to the variation of the (In)GaAs band edge energies with Bi composition x are given by: $E_{\mathrm{CB}}(x) = E_{\mathrm{CB}}^{(0)} - \alpha x$, $E_{\mathrm{HH}}(x) = E_{\mathrm{LH}}(x) = E_{\mathrm{VB}}^{(0)} + \kappa x$, and $E_{\mathrm{SO}}(x) = E_{\mathrm{SO}}^{(0)} - \gamma x$, for the CB, VB, and SO band edges, respectively (where the superscript (0) denotes the band edge energy of the unperturbed host matrix).

We note that the incorporation of Bi then introduces five band structure parameters in addition to those that describe the (In)GaAs host matrix, all of which are required to provide a reliable description

of the $GaBi_xAs_{1-x}$ band structure using a VBAC-based approach. (As we see in Section 10.3.3, this number increases to seven in the presence of pseudomorphic strain.) These parameters are (1–3) the virtual crystal contributions to the energies of the CB (α), VB (κ), and SO (γ) band edge energies, (4) the energy of the Bi-related localized states relative to the (In)GaAs VB edge (ΔE_{Bi}), and (5) the VBAC coupling strength (β). This number of free parameters is too large to provide an unambiguous fit to the available experimental band structure data, thereby emphasizing the importance of atomistic calculations in informing not only the understanding of the fundamental properties of Bi-containing alloys, but also the development and parametrization of sufficiently detailed phenomenological models to enable the investigation of heterostructures and devices. The 12-band **k·p** Hamiltonian derived by Broderick et al. using this approach is presented in full in Ref. [54]; the Bi-related band structure parameters derived for $GaBi_xAs_{1-x}$ and $InBi_xAs_{1-x}$ alloys on the basis of the TB analysis are listed in Table 10.1. We note that while this 12-band model contains fewer bands than the 14-band model originally introduced by Alberi et al. in Ref. [4], it ultimately provides a more accurate description of the band structure since, in addition to VBAC interactions, the contributions of strain and virtual crystal effects on the composition dependence of the band edge energies have been explicitly taken into account. Additionally, the 12-band model is fully parametrized on the basis of atomistic supercell calculations, without making recourse to fits to experimental data, thereby providing genuine predictive capability.

Figure 10.5a compares the results of calculations of the $Ga_{32}Bi_1As_{31}$ ($x = 3.125\%$) band structure in the vicinity of the zone center using the 12-band **k·p** (solid lines) and full TB (closed circles) Hamiltonians. Note that since the TB calculation is carried out on a 64-atom supercell, the appropriate values of ΔE_{Bi} and β for the supercell must be used in the **k·p** calculation in order to directly compare the band structures calculated using the two methods (cf. Figure 10.4). We observe excellent agreement between the **k·p** and TB band structures. It can be seen that (1) the VBAC interaction has pushed the alloy VB edge upward in energy, and (2) the CB and SO band edge energies in the alloy are well described by the Bi-induced virtual crystal shifts described above. The VBAC interaction, which describes the strong composition-dependent band gap bowing, accounts for the majority of the observed band gap reduction with increasing x. We also note that the presence of the VBAC interaction is verified by the presence of TB bands that correspond to the lower energy Bi-related ($E_-^{LH/HH}$) bands in the **k·p** calculation. These calculations therefore confirm that the 12-band Hamiltonian provides an accurate description the unusual band structure of (In)$GaBi_xAs_{1-x}$ alloys [42,54].

The main features of the band structure calculated using the 12-band Hamiltonian are in excellent agreement with those obtained from TB calculations performed on realistic disordered alloys across a wide range of Bi compositions, despite having been derived and parameterized for ordered alloys [54]. Figure 10.5b shows the variation of E_g (solid black line) and Δ_{SO} (dashed black line) as a function of Bi composition x at 300 K in $GaBi_xAs_{1-x}$, calculated using the 12-band **k·p** model. Comparing the results of these calculations to a range of experimental data from the literature, we see that the 12-band model produces a highly accurate description of the main features of the $GaBi_xAs_{1-x}$ band structure, and does so across a broad composition range. Overall, this analysis suggests that a suitably constructed and parametrized

TABLE 10.1 Bi-related Band Structure Parameters for the 12-band **k·p** Hamiltonian, Computed Directly Using TB Supercell Calculations for $GaBi_xAs_{1-x}$ and $InBi_xAs_{1-x}$ Alloys as Described in the Text

Alloy	ΔE_{Bi} (eV)	β (eV)	α (eV)	κ (eV)	γ (eV)
$GaBi_xAs_{1-x}$	−0.183	1.13	2.82	1.01	0.55
$InBi_xAs_{1-x}$	−0.217	0.92	2.60	1.03	0.02

Source: C.A. Broderick et al., *Semicond. Sci. Technol.*, 28, 125025, 2013.
The parameters ΔE_{Bi} and β specify the nature of the VBAC interaction, while α, κ, and γ, respectively, describe the Bi-induced virtual crystal (conventional alloy) shifts to the CB, VB, and SO band edge energies.

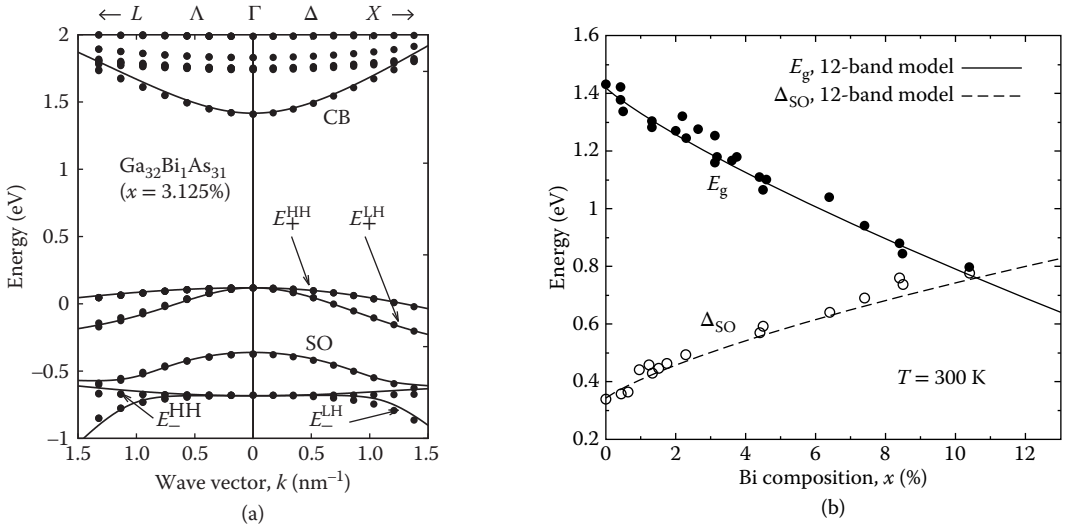

FIGURE 10.5 (a) Calculated band structure of an ordered, cubic $Ga_{32}Bi_1As_{31}$ (x = 3.125%, 64-atom) supercell, along the Λ- and Δ-directions, close to the center of the Brillouin zone, using the TB (closed circles) and 12-band **k·p** (solid lines) Hamiltonians. (Adapted from Broderick, C.A. et al., *Semicond. Sci. Technol.*, 28, 125025, 2013; C.A. Broderick et al., *Phys. Stat. Sol. B*, 250, 773, 2013.) The zero of energy is taken at the host matrix (GaAs) VB edge. (b) Comparison between theory and experiment for the variation of the band gap (E_g) and spin-orbit-splitting energy (Δ_{SO}) as a function of Bi composition x in $GaBi_xAs_{1-x}$ at 300 K. Experimental measurements of E_g are depicted by closed circles. (The experimental data are taken from S. Tixier et al., *Appl. Phys. Lett.*, 82, 2245, 2003; S. Francoeur et al., *Appl. Phys. Lett.*, 82, 3874, 2003; W. Huang et al., *J. Appl. Phys.*, 98, 053505, 2005; K. Alberi et al., *Phys. Rev. B*, 75, 045203, 2007; Z. Batool et al., *J. Appl. Phys.*, 111, 113108, 2012; J. Yoshida et al., *Jpn. J. Appl. Phys.*, 42, 371, 2003.) Experimental measurements of Δ_{SO} are depicted by open circles. (The experimental data are taken from K. Alberi et al., *Phys. Rev. B*, 75, 045203, 2007; Z. Batool et al., *J. Appl. Phys.*, 111, 113108, 2012; B. Fluegel et al., *Phys. Rev. Lett.*, 97, 067205, 2006.) Solid (dashed) lines show the variation of E_g (Δ_{SO}) calculated using the 12-band **k·p** model. (Adapted from C.A. Broderick et al., *Semicond. Sci. Technol.*, 28, 125025, 2013; C.A. Broderick et al., *Phys. Stat. Sol. B*, 250, 773, 2013; C.A. Broderick et al., *IEEE J. Sel. Top. Quantum Electron.*, 21, 1503313, 2015.)

VBAC model can be used to reliably describe the evolution of the main features of the band structure of (In)$GaBi_xAs_{1-x}$ alloys across a wide range of Bi compositions.

10.3.3 Calculation of Band Offsets in Bi-Containing Quantum Wells

Having considered the theoretical description of the bulk band structure of dilute bismide alloys, we turn our attention now to the calculation of the band offsets in pseudomorphically strained $GaBi_xAs_{1-x}$/(Al)GaAs QWs, as a prelude to our discussion of the electronic and optical properties of GaAs-based dilute bismide QW lasers in Section 10.4. Here, we outline the calculation of the band offsets at arbitrary Bi composition x using the 12-band **k·p** model described in Section 10.3.2. Full details of the analysis upon which this analysis of the QW band offsets is based can be found in Ref. [79].

In order to calculate the strained $GaBi_xAs_{1-x}$ band edge energies we consider the 12-band Hamiltonian including the Bir-Pikus matrix elements describing the impact of pseudomorphic strain on the band edge energies at Γ [81]. The steps in the calculation of the band offsets for a $GaBi_xAs_{1-x}$/GaAs QW are illustrated schematically in Figure 10.6. While we focus here on GaAs-based heterostructures, we note that this method for calculating the QW band offsets is generally applicable to alloys whose band structure can be described using the 12-band model, and so can readily be applied to Bi-containing alloys grown

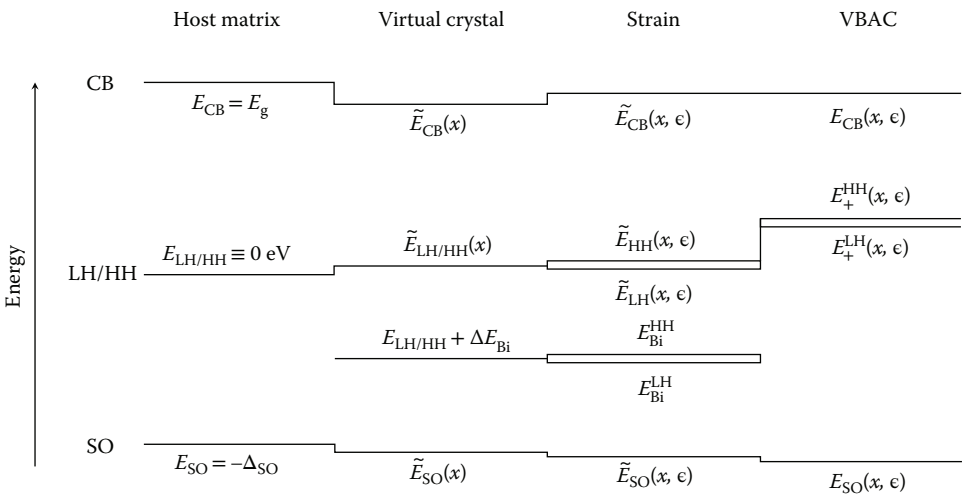

FIGURE 10.6 Schematic illustration of the calculation of the bulk $GaBi_xAs_{1-x}$/GaAs CB, LH, HH, and SO band offsets, using the 12-band **k·p** model of Section 10.3.2 and including the effects of pseudomorphic strain on the band edge energies. (Adapted from Broderick, C.A. et al., *Semicond. Sci. Technol.*, 30, 094009, 2015; G.M.T. Chai et al., *Semicond. Sci. Technol.*, 30, 094015, 2015.) The zero of energy is taken at the VB edge of the unstrained GaAs host matrix. The full details of each step in the calculation are outlined in the text.

using alternative barriers and/or substrates. For barrier materials other than GaAs, the calculation follows straightforwardly from that presented here: one needs simply to calculate the (In)$GaBi_xAs_{1-x}$ band edge energies as outlined below, with the barrier band edge energies then calculated with respect to the (In)GaAs host matrix in the usual way (using, e.g., the model solid theory [82]) in order to determine the band offsets.

The first step in the calculation is shown on the left side of Figure 10.6, in the portion labeled "Host matrix." Here, we begin with a GaAs host matrix (barrier material) and choose the zero of energy to lie at the unstrained VB edge. Second, in the section labeled "Virtual crystal," we include the conventional alloy contributions to the Bi-induced band edge energy shifts for a given Bi composition x as $\widetilde{E}_{CB}(x) = E_{CB} - \alpha x \equiv E_g - \alpha x$, $\widetilde{E}_{LH/HH}(x) = E_{LH/HH} + \kappa x \equiv \kappa x$, and $\widetilde{E}_{SO}(x) = E_{SO} - \gamma x \equiv -\Delta_{SO} - \gamma x$, where, as described above, incorporation of Bi has introduced a set of degenerate resonant impurity levels at energy ΔE_{Bi} relative to the host matrix VB edge.

Third, in the section labeled "Strain," we calculate the strain-induced shifts to the band edge energies, thereby taking account of the fact that an (In)$GaBi_xAs_{1-x}$ epilayer grown on a GaAs substrate will be under compressive pseudomorphic strain. Following the conventions of Krijn [83], we calculate these strain-induced shifts to the CB, HH, LH, and SO band edge energies as $\widetilde{E}_{CB}(x, \epsilon) = \widetilde{E}_{CB}(x) + \delta E_{CB}^{hy}$, $\widetilde{E}_{LH}(x, \epsilon) = \widetilde{E}_{LH}(x) + \delta E_{VB}^{hy} + \delta E_{VB}^{ax}$, $\widetilde{E}_{HH}(x, \epsilon) = \widetilde{E}_{HH}(x) + \delta E_{VB}^{hy} - \delta E_{VB}^{ax}$, and $\widetilde{E}_{SO}(x, \epsilon) = \widetilde{E}_{SO}(x) + \delta E_{VB}^{hy}$, where the energy shifts associated with the hydrostatic and axial components of the strain are $\delta E_{CB}^{hy} = a_c \left(\epsilon_{xx} + \epsilon_{yy} + \epsilon_{zz} \right)$, $\delta E_{VB}^{hy} = a_v \left(\epsilon_{xx} + \epsilon_{yy} + \epsilon_{zz} \right)$, and $\delta E_{VB}^{ax} = -\frac{b}{2} \left(\epsilon_{xx} + \epsilon_{yy} - 2\epsilon_{zz} \right)$ [81]. The non-zero components of the strain tensor given by $\epsilon_{xx} = \epsilon_{yy} = \frac{a_0 - a(x)}{a(x)}$ and $\epsilon_{zz} = -\frac{2C_{12}(x)}{C_{11}(x)}$ [83], where a_0 is the substrate lattice constant, and $a(x)$, $C_{11}(x)$, and $C_{12}(x)$ are the (In)$GaBi_xAs_{1-x}$ lattice and elastic constants.

In order to obtain $a(x)$, $C_{11}(x)$, and $C_{12}(x)$ for (In)$GaBi_xAs_{1-x}$ we interpolate linearly between those of the end-point binary compounds: we use the (Al)GaAs lattice and elastic constants recommended by Vurgaftman et al. [84], and for (In)GaBi we use those calculated *ab initio* by Ferhat and Zaoui [65]. Due to a lack of information in the literature regarding the band edge deformation potentials of (In)GaBi, we take a_c, a_v, and b for (In)$GaBi_xAs_{1-x}$ to be equal to those of the (In)GaAs host matrix, which is expected to be a good approximation at the dilute Bi compositions under consideration.

Due to the *p*-like symmetry of the Bi-related impurity states, there is a lifting of the degeneracy of the LH- and HH-like impurity states in the presence of pseudomorphic strain, resulting from the tetragonal distortion of the crystal lattice. Based on the results of TB calculations on large, ordered supercells, it is found that the energies of these states, $E_{\mathrm{Bi}}^{\mathrm{LH}}$ and $E_{\mathrm{Bi}}^{\mathrm{HH}}$, vary with strain as $E_{\mathrm{Bi}}^{\mathrm{LH}} = \Delta E_{\mathrm{Bi}} + \delta E_{\mathrm{Bi}}^{\mathrm{hy}} + \delta E_{\mathrm{Bi}}^{\mathrm{ax}}$ and $E_{\mathrm{Bi}}^{\mathrm{HH}} = \Delta E_{\mathrm{Bi}} + \delta E_{\mathrm{Bi}}^{\mathrm{hy}} - \delta E_{\mathrm{Bi}}^{\mathrm{ax}}$, with $\delta E_{\mathrm{Bi}}^{\mathrm{hy}}$ and $\delta E_{\mathrm{Bi}}^{\mathrm{ax}}$ given by $\delta E_{\mathrm{Bi}}^{\mathrm{hy}} = a_{\mathrm{Bi}}\left(\epsilon_{xx} + \epsilon_{yy} + \epsilon_{zz}\right)$ and $\delta E_{\mathrm{Bi}}^{\mathrm{ax}} = -\frac{b_{\mathrm{Bi}}}{2}\left(\epsilon_{xx} + \epsilon_{yy} - 2\epsilon_{zz}\right)$. The Bi-impurity state hydrostatic and axial deformation potentials—determined using large supercell TB calculations to track the evolution of the Bi-related localized states as functions of hydrostatic and axial strain—are $a_{\mathrm{Bi}} = -1.11$ eV and $b_{\mathrm{Bi}} = -1.71$ eV, respectively.

Finally, in the section labeled "VBAC," the VBAC interactions between the strained virtual crystal band edges and the Bi-related localized states are included to finally determine the band offsets. Examining the 12-band Hamiltonian [54,79] we see that in bulk (with $k = 0$), the CB is decoupled from the Bi-related states, so that $E_{\mathrm{CB}}(x,\epsilon) = \widetilde{E}_{\mathrm{CB}}(x,\epsilon)$, while the host matrix HH-band couples directly to $E_{\mathrm{Bi}}^{\mathrm{HH}}$, giving

$$E_{\pm}^{\mathrm{HH}}(x,\epsilon) = \frac{\widetilde{E}_{\mathrm{HH}}(x,\epsilon) + E_{\mathrm{Bi}}^{\mathrm{HH}}}{2} \pm \sqrt{\left(\frac{\widetilde{E}_{\mathrm{HH}}(x,\epsilon) - E_{\mathrm{Bi}}^{\mathrm{HH}}}{2}\right)^2 + |V_{\mathrm{Bi}}|^2}\,. \tag{10.4}$$

Since the axial component of the pseudomorphic strain couples the host matrix LH and SO bands [83], there exists a second-order strain-induced coupling between the LH-like Bi-related localized states and the host matrix SO band. As a result, the SO band cannot be treated independently and the LH, SO, and lower-energy LH-like VBAC band edges in the strained alloy are obtained as the eigenvalues $E_{\pm}^{\mathrm{LH}}(x,\epsilon)$ and $E_{\mathrm{SO}}(x,\epsilon)$ of the 3×3 Hermitian matrix

$$\begin{pmatrix} \widetilde{E}_{\mathrm{LH}}(x,\epsilon) & -\sqrt{2}\,\delta E_{\mathrm{VB}}^{\mathrm{ax}} & V_{\mathrm{Bi}} \\ & \widetilde{E}_{\mathrm{SO}}(x,\epsilon) & 0 \\ & & E_{\mathrm{Bi}}^{\mathrm{LH}} \end{pmatrix}, \tag{10.5}$$

which completes the calculation.

On the basis of detailed spectroscopic measurements performed on a series of $GaBi_xAs_{1-x}$/(Al)GaAs QW laser structures with $x \approx 2\%$ [79], the Bi-related parameters α and β of Table 10.1 have been refined in order to produce an optimized set of parameters for modeling *real* $GaBi_xAs_{1-x}$ QWs at low *x*, giving $\alpha = 2.63$ eV and $\beta = 1.45$ eV, with the remaining parameters unchanged from those listed in Table 10.1. These optimized Bi-related band structure parameters have been demonstrated to produce band offsets and transitions energies which are in good, quantitative agreement with those measured for a series of $GaBi_xAs_{1-x}$/(Al)GaAs QWs. This not only verifies the predictive capability of the 12-band model, but also emphasizes its potential for application to the calculation and analysis of the electronic and optical properties of dilute bismide QWs and related heterostructures.

10.4 $GaBi_xAs_{1-x}$/(Al)GaAs Quantum Well Lasers

Having focused so far on constructing a consistent approach to model the (In)$GaBi_xAs_{1-x}$ band structure, in this section we turn our attention to the theory and modeling of dilute bismide semiconductor lasers. In order to address the potential of $GaBi_xAs_{1-x}$ for the development of long-wavelength GaAs-based lasers, we introduce a theoretical model—introduced by Broderick et al. [15]—describe the impact of Bi on the performance of real and ideal QW laser devices. After outlining the theoretical model in Section 10.4.1, in Sections 10.4.2 and 10.4.3 we analyze the effect of Bi incorporation on the electronic and optical properties of an exemplar $GaBi_xAs_{1-x}$/(Al)GaAs QW, which has been chosen on the basis of the active region employed in the first electrically pumped dilute bismide laser [30]. Using the theoretical model we review

the Bi-induced changes to the band structure, DOS, optical transition matrix elements, and optical gain, and show how the choice of barrier material plays a key role in optimizing the performance of the device at low Bi composition.

Next, in Section 10.4.4, we review gain calculations for a $GaBi_xAs_{1-x}$/GaAs laser structure with high Bi composition ($x = 13\%$), designed to emit at 1.55 μm and to have $\Delta_{SO} > E_g$ in order to suppress CHSH Auger recombination and IVBA (cf. Figure 10.1). We highlight that the modal and differential gain at this composition is significantly higher than that at lower x due to the combined effects of increased CB offset and compressive strain at larger x. Finally, in Section 10.4.5 we review comparison of the results of these theoretical calculations directly to experimental measurements of the SE and gain [39]. These are the first measurements of SE and optical gain for a dilute bismide semiconductor laser and the only such comparison between theory and experiment to date for this material system.

The theoretical results reviewed in this section elucidate important trends in the properties and performance of $GaBi_xAs_{1-x}$ QW lasers grown on GaAs substrates, and confirm the potential of dilute bismide QW structures for the development of highly efficient 1.55-μm lasers. Comparison to experimental data displays the predictive capability of the theoretical model, highlighting its potential for use in the design and optimization of future dilute bismide devices.

10.4.1 Theoretical Model for Dilute Bismide Quantum Well Lasers

Here, we provide a brief outline of the theoretical model developed to study the electronic and optical properties of dilute bismide QW lasers. Full details of the model can be found in Ref. [15]. The theoretical model is based upon the 12-band **k·p** Hamiltonian presented in Section 10.3.2. The QW eigenstates are calculated in the envelope function approximation using a numerically efficient plane wave expansion method, similar to that widely employed in first-principles calculations. An overview of the plane wave expansion method within the context of **k·p** theory can be found in Ref. [85].

In the plane wave approach, the QW eigenstate $|\psi_n(k_\parallel, z)\rangle$ at position z and in-plane wave vector $k_\parallel$, having energy $E_n(k_\parallel)$, is written as

$$|\psi_n(k_\parallel, z)\rangle = \frac{1}{\sqrt{L}} \sum_{b=1}^{12} \sum_{m=-M}^{M} a_{nbm}(k_\parallel)\, e^{iG_m z} |u_b\rangle\,, \tag{10.6}$$

where L is the length of the calculational supercell ($-\frac{L}{2} \leq z \leq \frac{L}{2}$), b indexes the bands of the bulk **k·p** Hamiltonian, m indexes the plane waves, and the discrete wave vectors G_m are defined as $G_m = \frac{2m\pi}{L}$. $|u_b\rangle$ are the zone-center Bloch states of the bulk 12-band **k·p** Hamiltonian. We note that this description of the QW eigenstates, which consists of $2M + 1$ independent Fourier components, is formally exact as $M \to \infty$. However, for most practical applications, a basis set containing ~ 70 plane waves is sufficient to ensure that the calculated eigenstates have converged with respect to further increases in the basis set size [86].

By substituting Equation 10.6 into the multi-band Schrödinger equation, it is straightforward to derive analytical expressions for the matrix elements of the QW Hamiltonian in the plane wave basis set defined by the set of wave vectors G_m [85,86], leading to a reciprocal space Hamiltonian matrix of size $12(2M + 1) \times 12(2M + 1)$. In this manner, an independent Hamiltonian matrix is obtained at each in-plane wave vector $k_\parallel$, diagonalization of which yields the QW band structure $E_n(k_\parallel)$ and eigenstates $|\psi_n(k_\parallel, z)\rangle$. We note that the eigenstates defined by Equation 10.6 satisfy periodic boundary conditions over the calculational supercell, meaning that a QW will tend to undergo spurious interactions with identical QWs in neighboring supercells. This can readily be mitigated by choosing a sufficiently large supercell length L, and in practice this feature of the method means that it can be used to efficiently investigate the electronic and optical properties of superlattices.

Analysis of the band offsets (cf. Section 10.3.3) suggests that the small CB offset leads to very weak electron confinement in $GaBi_xAs_{1-x}$/GaAs QWs at low x. As such, hole-induced electrostatic confinement

of electron states is likely to play an important role in determining the electronic and optical properties of $GaBi_xAs_{1-x}$/GaAs QW lasers at low x. For selected laser structures (cf. Section 10.4.3), we therefore also consider self-consistent calculations of the electronic and optical properties, in which the QW band structure and eigenstates are calculated self-consistently by using Poisson's equation in order to determine the net carrier-induced electrostatic potential at each injected carrier density n. For the self-consistent calculations, the coupled multi-band Schrödinger–Poisson system is solved in reciprocal space, using the technique outlined in Ref. [86].

The atomistic supercall calculations reviewed in Section 10.2 clearly demonstrate that Bi-induced hybridization plays an important role in determining the character of the VB edge eigenstates. As such, in order to facilitate accurate calculation of the optical properties of dilute bismide laser structures, it is important to include the impact of Bi incorporation on the QW eigenstates. For this reason, in the theoretical model the QW band structure, envelope functions, and Hamiltonian (calculated in the axial approximation [87]) are used *directly* to compute the optical properties of a given laser structure. Key to this analysis is the calculation of the optical (momentum) transition matrix elements. Following Ref. [88], the transition matrix element describing optical recombination of an electron-hole pair at in-plane wave vector $k_{\|}$ is given in terms of the QW eigenstates described by Equation 10.6 as [88]

$$P_{n_c,n_v}\left(k_{\|}\right) = \frac{m_0}{\hbar}\left\langle \psi_{n_v}\left(k_{\|},z\right)\left|\frac{\partial\widehat{H}}{\partial k_{\|}}\cdot\widehat{e}\right|\psi_{n_c}\left(k_{\|},z\right)\right\rangle, \tag{10.7}$$

where $\widehat{e}$ is a unit vector denoting the polarization of the emitted photon, and n_c (n_v) indexes the conduction (valence) subbands. This approach to calculating the optical matrix elements then explicitly incorporates key band structure effects in the computation of the optical properties including: Bi-induced VB hybridization (VBAC) and localization, temperature- and injection-dependent carrier spillout from the QW(s), band mixing at non-zero in-plane wave vector, and pseudomorphic strain, all of which are necessary to quantitatively describe the SE and optical gain in real devices [15,39]. Note that when discussing optical transition "strengths" below we are referring to the quantity $\frac{1}{2m_0}|P_{n_c,n_v}\left(k_{\|}\right)|^2$, which has units of energy.

Using the optical matrix elements computed in this manner, the calculation of the material gain spectrum $g(\hbar\omega)$ follows the approach of Ref. [89], which consists of transforming the corresponding polarization component of the SE rate $r_{\text{spon}}(\hbar\omega)$ at fixed carrier density n according to

$$g(\hbar\omega) = \frac{3\pi^2\hbar c^2}{2n_r^2\omega^2}\left(1-\exp\left(\frac{\hbar\omega-\Delta F}{k_{\text{B}}T}\right)\right)r_{\text{spon}}(\hbar\omega), \tag{10.8}$$

where $\hbar\omega$ is the photon energy, ΔF is the quasi-Fermi level separation at temperature T in the presence of an injected carrier density n, and n_r is the refractive index of the optical mode in the active region. The SE rate r_{spon} is calculated as described in Refs. [15,89]. Here, we are concerned solely with transverse electric (TE)-polarized gain, which is of larger magnitude than the transverse magnetic (TM)-polarized gain in the compressively strained structures under investigation. As such, r_{spon} is understood to represent only the TE-polarized component of the total SE rate (which gives rise to the leading factor of $\frac{3}{2}$ in Equation 10.8). This approach to calculating the optical gain spectrum has the benefits that (1) it removes the anomalous absorption at energies below the QW band gap that can plague calculations undertaken using an energy and crystal momentum-independent interband relaxation time (homogeneous spectral linewidth), and (2) by definition, $g(\Delta F) = 0$, so that the transparency point on the high-energy side of the gain peak is maintained at the correct, thermodynamically consistent energy. As we will see, when we review the comparison of theory and experiment for the optical gain in Section 10.4.5, this latter characteristic of Equation 10.8 enables a distinct carrier density n in the theoretical calculations to be associated with each current density J at which the gain spectrum is measured, thereby facilitating *direct* comparison of theoretical calculations to experiment.

Based on theory-experiment comparisons, it is found that the spectral broadening of $GaBi_xAs_{1-x}$ QWs is well described using a hyperbolic secant lineshape of the form $S(\hbar\omega) = \frac{1}{\pi\delta}\text{sech}\left(\frac{E_0-\hbar\omega}{\delta}\right)$, where δ is the (homogeneous) spectral linewidth. We note that this lineshape function was previously found to describe well the SE spectra of 1.3-μm GaInNAs dilute nitride QW lasers [90]. For the calculations in Section 10.4.2 we set $\delta = 6.6$ meV, a typical value for III–V semiconductors such as InGaAs. This enables direct comparison between the calculated SE and gain in Bi-free and Bi-containing structures, and hence quantitative analysis on the impact of Bi incorporation on the optical properties. This value of the spectral linewidth is smaller than that observed in real $GaBi_xAs_{1-x}$ devices by a factor of approximately four [15]. Larger more realistic linewidths are included in several of the calculations presented in Sections 10.4.2 and 10.4.3 by inhomogeneously broadening the calculated spectra [15], and the spectral linewidth is directly analyzed in Section 10.4.5 when review the comparison to experimental measurements of the SE and modal gain spectra [39].

For the calculation of the threshold properties in Sections 10.4.2 and 10.4.3, internal optical (cavity) losses of $\alpha_i = 4\ \text{cm}^{-1}$ are assumed, lattice-matched $Al_{0.4}Ga_{0.6}As$ cladding layers are chosen to form the separate confinement heterostructure (SCH), and an overall cavity length of 1 mm is assumed for all of the devices considered [30]. The calculated facet reflectivity for this structure is $R = 0.35$. The confinement factor Γ of the fundamental (TE-polarized) optical mode was calculated for each laser structure using an effective index approach [91]. It was found for all laser structures considered that Γ is maximized for 150-nm-thick (Al)GaAs barriers. Alternative approaches, based on similar models of the band structure, have been proposed by other authors [14,92]. However, the theoretical model presented here benefits from detailed and unambiguous parameterization based on atomistic calculations, and has been confirmed to provide reliable predictions of key trends in the properties and performance of GaAs-based dilute bismide QW lasers [15], including descriptions of the measured SE and optical gain that are in quantitative agreement with experimental measurements on real devices [39].

10.4.2 Impact of Bi Incorporation on the Band Structure and Gain of Quantum Wells

In order to quantify the impact of Bi incorporation on the electronic and optical properties of laser structures having low Bi compositions, we compare the calculated gain characteristics of a $GaBi_xAs_{1-x}$ QW laser structure at low x with those of an equivalent laser structure containing a GaAs (Bi-free, $x = 0$) QW [15]. We focus here on Bi-containing laser structures having a single 6.9-nm-thick $GaBi_{0.021}As_{0.979}$ ($x = 2.1\%$) QW with $Al_{0.144}Ga_{0.856}As$ barriers [30,79].

The calculated CB and VB offsets are $\Delta E_{CB} = 150$ meV and $\Delta E_{HH/LH} = 65$ meV in the unstrained Bi-free QW [79]. The incorporation of 2.1% Bi introduces a compressive in-plane strain of 0.25% and, when VBAC effects are taken into account, this leads to calculated HH and LH band offsets $\Delta E_{HH} = 224$ meV and $\Delta E_{LH} = 208$ meV, and a CB offset $\Delta E_{CB} = 185$ meV [15,79]. The transition energy between the lowest-energy bound electron state ($e1$) and the highest-energy bound hole state ($hh1$) is calculated to be 1.480 eV in the Bi-free QW, which is reduced to 1.289 eV in the $x = 2.1\%$ QW. Of this 191 meV reduction in the QW band gap, only 31 meV is due to the Bi-induced reduction of the CB edge energy, while the remaining 160 meV results from the VBAC-induced upward shift of the VB edge energy.

Figure 10.7 shows the calculated VB structure and DOS for the $x = 0$ and 2.1% QWs. We focus our analysis of the band structure and DOS on the VB, since Bi incorporation has a comparatively minor impact on the CB (cf. Section 10.2). The zero of energy is taken at the energy of the $hh1$ state (i.e., the QW VB edge) in each case, to facilitate a direct comparison of the DOS. Three factors contribute to differences in the band structure and DOS between the two QWs: (1) the compressive strain in the Bi-containing structure leads to a slight increase in the energy separation between $hh1$ and the highest-energy bound LH state ($lh1$), leading in turn to a reduction in the DOS at energies from 0 to −20 meV, (2) the VBAC interaction pushes the VB dispersion upward in energy at fixed $k_{\|}$, increasing the DOS in the $x = 2.1\%$ QW relative

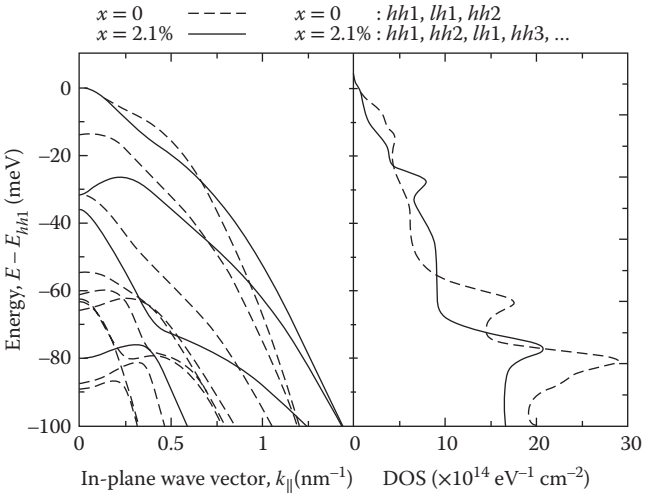

FIGURE 10.7 Calculated band structure (left panel) and density of states (DOS; right panel) in the vicinity of the VB edge for 6.9-nm-thick GaAs (Bi-free; dashed lines) and $GaBi_{0.021}As_{0.979}$ ($x = 2.1\%$; solid lines) QWs with $Al_{0.144}Ga_{0.856}As$ barriers. (Adapted from Broderick, C.A. et al., *IEEE J. Sel. Top. Quantum Electron.*, 21, 1503313, 2015.) Note that the zero of energy has been taken in each case at the energy of the highest bound hole state (*hh*1) to facilitate comparison between the two QWs.

to the Bi-free QW for energies between −20 and −60 meV, and (3) the barrier VB edge lies approximately 60 meV (200 meV) below *hh*1 in energy in the $x = 0$ ($x = 2.1\%$) QW, so that the calculated DOS in the Bi-free structure increases above that of the $x = 2.1\%$ QW for energies between −60 and −100 meV due to the large density of VB states in the AlGaAs barrier. We note that, overall, the calculated impact of Bi incorporation on the VB structure and DOS are relatively minor at low Bi compositions $x \approx 2\%$.

Having considered the impact of Bi incorporation on the QW band structure and DOS, we now turn our attention to the impact of strain and Bi-induced VBAC on the interband optical transition strengths, and then investigate the effect of Bi incorporation on the material gain. Figure 10.8a shows the TE-polarized optical transitions strengths, calculated as a function of $k_{\|}$ for the *e*1-*hh*1 and *e*1-*lh*1 transitions, for the Bi-free (dashed lines) and Bi-containing (solid lines) QWs. Incorporation of Bi reduces the optical transition strength at $k_{\|} = 0$ for transitions between the *e*1 and *hh*1/*lh*1 states. For example, the *e*1–*hh*1 transition strength at $k_{\|} = 0$ in the Bi-containing QW is approximately 67% of that in the Bi-free QW. The majority of the reduction in the optical transition strengths at $x = 2.1\%$ is due to the VBAC effect, which accounts for all but 1.1% of the calculated 33.4% reduction for the TE-polarized *e*1–*hh*1 transition in the $x = 2.1\%$ QW. This is due to the fact that the hybridization of the bound hole states in the QW with Bi-related localized states leads to a portion of the hole states consisting of an admixture of *p*-like Bi-related localized states which do not couple optically to the *s*-like states at the CB edge [54] (cf. Section 10.3.2).

Next, we investigate the variation of the peak material gain as a function of carrier density *n*, for temperatures $T = 100$, 200, and 300 K. The results of these calculations are shown in Figure 10.8b for the $x = 2.1\%$ and $x = 0$ laser structures using solid and dashed lines, respectively. Since the material gain at fixed *n* is inversely proportional to the photon energy, to facilitate a direct comparison of the gain characteristics of the two QW structures—which have different band gaps—all CB states in the $x = 0$ calculation have been shifted downward in energy by 191 meV in order to bring the *e*1–*hh*1 transition energies into coincidence for both QW structures. Therefore, the differences between the material gain for two laser structures shown in Figure 10.8b arise only due to the impact of Bi incorporation on the QW band structure and optical transition strengths. Based on a calculated optical confinement factor $\Gamma = 1.66\%$ for the

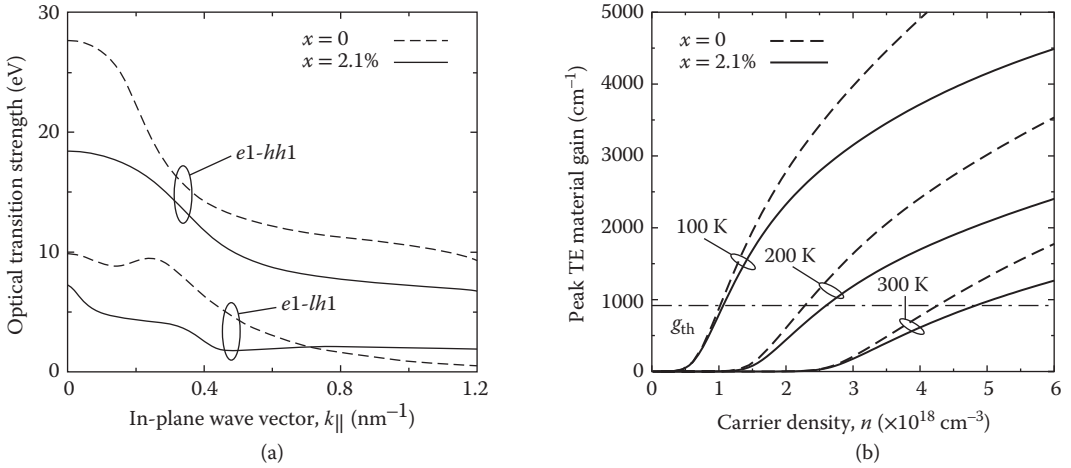

FIGURE 10.8 (a) Calculated TE-polarized optical transition strengths as a function of in-plane wave vector ($k_{\|}$), for optical transitions between the lowest-energy bound electron and highest-energy bound LH- and HH-like states for 6.9-nm-thick GaAs (Bi-free; dashed lines) and $GaBi_{0.021}As_{0.979}$ ($x = 2.1\%$; solid lines) QWs having $Al_{0.144}Ga_{0.856}As$ barriers. (Adapted from Broderick, C.A. et al., *IEEE J. Sel. Top. Quantum Electron.*, 21, 1503313, 2015.) (b) Calculated variation of the peak TE-polarized material gain as a function of carrier density, at temperatures of 100, 200, and 300 K for the same QWs as in (a). (Adapted from Broderick, C.A. et al., *IEEE J. Sel. Top. Quantum Electron.*, 21, 1503313, 2015.) The horizontal dash-dotted line denotes the calculated threshold material gain, $g_{th} = 919$ cm^{-1}.

$x = 2.1\%$ laser structure, we estimate a threshold material gain $g_{th} = 919$ cm^{-1} (denoted by a horizontal dash-dotted line in Figure 10.8b).

There is little change in the calculated transparency carrier density, n_{tr}, between the $x = 0$ and 2.1% laser structures for all temperatures considered. This reflects the small difference between the calculated VB edge DOS for the two QWs (cf. Figure 10.7). However, we find for $n > n_{tr}$ that (1) the gain saturates at a lower level with increasing n at $x = 2.1\%$, and (2) the threshold carrier density, n_{th}, is higher at $x = 2.1\%$ than in the Bi-free QW. Both of these features arise due primarily to the general reduction in the optical transition strengths in the Bi-containing QW structure. We also note that while the differential gain at threshold, $\frac{dg}{dn}$, is approximately equal for both the Bi-free and Bi-containing lasers structures at 100 K, $\frac{dg}{dn}$ for the Bi-containing QW decreases with respect to that of the Bi-free QW with increasing temperature. Since the modulation response frequency of a semiconductor laser is proportional to the square root of the differential gain [93,94], these calculations suggest that Bi incorporation can be expected to lead to a degradation of the dynamical performance at low Bi compositions compared to that of an equivalent Bi-free structure.

10.4.3 Optimization of $GaBi_xAs_{1-x}$/(Al)GaAs Laser Structures

Theoretical and experimental analysis has shown that $GaBi_xAs_{1-x}$/GaAs QWs have very small CB offsets at low x, estimated at only 55 meV for a QW with $x = 2.1\%$ [79]. Such a low CB offset is detrimental to laser operation, leading to low electron-hole spatial overlap and significant electron leakage from the QW at high temperatures and carrier densities, and is therefore a major factor limiting the material gain at low x. Electron confinement must then be improved by growing $GaBi_xAs_{1-x}$ QWs with $Al_yGa_{1-y}As$ barriers to increase the CB offset, but at the cost of reducing the optical confinement in a laser structure with a fixed Al composition y in in the cladding layers. This leads to a trade-off between the material gain (which increases with increasing y due to improved electron confinement) and the optical confinement (which decreases with increasing y due to a reduced refractive index contrast between the barrier and cladding layers). The theoretical model described above has been used to determine the Al composition

y in the barrier layers required to optimize the performance of a laser structure containing the same 6.9-nm-thick $GaBi_{0.021}As_{0.979}$ QW considered in the exemplar structure of Section 10.4.2. The modal gain of the laser structure, Γg, is the quantity optimized, in order to minimize n_{th}. The gain calculations for this optimization include (1) electrostatic effects, computed self-consistently via coupling the Poisson's equation, and (2) inhomogeneous broadening, in order to produce realistic estimates of the material gain in the laser structure [15], an approach which has been shown to be in excellent agreement with experiment [39,95].

Figure 10.9 shows the calculated variation of n_{th} and g_{th} with the barrier Al composition y for this series of laser structures. Examining first the variation of g_{th} with y (closed circles, solid line) we see that the degradation of the optical confinement when Al is incorporated in the barrier layers leads to an increase in g_{th} with increasing y. This increase in g_{th} is modest at low Al compositions $y < 10\%$, with the optical confinement factor at $y = 10\%$ being 90% of its maximum calculated value $\Gamma = 1.95\%$ at $y = 0$. For $y > 10\%$ Γ degrades rapidly with increasing y, reaching values as low as 1.24% at $y = 25\%$. This results in a large increase in g_{th}, since $g_{th} \propto \Gamma^{-1}$. The net effect of this trend on the laser performance can be seen by considering the variation of n_{th} with y (open circles, dashed line). Beginning with a QW having GaAs barriers ($y = 0$), a large threshold carrier density $n_{th} = 6.3 \times 10^{18}$ cm^{-3} is calculated. When Al is incorporated in the barrier layers n_{th} decreases rapidly with increasing y, reaching a minimum value for $10\% < y < 15\%$. For $y < 10\%$, we therefore conclude that the increase in the material gain at fixed carrier density arising from the enhanced CB offset and electron confinement is sufficient to overcome the associated reduction in Γ, leading to a strong reduction in n_{th}. However, for $y > 15\%$, we find that n_{th} increases rapidly with increasing y, reflecting that any further improvement in the material gain is insufficient to overcome the rapid further degradation of Γ and associated increase in g_{th}. These theoretical calculations therefore suggest that QWs having barrier Al compositions $y = 10\%$–15% offer enhanced performance as compared to a QW having GaAs barriers, for $x \approx 2\%$.

While the first electrically pumped $GaBi_xAs_{1-x}$ QW laser had a nominal barrier Al composition $y = 20\%$ [30], more recently improved device performance has been demonstrated following the structural optimization outlined here [96]. Specifically, measurements confirm that choosing $Al_{0.12}Ga_{0.88}As$ as opposed to $Al_{0.20}Ga_{0.80}As$ barriers leads to a decrease of the threshold current density, J_{th}, by a factor of

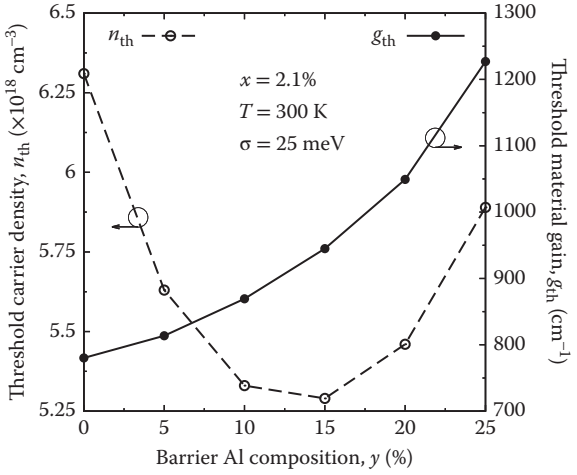

FIGURE 10.9 Calculated variation of the threshold carrier density (n_{th}; open circles, dashed lines) and threshold material gain (g_{th}; closed circles, solid lines) as a function of the barrier Al composition y for a series of 6.9-nm-thick $GaBi_{0.021}As_{0.979}/Al_yGa_{1-y}As$ QWs. (Adapted from Broderick, C.A. et al., *IEEE J. Sel. Top. Quantum Electron.*, 21, 1503313, 2015.) The calculated gain spectra, from which the values of n_{th} were determined, include an inhomogeneous broadening of 25 meV.

TABLE 10.2 Measured Room Temperature Emission Wavelength λ_{meas} and Threshold Current Density J_{th} for a Series of $GaBi_xAs_{1-x}$/$Al_yGa_{1-y}As$ Laser Structures Containing Nominally Identical 6.4-nm-Thick $GaBi_{0.022}As_{0.978}$ ($x = 2.2\%$) QWs

QW Material	Barrier Material	λ_{meas} (nm)	J_{th} (kA cm^{-2})
$GaBi_{0.022}As_{0.978}$	GaAs	938	7.5
$GaBi_{0.022}As_{0.978}$	$Al_{0.12}Ga_{0.88}As$	947	1.0–1.1
$GaBi_{0.022}As_{0.978}$	$Al_{0.20}Ga_{0.80}As$	947	1.5–1.6

Source: Marko et al., *J. Phys. D: Appl. Phys.*, 47, 345103, 2014.
The measured variation of J_{th} with barrier Al composition y closely follows the trend calculated for the threshold carrier density n_{th} (cf. Figure 10.9), verifying that J_{th} can be minimized by varying y to control the trade-off between the carrier and optical confinement.

approximately three. The results of these measurements, which emphasize the importance of AlGaAs barriers for the design and optimization of $GaBi_xAs_{1-x}$ QW lasers at low x, are summarized in Table 10.2. Full details of this optimization analysis can be found in Refs. [15] and [96]. Examining the experimental data in Table 10.2 we can clearly identify the trends predicted by the theoretical calculations. First, the laser structure with Al-free (GaAs) barriers offers poor material gain leading to high J_{th}. Second, incorporation of 12% Al in the barrier (close to the optimum composition suggested by the theoretical calculations) results in a strong reduction of J_{th} as a consequence of the enhanced modal gain. Finally, increasing the Al composition in the barrier from 12% to 20% brings no further improvement in the laser performance; an increase in J_{th} is observed, in line with the degradation of the optical confinement and model gain predicted by the theoretical calculations.

Since the ultimate goal is to grow $GaBi_xAs_{1-x}$ lasers with Bi compositions $x > 10\%$, this analysis has been repeated for a series of $GaBi_xAs_{1-x}$/(Al)GaAs QW laser structures as a function of x. In general, it is found that Al incorporation in the barrier layers is required to enhance the otherwise weak electron confinement for $x < 5\%$. For $x = 6\%$, the $GaBi_xAs_{1-x}$/GaAs CB offset is sufficiently large that incorporating Al in the barrier layers has a negligible impact on the material gain. We therefore conclude that AlGaAs barriers are required to optimize the performance of $GaBi_xAs_{1-x}$ QWs with $x < 6\%$. For higher Bi compositions, GaAs barriers should suffice to provide modal gain which exceeds that calculated for the $x = 2.1\%$ laser structure considered above at fixed n, provided that QWs with $x > 6\%$ can be grown with sufficiently high material, structural, and optical quality. This conclusion has proved useful, particularly for the growth of hybrid laser structures such as those presented in Ref. [35], where the cladding and part of the barrier layers are grown by metal-organic vapor phase epitaxy (MOVPE) and the active region by molecular beam epitaxy (MBE). This analysis suggests that, for sufficiently large x, GaAs can be used as the final layer grown before a laser substructure is transferred in either direction between MOVPE and MBE growth chambers, thereby minimizing the impact of AlGaAs-related degradation during transport and storage of partly grown laser structures.

10.4.4 Gain Characteristics of $GaBi_xAs_{1-x}$ Quantum Wells at $\lambda = 1.55$ μm

Having elucidated the consequences of Bi incorporation for the gain characteristics and performance of real $GaBi_xAs_{1-x}$ QW lasers at low x, we now turn our attention to an ideal 1.55-μm device at higher Bi composition. We again choose a 6.9-nm-thick $GaBi_xAs_{1-x}$ QW, in order to compare the calculated modal gain directly to that calculated for the $x = 2.1\%$ QW of Sections 10.4.2 and 10.4.3, and, in light of the analysis described in Section 10.4.3, we choose GaAs as the barrier material. For this 6.9-nm-wide $GaBi_xAs_{1-x}$/GaAs QW, we calculate that a Bi composition $x = 13\%$ produces an emission wavelength close to 1.55 μm. We note that the bulk band structure of the $GaBi_{0.13}As_{0.87}$ QW material has $\Delta_{SO} > E_g$, so that

it should be possible to use $GaBi_xAs_{1-x}$ alloys to develop, on a GaAs substrate, a 1.55-μm laser in which the CHSH Auger recombination and IVBA processes are suppressed (cf. Figure 10.1) [12,15].

Incorporation of 13% Bi in the QW produces an in-plane compressive strain of 1.53%, compared to 0.25% in a QW with $x = 2.1\%$. We calculate that the tendency of the VBAC interaction to increase the effective masses of the bound hole states in the QW is overcome by the strain-induced reduction of the in-plane hole effective masses [93], leading to a significant reduction of the DOS in the vicinity of the VB edge at $x = 13\%$ [15]. This strong reduction of the DOS at energies close to the QW VB edge should then enhance the performance of this high Bi composition QW compared to the low x laser structures considered above.

Compared to the value $\Gamma = 1.66\%$ calculated for the optical confinement factor of the $x = 2.1\%$ laser structure of Section 10.4.2, we calculate $\Gamma = 1.26\%$ in the $x = 13\%$ QW. The reduction in Γ results from the longer emission wavelength of the $x = 13\%$ structure and leads to an increase of the material gain at threshold g_{th}. We note that the degradation of the optical confinement a long wavelengths can be mitigated by increasing the barrier thickness. We have not pursued this approach here, as our intention is to analyze only the impact of changing the Bi composition in otherwise identical laser structures. Based on the calculated difference in Γ between the two laser structures we calculate $g_{th} = 1208\ cm^{-1}$ at $x = 13\%$, which is 32% larger than the value of $919\ cm^{-1}$ calculated for the $x = 2.1\%$ laser structure. However, the calculated CB offset of 245 meV in the $GaBi_{0.13}As_{0.87}/GaAs$ QW is 60 meV larger than that calculated for the $GaBi_{0.021}As_{0.979}/Al_{0.144}Ga_{0.856}As$ QW above, leading to enhanced electron confinement in the 1.55-μm QW without the need to incorporate Al in the barrier layers. The calculated optical transition strength at $k_{\|} = 0$ for the $e1$–$hh1$ optical transition in the $x = 13\%$ QW is approximately 90% of that calculated at $x = 2.1\%$, suggesting that the Bi-induced decrease of the optical transition strengths described above is most prominent at low x, but has little additional effect as x is increased. When combined with the favorable impact of strain on the VB structure and DOS at larger values of x, these effects together ensure that the material gain at fixed n is significantly improved in the $x = 13\%$ QW compared to the $x = 2.1\%$ laser structure considered above.

The $x = 13\%$ laser structure has a calculated threshold carrier density $n_{th} = 3.42 \times 10^{18}\ cm^{-3}$ at 300 K, which is approximately 60% of that calculated for the $x = 2.1\%$ laser structure. Additionally, for the $x = 13\%$ laser structure $\frac{dg}{dn} = 2.76 \times 10^{-16}\ cm^2$ for the differential (material) gain at threshold, which is approximately a factor of two larger than the value of $1.32 \times 10^{-16}\ cm^2$ calculated for the $x = 2.1\%$ laser structure. This analysis suggests that the marked improvement in the threshold characteristics of the device at $x = 13\%$ results primarily from the increased compressive strain in the QW. First, the strain-induced reduction in the DOS at the VB edge leads to an increase in the quasi-Fermi level separation at fixed n, meaning that population inversion can be achieved at lower carrier densities. Second, the reduced in-plane hole effective masses ensure that holes occupy VB states over a smaller range of $k_{\|}$ (i.e. a narrow $\mathbf{k}$-space distribution) at $x = 13\%$ than at $x = 2.1\%$, better matching the distribution of electrons in the CB and hence ensuring that more carriers are available to contribute to the lasing mode [97]. These factors ensure that a larger increase in material gain can be obtained for a given change in n at $x = 13\%$, leading to the calculated significant improvement in $\frac{dg}{dn}$.

Previous calculations for InP-based QW structures designed to emit at 1.55 μm suggest that the maximum $\frac{dg}{dn}$ obtainable using optimized quaternary InGaAsP QWs are $\sim 14 \times 10^{-16}\ cm^2$ for a multi-QW device containing four QWs [98]. While the calculations here suggest that GaAs-based $GaBi_xAs_{1-x}$ QWs offer slightly lower $\frac{dg}{dn}$ than conventional InP-based devices, we recall that this analysis has been performed on a prototypical 1.55-μm dilute bismide structure in which the waveguide has not been optimized. We therefore conclude that 1.55-μm $GaBi_xAs_{1-x}$ QW lasers have the potential to offer differential gain, and hence dynamical performance, that compares favorably with existing InP-based devices, particularly for structures incorporating multiple $GaBi_xAs_{1-x}$ QWs. We further note that the calculated value of $\frac{dg}{dn}$ for this 1.55-μm dilute bismide device exceeds that calculated for ideal 1.3-μm GaIn(N)As/GaAs QW structures [99].

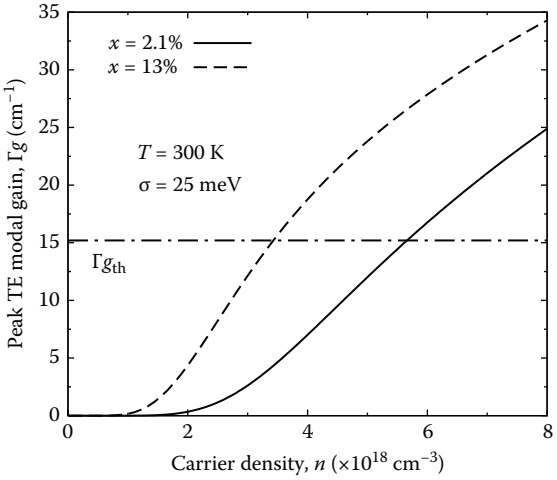

FIGURE 10.10 Calculated variation of the peak TE-polarized modal gain as a function of carrier density at 300 K for the $GaBi_{0.021}As_{0.979}/Al_{0.144}Ga_{0.856}As$ (solid line) and $GaBi_{0.13}As_{0.87}$/GaAs (dashed line) QWs described in the text. (Adapted from C.A. Broderick et al., *IEEE J. Sel. Top. Quantum Electron.*, 21, 1503313, 2015.) The $x = 2.1\%$ structure has a room temperature emission wavelength ~950 nm, while the $x = 13\%$ structure is designed to have $\Delta_{SO} > E_g$ and to emit at 1.55 μm. The horizontal dash-dotted line denotes the calculated threshold modal gain, $\Gamma g_{th} = 15.2\ cm^{-1}$.

Figure 10.10 shows the calculated variation of the modal gain as a function of n for the $x = 2.1\%$ (solid line) and $x = 13\%$ (dashed line) laser structures at 300 K. We see that, even in the presence of significant inhomogeneous broadening of the gain spectrum, the enhancement of the material gain described above is sufficient to overcome the reduction in Γ between $x = 2.1\%$ and 13%, leading to significantly improved modal gain at fixed n. Additionally, we note that the differential modal gain at threshold is significantly improved at $x = 13\%$, due to the fact that the calculated increase in the differential material gain at threshold is also sufficient to overcome the reduction in Γ between $x = 2.1\%$ and 13%. We recall that optimization of the waveguide can be undertaken for long-wavelength emission, which can be expected to further improve the performance of QW structures having high Bi compositions over that of low x structures operating at shorter wavelengths.

We therefore conclude overall that the performance of ideal $GaBi_xAs_{1-x}$ QW lasers at high Bi compositions ($x > 10\%$) should be significantly improved compared to that of shorter wavelength devices at lower x. The calculated improvement in the threshold characteristics with increasing x is primarily attributable to the associated increase in compressive strain in the QW when more Bi is incorporated. We also emphasize that the presence of a QW band structure in which $\Delta_{SO} > E_g$ promises to deliver even greater benefit for the laser operation, due to the elimination of the dominant non-radiative CHSH Auger recombination and IVBA losses.

10.4.5 Theory versus Experiment for First-Generation $GaBi_xAs_{1-x}$/(Al)GaAs Lasers

Having explored general trends in the electronic and optical properties of $GaBi_xAs_{1-x}$/(Al)GaAs QW lasers, we now compare the results of the theoretical calculations directly to experimental measurements of the SE and gain spectra. The experimental measurements were carried out using the segmented contact method, which allows measurement of the optical absorption, SE, and gain spectra of a multi-section device [100]. The multi-section devices used in this study were fabricated from a wafer on which a single QW $GaBi_xAs_{1-x}$/(Al)GaAs laser structure was grown. The laser structure is as described in Ref. [30]

and consists of $Al_{0.4}Ga_{0.6}As$ cladding layers, 150-nm-thick AlGaAs barriers and a $GaBi_xAs_{1-x}$ QW with a nominal thickness of 6.4 nm.

Two multisection devices were fabricated in order to carry out the segmented contact measurements: (1) a device fabricated from material located close to the edge of the wafer ("device 1"), and (2) a device fabricated from material located closer to the center of the wafer ("device 2"). Full details of the growth and fabrication of these devices, as well as the theory-experiment comparison to be outlined here, can be found in Ref. [39]. While the barrier has a nominal Al composition of 20%, based on the analysis of Ref. [79] we find that the actual composition is 14.4%—we use the latter in the theoretical calculations, as in Section 10.4.2.

Comparing the measured optical absorption spectra for the two multisection devices, it is found that (1) the optical (cavity) losses are in the range $\alpha_i = 10$–$15\ cm^{-1}$ at room temperature, and (2) the absorption edge measured for device 2 is red-shifted by approximately 20 nm compared to the that measured for device 1 [39]. The latter suggests the presence of Bi composition fluctuations across the wafer, with the Bi composition being slightly reduced toward the wafer edge. This is consistent with the measurements of Ref. [79], in which $GaBi_xAs_{1-x}$/(Al)GaAs laser structures having nominally identical $x = 2.2\%$ QWs-displayed variations of approximately 30 meV in the QW band gap (corresponding in theoretical calculations Bi compositional variations of up to $\pm$ 0.4%).

In order to investigate this behavior, we first analyze the measured SE spectra for the two devices. Figure 10.11a shows the SE spectra measured at threshold using the segmented contact method for device 1 (closed circles) and device 2 (closed squares). Both spectra have similar overall shape with the main difference between them being the wavelength of the SE peak, which is approximately 20 nm shorter in device 1. This shifted emission peak is consistent with the observed shift in the absorption edge between the two

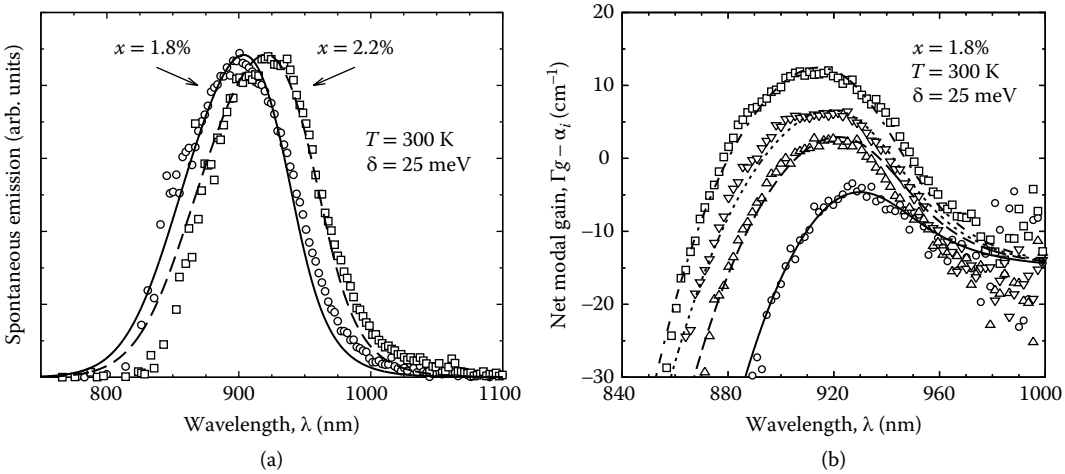

FIGURE 10.11 (a) Measured (using the segmented contact method; open symbols) and calculated (lines) SE spectra at threshold for a multisection device fabricated from material located close to the wafer edge (open circles; solid line) and fabricated from material located close to the center of the wafer (closed squares; dashed line). The ~20 nm shift in the SE peak wavelength between the two devices suggests the presence of Bi compositional variations of up to 0.4% across the wafer from which the devices were fabricated. (Adapted from Marko, I.P. et al., *Sci. Rep.*, 6, 28863, 2016.) (b) Measured (using the segmented contact method; closed symbols) and calculated (lines) net modal gain spectra for the device fabricated from material located at the wafer edge, at injected current densities $J = 0.7$, 1.4, 2.0, and 2.4 kA cm^{-2} (open circles, upright triangles, inverted triangles, and squares, respectively). The carrier densities used in the theoretical calculations (determined as outlined in the text) are $n = 5.12$, 7.11, 8.24, and $9.38 \times 10^{18}\ cm^{-3}$; the corresponding gain spectra are shown using solid, dashed, dotted, and dash-dotted lines, respectively. (Adapted from Marko, I.P. et al., *Sci. Rep.*, 6, 28863, 2016.)

devices. In order to quantify this variation in Bi composition across the wafer, and to facilitate theoretical analysis of the measured gain spectra for device 1, the theoretical model has been used to analyze the SE spectra shown in Figure 10.11a. The careful parameterization of the 12-band $\mathbf{k}\cdot\mathbf{p}$ model for the $GaBi_xAs_{1-x}$ band structure described in Section 10.3 enables us to treat the Bi composition x in the QW to be treated as the only free variable in the analysis of the SE emission peak. Through the theoretical calculations we find that the measured SE peak wavelength of 910 nm for device 1 corresponds to a Bi composition $x = 1.8\%$ in the QW (assuming that the QW thickness to be is fixed at its nominal value of 6.4 nm).

Next, by comparing the full calculated SE spectrum for device 1 (shown in Figure 10.11a using a solid line) to the experimental data, we find that the spectral broadening is best described by a hyperbolic secant lineshape with a homogeneous linewidth $\delta = 25$ meV. The theoretical calculation was also carried out at the threshold injection level, which was determined as the carrier density required to produce the TE-polarized threshold material gain $g_{th} = 1325$ cm^{-1} (computed using an effective index calculation to determine $\Gamma = 1.60\%$ for a QW containing 1.8% Bi, and using the measured threshold current density $J_{th} \approx 1.6$ kA cm^{-2} in similar devices [96]). Due to uncertainty in the absolute units and relative intensity of the SE measured for the two devices, the calculated SE spectrum has been normalized to the measured value at the SE peak in order to compare the theoretical and experimental data. We note that the measured and calculated SE spectra are in good overall agreement, with the spectral broadening observed in the experiment well described by a combination of the energy broadening at room temperature of the electron and hole (quasi-Fermi) distribution functions, as well as the hyperbolic secant line broadening.

In order to quantify the variation of the Bi composition across the wafer the SE spectrum for device 2 has also been calculated. The calculation of the SE spectrum in this case proceeds as above, this time with the Bi composition again being the only parameter allowed to vary. The result of this calculation is shown using a dashed line in Figure 10.11a. We find that (1) the measured SE peak at 932 nm is well described in the theoretical calculations, assuming that the QW contains the nominal Bi composition of 2.2%, and (2) the calculated SE spectrum for $x = 2.2\%$ is in good overall agreement with experiment for this device. Overall, these theoretical results suggest that the Bi composition in the QW is close to the nominal value of 2.2% near the center of the wafer, but is slightly reduced, to 1.8%, toward the wafer edges. We recall that this variation in Bi composition is consistent with that determined previously via spectroscopic measurements [79].

The closed symbols in Figure 10.11b show the (TE-polarized) net modal gain spectra, $\Gamma g - \alpha_i$, measured using the segmented contact method for device 1. These gain spectra, which were measured at room temperature for current densities $J = 0.7$, 1.4, 2.0, and 2.4 kA cm^{-2}, are shown using closed circles, upright triangles, inverted triangles and squares, respectively. The measured gain spectra are relatively broad, with a full width at half maximum of approximately 100 meV at a current density of 2 kA cm^{-2}, which is close to twice that observed for an InGaAs/GaAs SQW laser operating at a similar wavelength [101] and is most likely related to the strong Bi-induced inhomogeneous broadening that is characteristic of the optical spectra of $GaBi_xAs_{1-x}$ alloys [27,28]. Based on the $\alpha_i = 15$ cm^{-1} optical losses measured for device 1, the estimated peak modal gain at $J = 2$ kA cm^{-2} is $\Gamma g_{peak} = 24$ cm^{-1}. Using the calculated optical confinement factor for this laser structure, $\Gamma = 1.60\%$, we estimate a peak material gain $g_{peak} \approx 1500$ cm^{-1} at $J = 2$ kA cm^{-2}, which agrees well with the value of 1560 cm^{-1}, calculated using the theoretical model.

In order to compare the calculated gain spectra to experiment we must determine the carrier density n corresponding to each current density J at which each of the gain spectra depicted by open circles in Figure 10.11b was measured. In order to do so, we recall from Equation 10.8 that the transparency point at which there is zero material/modal gain on the high-energy side of the gain peak is given when the photon energy is equal to the quasi-Fermi level separation at a given level of injection (i.e., $g = 0$ for $\hbar\omega = \Delta F$). We therefore proceed by shifting the measured net modal gain spectra, $\Gamma g - \alpha_i$, upward by the $\alpha_i = 15$ cm^{-1} optical losses in order to obtain the absolute modal gain Γg in the device. Next, we use the transparency points on each of these absolute modal gain spectra to extract the quasi-Fermi level separation ΔF corresponding to each current density J in the experiment. Finally, using the theoretical model we calculate ΔF as a function of the injected carrier density n for the laser structure and, in doing so, determine the value of n which produces the extracted ΔF for each current density J. Following this procedure, we

find that the current densities $J = 0.7$, 1.4, 2.0, and 2.4 kA cm^{-2} at which the gain spectra were measured, correspond to quasi-Fermi level separations of 1.375, 1.409, 1.423, and 1.435 eV, which in turn correspond to carrier densities $n = 5.12$, 7.11, 8.24, and 9.38 $\times 10^{18}$ cm^{-3}, respectively, in the theoretical calculations. The TE-polarized component of the SE rate at each of these carrier was then calculated as outlined in Section 10.4.1, from which the theoretical net modal gain spectrum was calculated at each carrier density as $\Gamma g - \alpha_i$, with $\Gamma = 1.60\%$ (calculated) and $\alpha_i = 15$ cm^{-1} (measured), and where g is the TE-polarized material gain spectrum computed using Equation 10.8.

The results of these calculations are shown in Figure 10.11b using, in order of increasing current/carrier density, solid, dashed, dotted, and dash-dotted lines. We note that the theoretical gain spectra shown using solid, dashed and dotted lines include $e1$-$hh1$ optical transitions only while the dash-dotted line, corresponding to the highest current density of 2.4 kA cm^{-2}, also includes $e1$-$lh1$ optical transitions. Including $e1$-$hh1$ transitions only at the highest carrier density was found to underestimate the peak modal gain by ~10%, highlighting that TE-polarized optical recombination involving LH-like states contributes appreciably to the optical gain at higher levels of injection. This conclusion is consistent with further experimental analysis of the SE spectra [39]. Overall, we see that the theoretical spectra are in good, quantitative agreement with experiment: the calculated magnitude of the net modal gain is in excellent agreement with that measured using the segmented contact method across the full investigated range of current densities (confirming in particular the accuracy of the optical transition matrix elements computed within the framework of the 12-band **k**·**p** model), and the shape (in photon energy/wavelength) of the experimental gain spectrum is well reproduced at each current/carrier density by the theoretical model. This analysis has been repeated for a series of multisection and Fabry–Pérot laser devices having $x \approx 2\%$, demonstrating that the theoretical model is generally capable of quantitatively predicting the variation of g_{peak} with J for this new class of laser structures [39]. Overall, we note that the theoretical model presented in Section 10.4.1, which is directly underpinned by detailed analysis of the impact of Bi on the band structure (facilitated by atomistic supercell calculations), allows calculation of the device properties and performance that are generally in good, quantitative agreement with the available experimental data. The results presented here therefore confirm the predictive capability of the model, suggesting that it can be used as a reliable tool in the analysis, design and optimization of future devices.

10.5 Emerging Directions in Dilute Bismide Research

Having primarily focused so far in this chapter on describing developments in the growth, characterization, theory and applications of $GaBi_xAs_{1-x}$ alloys grown on GaAs substrates, in this section, we turn our attention to alternative, emerging directions in research on dilute bismide alloys. Sections 10.5.1 and 10.5.2 outline the basic band structure properties of the quaternary Bi-containing alloys $In_yGa_{1-y}Bi_xAs_{1-x}$ and $GaBi_xN_yAs_{1-x-y}$, grown on InP and GaAs substrates, respectively. By extending the **k**·**p** model for $GaBi_xAs_{1-x}$ alloys outlined in Sections 10.3.2 and 10.3.3, we demonstrate how the incorporation of Bi leads in each case to highly flexible alloy band structures and long emission wavelengths beyond those currently accessible using GaAs or InP substrates. We also briefly describe the development of $GaBi_xAs_{1-x}/GaN_yAs_{1-y}$ type-II QWs grown on GaAs and demonstrate that this novel class of strain-balanced heterostructures have significant promise for the development of photonic devices operating at and beyond 1.55 μm. $GaBi_xAs_{1-x}/GaN_yAs_{1-y}$ type-II QWs also have the potential to act as a route to extending the emission wavelength beyond that which has been demonstrated to date in $GaBi_xAs_{1-x}$/GaAs type-I QWs, which is currently limited by the challenges associated with the epitaxial growth of $GaBi_xAs_{1-x}$ QWs having sufficiently high optical quality and Bi composition to demonstrate electrically pumped lasing at 1.55 μm. In all three cases, we demonstrate the ability to engineer the band gap, spin-orbit-splitting energy, strain, and band offsets over wide ranges, highlighting the potential to use these new material concepts to deliver enhanced capabilities in near- and mid-infrared photonic devices, as well as (at shorter wavelengths) in photovoltaics.

In addition to these three emerging directions in dilute bismide research, we note that ongoing increase of interest in Bi-containing materials has led to initial demonstrations of a number of additional classes of alloys and heterostructures for a range of potential applications. While it is beyond the scope of this chapter to discuss all of these topics in detail, we briefly list prominent examples in Section 10.5.4 and refer the reader to the references provided therein for further information.

10.5.1 $In_yGa_{1-y}Bi_xAs_{1-x}$/InP Alloys

The vibrational-rotational spectra of many important gases are characterized by strong optical absorption at mid-infrared wavelengths. As such, there is broad scope for practical applications of semiconductor lasers operating in the 2- to 6-μm wavelength range, including: (1) chemical monitoring in industrial processes, (2) detection of environmental pollutants, (3) remote analysis of hazardous substances (including toxic gases and explosives), and (4) detection of biological markers in medical diagnostics. Despite this, mid-infrared semiconductor laser technology is generally less developed than the GaAs- and InP-based technologies that are employed in optical communications at wavelengths <2 μm. Despite the fact that much effort has been dedicated to developing mid-infrared emitters and detectors over the past two decades [102,103], there remains a need to develop new material and device concepts in order to overcome the limitations associated with current technologies. Existing mid-infrared semiconductor lasers typically (1) incorporate Sb-containing materials grown on nonideal and relatively expensive GaSb or InAs substrates, (2) suffer from difficulties associated with the growth of Sb-containing alloys, and/or (3) incorporate complicated heterostructures, such as in quantum (intraband) or interband cascade devices, where the laser structure must be formed of many layer repeats which require a high degree of design optimization and growth control.

In order to overcome these issues, as well as to improve device performance, there has been increasing effort directed toward developing Sb-free devices grown on InP substrates. From a practical perspective, it is desirable to develop InP-based mid-infrared devices, for several reasons. First, materials growth and processing is significantly more advanced for the InP platform than for GaSb or InAs, since the former has experienced heavy technological and commercial development for applications optical in communications. As such, the InP platform offers well-developed, reproducible, and cost-effective device fabrication. Second, from a material physics perspective, InP has several advantages over GaSb, including higher thermal conductivity and lower electrical resistance. Third, growth of InP-based materials stands to benefit from the ready availability of advanced optical components such as low-loss waveguides.

$In_yGa_{1-y}Bi_xAs_{1-x}$/InP alloys offer a promising alternative to existing approaches in this challenging wavelength range, and provide the possibility to develop interband diode lasers incorporating type-I QWs which (1) significantly extend the wavelength range accessible using InP substrates, (2) fully exploit the well-established aspects of the InP material platform [45], and (3) use the Bi-induced modifications to the band structure to suppress the Auger and IVBA loss mechanisms. Full details of our theoretical analysis of the $In_yGa_{1-y}Bi_xAs_{1-x}$ band structure and its suitability for applications in mid-infrared light-emitting devices, can be found in Refs. [13] and [42]. Figure 10.12a shows the variation of the band gap E_g and spin-orbit-splitting energy Δ_{SO} with Bi composition x in $In_{0.53}Ga_{0.47}Bi_xAs_{1-x}$ alloys grown pseudomorphically on InP. Solid (dashed) lines show the variation of E_g (Δ_{SO}) calculated using the 12-band $\mathbf{k}\cdot\mathbf{p}$ Hamiltonian described in Section 10.3, while closed (open) circles show the values of E_g (Δ_{SO}) measured using a combination of optical absorption, PL, and photomodulated reflectance spectroscopy [40,42]. We note that (1) the calculated variation of E_g and Δ_{SO} with x is in quantitative agreement with experiment, and (2) the theoretical and experimental data indicate that a band structure in which $\Delta_{SO} > E_g$ can be obtained at low Bi compositions $x > 3.5\%$ in compressively strained $In_{0.53}Ga_{0.47}Bi_xAs_{1-x}$/InP alloys, which is significantly lower than the ~10% required on GaAs (cf. Figure 10.3aa).

In order to evaluate the suitability of $In_yGa_{1-y}Bi_xAs_{1-x}$/InP alloys for applications in mid-infrared devices, we have used the 12-band $\mathbf{k}\cdot\mathbf{p}$ model to investigate general trends in the band structure as a function of the Bi and In compositions x and y [44,45]. Figure 10.12b shows a composition space map of the

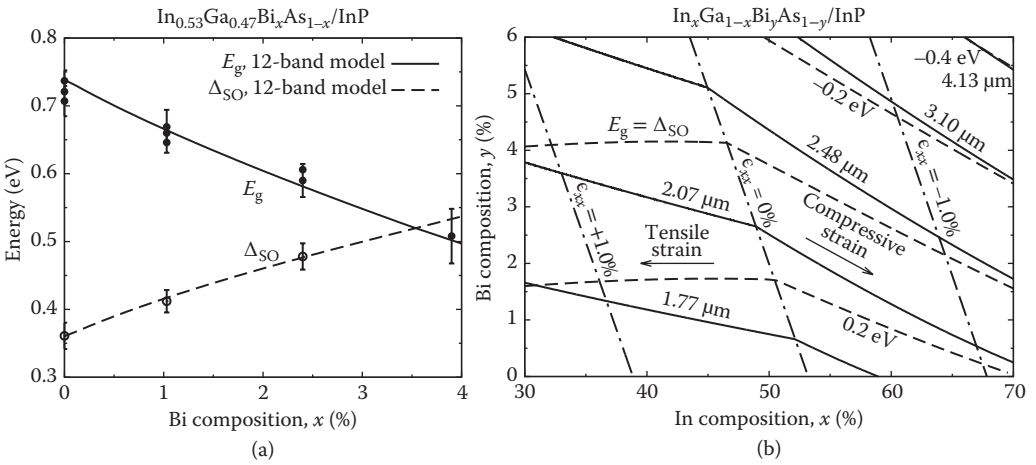

FIGURE 10.12 (a) Variation of the band gap (E_g) and spin-orbit-splitting energy (Δ_{SO}) as a function of Bi composition x, measured for $In_{0.53}Ga_{0.47}Bi_xAs_{1-x}$ alloys grown pseudomorphically on InP (closed and open circles, respectively), compared to calculations undertaken using the 12-band **k·p** Hamiltonian of Section 10.3 (solid and dashed lines, respectively). (Adapted from Marko, I.P. et al., *Appl. Phys. Lett.*, 101, 221108, 2012; C.A. Broderick et al., *Phys. Stat. Sol. B*, 250, 773, 2013.) (b) Calculated variation of E_g, and of the difference between the band gap and spin–orbit-splitting energy, $E_g - \Delta_{SO}$, as a function of Bi and In composition (x and y) for $In_yGa_{1-y}Bi_xAs_{1-x}$/InP alloys (Adapted from Broderick, C.A. et al., In *Proceedings of the 16th International Conference on Numerical Simulation of Optoelectronic Devices*, p. 47 (2016); C.A. Broderick et al., *Theory and design of $In_yGa_{1-y}As_{1-x}Bi_x$ mid-infrared semiconductor lasers: type-I quantum wells for emission beyond 3 μm on InP substrates*, submitted 2017.) Solid and dashed lines denote, respectively, paths in the composition space along which E_g and $E_g - \Delta_{SO}$ are constant. Dash-dotted lines denote paths in the composition space along which strain is constant. We see that $In_yGa_{1-y}Bi_xAs_{1-x}$ alloys can be grown either lattice-matched ($\epsilon_{xx} = 0$), or under compressive ($\epsilon_{xx} < 0$) or tensile ($\epsilon_{xx} > 0$) strain, and that compressively strained alloys enable emission at wavelengths >3 μm at relatively low Bi compositions. Alloys with compositions lying above the $E_g = \Delta_{SO}$ contour are alloys in which $\Delta_{SO} > E_g$, and in which suppression of the nonradiative Auger recombination and IVBA processes involving the SO band can hence be expected.

band gap (E_g), and of the difference between the band gap and spin-orbit-splitting energy ($E_g - \Delta_{SO}$), in pseudomorphically strained $In_yGa_{1-y}Bi_xAs_{1-x}$/InP, calculated using the 12-band model. The dash-dotted lines denote paths in the composition space along which the in-plane strain ϵ_{xx} is constant when the $In_yGa_{1-y}Bi_xAs_{1-x}$ alloy is grown pseudomorphically on an InP substrate and show that the alloy can be grown either lattice-matched ($\epsilon_{xx} = 0$) or under compressive ($\epsilon_{xx} < 0$) or tensile ($\epsilon_{xx} > 0$) strain. Solid lines denote alloy compositions for which E_g is constant. Examining the calculated variation of ϵ_{xx} and E_g with alloy composition, we observe that the band gap of $In_yGa_{1-y}Bi_xAs_{1-x}$ alloys can be varied over an extremely wide range, from ~1.55 to >4 μm, and that this broad spectral coverage can be obtained at modest strains $|\epsilon_{xx}| < 1.5\%$.

The dashed lines in Figure 10.12b denote alloy compositions for which $E_g - \Delta_{SO}$ is constant. Following the $\epsilon_{xx} = 0$ line, we calculate that (1) the amount of Bi required to bring about the $\Delta_{SO} > E_g$ band structure condition (all alloys lying above the $E_g = \Delta_{SO}$ contour) is close to 4% (a prediction in quantitative agreement with the results of a range of experimental measurements [40,43,54]), and (2) the amount of Bi required to achieve $\Delta_{SO} > E_g$ remains constant at approximately 4% in tensile strained alloys less In while, by contrast, it is significantly reduced in compressively strained alloys having higher In compositions. Compared to GaAs, $In_yGa_{1-y}As$/InP alloys have reduced band gaps and comparable spin-orbit-splitting energies, so that significantly less Bi is required to achieve the $\Delta_{SO} > E_g$ band structure condition in $In_yGa_{1-y}Bi_xAs_{1-x}$/InP alloys than in the $GaBi_xAs_{1-x}$/GaAs materials considered previously [13,40,42,43].

Therefore, in addition to providing the possibility of suppressing the CHSH Auger recombination and IVBA loss mechanisms in InP-based mid-infrared semiconductor lasers, the $In_yGa_{1-y}Bi_xAs_{1-x}$/InP material system also has the potential to provide an important proof of principle regarding the use of dilute bismide alloys to suppress the dominant loss mechanisms in GaAs-based semiconductor lasers operating at telecommunication wavelengths.

The difference in band structure between the compressive and tensile strained alloys shown in Figure 10.12b is readily understood: Bi, being larger than the As atoms it replaces, introduces compressive strain so that a significant fraction of In needs to be removed in order to obtain a tensile strained quaternary alloy containing Bi. This removal of In leads to a tensile strained InGaAs host matrix with an increased band gap, as well as a VB structure in which the VBAC coupling is weakened compared to the unstrained or compressively strained cases [13,42]. When these effects are combined, we find that significantly larger Bi compositions are required in tensile strained $In_yGa_{1-y}Bi_xAs_{1-x}$ in order to bring about a given band gap reduction for a fixed magnitude of the strain, as compared to the compressively strained case in which both In and Bi simultaneously contribute to the band gap reduction and compressive strain. Bearing this in mind, and examining Figure 10.12b from the perspective of laser design, we note that the composition region of interest are those lying above the $E_g = \Delta_{SO}$ contour, since it is expected that the Auger and IVBA processes involving the SO band will be suppressed in these alloys, while they are close to resonant for $2 < \lambda < 2.5$ μm, at which compositions $E_g \approx \Delta_{SO}$. Within the reduced composition range of alloys having $\Delta_{SO} > E_g$ we note that the alloy band gap covers a large range of wavelengths >2.5 μm and, significantly, that the 2.5- to >4-μm spectral range can be accessed in alloys having modest compressive strains and Bi compositions.

Our theoretical analysis of $In_yGa_{1-y}Bi_xAs_{1-x}$/InP laser structures [45] suggests that there is broad scope to use this material system to realize compressively strained type-I QWs having emission wavelengths in the 2.5- to 5 μm range. Furthermore, our analysis indicates that the physical characteristics of $In_yGa_{1-y}Bi_xAs_{1-x}$/InP QWs are intrinsically superior to existing GaSb-based heterostructures in the same wavelength range. First, as outlined above, theoretical calculations and experimental measurements have confirmed that $\Delta_{SO} > E_g$ can be achieved in $In_{0.53}Ga_{0.47}Bi_xAs_{1-x}$/InP alloys for Bi compositions as low as $x \approx 4\%$ [13,40,42,43]. This means that it should be possible to suppress the CHSH Auger recombination and IVBA processes and bring about highly efficient, temperature-stable operation in mid-infrared semiconductor lasers. Second, our calculations indicate that large type-I band offsets can readily be engineered in $In_yGa_{1-y}Bi_xAs_{1-x}$ QWs having unstrained ternary $In_{0.53}Ga_{0.47}As$ barriers. This has two significant benefits: (1) the presence of large CB and VB offsets will mitigate carrier leakage, thereby overcoming a key factor limiting the performance of GaSb-based devices at high temperature, and (2) the ability to grow lattice-matched, ternary InGaAs barriers, without the need to incorporate Al or P to enhance the carrier confinement, promises to simplify the growth of these laser structures. We note that (3) also means that it is not necessary to introduce a trade-off between the carrier and optical confinement in these structures (cf. Section 10.4.3). Furthermore, these ternary barrier materials are the same unstrained InGaAs buffer layers upon which $In_yGa_{1-y}Bi_xAs_{1-x}$ epilayers have typically been grown. Overall, this means that $In_yGa_{1-y}Bi_xAs_{1-x}$ QWs can be used to develop mid-infrared lasers in which the remainder of the laser structure is essentially identical to those commonly employed in the well-established 1.3- and 1.55-μm InP-based devices. As such, $In_yGa_{1-y}Bi_xAs_{1-x}$/InP alloys offer the possibility not only to provide type-I InP-based QW diode lasers operating at wavelengths >3 μm, but also promise to deliver several key advantages over competing mid-infrared device concepts.

While there have been some investigations of $In_yGa_{1-y}Bi_xAs_{1-x}$ alloys grown on InP substrates, this material system has to date been the subject of much less research attention than the GaAs-based $GaBi_xAs_{1-x}$ alloys discussed above. Initial growth of $In_yGa_{1-y}Bi_xAs_{1-x}$/InP materials was reported by Feng et al. [104,105], and has since been established by several other groups [106–109]. Enhanced understanding of the incorporation of Bi in InGaAs has led to growth of $In_yGa_{1-y}Bi_xAs_{1-x}$ alloys with Bi compositions as high as 7% [108], and recent experimental and theoretical analysis of a sample containing 5.8% Bi has verified the presence of a $\Delta_{SO} > E_g$ band structure up to and above room temperature [43]. Growth of

$In_yGa_{1-y}Bi_xAs_{1-x}$ alloys has, however, to date been limited to bulk-like epitaxial layers and so, in order to develop this material system for applications in mid-infrared semiconductor lasers, future studies will need to focus on (1) the growth of thin, strained QWs having high optical quality, and (2) providing evidence of reduced CHSH Auger recombination and IVBA in materials and heterostructures having $\Delta_{SO} > E_g$. While research on $In_yGa_{1-y}Bi_xAs_{1-x}$/InP alloys still faces challenges related to the establishment of growth and fabrication of high-quality materials and devices, recent successes in the development of $GaBi_xAs_{1-x}$ alloys, including the demonstration of electrically pumped lasers [30,35,96], suggest that these issues can be overcome with dedicated research effort. We therefore conclude that $In_yGa_{1-y}Bi_xAs_{1-x}$ alloys are a promising candidate for the development of highly efficient, temperature-stable semiconductor lasers, grown on InP substrates and operating in the 3- to 5-μm wavelength range.

10.5.2 $GaBi_xN_yAs_{1-x-y}$/GaAs Alloys

An alternative quaternary Bi-containing alloy which has begun to attract research interest is $GaBi_xN_yAs_{1-x-y}$. Initial theoretical analysis has demonstrated that the band structure of $GaBi_xN_yAs_{1-x-y}$ alloys grown on GaAs substrates is extremely flexible, providing significant potential to develop materials suited to a range of applications in the near- and mid-infrared [12,41,54]. First, since Bi is larger than the As atoms it replaces in GaAs to form $GaBi_xAs_{1-x}$, it introduces compressive strain when grown on a GaAs substrate. Likewise, N atoms, which are smaller than As, introduce tensile strain. As a result, independent control of the Bi and N compositions in $GaBi_xN_yAs_{1-x-y}$ allows precise strain engineering with respect to a GaAs substrate. Second, Bi or N incorporation in GaAs is known to cause a rapid reduction of the band gap with increasing composition. Theoretical calculations [54] suggest that coalloying Bi and N in GaAs to form $GaBi_xN_yAs_{1-x-y}$ enhances these effects and leads to an extremely large reduction in the band gap with increasing Bi and N composition, which is larger than that observed in either $GaBi_xAs_{1-x}$ or GaN_xAs_{1-x} alloys. As a result, the band gaps of $GaBi_xN_yAs_{1-x-y}$ alloys are capable of covering a large spectral range in the near- and mid-infrared, and come with the added benefit of being able to independently control the strain and band gap. These properties make $GaBi_xN_yAs_{1-x-y}$ alloys of particular interest for applications in multi-junction solar cells [21], since they offer the potential to provide optical absorption across an extremely wide range of wavelengths while remaining lattice-matched to either a GaAs or Ge substrate. In addition to the ability to readily control the strain in these alloys, the impact of Bi on the VB structure are retained, meaning that it should also be possible to achieve a band structure in which $\Delta_{SO} > E_g$ in order to suppress Auger and IVBA processes involving the SO band.

Detailed atomistic calculations suggest that the effects of Bi and N on the band structure are effectively decoupled [54], with the modified CB (VB) structure describable primarily in terms of the well-known effects of N-(Bi-)related localized resonant states. On this basis, Broderick et al. derived a 14-band $\mathbf{k}\cdot\mathbf{p}$ Hamiltonian to describe the band structure of $GaBi_xN_yAs_{1-x-y}$ alloys [54]. In this model the effects of N and Bi incorporation on the CB and VB structure are described independently in terms of band-anticrossing interactions between N-(Bi-)related localized states and the extended states of the GaAs host matrix CB (VB) edge (cf. Section 10.3.1), with additional virtual crystal and strain-induced modifications to the CB, VB, and SO band edge energies (cf. Section 10.3.2). The fact that N (Bi) primarily affects the CB (VB) then means that the $GaBi_xN_yAs_{1-x-y}$/GaAs band offsets are effectively independently controllable, so that there is significant potential to engineer the properties of QWs and related heterostructures. Incorporating N brings about a large reduction in the CB edge energy, resulting in a large CB offset and a small VB offset with respect to GaAs. The converse is true of Bi incorporation, which leads to a small CB offset and a large VB offset. Combining these effects means that the CB and VB offsets can be engineered relatively independently of one another in $GaBi_xN_yAs_{1-x-y}$, opening up significant potential to develop $GaBi_xN_yAs_{1-x-y}$ QWs providing tunable long-wavelength emission and absorption across a broad spectral range. Full details of our theoretical analysis of the $GaBi_xN_yAs_{1-x-y}$ band structure can be found in Refs. [12], [41] and [54].

Figure 10.13 shows a composition space map of the band gap (E_g), and of the difference between the band gap and spin–orbit-splitting energy ($E_g - \Delta_{SO}$), in pseudomorphically strained $GaBi_xN_yAs_{1-x-y}$/GaAs, calculated using the 14-band **k·p** Hamiltonian described in Ref. [54]. The dashed lines denote paths in the composition space along which the in-plane strain is constant when the $GaBi_xN_yAs_{1-x-y}$ alloy is grown pseudomorphically on a GaAs substrate, and show that the alloy can be grown either lattice-matched ($\epsilon_{xx} = 0$), or under compressive ($\epsilon_{xx} < 0$) or tensile ($\epsilon_{xx} > 0$) strain. Solid lines denote alloy compositions for which E_g is constant. Examining the calculated variation of ϵ_{xx} and E_g with alloy composition we observe that the band gap of $GaBi_xN_yAs_{1-x-y}$ alloys can be varied over an extremely wide range, from that of GaAs (~850 nm) through the near-infrared, and out to wavelengths >3 μm in the mid-infrared. We also note that this broad spectral coverage can be obtained in alloys that are lattice-matched to GaAs, or in alloys having modest tensile or compressive strains with $|\epsilon_{xx}| < 1\%$. The dashed lines in Figure 10.13 denote alloy compositions for which $E_g - \Delta_{SO}$ is constant. The fact that N incorporation brings about a large band gap reduction while leaving the VB structure almost unaltered (apart from minor strain-induced changes) means that co-alloying N and Bi has the potential to bring about an $\Delta_{SO} > E_g$ band structure at lower Bi compositions than in the N-free $GaBi_xAs_{1-x}$ alloy. Following the $\epsilon_{xx} = 0$ line we calculate that the amount of Bi required to bring about the $\Delta_{SO} > E_g$ band structure condition (all alloys lying to the right of the $E_g = \Delta_{SO}$ contour) is close to 7% in the lattice-matched case, which is approximately 30% less than that required in $GaBi_xAs_{1-x}$. In addition to the aforementioned potential of this material system for applications in multijunction solar cells, we also note that the wide composition range for which $\Delta_{SO} > E_g$ indicates that there is also potential to grow $GaBi_xN_yAs_{1-x-y}$ QWs for >2-μm emission, bringing about the possibility of developing GaAs-based mid-infrared diode lasers with suppressed CHSH Auger recombination and IVBA. This also provides an interesting opportunity to develop monolithic GaAs-based VCSELs and related devices operating in the mid-infrared.

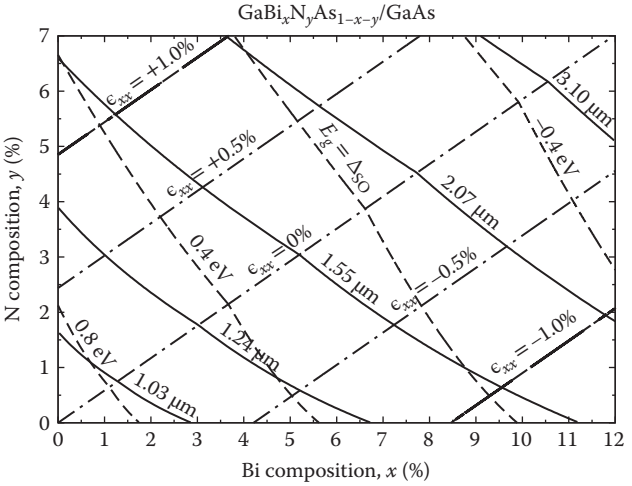

FIGURE 10.13 Calculated variation of the band gap (E_g), and of the difference between the band gap and spin-orbit-splitting energy ($E_g - \Delta_{SO}$), as a function of Bi and N compositions (x and y) for $GaBi_xN_yAs_{1-x-y}$ alloys grown pseudomorphically on GaAs compositions. (Adapted from Broderick, C.A. et al., *Semicond. Sci. Technol.*, 27, 094011, 2012.) Solid and dashed lines denote, respectively, paths in the composition space along which E_g and $E_g - \Delta_{SO}$ are constant. Dash-dotted lines denote paths in the composition space along which strain is constant. We see that $GaBi_xN_yAs_{1-x-y}$ alloys can be grown either lattice-matched ($\epsilon_{xx} = 0$), or under compressive ($\epsilon_{xx} < 0$) or tensile ($\epsilon_{xx} > 0$) strain, and that the band gap can be varied over an extremely wide spectral range in the near- and mid-infrared. Alloys with compositions lying to the right of the $E_g = \Delta_{SO}$ contour are alloys in which $\Delta_{SO} > E_g$, and in which suppression of the non-radiative Auger recombination and IVBA processes involving the SO band can hence be expected.

While there is increasing interest in the development of $GaBi_xN_yAs_{1-x-y}$ alloys for applications in the areas described above, to date there have been limited growth studies of this novel material system. Initial growth of $GaBi_xN_yAs_{1-x-y}$ bulk-like epitaxial layers was undertaken via MBE in the mid-2000s [110,111], while more recently epitaxial layers and QWs have been grown using MOVPE [112]. MOVPE growth has realized co-alloying of Bi and N compositions up to $x = 3.5\%$ and $y = 2.7\%$, but PL has only been observed for samples with compositions up to $x = 1.8\%$ and $y = 1.8\%$. However, the observed PL at these compositions demonstrated an alloy band gap close to 1 eV, confirming that co-alloying of Bi and N can bring about an extreme reduction of the band gap at relatively low compositions [112]. Meanwhile, MBE growth of $GaBi_xN_yAs_{1-x-y}$ alloys has demonstrated electroluminescence at wavelengths out to 1.3 μm [113], and Bi and N incorporations up to $x = 4.5\%$ and $y = 8.0\%$ have been obtained [110]. However, no measurements of the band gap have been performed on these higher composition materials, with the Bi and N compositions having been deduced on the basis of x-ray diffraction measurements. Despite the significant challenges presented by the growth of this highly mismatched quaternary alloy, initial investigations have been promising, and the realization of a 1-eV band gap at such low compositions merits further investigations in order to refine the growth of $GaBi_xN_yAs_{1-x-y}$ alloys to exploit their significant potential for practical applications.

10.5.3 $GaBi_xAs/GaN_yAs_{1-y}$ Type-II Quantum Wells on GaAs

In addition to the direct co-alloying Bi and N in GaAs to form $GaBi_xN_yAs_{1-x-y}$, Bi and N incorporation in GaAs opens up additional possibilities for band structure engineering and the development of novel heterostructures. For example, GaN_xAs_{1-x} ($GaBi_xAs_{1-x}$) can be used to form tensile (compressively) strained type-I QWs it possible to grow strain-balanced type-II QWs and superlattices in which electrons (holes) are confined in the GaN_xAs_{1-x} ($GaBi_xAs_{1-x}$) layer(s) [46,114]. Such structures offer a number of advantages for applications in semiconductor lasers, photodetectors, and solar cells, including the possibility of facilitating optical emission and absorption across a wide range of wavelengths, as well as exploiting a high degree of control over the built-in strain, carrier transport, radiative lifetimes and non-radiative recombination rates [46]. A range of type-II QWs grown on GaAs using different combinations of III-V alloys are currently under investigation for long-wavelength applications. However, strain-balancing of type-II QWs on GaAs has not been possible using the $(In)Ga(N)As_{1-x}Sb_x/In_yGa_{1-y}As$ or $GaAs_{1-x}Bi_x/In_yGa_{1-y}As$ structures that have been investigated recently [115–118], since the constituent alloys are all compressively strained when grown on GaAs. By comparison, $GaAs_{1-x}Bi_x/GaN_yAs_{1-y}$ type-II structures are highly engineerable, can be grown with little or no net strain relative to a GaAs substrate, and significant potential for applications in GaAs-based near- and mid-infrared photonic devices [46] and (at shorter wavelengths) in photovoltaics [114].

The band alignment, band offsets, and emission wavelengths achievable using $GaAs_{1-x}Bi_x/GaN_yAs_{1-y}$ structures on GaAs are summarized in Figure 10.14a. Dashed lines denote Bi and N compositions x and y for which the CB offset ΔE_{CB} and VB offset ΔE_{VB} are equal in $GaAs_{1-x}Bi_x$ and GaN_yAs_{1-y}. Above (below) the lower dashed line ΔE_{CB} is larger (smaller) in GaN_yAs_{1-y} than in $GaAs_{1-x}Bi_x$, while to the right (left) of the upper dashed line ΔE_{VB} is larger (smaller) in $GaAs_{1-x}Bi_x$ than in GaN_yAs_{1-y}. This divides the composition space into three regions. Regions A and C correspond to type-I band alignment, with electrons and holes both confined within either the GaN_yAs_{1-y} (A) or $GaAs_{1-x}Bi_x$ (C) layers. Region B corresponds to type-II band alignment, with holes (electrons) confined in the $GaAs_{1-x}Bi_x$ (GaN_yAs_{1-y}) layers. Closed circles denote increases of 50 meV in the respective band offsets, beginning from zero at $x = y = 0$ (GaAs). Solid (dash-dotted) lines in region B denote alloy compositions for which the band gap in a bulk (QW) type-II structure—between the GaN_yAs_{1-y} CB and $GaAs_{1-x}Bi_x$ VB—is constant. These calculations demonstrate that type-II $GaAs_{1-x}Bi_x/GaN_yAs_{1-y}$ QWs grown on GaAs have the potential to cover an extremely broad spectral range, through the near-infrared to mid-infrared wavelengths in excess of 3 μm [46].

Broderick et al. [46] recently presented the first demonstration of this new class of III-V heterostructures. The black and gray lines in Figure 10.14b show, respectively, the room temperature PL and optical absorption measured for a prototypical type-II $GaBi_xAs_{1-x}/GaN_yAs_{1-y}$ structure containing five QWs, having Bi and N compositions $x = 3.3\%$ and $y = 5.2\%$, and with $GaBi_xAs_{1-x}$ and GaN_yAs_{1-y} layer thicknesses of 10.5 and 9.2 nm [46]. This structure was grown using MOVPE and characterized using high-resolution x-ray diffraction measurements, which indicated high structural quality of the strained type-II QWs. Examining Figures 10.14a and 10.14b we note that, despite the intrinsically low optical efficiency of the type-II structure (resulting from the low electron–hole overlap under optical excitation), the measured PL peak at 0.72 eV agrees well with the *e*1-*hh*1 transition energy of 0.74 eV calculated using a carefully parameterized 14-band **k·p** Hamiltonian [54] in conjunction with the plane wave expansion method described in Section 10.4.1 [46]. We note that this band gap corresponds to an emission wavelength ~1.7μm, which demonstrates that $GaBi_xAs_{1-x}/GaN_yAs_{1-y}$ type-II QWs grown on GaAs can readily provide emission at and beyond 1.55 μm with significantly reduced Bi compositions compared to those required in type-I $GaBi_xAs_{1-x}$/(Al)GaAs structures (cf. Section 10.4). Indeed, it is promising to observe clear room temperature PL and optical absorption in initial growth of a prototype structure, particularly given the significant scope to optimize the layer ordering and thicknesses to enhance the optical efficiency [46]. Theoretical investigations are ongoing to quantify the electronic and optical properties of this new class of GaAs-based heterostructures, and refinement of epitaxial growth is expected to lead to the development of structures with improved optical efficiency for potential applications in QW solar cells, as well as in semiconductor lasers and photodetectors.

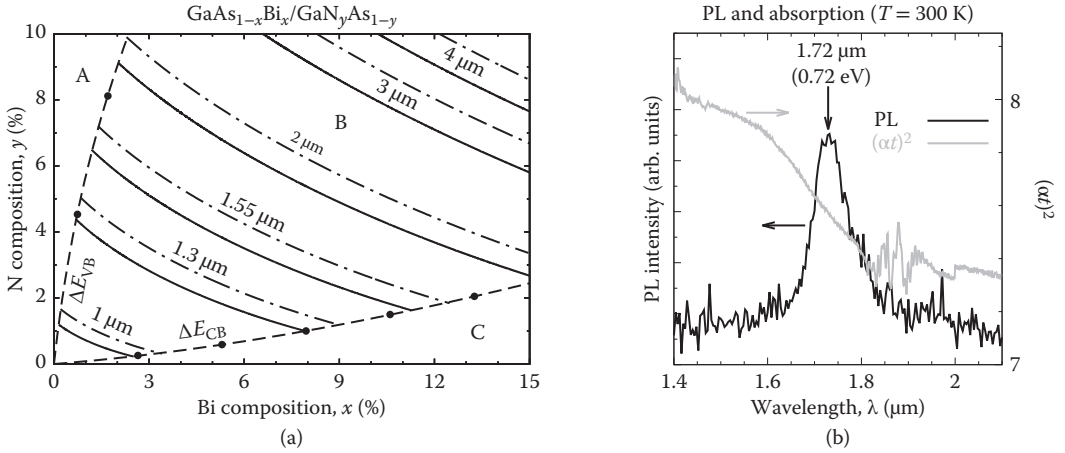

FIGURE 10.14 (a) Composition space map showing regions of type-I (A and C) and type-II (B) band alignment, as well as the emission wavelengths accessible using $GaAs_{1-x}Bi_x/GaN_yAs_{1-y}$ type-II QWs grown on GaAs. Solid (dash-dotted) lines denote paths in the composition space along which the type-II band gap in bulk (QW) $GaAs_{1-x}Bi_x/GaN_yAs_{1-y}$ is constant. Dashed lines denote paths in the composition space along which the CB and VB offsets ΔE_{CB} and ΔE_{VB} are equal in $GaAs_{1-x}Bi_x$ and GaN_yAs_{1-y}. Closed circles denote 50 meV increases in ΔE_{CB} and ΔE_{VB} starting from zero at $x = y = 0$. (Adapted from Broderick, C.A. et al., $GaAs_{1-x}Bi_x/GaN_yAs_{1-y}$ type-II quantum wells: Novel strain-balanced heterostructures for GaAs-based near- and mid-infrared photonics, *Sci. Rep.*, 7, 46371, 2017) (b) Room temperature photoluminescence (PL) spectrum (black line) and squared product $(\alpha t)^2$ of the optical absorption α and propagation length *t* (from transmission measurements; gray line) measured for a prototypical $GaAs_{0.967}Bi_{0.033}/GaN_{0.062}As_{0.938}$ type-II QW structure, grown by MOVPE. (Adapted from C.A. Broderick et al., Type-II strain-balanced superlattices on GaAs: novel heterostructure for photonics and photovoltaics, in *Proceedings of the 17th International Conference on Numerical Simulation of Optoelectronic Devices*, 2017)

10.5.4 Further Directions: Quaternary Alloys, Type-II Structures, and Narrow-Gap Materials

Here, we briefly outline additional emerging directions in resarch on dilute bismide alloys by describing some recent work on materials and heterostructures distinct from, but related to, those discussed above. We refer the reader to the reference provided for further information.

Recently, the quaternary dilute bismide alloy $GaBi_xAs_{1-x-y}P_y$ has been proposed as an alternative material system to $GaBi_xN_yAs_{1-x-y}$ that can provide a 1 eV band gap while remaining lattice matched to a GaAs substrate (cf. Section 10.5.2). In contrast to co-alloying Bi with N to form $GaBi_xN_yAs_{1-x-y}$, incorporation of phosphorus (P), which also compensates the compressive strain brought about by Bi incorporation, tends to increase the band gap at fixed Bi composition. This means that larger Bi compositions will be required to achieve lattice-matched alloys having a given band gap compared to those required in $GaBi_xN_yAs_{1-x-y}$ alloys. However, initial investigations of $GaBi_xP_yAs_{1-x-y}$ have demonstrated that incorporation of a small amount of P significantly enhances Bi incorporation during MOVPE growth [119]. Although it is likely to be related to the P-induced reduction of compressive strain in the quaternary alloy, the precise mechanism by which co-alloying P and Bi enhances Bi incorporation has yet to be definitively determined. Related structures have been grown which include compressively strained $GaBi_xAs_{1-x}$ QWs having tensile strained $GaAs_{1-y}P_y$ barriers, with the first results for this new class of GaAs-based strain-compensated laser structures having been reported recently in Ref. [120]. Investigations of quaternary dilute bismide alloys containing P (or N) are at a relatively early stage, but initial results have been promising and investigations of these novel semiconductor alloys are ongoing.

Further examples of recently emerging directions in dilute bismide research are the growth of type-II and metamorphic heterostructures. In Section 10.5.3 we described that $GaBi_xAs_{1-x}/GaN_yAs_{1-y}$ type-II QWs can be used to achieve long emission wavelengths on GaAs substrates. Recently, GaAs-based $In_yGa_{1-y}As/GaBi_xAs_{1-x}$ type-II QWs have been suggested as an alternative approach to achieving this goal. These heterostructures are similar in principle to the $GaBi_xAs_{1-x}/GaN_yAs_{1-y}$ type-II QWs discussed above, but differ in two important aspects. First, the electron-confining layer(s) of the structure are formed of the conventional alloy $In_yGa_{1-y}As$. The ease and reproducibility with which high-quality $In_yGa_{1-y}As$ strained layers can be grown on GaAs is likely to be of benefit for the realization of electrically pumped lasers and related devices. Second, while the electron-confining GaN_yAs_{1-y} layer(s) in a type-II $GaBi_xAs_{1-x}/GaN_yAs_{1-y}$ structure are tensile strained, the $In_yGa_{1-y}As$ layer(s) in a type-II $In_yGa_{1-y}As/GaBi_xAs_{1-x}$ structure are compressively strained. This means that it is not possible to grow strain-balanced $In_yGa_{1-y}As/GaBi_xAs_{1-x}$ type-II structures on GaAs, which limits the ability to grow large numbers of QWs due to the accumulation of compressive strain in the device active region. This may limit the ability to grow device structure with sufficiently high optical efficiency. MBE growth of $In_yGa_{1-y}As/GaBi_xAs_{1-x}$ type-II structures has recently been established by Pan et al. [115], who reported low temperature PL at a wavelength close to 1.1 μm, thereby providing an initial confirmation of the potential of these structure to overcome the limitations associated with the growth of $GaBi_xAs_{1-x}$/GaAs type-I QWs having high Bi compositions ($\sim 10\%$).

The first metamorphic heterostructures based on dilute bismide alloys were presented recently by Gu et al. [121], who demonstrated MBE growth of InP-based type-I structures incorporating $Al_yIn_{1-y}As$ metamorphic buffer layers and $InBi_xAs_{1-x}$ QWs. The aim of these structures is the same as that outlined for $In_yGa_{1-y}Bi_xAs_{1-x}$/InP in Section 10.5.1: to extend the emission wavelength in InP-based QWs beyond 3 μm. Current approaches to obtaining emission at wavelengths >3 μm on InP include InAs metamorphic QWs grown on $Al_yIn_{1-y}As$ metamorphic buffer layers. It is hoped that the strong reduction of the $InBi_xAs_{1-x}$ band gap with increasing Bi composition [54] should allow the wavelength to be extended significantly beyond that obtainable using existing devices [19]. The initial work of Gu et al. demonstrated low temperature PL at wavelengths beyond 3.1 μm, confirming the validity of this approach. However, further effort, is required to quantify and exploit the potential of this class of dilute bismide heterostructure for applications in mid-infrared photonic devices.

Finally, we note that there has also been a steady growth of interest in narrow-gap dilute bismide alloys such as (In)GaBi$_x$Sb$_{1-x}$. These alloys, typically grown on GaSb or InAs substrates, offer the potential to use Bi to reduce the already narrow (In)GaSb band gap, and hence cover an extremely broad range of wavelengths in the mid-infrared and beyond. As such, while investigations of Sb-containing dilute bismide alloys are at a relatively early stage, (In)GaBi$_x$Sb$_{1-x}$ alloys are attracting interest for the development of type-I and -II QW lasers and photodetectors throughout the full mid-infrared spectral range [17,122]. Furthermore, due to their intrinsically large spin-orbit coupling, in addition to the potential of using Bi incorporation to introduce large, tunable Rashba spin-splittings, interest in narrow-gap dilute bismide alloys for applications in spintronics is likely to attract significant attention in the coming years [25].

10.6 Summary and Conclusions

In this chapter we have provided an introduction to the theory and simulation of dilute bismide materials and devices, which we have framed within the context of recent and ongoing research on GaBi$_x$As$_{1-x}$ and related Bi-containing alloys. Beginning with a general overview of the electronic properties of dilute bismide alloys, we highlighted the potential to exploit the Bi-induced modifications of the band structure of conventional III–V semiconductors to develop photonic, photovoltaic, and spintronic devices with enhanced performance and capabilities. In particular, we discussed that the incorporation of Bi in (In)GaAs leads to a significant reduction in the material band gap (E_g), which is accompanied by a strong enhancement of the spin-orbit-splitting energy (Δ_{SO}), both of which are characterized by strong composition–dependent bowing. We demonstrated that it is possible, through Bi incorporation, to engineer a band structure in which $\Delta_{SO} > E_g$ at telecommunication wavelengths, which is expected to lead to suppression of the dominant nonradiative Auger recombination and IVBA processes (involving the SO band) that limit the high-temperature performance of existing InP-based 1.55-μm semiconductor lasers. As such, we concluded that GaBi$_x$As$_{1-x}$ alloys have the potential to deliver the long sought-after goal of realizing uncooled operation of highly efficient and temperature-stable telecom lasers, which promises to deliver significant energy savings in next-generation optical communications networks. Growth of long-wavelength devices on GaAs also presents the opportunity to exploit vertical-cavity architectures, meaning that dilute bismide alloys have significant potential to extend the benefits of the GaAs platform to telecommunications and mid-infrared wavelengths and, hence, facilitate the development of advanced photonics technologies, with enhanced performance and new capabilities.

Taking into account the fact that Bi incorporated into (In)GaAs acts as an isovalent impurity, in Section 10.2 we highlighted the need for detailed atomistic models to provide a quantitative understanding of the electronic properties. On this basis, we provided a review of the atomistic methods which have been developed for, and applied to, the study of GaBi$_x$As$_{1-x}$ and related alloys. Having identified the theoretical methods and models that are capable of describing the impact of Bi on the electronic structure, we then compared the results of atomistic TB calculations for realistic, disordered GaBi$_x$As$_{1-x}$ alloys to experiment. We showed that a TB approach is capable of describing, in a quantitative manner, the strong Bi-induced reduction (increase) in the band gap (spin-orbit-splitting energy), across the full composition range for which experimental data are available. We further demonstrated that a TB approach quantitatively predicts the strong Bi-induced increase in the magnitude of the electron effective g factor in pseudomorphically strained GaBi$_x$As$_{1-x}$/GaAs alloys, thereby providing detailed insight into roles played by strain and short-range alloy disorder in determining the details of the strongly perturbed VB structure and confirming the predictive capability of the theory.

In Section 10.3 we turned our attention to deriving a continuum $\mathbf{k}\cdot\mathbf{p}$ model which is suited to modeling the electronic and optical properties of dilute bismide heterostructures. Using an atomistic TB model, we reached two significant conclusions. First, we reviewed that the main features of the band structure of (In)GaBi$_x$As$_{1-x}$ alloys are well described in terms of a phenomenological Bi composition–dependent VBAC interaction between the extended VB edge states of the host matrix semiconductor and localized

Bi-related impurity states which are resonant with the host matrix VB. Second, we demonstrated how the TB method can be used to directly determine the energies of the Bi-related impurity states, as well as the coupling strength between the Bi-related impurity states and host matrix VB edge states, thereby removing the parametric ambiguity associated with band-anticrossing models parameterized solely with reference to experimental measurements. On this basis, we reviewed the derivation of a 12-band **k·p** Hamiltonian for (In)$GaBi_xAs_{1-x}$ directly from a series of detailed atomistic supercell calculations, provided a consistent set of parameters for performing calculations on real alloys, and used the **k·p** model to explicitly derive expressions for the band edge energies and offsets in pseudomorphically strained (In)$GaBi_xAs_{1-x}$ alloys and QWs, respectively.

Next, in Section 10.4, we discussed the theory and modeling of dilute bismide QW lasers grown on GaAs substrates. We applied the 12-band **k·p** model to the study of $GaBi_xAs_{1-x}$/(Al)GaAs QW heterostructures and quantified the impact of Bi incorporation on the (1) VB structure and DOS, (2) optical gain, and (3) threshold characteristics of realistic device structures. By considering the incorporation of Al in the barrier layers to improve electron confinement in the QW, we demonstrated the presence of a trade-off between the carrier and optical confinement which allows the design of $GaBi_xAs_{1-x}$/(Al)GaAs QWs at low Bi compositions $x < 6\%$ to be engineered in order to minimize the threshold current density, a prediction which has been verified experimentally. Repeating this analysis as a function of the QW Bi composition x, we demonstrated that the need to incorporate Al in the barrier layers is removed for $x > 6\%$, which should lead to improved device performance as the Bi composition in GaAs-based laser structures is increased. We extended this analysis to consider the gain characteristics of a $GaBi_xAs_{1-x}$ laser structure designed to emit at 1.55 μm and showed that QWs having higher Bi compositions stand to benefit not only from suppressed Auger recombination and IVBA, but also from the effects of Bi-induced compressive strain, with the latter leading to a favorable VB structure that should, in an ideal device, deliver low threshold current density and high differential gain. Having confirmed the potential of $GaBi_xAs_{1-x}$ QWs for the development of highly efficient and temperature-stable 1.55-μm GaAs-based semiconductor lasers, we then turned our attention to a detailed comparison between theory and experiment for the first generation of electrically pumped $GaBi_xAs_{1-x}$ laser devices. We demonstrated that the theoretical approach outlined in Sections 10.3 and 10.4 is capable of quantitatively describing the SE and optical gain of this new class of GaAs-based semiconductor lasers, confirming its predictive capability and highlighting its potential for use in the design and optimization of future $GaBi_xAs_{1-x}$ lasers and related devices.

Finally, in Section 10.5, we provided an overview of emerging directions in dilute bismide research, including quaternary Bi-containing alloys grown on InP and GaAs substrates, type-II QWs, and narrow-gap materials. We described the potential of $In_yGa_{1-y}Bi_xAs_{1-x}$ and $GaBi_xN_yAs_{1-x-y}$ alloys for respective applications in InP-based mid-infrared photonics, and in multi-junction solar cells. In both cases, we showed that Bi-containing quaternary alloys enable a large degree of control over the band structure, allowing for flexible engineering of the band gap, spin-orbit-splitting energy, band offsets, and strain over wide ranges, thereby making it possible to design new photonic and photovoltaic materials and device structures with enhanced capabilities in the near- and mid-infrared. We described that $GaBi_xAs_{1-x}$/GaN_yAs_{1-y} type-II structures grown on GaAs offer a route to emission on GaAs at wavelengths out to and beyond 3 μm, in strain-balanced QWs and superlattices. We further highlighted emerging work on $In_yGa_{1-y}As$/$GaBi_xAs_{1-x}$ type-II QWs and metamorphic $InBi_xAs_{1-x}$ QWs, which, respectively, offer alternative routes to extending the spectral ranges accessible using GaAs and InP substrates. Finally, we described growing interest in narrow-gap (In)$GaBi_xSb_{1-x}$ alloys, which have significant potential for applications across the entirety of the mid-infrared as well as in spintronic devices.

Overall, we conclude that dilute bismide alloys are a rich and rapidly growing area of semiconductor research in which there remain a range of materials physics and device engineering challenges that must be overcome in order to realize the significant potential of this new class of III–V semiconductor materials for practical applications. The theoretical models and calculations reviewed in this chapter constitute the most complete analysis to date of the electronic, optical, and spin properties of dilute bismide alloys, and are expected to serve as a platform to support, interpret, and stimulate ongoing and future research on Bi-containing materials and devices.

Acknowledgments

The writing of this chapter was supported by the Engineering and Physical Sciences Research Council, UK (project nos. EP/K029665/1, EP/H005587/01 and EP/H050787/1), and by Science Foundation Ireland (project no. 15/IA/3082). The research described in this chapter was further supported by the European Commission Seventh Framework Programme project BIANCHO (project no. FP7-257974), by the Irish Research Council (RS/2010/2766), and by Science Foundation Ireland (project no. 10/IN.1/I2994). The authors thank Dr. M. Usman (University of Melbourne, Australia), Prof. K. Volz (Philipps-Universität Marburg, Germany), Prof. A. Krotkus (Center for Physical Sciences and Technology, Lithuania), Prof. J. P. R. David (University of Sheffield, UK), Prof. T. Tiedje (University of Victoria, Canada), Dr. S. Mazzucato (Institut National des Sciences Appliquées, Toulouse, France), Prof. J. M. O. Zide (University of Delaware, USA), and their colleagues, for supplying materials, devices, and experimental data to facilitate aspects of this research, as well as for many useful discussions on the physics of dilute bismide materials and devices. C.A.B. thanks Prof. J. M. Rorison (University of Bristol, UK) for useful feedback on an earlier version of this chapter. The authors also thank their colleagues at their respective institutions for the contributions they have made to the research described in this chapter.

References

1. S. Tixier, M. Adamcyk, T. Tiedje, S. Francoeur, A. Mascarenhas, P. Wei, and F. Schiettekatte. Molecular beam epitaxy growth of $GaAs_{1-x}Bi_x$. *Appl. Phys. Lett.*, 82:2245, 2003.
2. S. Francoeur, M. J. Seong, A. Mascarenhas, S. Tixier, M. Adamcyk, and T. Tiedje. Band gap of $GaAs_{1-x}Bi_x$, $0 < x < 3.6\%$. *Appl. Phys. Lett.*, 82:3874, 2003.
3. W. Huang, K. Oe, G. Feng, and M. Yoshimoto. Molecular-beam epitaxy and characteristics of $GaN_yAs_{1-x}Bi_x$. *J. Appl. Phys.*, 98:053505, 2005.
4. K. Alberi, J. Wu, W. Walukiewicz, K. M. Yu, O. D. Dubon, S. P. Watkinsa, C. X. Wang, X. Liu, Y. J. Cho, and J. Furdyna. Valence-band anticrossing in mismatched III-V semiconductor alloys. *Phys. Rev. B*, 75:045203, 2007.
5. M. Usman, C. A. Broderick, A. Lindsay, and E. P. O'Reilly. Tight-binding analysis of the electronic structure of dilute bismide alloys of GaP and GaAs. *Phys. Rev. B*, 84:245202, 2011.
6. Z. Batool, K. Hild, T. J. C. Hosea, X. Lu, T. Tiedje, and S. J. Sweeney. The electronic band structure of GaBiAs/GaAs layers: Influence of strain and band anti-crossing. *J. Appl. Phys.*, 111:113108, 2012.
7. B. Fluegel, S. Francoeur, A. Mascarenhas, S. Tixier, E. C. Young, and T. Tiedje. Giant spin-orbit bowing in $GaAs_{1-x}Bi_x$. *Phys. Rev. Lett.*, 97:067205, 2006.
8. W. Shan, W. Walukiewicz, J. W. Ager, E. E. Haller, J. F. Geisz, D. J. Friedman, J. M. Olson, and S. R. Kurtz. Band anticrossing in GaInNAs alloys. *Phys. Rev. Lett.*, 82:1221, 1999.
9. S. J. Sweeney. Light emitting semiconductor device. Patent no. WO2010149978 A1 (filed: June 25, 2010, issued: December 29, 2010).
10. S. J. Sweeney. Bismide-alloys for higher efficiency infrared semiconductor lasers. In *Proceedings of the 22nd IEEE International Semiconductor Laser Conference (ISLC)*, p. 111, 2010.
11. S. J. Sweeney, Z. Batool, K. Hild, S. R. Jin, and T. J. C. Hosea. The potential role of bismide alloys in future photonic devices. In *Proceedings of the 13th International Conference on Transparent Optical Networks*, p. 1, 2011.
12. C. A. Broderick, M. Usman, S. J. Sweeney, and E. P. O'Reilly. Band engineering in dilute nitride and bismide semiconductor lasers. *Semicond. Sci. Technol.*, 27:094011, 2012.
13. S. Sweeney, S. R. Jin. Bismide–nitride alloys: Promising for ecient light emitting devices in the nearand mid-infrared. *J. Appl. Phys.* 113, 043110, 2013.
14. S. Imhof, C. Bückers, A. Thränhardt, J. Hader, J. V. Moloney, and S. W. Koch. Microscopic theory of the optical properties of Ga(AsBi)/GaAs quantum wells. *Semicond. Sci. Technol.*, 23:125009, 2008.

15. C. A. Broderick, P. E. Harnedy, and E. P. O'Reilly. Theory of the electronic and optical properties of dilute bismide quantum well lasers. *IEEE J. Sel. Top. Quantum Electron.*, 21:1503313, 2015.
16. A. Geižutis, V. Pačebutas, R. Butkutė, P. Svidovskya, V. Strazdienėa, and A. Krotkus. Growth and characterization of UTC photo-diodes containing $GaAs_{1-x}Bi_x$ absorber layer., *Solid State Electron.*, 99:101, 2014.
17. J. J. Lee, J. D. Kim, and M. Razeghi. Growth and characterization of InSbBi for long wavelength infrared photodetectors. *Appl. Phys. Lett.*, 70:3266, 1997.
18. I. C. Sandall, F. Bastiman, B. White, R. Richards, D. Mendes, J. P. R. David, and C. H. Tan. Demonstration of InAsBi photoresponse beyond 3.5 μm. *Appl. Phys. Lett.*, 104:171109, 2014.
19. Y. Gu, Y. G. Zhang, X. Y. Chen, Y. J. Ma, S. P. Xi, B. Du, and H. Li. Nearly lattice-matched short-wave infrared InGaAsBi photodetectors on InP. *Appl. Phys. Lett.*, 108:032102, 2016.
20. C. J. Hunter, F. Bastiman, A. R. Mohmad, R. Richards, J. S. Ng, S. J. Sweeney, and J. P. R. David. Absorption characteristics of $GaAs_{1-x}Bi_x$/GaAs diodes in the near-infrared. *IEEE Photon. Tech. Lett.*, 24:2191, 2012.
21. S. J. Sweeney, K. Hild, and S. R. Jin. The potential of GaAsBiN for multi-junction solar cells. In *Proceedings of the 39th IEEE Photovoltaic Specialists Conference*, p. 2474, 2013.
22. T. Thomas, A. Mellor, N. P. Hylton, M. Führer, D. Alonso-Álvarez, A. Braun, N. J. Ekins-Daukes, J. P. R. David, and S. J. Sweeney. Requirements for a GaAsBi 1 eV sub-cell in a GaAs-based multi-junction solar cell. *Semicond. Sci. Technol.*, 30:094010, 2015.
23. S. J. Sweeney, and K. Hild. Light receiving device. Patent no. 20160149060 (filed: June 19, 2014, issued: May 26, 2016).
24. S. Mazzucato, T. T. Zhang, H. Carrère, D. Lagarde, P. Boonpeng, A. Arnoult, G. Lacoste, A. Balocchi, T. Amand, C. Fontaine, and X. Marie. Electron spin dynamics and g-factor in GaAsBi. *Appl. Phys. Lett.*, 102:252107, 2013.
25. R. A. Simmons, S. R. Jin, S. J. Sweeney, and S. K. Clowes. Enhancement of the Rashba interaction in GaAs/AlGaAs quantum wells due to the incorporation of bismuth. *Appl. Phys. Lett.*, 107:142401, 2015.
26. P. Dongmo, Y. Zhong, P. Attia, C. Bomberger, R. Cheaito, J. F. Ihlefeld, P. E. Hopkins, and J. M. O. Zide. Enhanced room temperature electronic and thermoelectric properties of the dilute bismuthide InGaBiAs. *J. Appl. Phys.*, 112:093710, 2012.
27. M. Usman, C. A. Broderick, Z. Batool, K. Hild, T. J. C. Hosea, S. J. Sweeney, and E. P. O'Reilly. Impact of alloy disorder on the band structure of compressively strained $GaBi_xAs_{1-x}$. *Phys. Rev. B*, 87:115104, 2013.
28. M. Usman, and E. P. O'Reilly. Atomistic tight-binding study of the electronic structure and interband optical transitions in $GaBi_xAs_{1-x}$/GaAs quantum wells. *Appl. Phys. Lett.*, 104:071103, 2014.
29. C. A. Broderick, S. Mazzucato, H. Carrère, T. Amand, H. Makhloufi, A. Arnoult, C. Fontaine, O. Donmez, A. Erol, M. Usman, E. P. O'Reilly, and X. Marie. Anisotropic electron g factor as a probe of the electronic structure of $GaBi_xAs_{1-x}$/GaAs epilayers. *Phys. Rev. B*, 90:195301, 2014.
30. P. Ludewig, N. Knaub, N. Hossain, S. Reinhard, L. Natterman, I. P. Marko, S. R. Jin, K. Hild, S. Chatterjee, W. Stolz, S. J. Sweeney, and K. Volz. Electrical injection Ga(AsBi)/(AlGa)As single quantum well laser. *Appl. Phys. Lett.*, 102:242115, 2013.
31. S. J. Sweeney, A. F. Phillips, A. R. Adams, E. P. O'Reilly, and P. J. A. Thijs. The effect of temperature dependent processes on the performance of 1.5-μm compressively strained InGaAs(P) MQW semiconductor diode lasers. *IEEE Photon. Tech. Lett.*, 10:1076, 1998.
32. S. J. Sweeney, A. R. Adams, M. Silver, E. P. O'Reilly, J. R. Watling, A. B. Walker, and P. J. A. Thijs. Dependence of threshold current on QW position and on pressure in 1.5-μm InGaAs(P) lasers. *Phys. Stat. Sol. B*, 211:525, 1999.
33. A. R. Adams. Band-structure engineering for low-threshold high-efficiency semiconductor lasers. *Electron. Lett.*, 22:249, 1986.

34. Y. Tominaga, K. Oe, and M. Yoshimoto. Low temperature dependence of oscillation wavelength in $GaAs_xBi_{1-x}$ laser by photo-pumping. *Appl. Phys. Express*, 3:062201, 2010.
35. R. Butkutė, A. Geižutis, V. Pačebutas, B. Čechavičius, V. Bukauskas, R. Kundrotas, P. Ludewig, K. Volz, and A. Krotkus. Multi-quantum well Ga(AsBi)/GaAs laser diodes with more than 6% of bismuth. *Electron. Lett.*, 50:1155, 2014.
36. I. P. Marko, S. R. Jin, K. Hild, Z. Batool, Z. L. Bushell, P. Ludewig, W. Stolz, K. Volz, V. Pačebutas, A. Geižutis, A. Krotkus, and S. J. Sweeney. Properties of hybrid MOVPE/MBE grown GaAsBi/GaAs-based near-infrared emitting quantum well lasers. *Semicond. Sci. Technol.*, 30:094008, 2015.
37. P. Ludewig, L. Natterman, W. Stolz, and K. Volz. MOVPE growth mechanisms of dilute bismide III/V alloys. *Semicond. Sci. Technol.*, 30:094017, 2015.
38. R. D. Richards, F. Bastiman, C. J. Hunter, D. F. Mendes, A. R. Mohmad, J. S. Roberts, and J. P. R. David. Molecular beam epitaxy growth of GaAsBi using As_2 and As_4. *J. Cryst. Growth*, 390:120, 2014.
39. I. P. Marko, C. A. Broderick, S. R. Jin, P. Ludewig, W. Stolz, K. Volz, J. M. Rorison, E. P. O'Reilly, and S. J. Sweeney. Optical gain in GaAsBi/GaAs quantum well diode lasers. *Sci. Rep.*, 6:28863, 2016.
40. I. P. Marko, Z. Batool, K. Hild, S. R. Jin, N. Hossain, T. J. C. Hosea, J. P. Petropoulos, Y. Zhong, P. B. Dongmo, J. M. O. Zide, and S. J. Sweeney. Temperature and Bi-concentration dependence of the bandgap and spin–orbit splitting in InGaBiAs/InP semiconductors for mid-infrared applications. *Appl. Phys. Lett.*, 101:221108, 2012.
41. S. R. Jin, and S. J. Sweeney. InGaAsBi alloys on InP for effcient near-and mid-infrared light emitting devices. *J. Appl. Phys*. 114:213103, 2013.
42. C. A. Broderick, M. Usman, and E. P. O'Reilly. 12-band $\mathbf{k}\cdot\mathbf{p}$ model for dilute bismide alloys of (In)GaAs derived from supercell calculations. *Phys. Stat. Sol. B*, 250:773, 2013.
43. G. M. T. Chai, C. A. Broderick, S. R. Jin, J. P. Petropoulous, Y. Zong, P. B. Dongmo, J. M. O. Zide, E. P. O'Reilly, S. J. Sweeney, and T. J. C. Hosea. Experimental and modelling study of InGaBiAs/InP alloys with up to 5.8% Bi, and with $\Delta_{SO} > E_g$. *Semicond. Sci. Technol.*, 30:094015, 2015.
44. C. A. Broderick, W. Xiong, and J. M. Rorison. Theory of InGaBiAs dilute bismide alloys for highly efficient InP-based mid-infrared semiconductor lasers. In *Proceedings of the 16th International Conference on Numerical Simulation of Optoelectronic Devices*, p. 47, 2016.
45. C. A. Broderick, W. Xiong, S. J. Sweeney, E. P. O'Reilly, J. M. Rorison, Theory and design of $In_yGa_{1-y}As_{1-x}Bi_x$ mid-infrared semiconductor lasers: type-I quantum wells for emission beyond 3 μm on InP substrates. Submitted, 2017.
46. C. A. Broderick, S. R. Jin, I. P. Marko, K. Hild, P. Ludewig, Z. L. Bushell, W. Stolz, J. M. Rorison, E. P. O'Reilly, K. Volz, and S. J. Sweeney. $GaAs_{1-x}Bi_x/GaN_yAs_{1-y}$ type-II quantum wells: Novel strain-balanced heterostructures for GaAs-based near- and mid-infrared photonics. *Sci. Rep.* 7, 46371, 2017.
47. A. Janotti, S. H. Wei, and S. B. Zhang. Theoretical study of the effects of isovalent coalloying of Bi and N in GaAs. *Phys. Rev. B*, 65:115203, 2002.
48. P. Carrier, and S. H. Wei. Calculated spin-orbit splitting of all diamondlike and zinc-blende semiconductors: Effects of $p_{1/2}$ local orbitals and chemical trends. *Phys. Rev. B*, 70:035212, 2004.
49. Y. Zhang, A. Mascarenhas, and L. W. Wang. Similar and dissimilar aspects of III–V semiconductors containing Bi versus N. *Phys. Rev. B*, 71:155201, 2005.
50. E. P. O'Reilly, A. Lindsay, P. J. Klar, A. Polimeni, and M. Capizzi. Trends in the electronic structure of dilute nitride alloys. *Semicond. Sci. Technol.*, 24:033001, 2009.
51. P. Vogl, H. P. Hjalmarson, and J. D. Dow. A semi-empirical tight-binding theory of the electronic structure of semiconductors. *J. Phys. Chem. Solids*, 44:365, 1983.
52. E. P. O'Reilly, A. Lindsay, S. Tomić, and M. Kamal-Saadi. Tight-binding and $\mathbf{k}\cdot\mathbf{p}$ models for the electronic structure of Ga(In)NAs and related alloys. *Semicond. Sci. Technol.*, 17:870, 2002.
53. A. Lindsay, and E. P. O'Reilly. Unification of the band-anticrossing and cluster-state models of dilute nitride semiconductor alloys. *Phys. Rev. Lett.*, 93:196402, 2004.
54. C. A. Broderick, M. Usman, and E. P. O'Reilly. Derivation of 12- and 14-band $\mathbf{k}\cdot\mathbf{p}$ Hamiltonians for dilute bismide and bismide–nitride alloys. *Semicond. Sci. Technol.*, 28:125025, 2013.

55. F. A. Trumbore, M. Gershenzon, and D. G. Thomas. Luminescence due to the isoelectronic substitution of bismuth for phosphorus in gallium phosphide. *Appl. Phys. Lett.*, 9:4, 1966.
56. H. X. Deng, J. Li, S. S. Li, H. Peng, J. B. Xia, L. W. Wang, and S. H. Wei. Band crossing in isovalent semiconductor alloys with large size mismatch: First-principles calculations of the electronic structure of Bi and N incorporated GaAs. *Phys. Rev. B*, 82:193204, 2010.
57. R. S. Joshya, A. J. Ptak, R. France, A. Mascarenhas, and R. N. Kini. Resonant state due to Bi in the dilute bismide alloy $GaAs_{1-x}Bi_x$. *Phys. Rev. B*, 90:165203, 2014.
58. K. Alberi, D. A. Beaton, and A. Mascarenhas. Direction observation of the E_- resonant state in $GaAs_{1-x}Bi_x$. *Phys. Rev. B*, 92:241201, 2015.
59. V. Virkkala, V. Havu, F. Tuomisto, and M. J. Puska. Modeling Bi-induced changes in the electronic structure of $GaAs_{1-x}Bi_x$ alloys. *Phys. Rev. B*, 88:235201, 2013.
60. O. Rubel, A. Bokhanchuk, S. J. Ahmed, and E. Assmann. Unfolding the band structure of disordered solids: From bound states to high-mobility Kane fermions. *Phys. Rev. B*, 90:115202, 2014.
61. J. Hader, J. V. Moloney, O. Rubel, S. C. Badescu, S. Johnson, and S. W. Koch. Microscopic modelling of opto-electronic properties of dilute bismide materials for the mid-IR. *Proc. SPIE*, 9767:976709, 2016.
62. M. Mbarki and A. Rebey. First-principles calculation of the physical properties of $GaAs_{1-x}Bi_x$ alloys. *Semicond. Sci. Technol.*, 26:105020, 2011.
63. R. Kudraweic, J. Kopaczek, M. P. Polak, P. Scharoch, M. Gladysiewicz, R. D. Richards, F. Bastiman, and J. P. R. David. Experimental and theoretical studies of band gap alignment in $GaAs_{1-x}Bi_x$/GaAs quantum wells. *J. Appl. Phys.*, 116:233508, 2014.
64. M. P. Polak and P. Scharoch, and R. Kudraweic. First-principles calculations of bismuth-induced changes in the band structure of dilute Ga–V–Bi and In–V–Bi alloys: Chemical trends versus experimental data. *Semicond. Sci. Technol.*, 30:094001, 2015.
65. M. Ferhat and A. Zaoui. Structural and electronic properties of III–V bismuth compounds. *Phys. Rev. B*, 73:115107, 2006.
66. A. Belabbes, A. Zaoui, and M. Ferhat. Lattice dynamics study of bismuth III–V compounds. *J. Phys.: Condens. Matter*, 20:415221, 2008.
67. H. Jacobsen, B. Puchala, T. F. Keuch, and D. Morgan. Ab initio study of the strain dependent thermodynamics of Bi doping in GaAs. *Phys. Rev. B*, 86:085207, 2012.
68. G. Luo, S. Yang, J. Li, M. Arjmand, I. Szlufarska, A. S. Brown, T. F. Keuch, and D. Morgan. First-principles studies on molecular beam epitaxy growth of $GaAs_{1-x}Bi_x$. *Phys. Rev. B*, 92:035415, 2015.
69. M. P. J. Punkinnen, P. Laukkanen, M. Kuzmin, H. Levämäki, J. Løang, M. Tuominen, M. Yasir, J. Dahl, S. Lu, E. K. Delczeg-Czirjak, L. Vitos, and K. Kokko. Does Bi form clusters in $GaAs_{1-x}Bi_x$ alloys. *Semicond. Sci. Technol.*, 29:115007, 2014.
70. M. P. J. Punkinnen, A. Lahti, P. Laukkanen, M. Kuzmin, M. Tuominen, M. Yasir, J. Dahl, J. Makela, H. L. Zhang, L. Vitos, and K. Kokko. Thermodynamics of pseudobinary $GaAs_{1-x}Bi_x$ ($0 \leq x \leq 1$) alloys studied by different exchange-correlation functionals, special quasi-random structures and Monte Carlo simulations. *Computat. Condens. Matter*, 5:7, 2015
71. P. N. Keating. Effect of invariance requirements on the elastic strain energy of crystals with application to the diamond structure. *Phys. Rev.*, 145:637, 1966.
72. O. L. Lazarenkova, P. von Allmen, F. Oyafuso, S. Lee, and G. Klimeck. Effect of anharmonicity of the strain energy on band offsets in semiconductor nanostructures. *Appl. Phys. Lett.*, 85:4193, 2004.
73. J. C. Slater and G. F. Koster. Simplified LCAO method for the periodic potential problem. *Phys. Rev.*, 94:1498, 1954.
74. J. Yoshida, T. Kita, O. Wada, and K. Oe. Temperature dependence of $GaAs_{1-x}Bi_x$ band gap studied by photo-reflectance spectroscopy. *Jpn. J. Appl. Phys.*, 42:371, 2003.
75. G. Hendorfer and J. Schneider. g-factor and effective mass anisotropies in pseudomorphic strained layers. *Semicond. Sci. Technol.*, 6:595, 1991.
76. P. L. Jeune, D. Robart, X. Marie, T. Amand, M. Brousseau, J. Barrau, V. Kalevich, and D. Rodichev. Anisotropy of the electron Landé g-factor in quantum wells. *Semicond. Sci. Technol.*, 12:380, 1997.

77. M. Oestreich and S. Hallstein, W. W. Rühle. Spin quantum beats in semiconductors. *IEEE J. Sel. Top. Quantum Electron.*, 2:747, 1996.
78. A. P. Heberle, W. W. Rühle, and K. Ploog. Quantum beats of electron Larmor precession in GaAs wells. *Phys. Rev. Lett.*, 72:3887, 1994.
79. C. A. Broderick, P. E. Harnedy, P. Ludewig, Z. L. Bushell, K. Volz, R. J. Manning, and E. P. O'Reilly. Determination of type-I band offsets in $GaBi_xAs_{1-x}$ quantum wells using polarisation-resolved photovoltage spectroscopy and 12-band **k·p** calculations. *Semicond. Sci. Technol.*, 30:094009, 2015.
80. T. B. Bahder. 8-Band **k·p** model of strained zinc-blende crystals. *Phys. Rev. B*, 41:11992, 1990.
81. G. L. Bir, G. E. Pikus. *Symmetry and Strain-Induced Effects in Semiconductors* New York: Wiley, 1974.
82. C. G. Van de Walle. Band lineups and deformation potentials in the model-solid theory. *Phys. Rev. B*, 39:1871, 1989.
83. M. P. C. M. Krijn. Heterojunction band offsets and effective masses in III–V quaternary alloys. *Semicond. Sci. Technol.*, 6:27, 1991.
84. I. Vurgaftman, J. R. Meyer, and L. R. Ram-Mohan. Band parameters for III–V compound semiconductors and their alloys. *J. Appl. Phys.*, 89:5815, 2001.
85. M. Ehrhardt and T. Koprucki, editors. *Multiband Effective Mass Approximations: Advanced Mathematical Models and Numerical Techniques*. Berlin: Springer, 2014.
86. S. B. Healy, and E. P. O'Reilly. Influence of electrostatic confinement on optical gain in GaInNAs quantum-well lasers. *IEEE J. Quant. Electron.*, 42:608, 2006.
87. A. T. Meney, B. Gonul, and E. P. O'Reilly. Evaluation of various approximations used in the envelope-function method. *Phys. Rev. B*, 50:10893, 1994.
88. F. Szmulowicz. Derivation of a general expression for the momentum matrix elements within the envelope-function approximation. *Phys. Rev. B*, 51:1613, 1995.
89. C. S. Chang, S. L. Chuang, J. R. Minch, W. C. W. Fang, Y. K. Chen, and T. Tanbun-Ek. Amplified spontaneous emission spectroscopy in strained quantum-well lasers. *IEEE J. Sel. Top. Quantum Electron.*, 1:1100, 1995.
90. S. Tomić, E. P. O'Reilly, R. Fehse, S. J. Sweeney, A. R. Adams, A. D. Andreev, S. A. Choulis, T. J. C. Hosea, and H. Riechert. Theoretical and experimental analysis of 1.3-μm InGaAsN/GaAs lasers. *IEEE J. Sel. Top. Quantum Electron.*, 9:1228, 2003.
91. K. Kawano and T. Kitoh. *Introduction to Optical Waveguide Analysis: Solving Maxwell's Equation and the Schrödinger Equation*. New York, NY: Wiley, 2001.
92. M. Gladysiewicz, R. Kudrawiec, and M. S. Wartak. 8-Band and 14-band **k·p** modeling of electronic band structure and material gain in Ga(In)AsBi quantum wells grown on GaAs and InP substrates. *J. Appl. Phys.*, 118:055702, 2015.
93. E. P. O'Reilly and A. R. Adams. Band-structure engineering in strained semiconductor lasers. *IEEE J. Quant. Electron.*, 30:366, 1994.
94. S. L. Chuang. *Physics of Photonic Devices*. New York, NY: Wiley, 2009.
95. C. A. Broderick, I. P. Marko, J. M. Rorison, S. J. Sweeney, and E. P. O'Reilly. GaAs-based dilute bismide semiconductor lasers: Theory vs. experiment. In *Proceedings of the 16th International Conference on Numerical Simulation of Optoelectronic Devices*, p. 209, 2016.
96. I. P. Marko, P. Ludewig, Z. L. Bushell, S. R. Jin, K. Hild, Z. Batool, S. Reinhard, L. Natterman, W. Stolz, K. Volz, and S. J. Sweeney. Physical properties and optimization of GaBiAs/(Al)GaAs based near-infrared laser diodes grown by MOVPE with up to 4.4% Bi. *J. Phys. D: Appl. Phys.*, 47:345103, 2014.
97. A. Ghiti, M. Silver, and E. P. O'Reilly. Low threshold current and high differential gain in ideal tensile- and compressive-strained quantum-well lasers. *J. Appl. Phys.*, 71:4626, 1992.
98. M. Silver and E. P. O'Reilly. Optimization of long wavelength InGaAsP strained quantum-well lasers. *IEEE J. Quant. Electron.*, 31:1193, 1995.
99. S. Tomić and E. P. O'Reilly. Optimization of material parameters in 1.3-μm InGaAsN-GaAs lasers. *IEEE Photon. Tech. Lett.*, 15:6, 2003.

100. P. Blood, G. M. Lewis, P. M. Smowton, H. Summers, J. Thomson and J. Lutti. Characterization of semiconductor laser gain media by the segmented contact method. *IEEE J. Sel. Top. Quantum Electron.*, 9:1275, 2003.
101. D. J. Bossert and D. Gallant. Gain, refractive index, and α-parameter in InGaAs–GaAs SQW broad-area lasers. *IEEE Photon. Tech. Lett.*, 8:322, 1996.
102. A. Bauer, K. Rößner, T. Lehnhardt, M. Kamp, S. Höfling, L. Worschech, and A. Forchel. Mid-infrared semiconductor heterostructure lasers for gas sensing applications. *Semicond. Sci. Technol.*, 26:014032, 2011.
103. E. Tournié and A. N. Baranov. Mid-infrared semiconductor lasers: A review. *Semiconduct. Semimet.*, 86:183, 2012.
104. G. Feng, M. Yoshimoto, K. Oe, A. Chayahara, and Y. Horino. New III–V semiconductor InGaAsBi alloy grown by molecular beam epitaxy. *Jpn. J. Appl. Phys.*, 44:L1161, 2005.
105. G. Feng, K. Oe, and M. Yoshimoto. Bismuth containing III–V quaternary alloy InGaAsBi grown by MBE. *Phys. Stat. Sol. A*, 203:2760, 2006.
106. J. P. Petropoulos, Y. Zhong, and J. M. O. Zide. Optical and electrical characterization of InGaAsBi for use as a mid-infrared optoelectronic material. *Appl. Phys. Lett.*, 99:031110, 2011.
107. Y. Zhong, P. B. Dongmo, J. P. Petropoulos, and J. M. O. Zide. Effects of molecular beam epitaxy growth conditions on composition and optical properties of $In_xGa_{1-x}Bi_yAs_{1-y}$. *Appl. Phys. Lett.*, 100:112110, 2012.
108. J. Devenson, V. Pačebutas, R. Butkutė, A. Baranov, and A. Krotkus. Structure and optical properties of InGaAsBi with up to 7% bismuth. *Appl. Phys. Express*, 5:015503, 2012.
109. S. Zhou, M. Qi, L. Ai, A. Xu, and S. Wang. Effects of buffer layer preparation and Bi concentration on InGaAsBi layers grown by gas source molecular beam epitaxy. *Semicond. Sci. Technol.*, 30:125001, 2015.
110. M. Yoshimoto, W. Huang, Y. Takehara, J. Saraie, A. Chayahara, Y. Horino, and K. Oe. New semiconductor GaNAsBi alloy grown by molecular beam epitaxy. *Jpn. J. Appl. Phys.*, 43:L845, 2004.
111. S. Tixier, S. E. Webster, E. C. Young, T. Tiedje, S. Francoeur, A. Mascarenhas, P. Wei, and F. Schiettekatte. Band gaps for the dilute quaternary alloys $GaN_xAs_{1-x-y}Bi_y$ and $Ga_{1-y}In_yN_xAs_{1-x}$. *Appl. Phys. Lett.*, 86:112113, 2005.
112. Z. L. Bushell, P. Ludewig, N. Knaub, Z. Batool, K. Hild, W. Stolz, S. J. Sweeney, and K. Volz. Growth and characterisation of Ga(AsNBi) alloy by metal-organic vapour phase epitaxy. *J. Cryst. Growth*, 396:79, 2014.
113. M. Yoshimoto, W. Huang, G. Feng, Y. Tanaka, and K. Oe. Molecular-beam epitaxy of GaNAsBi layer for temperature-insensitive wavelength emission. *J. Cryst. Growth*, 301–302:975, 2007.
114. J. Hwang, J. D. Phillips. Band structure of strain-balanced GaAsBi/GaAsN superlattices on GaAs. *Phys. Rev. B*, 83:195327, 2011.
115. W. Pan, L. Zhang, L. Zhu, Y. Li, X. Chen, X. Wu, F. Zhang, J. Shao, and S. Wang. Optical properties and band bending in InGaAs/GaAsBi/InGaAs type-II quantum well grown by gas source molecular beam epitaxy. *J. Appl. Phys.*, 120:105702, 2016.
116. J. F. Klem, O. Blum, S. R. Kurtz, I. J. Fritz, and K. D. Choquette. GaAsSb/InGaAs type-II quantum wells for long-wavelength lasers on GaAs substrates. *J. Vac. Sci. Technol. B*, 18:1605, 2000.
117. J. Y. Yeh, L. J. Mawst, A. A. Khandekar, T. F. Kuech, I. Vurgaftman, J. R. Meyer, and N. Tansu. Long wavelength emission of InGaAsN/GaAsSb type II "W" quantum wells. *Appl. Phys. Lett.*, 88:051115, 2006.
118. C. Berger, C. Möller, P. Hens, C. Fuchs, W. Stolz, S. W. Koch, A. R. Perez, J. Hader, and J. V. Moloney. Novel type-II material system for laser applications in the near-infrared regime. *AIP Adv.*, 5:047105, 2015.
119. K. Forghani, Y. Guan, M. Losurdo, G. Luo, D. Morgan, S. E. Babcock, A. S. Brown, L. J. Mawst, and T. F. Kuech. $GaAs_{1-y-z}P_yBi_z$, an alternative reduced band gap alloy system lattice-matched to GaAs. *Appl. Phys. Lett.*, 105:111101, 2014.

120. H. Kim, K. Forghani, Y. Guan, G. Luo, A. Anand, D. Morgan, T. F. Keuch, L. J. Mawst, Z. R. Lingley, B. J. Foran, and Y. Sin. Strain-compensated $GaAs_{1-y-z}P_y/GaAs_{1-x}Bi_x/GaAs_{1-y}P_y$ quantum wells for laser applications. *Semicond. Sci. Technol.*, 30:094011, 2015.
121. Y. Gu, Y. G. Zhang, X. Y. Chen, Y. J. Ma, S. P. Xi, B. Du, W. Y. Ji, and Y. H. Shi. Metamorphic $InAs_{1-x}Bi_x/In_{0.83}Al_{0.17}As$ quantum well structures on InP for mid-infrared emission. *Appl. Phys. Lett.*, 109:122102, 2016.
122. M. K. Rajpalke, W. M. Linhart, M. Birkett, K. M. Yu, D. O. Scanlon, J. Buckeridge, T. S. Jones, M. J. Ashwin, and T. D. Veal. Growth and properties of GaSbBi alloys. *Appl. Phys. Lett.*, 103:142106, 2013.

III

Nanostructures

11

Quantum Wells

Seoung-Hwan Park
and
Doyeol Ahn

11.1 Introduction

A quantum well (QW) is a heterostructure in which one thin-well layer is surrounded by two barrier layers. This layer is so thin that both electrons and holes are quantized. The electronic and the optical properties of quantized states offer new opportunities in developing practical devices, such as QW infrared photodetectors, quantum cascade lasers, all-optical switches, modulators, and many others [1–4]. Hence, it is very important to obtain eigenvalues and wave functions for the design of the active region in these optoelectronic devices based on QW structures. In this chapter, we review theoretical formalism to calculate eigenvalues and wave functions of (001)-oriented zinc-blende and (0001)-oriented wurtzite QW structures [5–10]. We block diagonalize zinc-blende and wurtzite Luttinger–Kohn 6×6 Hamiltonians for the valence bands to two 3×3 Hamiltonians, which have analytical solutions for eigenvalues and eigenvectors. We derive several important forms such as interband optical matrix elements and optical gains [11–15]. Also, as a numerical example, we calculate eigenvalues and wave functions for zinc-blende and wurtzite Hamiltonians using a finite-difference method (FDM) [4]. On the basis of this information, we discuss crystal orientation effects on electronic and optical properties of strained zinc-blende and wurtzite QW structures, including the Hamiltonian for nonpolar wurtzite QW structures.

11.2 Band Structures of Bulk Semiconductors

11.2.1 Zinc-Blende Hamiltonian of the (001) Orientation

11.2.1.1 6 × 6 Hamiltonian for the Valence Band

The Luttinger–Kohn Hamiltonian for the valence band of the (001)-oriented zinc-blende semiconductor is given by [2,4,5]

$$H^{\mathrm{LK}}(\vec{k},\bar{\varepsilon}) = -\begin{pmatrix} P+Q & -S & R & 0 & -S/\sqrt{2} & \sqrt{2}R \\ -S^{\dagger} & P-Q & 0 & R & -\sqrt{2}Q & \sqrt{3/2}S \\ R^{\dagger} & 0 & P-Q & 0 & \sqrt{3/2}S^{\dagger} & \sqrt{2}Q \\ 0 & R^{+} & S^{\dagger} & P+Q & -\sqrt{2}R^{\dagger} & -S^{\dagger}/\sqrt{2} \\ -S^{\dagger}/\sqrt{2} & 0 & \sqrt{3/2}S & -\sqrt{2}R & P+\Delta & 0 \\ \sqrt{2}R^{\dagger} & \sqrt{3/2}S^{\dagger} & \sqrt{2}Q^{\dagger} & -S/\sqrt{2} & 0 & P+\Delta \end{pmatrix}, \tag{11.1}$$

where

$$\begin{aligned} &P = P_k + P_\varepsilon, \quad Q = Q_k + Q_{\varepsilon,} \\ &R = R_k + R_\varepsilon, \quad S = S_k + S_\varepsilon, \\ &P_k = \left(\frac{\hbar^2}{2m_0}\right)\gamma_1(k_x^2 + k_x^2 + k_z^2), \quad Q_k = \left(\frac{\hbar^2}{2m_0}\right)\gamma_2(k_z^2 + k_x^2 - 2k_z^2) \\ &R_k = \left(\frac{\hbar^2}{2m_0}\right)\sqrt{3}[-\gamma_2(k_x^2 - k_x^2) + 2i\gamma_3 k_x k_y], \quad S_k = \left(\frac{\hbar^2}{2m_0}\right)2\sqrt{3}\gamma_3(k_x - ik_y)k_z, \\ &P_\varepsilon = -a_v(\varepsilon_{xx} + \varepsilon_{yy} + \varepsilon_{zz}), \quad Q_\varepsilon = -\frac{b}{2}(\varepsilon_{xx} + \varepsilon_{yy} - 2\varepsilon_{zz}), \\ &R_\varepsilon = \frac{\sqrt{3}}{2}b(\varepsilon_{xx} - \varepsilon_{yy}) - id\varepsilon_{xy}, \quad S_\varepsilon = -d(\varepsilon_{zx} - i\varepsilon_{yz}), \end{aligned} \tag{11.2}$$

and $\bar{\varepsilon} = (\varepsilon_{ij})$ for $i, j = x, y, z$ is a symmetric strain tensor; γ_1, γ_2, and γ_3 are the Luttinger parameters; a_v, b, and d are the Bir–Pikus deformation potentials; Δ is the spin–orbit split-off energy; m_0 is the free electron mass; $\hbar$ is Planck's constant divided by 2π; and k_i is the wave vector. The superscript † means taking both transpose (~) and complex conjugate (*). In the Hamiltonian H, we restricted ourselves to the biaxial strain case for simplicity, namely,

$$\varepsilon_{xx} = \varepsilon_{yy} \neq \varepsilon_{zz}, \quad \varepsilon_{ij} = 0 \quad \text{for} \quad i \neq j. \tag{11.3}$$

For the case of a strained-layer semiconductor pseudomorphically grown on a (001)-oriented substrate,

$$\varepsilon_{xx} = \varepsilon_{yy} = \frac{a_s^z - a_l^z}{a_l^z}, \quad \varepsilon_{zz} = -\frac{2C_{12}}{C_{11}}\varepsilon_{xx}, \tag{11.4}$$

where a_s^z and a_l^z are the lattice constants of the substrate (s) and the layer (l) material, and C_{11} and C_{12} are the stiffness constants for the zinc-blende structure. The bases for the Hamiltonian are

$$|u_1\rangle = \left|\frac{3}{2}, \frac{3}{2}\right\rangle = -\frac{1}{\sqrt{2}}\left|(X + iY)\uparrow\right\rangle,$$

$$
\begin{aligned}
\left|u_2\right\rangle &= \left|\frac{3}{2},\frac{1}{2}\right\rangle = \frac{1}{\sqrt{6}}\left|-(X-iY)\downarrow +2Z\uparrow\right\rangle,\\
\left|u_3\right\rangle &= \left|\frac{3}{2},-\frac{1}{2}\right\rangle = \frac{1}{\sqrt{6}}\left|(X-iY)\uparrow +2Z\downarrow\right\rangle,\\
\left|u_4\right\rangle &= \left|\frac{3}{2},-\frac{3}{2}\right\rangle = \frac{1}{\sqrt{2}}\left|(X-iY)\downarrow\right\rangle,\\
\left|u_5\right\rangle &= \left|\frac{1}{2},\frac{1}{2}\right\rangle = \frac{1}{\sqrt{3}}\left|(X+iY)\downarrow +Z\uparrow\right\rangle,\\
\left|u_6\right\rangle &= \left|\frac{1}{2},-\frac{1}{2}\right\rangle = \frac{1}{\sqrt{3}}\left|(X-iY)\uparrow -Z\downarrow\right\rangle.
\end{aligned}
\tag{11.5}
$$

11.2.1.2 Block-Diagonalized 3 × 3 Hamiltonian

Under the axial approximation [2], we write the R_k term

$$
\begin{aligned}
R_k &= -\left(\frac{\hbar^2}{2m_0}\right)\sqrt{3}\left[\frac{\gamma_2+\gamma_2}{2}(k_x-ik_x)^2+\frac{\gamma_2-\gamma_3}{2}(k_x+ik_x)^2\right]\\
&\cong -\left(\frac{\hbar^2}{2m_0}\right)\sqrt{3}\left[\bar{\gamma}(k_x-ik_x)^2\right],
\end{aligned}
\tag{11.6}
$$

where $\bar{\gamma}=\frac{\gamma_2+\gamma_3}{2}$. In this approximation, we assume $\gamma_2\cong\gamma_3$ in the R_k term only, whereas we still use γ_2 and γ_3 in the other terms. When we define the angle φ by

$$
k_x+ik_x = k_{||}e^{i\varphi},
\tag{11.7}
$$

we can write

$$
R_k = R_\rho e^{-2i\varphi},\quad S_k = S_\rho e^{-i\varphi},\quad R_\rho = -\left(\frac{\hbar^2}{2m_0}\right)\sqrt{3}\bar{\gamma}k_{||}^2,\quad S_\rho = \left(\frac{\hbar^2}{2m_0}\right)2\sqrt{3}\gamma_3 k_{||}k_z,
\tag{11.8}
$$

where $k_{||}=\sqrt{k_x^2+k_x^2}$. Then, the 6×6 Hamiltonian can be block-diagonalized into two 3×3 Hamiltonians by using the transformation matrix U

$$
H = UH^{\mathrm{LK}}U^{+}
$$

$$
= -\begin{pmatrix}
P+Q & -R_\rho-iS_\rho & -\sqrt{2}R_\rho+i\sqrt{\frac{1}{2}}S_\rho & 0 & 0 & 0\\
-R_\rho+iS_\rho & P-Q & \sqrt{2}Q+i\sqrt{\frac{3}{2}}S_\rho & 0 & 0 & 0\\
-\sqrt{2}R_\rho-i\sqrt{\frac{1}{2}}S_\rho & \sqrt{2}Q-i\sqrt{\frac{3}{2}}S_\rho & P+\Delta & 0 & 0 & 0\\
0 & 0 & 0 & P+Q & -R_\rho+iS_\rho & -\sqrt{2}R_\rho-i\sqrt{\frac{1}{2}}S_\rho\\
0 & 0 & 0 & -R_\rho-iS_\rho & P-Q & \sqrt{2}Q-i\sqrt{\frac{3}{2}}S_\rho\\
0 & 0 & 0 & -\sqrt{2}R_\rho+i\sqrt{\frac{1}{2}}S_\rho & \sqrt{2}Q+i\sqrt{\frac{3}{2}}S_\rho & P+\Delta
\end{pmatrix}
\tag{11.9}
$$

and

$$U = \begin{pmatrix} \alpha & 0 & 0 & i\alpha^* & 0 & 0 \\ 0 & -i\beta & -\beta^* & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & i\beta & -\beta^* \\ \alpha & 0 & 0 & -i\alpha^* & 0 & 0 \\ 0 & i\beta & -\beta^* & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -i\beta & -\beta^* \end{pmatrix}, \tag{11.10}$$

where

$$\alpha = \frac{1}{\sqrt{2}} e^{i\frac{3}{2}\varphi}, \quad \beta = \frac{1}{\sqrt{2}} e^{i\frac{1}{2}\varphi}. \tag{11.11}$$

A new basis set is given by using the basis transformation, $|i> = \sum T_{ij}|u_j>$, where $T = U^*$. That is,

$$\begin{aligned}
|1\rangle &= -\frac{1}{2}\Big|(X+iY)\uparrow\Big\rangle e^{-i\frac{3}{2}\varphi} - i\frac{1}{2}\Big|(X-iY)\downarrow\Big\rangle e^{-i\frac{3}{2}\varphi} \\
|2\rangle &= i\frac{1}{2\sqrt{3}}\Big|-(X-iY)\downarrow +2Z\uparrow\Big\rangle e^{-i\frac{1}{2}\varphi} - \frac{1}{2\sqrt{3}}\Big|(X-iY)\uparrow +2Z\downarrow\Big\rangle e^{i\frac{1}{2}\varphi} \\
|3\rangle &= -i\frac{1}{\sqrt{6}}\Big|(X+iY)\downarrow +Z\uparrow\Big\rangle e^{-i\frac{1}{2}\varphi} - \frac{1}{\sqrt{6}}\Big|(X-iY)\uparrow -Z\downarrow\Big\rangle e^{i\frac{1}{2}\varphi} \\
|4\rangle &= -\frac{1}{2}\Big|(X+iY)\uparrow\Big\rangle e^{-i\frac{3}{2}\varphi} + i\frac{1}{2}\Big|(X-iY)\downarrow\Big\rangle e^{i\frac{3}{2}\varphi} \\
|5\rangle &= -i\frac{1}{2\sqrt{3}}\Big|-(X-iY)\downarrow +2Z\uparrow\Big\rangle e^{-i\frac{1}{2}\varphi} - \frac{1}{2\sqrt{3}}\Big|(X-iY)\uparrow +2Z\downarrow\Big\rangle e^{i\frac{1}{2}\varphi} \\
|6\rangle &= i\frac{1}{\sqrt{6}}\Big|(X+iY)\downarrow +Z\uparrow\Big\rangle e^{-i\frac{1}{2}\varphi} - \frac{1}{\sqrt{6}}\Big|(X-iY)\uparrow -Z\downarrow\Big\rangle e^{i\frac{1}{2}\varphi}.
\end{aligned} \tag{11.12}$$

11.2.2 Wurtzite Hamiltonian of the (0001) Orientation

11.2.2.1 6 × 6 Hamiltonian for the Valence Band

The *c*-plane Hamiltonian for the valence band of the (0001)-oriented wurtzite semiconductor in $\{|1\rangle, |2\rangle, |3\rangle, |4\rangle, |5\rangle, |6\rangle\}$ bases is given by [4,6]

$$H(\vec{k}, \bar{\bar{\varepsilon}}) = \begin{pmatrix} F & -K^* & -H^* & 0 & 0 & 0 \\ -K & G & H & 0 & 0 & \Delta \\ -H & H* & \lambda & 0 & \Delta & 0 \\ 0 & 0 & 0 & F & -K & H \\ 0 & 0 & \Delta & -K* & G & -H^* \\ 0 & \Delta & 0 & H^* & -H & \lambda \end{pmatrix}, \tag{11.13}$$

where

$$\begin{aligned}
F &= \Delta_1 + \Delta_2 + \lambda + \vartheta, \\
G &= \Delta_1 - \Delta_2 + \lambda + \vartheta, \\
\lambda &= \frac{\hbar^2}{2m_0}\left[A_1 k_z^2 + A_2\left(k_x^2 + k_y^2\right)\right] + D_1\varepsilon_{zz} + D_2\left(\varepsilon_{xx} + \varepsilon_{yy}\right),
\end{aligned}$$

$$\vartheta = \frac{\hbar^2}{2m_0}\left[A_3 k_z^2 + A_4\left(k_x^2 + k_y^2\right)\right] + D_3\varepsilon_{zz} + D_4\left(\varepsilon_{xx} + \varepsilon_{yy}\right), \tag{11.14}$$

$$K = \frac{\hbar^2}{2m_0}A_5\left(k_x + ik_y\right)^2 + D_5\left(\varepsilon_{xx} - \varepsilon_{yy}\right),$$

$$H = \frac{\hbar^2}{2m_0}A_6\left(k_x + ik_y\right)k_z + D_6\left(\varepsilon_{xz} + i\varepsilon_{yz}\right),$$

$$\Delta = \sqrt{2}\Delta_3,$$

and the A_i's are the valence-band effective-mass parameters analogous to the Luttinger parameters for the zinc-blende semiconductors, the D_i's are the deformation potentials for wurtzite semiconductors, Δ_1 is the crystal-field split energy, and Δ_2 and Δ_3 are the spin–orbit interaction energies. Under the cubic approximation [6,10], the following relations hold for the parameters A_i's and D_i's:

$$A_1 - A_2 = -A_3 = 2A_4, \quad A_3 + 4A_5 = \sqrt{2}A_6, \quad D_1 - D_2 = -D_3 = 2D_4, \quad D_3 + 4D_5 = \sqrt{2}D_6. \tag{11.15}$$

Here, we restricted ourselves to the biaxial strain case for simplicity, namely,

$$\varepsilon_{xx} = \varepsilon_{yy} \neq \varepsilon_{zz}, \quad \varepsilon_{ij} = 0 \quad \text{for} \quad i \neq j. \tag{11.16}$$

For the case of a strained-layer semiconductor pseudomorphically grown on a (0001)-oriented substrate,

$$\varepsilon_{xx} = \varepsilon_{yy} = \frac{a_s^w - a_l^w}{a_l^w}, \quad \varepsilon_{zz} = -\frac{2C_{13}}{C_{33}}\varepsilon_{xx}, \tag{11.17}$$

where a_s^w and a_l^w are the lattice constants of the substrate (s) and the layer (l) material, and C_{13} and C_{33} are the stiffness constants for the wurtzite structure. The bases for the Hamiltonian are

$$\begin{aligned}
\left|u_1\right\rangle &= -\frac{1}{\sqrt{2}}\left|(X + iY)\uparrow\right\rangle, \\
\left|u_2\right\rangle &= \frac{1}{\sqrt{2}}\left|(X - iY)\uparrow\right\rangle, \\
\left|u_3\right\rangle &= \left|Z\uparrow\right\rangle, \\
\left|u_4\right\rangle &= \frac{1}{\sqrt{2}}\left|(X - iY)\downarrow\right\rangle, \\
\left|u_5\right\rangle &= -\frac{1}{\sqrt{2}}\left|(X + iY)\downarrow\right\rangle, \\
\left|u_6\right\rangle &= \left|Z\downarrow\right\rangle.
\end{aligned} \tag{11.18}$$

11.2.2.2 Block-Diagonalized 3 × 3 Hamiltonian

When we define the angle φ by

$$k_x + ik_x = k_{||}e^{i\varphi}, \tag{11.19}$$

we can write

$$K = K_t e^{2i\varphi}, \quad H = H_t e^{i\varphi}, \quad K_t = \left(\frac{\hbar^2}{2m_0}\right) A_5 k_{||}^2, \quad H_t = \left(\frac{\hbar^2}{2m_0}\right) A_6 k_{||} k_z. \tag{11.20}$$

Then, the 6×6 Hamiltonian can be block-diagonalized into two 3×3 Hamiltonians following a similar procedure to that of the zinc-blende structure by using the transformation matrix U [4,6,7]

$$H' = UHU^{+} = \begin{pmatrix} F & K_t & -iH_t & 0 & 0 & 0 \\ K_t & G & \Delta - iH_t & 0 & 0 & 0 \\ iH_t & \Delta + iH_t & \lambda & 0 & 0 & 0 \\ 0 & 0 & 0 & F & K_t & iH_t \\ 0 & 0 & 0 & K_t & G & \Delta + iH_t \\ 0 & 0 & 0 & -iH_t & \Delta - iH_t & \lambda \end{pmatrix}, \tag{11.21}$$

and

$$U = \begin{pmatrix} \alpha & 0 & 0 & \alpha^* & 0 & 0 \\ 0 & \beta^* & 0 & 0 & \beta & 0 \\ 0 & 0 & \beta & 0 & 0 & \beta^* \\ \alpha & 0 & 0 & -\alpha^* & 0 & 0 \\ 0 & \beta^* & 0 & 0 & -\beta & 0 \\ 0 & 0 & -\beta & 0 & 0 & \beta^* \end{pmatrix} \tag{11.22}$$

where

$$\alpha = \frac{1}{\sqrt{2}} e^{i\left(\frac{3}{4}\pi + \frac{3}{2}\varphi\right)}, \quad \beta = \frac{1}{\sqrt{2}} e^{i\left(\frac{1}{4}\pi + \frac{1}{2}\varphi\right)}. \tag{11.23}$$

Also, a new basis set is given by using the basis transformation, $|i> = \sum T_{ij} |u_j>$, where $T = U^*$. That is,

$$\begin{aligned}
|1\rangle &= -\frac{1}{\sqrt{2}} \left|(X + iY) \uparrow\right\rangle \frac{1}{\sqrt{2}} e^{i\left(\frac{3}{4}\pi + \frac{3}{2}\varphi\right)} + \frac{1}{\sqrt{2}} \left|(X - iY) \downarrow\right\rangle \frac{1}{\sqrt{2}} e^{-i\left(\frac{3}{4}\pi + \frac{3}{2}\varphi\right)} \\
|2\rangle &= \frac{1}{\sqrt{2}} \left|(X - iY) \uparrow\right\rangle \frac{1}{\sqrt{2}} e^{-i\left(\frac{1}{4}\pi + \frac{1}{2}\varphi\right)} - \frac{1}{\sqrt{2}} \left|(X + iY) \downarrow\right\rangle \frac{1}{\sqrt{2}} e^{i\left(\frac{1}{4}\pi + \frac{1}{2}\varphi\right)} \\
|3\rangle &= \left|Z \uparrow\right\rangle \frac{1}{\sqrt{2}} e^{i\left(\frac{1}{4}\pi + \frac{1}{2}\varphi\right)} + \left|Z \downarrow\right\rangle e^{i\frac{1}{2}\varphi} \frac{1}{\sqrt{2}} e^{-i\left(\frac{1}{4}\pi + \frac{1}{2}\varphi\right)} \\
|4\rangle &= -\frac{1}{\sqrt{2}} \left|(X + iY) \uparrow\right\rangle \frac{1}{\sqrt{2}} e^{i\left(\frac{3}{4}\pi + \frac{3}{2}\varphi\right)} - \frac{1}{\sqrt{2}} \left|(X - iY) \downarrow\right\rangle \frac{1}{\sqrt{2}} e^{-i\left(\frac{3}{4}\pi + \frac{3}{2}\varphi\right)} \\
|5\rangle &= \frac{1}{\sqrt{2}} \left|(X - iY) \uparrow\right\rangle \frac{1}{\sqrt{2}} e^{-i\left(\frac{1}{4}\pi + \frac{1}{2}\varphi\right)} + \frac{1}{\sqrt{2}} \left|(X + iY) \downarrow\right\rangle \frac{1}{\sqrt{2}} e^{i\left(\frac{1}{4}\pi + \frac{1}{2}\varphi\right)} \\
|6\rangle &= -\left|Z \uparrow\right\rangle \frac{1}{\sqrt{2}} e^{i\left(\frac{1}{4}\pi + \frac{1}{2}\varphi\right)} + \left|Z \downarrow\right\rangle e^{i\frac{1}{2}\varphi} \frac{1}{\sqrt{2}} e^{-i\left(\frac{1}{4}\pi + \frac{1}{2}\varphi\right)}.
\end{aligned} \tag{11.24}$$

11.3 Band Structures of Strained-Layer QW

11.3.1 Zinc-Blende Semiconductor

Here, we consider a strained-layer QW structure, assuming that the growth direction is along the z-axis and the strain caused by lattice mismatch is entirely elastically accommodated in the QW.

11.3.1.1 Conduction Band

For the unstrained QW, the effective mass theory for the conduction band is obtained from the dispersion relation

$$E(k) = \frac{\hbar^2 k^2}{2m_{\mathrm{e}}}, \tag{11.25}$$

where the effective mass of the electron in the conduction band is $m_{\mathrm{e}} = m_{\mathrm{b}}$ in the barrier region and $m_{\mathrm{e}} = m_{\mathrm{w}}$ in the QW. The potential for the unstrained QW is given by

$$V_{\mathrm{c}}(z) = \begin{cases} \Delta E_{\mathrm{c}} & |z| > \frac{L_{\mathrm{w}}}{2} \\ 0 & |z| \leq \frac{L_{\mathrm{w}}}{2} \end{cases}, \tag{11.26}$$

where ΔE_{c} is the conduction-band offset. The energies are all measured from the conduction-band edge of unstrained QW. For a strained QW, the effective mass equation for a single band is

$$\left[-\frac{\hbar^2}{2m_{\mathrm{e}}^{\mathrm{c}}}\frac{\partial^2}{\partial z^2} + \frac{\hbar^2}{2m_{\mathrm{e}}^{\mathrm{c}}}\nabla_t^2 + V(z) + a_{\mathrm{c}}(\varepsilon_{xx} + \varepsilon_{yy} + \varepsilon_{zz})\right]\psi(\vec{r}) = E\psi(\vec{r}), \tag{11.27}$$

where $m_{\mathrm{e}}^{\mathrm{c}}$ is the effective mass in the conduction band for zinc-blende structure and a_{c} is the conduction-band deformation potential. In general, the wave function $\psi(\vec{r})$ can be written

$$\psi^{c\eta}(\vec{r}) = \frac{e^{i(k_x x + k_y y)}}{\sqrt{A}} f(z)|S, \eta> \tag{11.28}$$

and

$$\left[-\frac{\hbar^2}{2m_{\mathrm{e}}^{\mathrm{c}}}\frac{\partial^2}{\partial z^2} + V(z) + a_{\mathrm{c}}(\varepsilon_{xx} + \varepsilon_{yy} + \varepsilon_{zz})\right] f(z) = \left(E(k_t) - \frac{\hbar^2 k_t^2}{2m_{\mathrm{e}}^{\mathrm{c}}}\right) f(z), \tag{11.29}$$

where η is the electron spin and $|S>$ is the basis function near the zone center in the conduction band. The eigenvalues and eigenfunctions in the conduction band are obtained from the earlier equation. When we ignore the k_t dependence of $f(z)$, Equation 11.29 is solved at $k_t = 0$ for the nth sub-band energy $E_n(0)$ and we have $E_n(k_t) = E_n(0) + \hbar^2 k_t^2 / 2m^*$.

11.3.1.2 Valence Band

For the unstrained QW, a QW potential is given by

$$V_{\mathrm{h}}(z) = \begin{cases} -\Delta E_{\mathrm{v}} & |z| > \frac{L_{\mathrm{w}}}{2} \\ 0 & |z| \leq \frac{L_{\mathrm{w}}}{2} \end{cases}, \tag{11.30}$$

where ΔE_v is the valence-band offset. The eigenvalues and eigenfunctions for the strained QW can be obtained by solving the effective mass equation for a QW potential $V_h(z)$ given earlier:

$$\left[H^{LK}\left(k_x, k_y, k_z = -i\frac{\partial}{\partial z}\right) + V_h(z)\right] \cdot \begin{bmatrix} F_1 \\ F_2 \\ F_3 \\ F_4 \\ F_5 \\ F_6 \end{bmatrix} = E \begin{bmatrix} F_1 \\ F_2 \\ F_3 \\ F_4 \\ F_5 \\ F_6 \end{bmatrix}, \tag{11.31}$$

where H^{LK} is the 6×6 Hamiltonian given by Equation 11.1 and the envelope functions F_1, F_2, F_3, F_4, F_5, and F_6 can be written in the vector form

$$\vec{F}_{\vec{k}}(\vec{r}) = \begin{bmatrix} F_1 \\ F_2 \\ F_3 \\ F_4 \\ F_5 \\ F_6 \end{bmatrix} = \begin{bmatrix} g_{3/2,3/2}(k_x, k_y, z) \\ g_{3/2,1/2}(k_x, k_y, z) \\ g_{3/2,-1/2}(k_x, k_y, z) \\ g_{3/2,03/2}(k_x, k_y, z) \\ g_{1/2,1/2}(k_x, k_y, z) \\ g_{1/2,-1/2}(k_x, k_y, z) \end{bmatrix} \frac{e^{i(k_x x + k_y y)}}{\sqrt{A}}. \tag{11.32}$$

The wave function in component form is expressed as

$$\begin{aligned} \psi_{\vec{k}}(\vec{r}) &= F_1 \left|\frac{3}{2}, \frac{3}{2}\right\rangle + F_2 \left|\frac{3}{2}, \frac{1}{2}\right\rangle + F_3 \left|\frac{3}{2}, -\frac{1}{2}\right\rangle + F_4 \left|\frac{3}{2}, -\frac{3}{2}\right\rangle + F_5 \left|\frac{1}{2}, \frac{1}{2}\right\rangle + F_6 \left|\frac{1}{2}, -\frac{1}{2}\right\rangle \\ &= \frac{e^{i(k_x x + k_y y)}}{\sqrt{A}} \sum_{\upsilon} g_{3/2,\upsilon}(k_x, k_y, z) \left|\frac{3}{2}, \upsilon\right\rangle, \end{aligned} \tag{11.33}$$

where $\upsilon = \frac{3}{2}, \frac{1}{2}, -\frac{1}{2}, -\frac{3}{2}, \frac{1}{2}$, and $-\frac{1}{2}$. Denoting $\vec{k}_t = k_x\hat{x} + k_y\hat{y}$, we can write

$$\left[H^{LK}\left(\vec{k}_t, k_z = -i\frac{\partial}{\partial z}\right) + V_h(z)\right] \cdot \begin{bmatrix} g_{3/2,3/2}(\vec{k}_t, z) \\ g_{3/2,1/2}(\vec{k}_t, z) \\ g_{3/2,-1/2}(\vec{k}_t, z) \\ g_{3/2,03/2}(\vec{k}_t, z) \\ g_{1/2,1/2}(\vec{k}_t, z) \\ g_{1/2,-1/2}(\vec{k}_t, z) \end{bmatrix} = E(\vec{k}_t) \begin{bmatrix} g_{3/2,3/2}(\vec{k}_t, z) \\ g_{3/2,1/2}(\vec{k}_t, z) \\ g_{3/2,-1/2}(\vec{k}_t, z) \\ g_{3/2,03/2}(\vec{k}_t, z) \\ g_{1/2,1/2}(\vec{k}_t, z) \\ g_{1/2,-1/2}(\vec{k}_t, z) \end{bmatrix}. \tag{11.34}$$

However, it is convenient to solve eigenvalues and eigenfunctions with a block-diagonalized 3×3 Hamiltonian for a numerical calculation. Let us look at the upper 3×3 Hamiltonian in Equation 11.9. The wave functions can be written as

$$\psi^U(\vec{k}_t, \vec{r}) = \frac{e^{i\vec{k}_t \cdot \vec{r}_t}}{\sqrt{A}} \left[g^{(1)}(k_z, z)\,|1\rangle + g^{(2)}(k_z, z)\,|2\rangle + g^{(3)}(k_z, z)\,|3\rangle\right]. \tag{11.35}$$

The wave function satisfies the upper Hamiltonian equation

$$-\begin{pmatrix} P+Q & -R_\rho - iS_\rho & -\sqrt{2}R_\rho + i\sqrt{\frac{1}{2}}S_\rho \\ -R_\rho + iS_\rho & P-Q & \sqrt{2}Q + i\sqrt{\frac{3}{2}}S_\rho \\ -\sqrt{2}R_\rho - i\sqrt{\frac{1}{2}}S_\rho & \sqrt{2}Q - i\sqrt{\frac{3}{2}}S_\rho & P+\Delta \end{pmatrix}\begin{bmatrix} g^{(1)}(\vec{k}_t,z) \\ g^{(2)}(\vec{k}_t,z) \\ g^{(3)}(\vec{k}_t,z) \end{bmatrix}$$

$$+\begin{pmatrix} V_h(z) & 0 & 0_\rho \\ 0 & V_h(z) & 0 \\ 0 & 0 & V_h(z) \end{pmatrix}\begin{bmatrix} g^{(1)}(\vec{k}_t,z) \\ g^{(2)}(\vec{k}_t,z) \\ g^{(3)}(\vec{k}_t,z) \end{bmatrix} = E(\vec{k}_t)\begin{bmatrix} g^{(1)}(\vec{k}_t,z) \\ g^{(2)}(\vec{k}_t,z) \\ g^{(3)}(\vec{k}_t,z) \end{bmatrix}. \tag{11.36}$$

A similar procedure holds for the lower 3×3 Hamiltonian. That is, the wave function is given by

$$\psi^L(\vec{k}_t,\vec{r}) = \frac{e^{i\vec{k}_t\cdot\vec{r}_t}}{\sqrt{A}}\left[g^{(4)}(k_z,z)\,|\rangle + g^{(5)}(k_z,z)\,|\rangle + g^{(6)}(k_z,z)\,|\rangle\right], \tag{11.37}$$

which satisfies the following 3×3 Hamiltonian equation

$$-\begin{pmatrix} P+Q & -R_\rho + iS_\rho & -\sqrt{2}R_\rho - i\sqrt{\frac{1}{2}}S_\rho \\ -R_\rho - iS_\rho & P-Q & \sqrt{2}Q - i\sqrt{\frac{3}{2}}S_\rho \\ -\sqrt{2}R_\rho + i\sqrt{\frac{1}{2}}S_\rho & \sqrt{2}Q + i\sqrt{\frac{3}{2}}S_\rho & P+\Delta \end{pmatrix}\begin{bmatrix} g^{(4)}(\vec{k}_t,z) \\ g^{(5)}(\vec{k}_t,z) \\ g^{(6)}(\vec{k}_t,z) \end{bmatrix}$$

$$+\begin{pmatrix} V_h(z) & 0 & 0_\rho \\ 0 & V_h(z) & 0 \\ 0 & 0 & V_h(z) \end{pmatrix}\begin{bmatrix} g^{(1)}(\vec{k}_t,z) \\ g^{(2)}(\vec{k}_t,z) \\ g^{(3)}(\vec{k}_t,z) \end{bmatrix} = E(\vec{k}_t)\begin{bmatrix} g^{(4)}(\vec{k}_t,z) \\ g^{(5)}(\vec{k}_t,z) \\ g^{(6)}(\vec{k}_t,z) \end{bmatrix}. \tag{11.38}$$

11.3.2 Wurtzite Semiconductor

We consider nitride-based semiconductors as an example of wurtzite materials. We assume that a strained-layer semiconductor is pseudomorphically grown along the (0001) direction (c-axis) on a (0001)-oriented substrate [16]. The wurtzite structure differs from the zinc-blende structure in several aspects. First, in the (0001) wurtzite GaN-based QW structures, there exists a large internal field due to the strain-induced piezoelectric (PZ) and spontaneous (SP) polarizations [17–21]. Second, there are energy splittings in the valence band such as a crystal field splitting and the splitting due to spin–orbit interaction [6].

11.3.2.1 Internal Field

For binary AB compounds such as a GaN wurtzite structure, the sequence of the atomic layers of the constituents A and B is reversed along the [0001] and [000$\bar{1}$] directions. Defining the $+z$ direction with a vector pointing from a Ga atom to the nearest neighbor N atom, the Ga face means that Ga is on the top position of the {0001} bilayer, corresponding to the [0001] polarity. The induced piezoelectric polarization is given by [20]

$$P_z^{\mathrm{PZ}} = e_{31}(\varepsilon_{xx}+\varepsilon_{yy}) + e_{33}\varepsilon_{zz} = 2\left(e_{31} - \frac{C_{13}}{C_{33}}e_{33}\right)\varepsilon_{xx}, \tag{11.39}$$

where e_{ij} are piezoelectric constants. In the case of AlGaN system, for example, the piezoelectric polarization is negative for tensile strain and positive for compressive strain because $2\left(e_{31} - \frac{C_{13}}{C_{33}}e_{33}\right) < 0$ for the whole range of Al compositions. Also, there exists a spontaneous polarization along the c-axis. That is, the spontaneous polarization for GaN (or AlN) has been found to be negative, meaning that the polarization for Ga-face (or Al-face) is pointing toward the substrate. As a result, in the AlGaN system, the alignment of the piezoelectric polarization and the spontaneous polarization is parallel in the case of tensile strain and antiparallel in the case of compressive strain [18]. In the case of multiple QW (MQW) structure, the internal field F_z is determined from the periodic boundary condition and the difference between the sum of spontaneous and piezoelectric polarizations of the well and barrier layers. The continuity of the displacement vector (D_z) normal to the surface gives

$$D_z = \varepsilon_{\mathrm{w}} F_{\mathrm{w}} + P_{\mathrm{w}} = \varepsilon_{\mathrm{b}} F_{\mathrm{b}} + P_{\mathrm{b}}, \tag{11.40}$$

where the subscripts w and b mean the well region and the barrier region, respectively, $P_{\mathrm{w}} = P_{\mathrm{w}}^{\mathrm{PZ}} + P_{\mathrm{w}}^{\mathrm{SP}}$, and $P_{\mathrm{b}} = P_{\mathrm{b}}^{\mathrm{PZ}} + P_{\mathrm{b}}^{\mathrm{SP}}$. We consider the periodic boundary condition that the net voltage drop over one period is zero. That is,

$$F_{\mathrm{w}} L_{\mathrm{w}} + F_{\mathrm{b}} L_{\mathrm{b}} = 0, \tag{11.41}$$

where L_{w} and L_{b} are the well width and the barrier width, respectively. Then, we find the electric fields in the well and the barrier:

$$\begin{aligned} F_{\mathrm{w}} &= \frac{L_{\mathrm{b}}}{\varepsilon_{\mathrm{b}} L_{\mathrm{w}} + \varepsilon_{\mathrm{w}} L_{\mathrm{b}}}(P_{\mathrm{b}} - P_{\mathrm{w}}) \\ F_{\mathrm{b}} &= \frac{-L_{\mathrm{w}}}{\varepsilon_{\mathrm{b}} L_{\mathrm{w}} + \varepsilon_{\mathrm{w}} L_{\mathrm{b}}}(P_{\mathrm{b}} - P_{\mathrm{w}}). \end{aligned} \tag{11.42}$$

Hence, the self-consistent solution, which solves the Schrödinger equation and Poisson equation simultaneously, is necessary for the wurtzite QWs.

11.3.2.2 Energy Splitting

At the zone center ($k_x = k_y = k_z = 0$), we can obtain the following doubly degenerate band-edge energies

$$\begin{aligned} E_1 &= \Delta_1 + \Delta_2 \\ E_2 &= \frac{\Delta_1 - \Delta_2}{2} + \sqrt{\left(\frac{\Delta_1 - \Delta_2}{2}\right)^2 + 2\Delta_3^2} \\ E_3 &= \frac{\Delta_1 - \Delta_2}{2} - \sqrt{\left(\frac{\Delta_1 - \Delta_2}{2}\right)^2 + 2\Delta_3^2}. \end{aligned} \tag{11.43}$$

Without the spin–orbit interaction, $\Delta_2 = \Delta_3 = 0$, we have the top two degenerate bands and the lower band as the reference level E_{v}. That is,

$$\begin{aligned} E_1 = E_2 &= E_{\mathrm{v}} + \Delta_1 = \Delta_1 \\ E_3 &= E_v = 0. \end{aligned} \tag{11.44}$$

Here, we set the reference energy E_v to 0. Then, with the spin–orbit interaction, the top valence-band energy is $E_1 = \Delta_1 + \Delta_2$, and the conduction-band edge is given by adding the bandgap energy E_g to E_v:

$$E_c = E_g + \Delta_1 + \Delta_2. \tag{11.45}$$

11.3.2.3 Self-Consistent Calculations with the Screening of Eigenvalues and Eigenfunctions in the Conduction Band and the Valence Band

The total potential profiles for the electrons and the holes are

$$V_c(z) = \begin{cases} E_c + \Delta E_c + |e| F_b z - |e| \varphi(z) & |z| \succ \frac{L_w}{2} \\ E_c + |e| F_w z - |e| \varphi(z) & |z| \leq \frac{L_w}{2} \end{cases} \tag{11.46}$$

and

$$V_h(z) = \begin{cases} -\Delta E_v + |e| F_b z - |e| \varphi(z) & |z| \succ \frac{L_w}{2} \\ |e| F_w z - |e| \varphi(z) & |z| \leq \frac{L_w}{2} \end{cases}. \tag{11.47}$$

Here, $\varphi(z)$ is the screening potential induced by the charged carriers and satisfies the Poisson equation

$$\frac{d}{dz}\left(\varepsilon(z)\frac{d}{dz}\right)\varphi(z) = -|e|[p(z) - n(z)], \tag{11.48}$$

where $\varepsilon(z)$ is the dielectric constant and we assume that there is no doping in the well and the barrier. The electron and the hole concentrations, $n(z)$ and $p(z)$, are related to the wave functions of the nth conduction sub-band and the mth valence sub-band by

$$n(z) = \frac{kTm_e}{\pi\hbar^2} \sum_n |f_n(z)|^2 \ln\left(1 + e^{[E_{fc} - E_{cn}]/kT}\right) \tag{11.49}$$

and

$$p(z) = \sum_{\sigma=U,L} \sum_m \int dk_{||} \frac{k_{||}}{2\pi} \sum_\nu |g^{\sigma(\nu)}_{mk_{||}}(z)|^2 \frac{1}{1 + e^{[E_{fv} - E_{vm}(k_{||})]/kT}}, \tag{11.50}$$

where n and m are the quantized sub-band indices for the conduction and the valence bands, E_{fc} and E_{fv} are the quasi-Fermi levels of the electrons and the holes, respectively, and E_{cn} and $f_n(z)$ are the quantized energy level of the electrons at a band-edge and eigenfunctions, respectively. The eigenvalues and eigenfunctions in the conduction band are obtained by solving

$$\left[-\frac{\hbar^2}{2m^w_{ez}}\frac{\partial^2}{\partial z^2} + V_c(z) + a^w_{ct}(\varepsilon_{xx} + \varepsilon_{yy}) + a^w_{cz}\varepsilon_{zz}\right] f(z) = \left(E(k_{||}) - \frac{\hbar^2 k^2_{||}}{2m^w_{et}}\right) f(z), \tag{11.51}$$

where a^w_{ct} and $a^w_{cz}$are the conduction-band deformation potentials along the c-axis and perpendicular to the c axis, respectively. Usually, it is convenient to obtain eigenvalues in the conduction band with $E_c = 0$ in Equation 11.46. The term $(E_g + \Delta_1 + \Delta_2)$ can be added when we calculate optical properties such as spontaneous emission coefficient and optical gain.

Also, $E_{\mathrm{vm}}(k_{||})$ is the energy for the mth sub-band in the valence band, σ denotes the upper (U) and the lower (L) blocks of the Hamiltonian, $k_{||}$ is the in-plane wave vector, ν refers to the new bases for the Hamiltonian, and $g_{\mathrm{mk}_{||}}^{\sigma(\nu)}(z)$ is the envelope function in the valence band.

The eigenvalues and eigenfunctions for the upper Hamiltonian equation of the valence band are obtained by solving

$$-\begin{pmatrix} F & K_t & -iH_t \\ K_t & G & \Delta - iH_t \\ iH_t & \Delta + iH_t & P + \Delta \end{pmatrix} \begin{bmatrix} g^{(1)}(\vec{k}_t, z) \\ g^{(2)}(\vec{k}_t, z) \\ g^{(3)}(\vec{k}_t, z) \end{bmatrix} + \begin{pmatrix} V_{\mathrm{h}}(z) & 0 & 0_{\rho} \\ 0 & V_{\mathrm{h}}(z) & 0 \\ 0 & 0 & V_{\mathrm{h}}(z) \end{pmatrix} \begin{bmatrix} g^{(1)}(\vec{k}_t, z) \\ g^{(2)}(\vec{k}_t, z) \\ g^{(3)}(\vec{k}_t, z) \end{bmatrix} = E(\vec{k}_t) \begin{bmatrix} g^{(1)}(\vec{k}_t, z) \\ g^{(2)}(\vec{k}_t, z) \\ g^{(3)}(\vec{k}_t, z) \end{bmatrix}. \tag{11.52}$$

A similar procedure holds for the lower 3×3 Hamiltonian, which satisfies the following 3×3 Hamiltonian equation

$$-\begin{pmatrix} F & K_t & iH_t \\ K_t & G & \Delta + iH_t \\ -iH_t & \Delta - iH_t & P + \Delta \end{pmatrix} \begin{bmatrix} g^{(4)}(\vec{k}_t, z) \\ g^{(5)}(\vec{k}_t, z) \\ g^{(6)}(\vec{k}_t, z) \end{bmatrix} + \begin{pmatrix} V_{\mathrm{h}}(z) & 0 & 0_{\rho} \\ 0 & V_{\mathrm{h}}(z) & 0 \\ 0 & 0 & V_{\mathrm{h}}(z) \end{pmatrix} \begin{bmatrix} g^{(1)}(\vec{k}_t, z) \\ g^{(2)}(\vec{k}_t, z) \\ g^{(3)}(\vec{k}_t, z) \end{bmatrix} = E(\vec{k}_t) \begin{bmatrix} g^{(4)}(\vec{k}_t, z) \\ g^{(5)}(\vec{k}_t, z) \\ g^{(6)}(\vec{k}_t, z) \end{bmatrix}. \tag{11.53}$$

The potential $\varphi(z)$ is obtained by integration

$$\varphi(z) = -\int_{-L/2}^{z} E(z')\mathrm{d}z', \tag{11.54}$$

where

$$E(z) = \int_{-L/2}^{z} \frac{1}{\varepsilon(z)}\rho(z')\mathrm{d}z'. \tag{11.55}$$

The procedures for the self-consistent calculations consist of the following steps:

1. Start with the potential profiles V_{c} and V_{h} with $\varphi^{(0)}(z) = 0$ in Equations 11.46 and 11.47.
2. Solve the Schrödinger equation (for electrons) and the block-diagonalized Hamiltonian (for holes) with the potential profiles $\varphi^{(n-1)}(z)$ in step (1) to obtain band structures and wave functions.
3. For a given carrier density, obtain the Fermi energies from Equations 11.49 and 11.50 by using the band structures and the charge distribution by using the wave functions.
4. Solve Poisson's equation to find $\varphi^{(n)}(z)$.

   Check if $\varphi^{(n)}(z)$ converges to $\varphi^{(n-1)}(z)$. If not, set $\varphi^{(n)}(z) = w\varphi^{(n)}(z) + (1-w)\varphi^{(n-1)}(z)$, $n = n+1$; then, return to step (2). If yes, the band structures and the wave functions obtained with $\varphi^{(n-1)}(z)$ are solutions. An adjustable parameter w ($0 < w < 1$) is typically set to be 0.5 at low carrier densities. With increasing carrier densities, a smaller value of w is needed for rapid convergence.

11.4 Optical Matrix Elements

11.4.1 Zinc-Blende Structure

The optical momentum matrix elements are defined as

$$|\hat{\varepsilon}\cdot M_{lm}^{\eta\sigma}(k_{||})|^2 = |<\psi_l^{\eta}|\hat{\varepsilon}\cdot\vec{p}|\psi_m^{\nu\sigma}>|^2, \tag{11.56}$$

which represents the interband transition probability between electrons and hole. Using the expressions given in Equations 11.28 and 11.35 and taking the φ integration of the momentum matrix elements, we obtain the following momentum matrix elements for the upper Hamiltonian.

Transverse electric (TE) polarization ($\hat{\varepsilon}=\hat{x}$ or $\hat{\varepsilon}=\hat{y}$):

$$\begin{aligned}
|M_{\mathrm{TE}}|^2 &= |<\psi_l^{\eta}|p_x|\psi_m^{\nu\sigma}>|^2 = |<\psi_l^{\eta}|p_y|\psi_m^{\nu\sigma}>|^2 \\
&= \delta_{\vec{k'}_t,\vec{k}_t}<f_n|g^{(1)}><S\uparrow|p_x|1>+<f_n|g^{(2)}><S\uparrow|p_x|2>+<f_n|g^{(3)}><S\uparrow|p_x|3>|^2 \\
&= \delta_{\vec{k'}_t,\vec{k}_t}|<S\uparrow|p_x|X>|^2|<f_n|g^{(1)}>\left(-\frac{1}{2}e^{-i\frac{3}{2}\varphi}\right) \\
&\quad +<f_n|g^{(2)}>\left(-\frac{1}{2\sqrt{3}}e^{i\frac{1}{2}\varphi}\right)+<f_n|g^{(3)}>\left(-\frac{1}{\sqrt{6}}e^{i\frac{1}{2}\varphi}\right)|^2 \\
&= \delta_{\vec{k'}_t,\vec{k}_t}|<S\uparrow|p_x|X>|^2\frac{1}{12} \\
&\quad\times\left\{\begin{array}{l} 3<f_n|g^{(1)}>^2+[<f_n|g^{(2)}>+\sqrt{2}<f_n|g^{(3)}>]^2 \\ +2\sqrt{3}\cos 2\phi<f_n|g^{(1)}><f_n|g^{(2)}>+2\sqrt{6}\cos 2\phi<f_n|g^{(1)}><f_n|g^{(3)}>\end{array}\right\}.
\end{aligned} \tag{11.57}$$

When we calculate the absorption coefficient, the summation over k_t is given by

$$\frac{1}{V}\sum_{k_t}=\frac{1}{L_z}\int_0^{\infty}\frac{k_t}{2\pi}\mathrm{d}k_t\int_0^{2\pi}\frac{\mathrm{d}\varphi}{2\pi}. \tag{11.58}$$

Thus, the term containing the $\cos 2\varphi$ factor does not contribute to the absorption coefficient because the integration over φ vanishes. We obtain

$$|M_{\mathrm{TE}}|^2=\delta_{\vec{k}'_t,\vec{k}_t}|<S\uparrow|p_x|X>|^2\frac{1}{12}\left\{3<f_n|g^{(1)}>^2+[<f_n|g^{(2)}>+\sqrt{2}<f_n|g^{(3)}>]^2\right\}. \tag{11.59}$$

Transverse magnetic (TM) polarization ($\hat{\varepsilon}=\hat{z}$):

$$\begin{aligned}
|M_{\mathrm{TM}}|^2 &= |<\psi_l^{\eta}|p_z|\psi_m^{\nu\sigma}>|^2 \\
&= \delta_{\vec{k'}_t,\vec{k}_t}|<f_n|g^{(1)}><S\uparrow|p_z|1>+<f_n|g^{(2)}><S\uparrow|p_z|2>+<f_n|g^{(3)}><S\uparrow|p_z|3>|^2 \\
&= \delta_{\vec{k'}_t,\vec{k}_t}|<S\uparrow|p_z|Z>|^2|<f_n|g^{(2)}>i\frac{1}{\sqrt{3}}e^{-i\frac{1}{2}\varphi}-<f_n|g^{(3)}>i\frac{1}{\sqrt{6}}e^{-i\frac{1}{2}\varphi}|^2 \\
&= \delta_{\vec{k'}_t,\vec{k}_t}|<S\uparrow|p_z|Z>|^2\frac{1}{3}\left|<f_n|g^{(2)}>-<f_n|g^{(3)}>\frac{1}{\sqrt{2}}\right|.
\end{aligned} \tag{11.60}$$

Here,

$$|<S\uparrow|p_x|X>|^2 = |<S\uparrow|p_z|Z>|^2 = \frac{m_o}{2}\left(\frac{m_o}{m_e}-1\right)\frac{(E_g+\Delta)E_g}{(E_g+2\Delta/3)}. \tag{11.61}$$

The momentum matrix elements for the lower Hamiltonian can be also obtained by a similar procedure:

$$|M_{TE}|^2 = \delta_{\vec{k}_t',\vec{k}_t}|<S\downarrow|p_x|X>|^2\frac{1}{12}\left\{3<f_n|g^{(6)}>^2+[<f_n|g^{(4)}>+\sqrt{2}<f_n|g^{(5)}>]^2\right\} \tag{11.62}$$

and

$$|M_{TM}|^2 = \delta_{\vec{k}_t',\vec{k}_t}|<S\downarrow|p_z|Z>|^2\frac{1}{3}\left|<f_n|g^{(5)}>-<f_n|g^{(4)}>\frac{1}{\sqrt{2}}\right|^2. \tag{11.63}$$

11.4.2 Wurtzite Structure

Similarly to the zinc-blende case, we can obtain the optical momentum matrix elements as follows.

TE polarization ($\hat{\varepsilon}=\hat{x}$ or $\hat{\varepsilon}=\hat{y}$):

$$\begin{aligned}
|M_{TE}|^2 &= |<\psi_l^\eta|p_x|\psi_m^{\nu\sigma}>|^2 = |<\psi_l^\eta|p_y|\psi_m^{\nu\sigma}>|^2\\
&= \delta_{\vec{k'}_t,\vec{k}_t}|<f_n|g^{(1)}><S\uparrow|p_x|1>+<f_n|g^{(2)}><S\uparrow|p_x|2>+<f_n|g^{(3)}><S\uparrow|p_x|3>|^2\\
&= \delta_{\vec{k'}_t,\vec{k}_t}|<S\uparrow|p_x|X>|^2\left|-\frac{1}{2}<f_n|g^{(1)}>e^{i\left(\frac{3}{4}\pi+\frac{3}{2}\varphi\right)}+<f_n|g^{(2)}>\frac{1}{2}e^{-i\left(\frac{1}{4}\pi+\frac{1}{2}\varphi\right)}\right|^2\\
&= \delta_{\vec{k'}_t,\vec{k}_t}|<S\uparrow|p_x|X>|^2\frac{1}{4}\left|<f_n|g^{(1)}>^2+<f_n|g^{(2)}>^2+2\cos\varphi<f_n|g^{(1)}><f_n|g^{(2)}>\right|.
\end{aligned} \tag{11.64}$$

Finally,

$$|M_{TE}|^2 = |\delta_{\vec{k}_t',\vec{k}_t}|<S\uparrow|p_x|X>|^2\frac{1}{4}\left\{<f_n|g^{(1)}>^2+<f_n|g^{(2)}>^2\right\}. \tag{11.65}$$

TM polarization ($\hat{\varepsilon}=\hat{z}$):

$$\begin{aligned}
|M_{TM}|^2 &= |<\psi_l^\eta|p_z|\psi_m^{\nu\sigma}>|^2\\
&= \delta_{\vec{k'}_t,\vec{k}_t}|<f_n|g^{(1)}><S\uparrow|p_z|1>+<f_n|g^{(2)}><S\uparrow|p_z|2>+<f_n|g^{(3)}><S\uparrow|p_z|3>|^2\\
&= \delta_{\vec{k'}_t,\vec{k}_t}<S\uparrow|p_z|Z>^2|<f_n|g^{(3)}>\frac{1}{\sqrt{2}}e^{i\left(\frac{1}{4}\pi+\frac{1}{2}\varphi\right)}|^2\\
&= \delta_{\vec{k'}_t,\vec{k}_t}<S\uparrow|p_z|Z>^2\frac{1}{2}<f_n|g^{(3)}>^2.
\end{aligned} \tag{11.66}$$

The momentum matrix elements for the lower Hamiltonian also can be obtained by a similar procedure:

$$|M_{TE}|^2 = |\delta_{\vec{k'}_t,\vec{k}_t}|<S\downarrow|p_x|X>|^2\frac{1}{4}\left\{<f_n|g^{(4)}>^2+<f_n|g^{(5)}>^2\right\} \tag{11.67}$$

and

$$|M_{TM}|^2 = \delta_{\vec{k'}_t,\vec{k}_t}<S\downarrow|p_z|Z>^2\frac{1}{2}<f_n|g^{(6)}>^2. \tag{11.68}$$

11.5 Band Structures of Bulk Semiconductors with an Arbitrary Crystal Orientation

11.5.1 Zinc-Blende Structure

The Hamiltonian for an arbitrary crystal orientation can be obtained using a rotation matrix

$$U = \begin{pmatrix} \cos\theta\cos\varphi & \cos\theta\sin\varphi & -\sin\theta \\ -\sin\varphi & \cos\varphi & 0 \\ \sin\theta\cos\varphi & \sin\theta\sin\varphi & \cos\theta \end{pmatrix}, \tag{11.69}$$

The rotation with the Euler angles θ and φ transforms physical quantities from the (x, y, z) coordinates to the (x', y', z') coordinates. Figure 11.1 shows configuration of the [*lmn*]-oriented coordinate system to the conventional [001]-oriented coordinate system. The relation between the coordinate systems for vectors and tensors is expressed as

$$\begin{aligned} k_\alpha &= U_{i\alpha} k'_i, \\ \varepsilon_{\alpha\beta} &= U_{i\alpha} U_{j\beta} \varepsilon'_{ij}, \end{aligned} \tag{11.70}$$

where the summation over repeated indices is assumed. Our object is to obtain $P_{k'}$, $Q_{k'}$, $R_{k'}$, and $S_{k'}$ in the (x', y', z') coordinates. First we obtain the 4×4 Hamiltonian for general crystal orientation by using invariant method and then extending these results to obtain the 6×6 Hamiltonian. Here, for simplicity, we consider the case of the QWs with $(11n)$ orientations with $\theta = \arctan(\sqrt{2}/n)$ and $\varphi = \pi/4$. For example, (110) means $\theta = \pi/2$ and $\varphi = \pi/4$. The 4×4 Hamiltonian to describe the interaction between the heavy-hole and light-hole bands along with the Bir–Pikus Hamiltonian for strain can be formally written as

$$H_t^{\mathrm{v}} = H_0^{\mathrm{v}}(\vec{k}) + H_{\mathrm{s}}^{\mathrm{v}}, \tag{11.71}$$

with $H_0^{\mathrm{v}}(\vec{k})$ representing the valence-band Hamiltonian and $H_{\mathrm{s}}^{\mathrm{v}}$ the valence-band strain or Bir–Pikus Hamiltonian. The Luttinger formulation of the most general Hamiltonian for the (001) crystal orientation

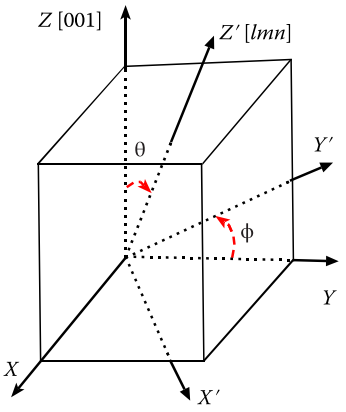

FIGURE 11.1 Configuration of the [*lmn*]-oriented coordinate system to the conventional [001]-oriented coordinate system.

is given by [22,23]

$$\begin{aligned} H_0^{\mathrm{v}} &= \frac{\hbar^2}{2m_{\mathrm{o}}}(\gamma_1 k^2 I_4 - 2\gamma_2[(J_x^2 - 1/3J^2)k_x^2 + (J_y^2 - 1/3J^2)k_y^2 + (J_z^2 - 1/3J^2)k_z^2] \\ &\quad - 4\gamma_3[\{J_xJ_y\}\{k_xk_y\} + \{J_yJ_z\}\{k_yk_z\} + \{J_zJ_x\}\{k_zk_x\}]), \\ H_{\mathrm{s}}^{\mathrm{v}} &= -a_v(\varepsilon_{xx} + \varepsilon_{yy} + \varepsilon_{zz})I_4 + b[(J_x^2 - 1/3J^2)\varepsilon_{xx} + (J_y^2 - 1/3J^2)\varepsilon_{yy} + (J_z^2 - 1/3J^2)\varepsilon_{zz}] \\ &\quad + 2d/\sqrt{3}[\{J_xJ_y\}\{\varepsilon_{xx}\} + \{J_yJ_z\}\{\varepsilon_{yy}\} + \{J_zJ_x\}\{\varepsilon_{zz}\}]), \end{aligned} \tag{11.72}$$

where the J_i's are the angular momentum matrices for a state with spin 3/2; I_4 is a 4×4 identity matrix; and $\{k_xk_y\} = (k_xk_y + k_yk_x)/2$. The angular momentum matrices are any three matrices, $(J_xJ_yJ_z)$, which satisfy the rules of commutation for angular momentum. These matrices given by Luttinger are as follows:

$$J_x = \begin{pmatrix} 0 & 0 & \sqrt{3}/2 & 0 \\ 0 & 0 & 1 & \sqrt{3}/2 \\ \sqrt{3}/2 & 1 & 0 & 0 \\ 0 & \sqrt{3}/2 & 0 & 0 \end{pmatrix}, \quad J_y = \begin{pmatrix} 0 & 0 & -i\sqrt{3}/2 & 0 \\ 0 & 0 & i & -i\sqrt{3}/2 \\ i\sqrt{3}/2 & -i & 0 & 0 \\ 0 & i\sqrt{3}/2 & 0 & 0 \end{pmatrix},$$

$$J_z = \begin{pmatrix} 3/2 & 0 & 0 & 0 \\ 0 & -1/2 & 0 & 0 \\ 0 & 0 & 1/2 & 0 \\ 0 & 0 & 0 & -3/2 \end{pmatrix}, \quad I_4 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{11.73}$$

We can obtain the (001)-oriented Hamiltonian by substituting Equation 11.73 into 11.72:

$$H_0^{\mathrm{v}} = \begin{pmatrix} P+Q & R & -S & 0 \\ R^* & P-Q & 0 & S^* \\ -S^* & 0 & P-Q & R \\ 0 & S^* & R^* & P+Q \end{pmatrix}. \tag{11.74}$$

This is exactly the same from as the upper 4×4 Hamiltonian given in Equation 11.1. However, note that, in this case, the ordering of the basis functions for the choice of the J_i's is $\left|\frac{3}{2},\frac{3}{2}\right\rangle, \left|\frac{3}{2},-\frac{1}{2}\right\rangle, \left|\frac{3}{2},\frac{1}{2}\right\rangle, \left|\frac{3}{2},-\frac{3}{2}\right\rangle$. Next, we consider the case of QWs with $(11n)$ orientations with $\theta = \arctan(\sqrt{2}/n)$ and $\varphi = \pi/4$. Then, from Equations 11.69 and 11.70,

$$\begin{aligned} P_{k'} &= \left(\frac{\hbar^2}{2m_0}\right)\gamma_1(k_{x'}^2 + k_{y'}^2 + k_{z'}^2), \\ Q_{k'} &= \left(\frac{\hbar^2}{2m_0}\right)\left\{\left[\frac{(n^2-4)(n^2-1)}{(n^2+2)^2}\gamma_2 + \frac{9n^2}{(n^2+2)^2}\gamma_3\right]k_{x'}^2 + \left[\frac{(n^2+2)(n^2-1)}{(n^2+2)^2}\gamma_2 + \frac{3}{(n^2+2)}\gamma_3\right]k_{y'}^2\right. \\ &\quad \left. + \left[-2\frac{(n^2-1)^2}{(n^2+2)^2}\gamma_2 - \frac{6+12n^2}{(n^2+2)}\gamma_3\right]k_{z'}^2\right\}, \\ R_{k'} &= \left(\frac{\hbar^2}{2m_0}\right)\left\{\sqrt{3}\left[\frac{(n^2-4)}{(n^2+2)^2}\gamma_2 - \frac{5n^2+n^4}{(n^2+2)^2}\gamma_3\right]k_{x'}^2 + \sqrt{3}\left[\frac{1}{(n^2+2)}\gamma_2 + \frac{1+n^2}{(n^2+2)^2}\gamma_3\right]k_{y'}^2\right. \end{aligned}$$

$$
\begin{aligned}
&+2\sqrt{3}\left[-\frac{(n^2-1)}{(n^2+2)^2}(\gamma_2-\gamma_3)\right]k_{z'}^2+i2\sqrt{3}\left[\frac{(n^2\gamma_2+2\gamma_3)}{(n^2+2)}\right]k_{x'}k_{y'}\\
&+2\sqrt{6}\left[i\frac{n}{(n^2+2)}(\gamma_2-\gamma_3)\right]k_{y'}k_{z'}+6\sqrt{6}\left[\frac{n}{(n^2+2)^2}(\gamma_2-\gamma_3)\right]k_{x'}k_{z'}\Bigg\},\\
-S_{k'}=&\left(\frac{\hbar^2}{2m_0}\right)\Bigg\{\left[\sqrt{6}\frac{n(4-n^2)}{(n^2+2)^2}(\gamma_2-\gamma_3)\right]k_{x'}^2+\left[\sqrt{6}\frac{n}{(n^2+2)}(\gamma_3-\gamma_2)\right]k_{y'}^2\\
&+\left[2\sqrt{6}\frac{n(n^2-1)}{(n^2+2)^2}(\gamma_2-\gamma_3)\right]k_{z'}^2+i2\sqrt{6}\left[\frac{n}{(n^2+2)}(\gamma_2-\gamma_3)\right]k_{x'}k_{y'}\\
&+i2\sqrt{3}\left[\frac{2\gamma_2+n^2\gamma_3}{(n^2+2)}\right]k_{y'}k_{z'}+\sqrt{3}\left[-\frac{12n^2}{(n^2+2)^2}\gamma_2-2\frac{(4-2n^2+n^4)}{(n^2+2)^2}\gamma_3\right]k_{x'}k_{z'}\Bigg\},\\
P_{\varepsilon'}=&-a_{v1}(\varepsilon_{x'x'}+\varepsilon_{y'y'}+\varepsilon_{z'z'}),\\
Q_{\varepsilon'}=&-\frac{b(n^2-1)(-\Gamma+3\sqrt{2}\varepsilon_{x'z'}n+\Gamma n^2)+d\sqrt{3}[\Gamma+\sqrt{2}\varepsilon_{x'z'}n+n^2(2\Gamma-\sqrt{2}\varepsilon_{x'z'}n)]}{(n^2+2)^2},\\
R_{\varepsilon'}=&-\frac{b\sqrt{3}(-\Gamma+3\sqrt{2}\varepsilon_{x'z'}n+\Gamma n^2)+d(\Gamma-3\sqrt{2}\varepsilon_{x'z'}n-\Gamma n^2)]}{(n^2+2)^2},\\
-S_{\varepsilon'}=&-\frac{b\sqrt{3}n(\sqrt{2}\Gamma-6\varepsilon_{x'z'}n-\sqrt{2}\Gamma n^2)+d[\sqrt{2}n(n^2-1)\Gamma+\varepsilon_{x'z'}(-4+2n^2-n^4)]}{(n^2+2)^2},
\end{aligned}
\tag{11.75}
$$

where $\Gamma=\varepsilon_{||}-\varepsilon_{z'z'}$, $\varepsilon_{||}=(a_s-a_e)/a_e$, and a_s and a_e are lattice constants for the substrate and the epilayer materials, respectively. The biaxial strain components for a general crystal orientation are determined from the condition that the layer is grown pseudomorphically and these strain coefficients should minimize the strain energy of the layer simultaneously [24,25]. That is,

$$
\varepsilon_{x'x'}=\varepsilon_{y'y'}=\varepsilon_{||},\quad \varepsilon_{x'y'}=\varepsilon_{y'z'}=0,\quad \varepsilon_{z'z'}=-2\frac{K_3}{\sqrt{2}K_2}\varepsilon_{||},\quad \varepsilon_{x'z'}=-2\frac{K_1}{\sqrt{2}K_2}\varepsilon_{||},
\tag{11.76}
$$

where

$$
\begin{aligned}
K_1&=(C_{11}+2C_{12})(-C_{11}+C_{12}+2C_{44})n(n^2-1),\\
K_2&=2C_{11}C_{44}+2C_{12}C_{44}+2C_{44}^2+(C_{11}^2+C_{11}C_{12}-2C_{12}^2+2C_{11}C_{44}-4C_{12}C_{44}+C_{11}C_{44}n^2)n^2,\\
K_2&=-2\left[C_{11}C_{44}+3C_{12}C_{44}-2C_{44}^2+(C_{11}^2+C_{11}C_{12}-2C_{12}^2-C_{11}C_{44}+C_{12}C_{44}n^2)n^2\right],
\end{aligned}
\tag{11.77}
$$

and C_{ij} are the stiffness constants in the strained epilayers. Then, the 4×4 Hamiltonian of $(11n)$-oriented zinc-blende crystal is given by

$$
H'_{4\times4}=\begin{pmatrix} P'+Q' & R' & -S' & 0\\ R'^* & P'-Q' & 0 & S'^*\\ -S'^* & 0 & P'-Q' & R'\\ 0 & S'^* & R'^* & P'+Q' \end{pmatrix},
\tag{11.78}
$$

where

$$
P'=P_{k'}+P_{\varepsilon'},\quad Q'=Q_{k'}+Q_{\varepsilon'},\quad R'=R_{k'}+R_{\varepsilon'},\quad S'=S_{k'}+S_{\varepsilon'}.
\tag{11.79}
$$

Off-diagonal strains in zinc-blende structure semiconductors induce a polarization given by [25]

$$P_i^{\mathrm{S}} = 2e_{14}\varepsilon_{jk}, \tag{11.80}$$

where P_i^{S} is the induced polarization, e_{14} is the piezoelectric constant, and $i, j, k = x, y, z$. Here, i, j, k are in cyclic order. A strained layer with a [001] growth direction has only diagonal strains. Thus, the (001)-oriented layer will not have strain-induced piezoelectric polarization fields. However, strained layers with any other growth direction will have piezoelectric polarization fields. The piezoelectric polarization along the growth direction is important because the electric field in the QW originates from polarization charges at heterojunction interfaces. The electric field induced perpendicular to the growth direction due to the strain can be calculated as

$$E_\perp = -\frac{\vec{P}^{\mathrm{S}} \cdot \hat{u}}{\varepsilon}, \tag{11.81}$$

where the unit vector $\hat{u}$ along the growth direction is given by $\hat{u} = \sqrt{1/(n^2+2)}(\hat{i} + \hat{j} + n\hat{k})$. For an $[11n]$ growth direction,

$$E_\perp = -\frac{2e_{14}}{\sqrt{n^2+2}}(n^2\varepsilon_{xy} + \varepsilon_{yz} + \varepsilon_{zx}), \tag{11.82}$$

where

$$\begin{aligned}
\varepsilon_{xy} &= \frac{1}{(n^2+2)}\left(-\varepsilon_{x'x'} + \sqrt{2}n\varepsilon_{x'z'} + 2n\varepsilon_{z'z'}\right),\\
\varepsilon_{yz} &= \frac{1}{2(n^2+2)}\left(-2n\varepsilon_{x'x'} - \sqrt{2}(2-n^2)\varepsilon_{x'z'} + 2n\varepsilon_{z'z'}\right)\\
\varepsilon_{zx} &= \varepsilon_{yz}.
\end{aligned} \tag{11.83}$$

It is straightforward to obtain the 6×6 Hamiltonian of $(11n)$-oriented zinc-blende crystal using components P', Q', R', and S' in the (x', y', z') coordinates. That is, the 6×6 Hamiltonian for the valence-band structure in the (x', y', z') coordinates can be written as

$$H'_{6\times 6} = -\begin{pmatrix}
P'+Q' & -S' & R' & 0 & -S'/\sqrt{2} & \sqrt{2}R' \\
-S'^{\dagger} & P'-Q' & 0 & R' & -\sqrt{2}Q' & \sqrt{3/2}S' \\
R'^{\dagger} & 0 & P'-Q' & 0 & \sqrt{3/2}S'^{\dagger} & \sqrt{2}Q' \\
0 & R'^{+} & S'^{\dagger} & P'+Q' & -\sqrt{2}R'^{\dagger} & -S'^{\dagger}/\sqrt{2} \\
-S'^{\dagger}/\sqrt{2} & 0 & \sqrt{3/2}S' & -\sqrt{2}R' & P'+\Delta & 0 \\
\sqrt{2}R'^{\dagger} & \sqrt{3/2}S'^{\dagger} & \sqrt{2}Q' & -S'/\sqrt{2} & 0 & P'+\Delta
\end{pmatrix}
\begin{matrix}
|3/2,3/2\rangle' \\ |3/2,1/2\rangle' \\ |3/2,-1/2\rangle' \\ |3/2,-3/2\rangle' \\ |1/2,1/2\rangle' \\ |1/2,-1/2\rangle'
\end{matrix}. \tag{11.84}$$

Here, the basis set in the (x', y', z') coordinates consists of the following basis functions:

$$\begin{aligned}
\left|u_1\right\rangle' &= \left|\frac{3}{2},\frac{3}{2}\right\rangle' = -\frac{1}{\sqrt{2}}\left|\left(X'+iY'\right)\uparrow'\right\rangle,\\
\left|u_2\right\rangle' &= \left|\frac{3}{2},\frac{1}{2}\right\rangle' = \frac{1}{\sqrt{6}}\left|-\left(X'-iY'\right)\downarrow' + 2Z'\uparrow'\right\rangle,\\
\left|u_3\right\rangle' &= \left|\frac{3}{2},-\frac{1}{2}\right\rangle' = \frac{1}{\sqrt{6}}\left|\left(X'-iY'\right)\uparrow' + 2Z'\downarrow'\right\rangle,
\end{aligned}$$

$$|u_4\rangle' = \left|\frac{3}{2}, -\frac{3}{2}\right\rangle' = \frac{1}{\sqrt{2}}\left|\left(X' - iY'\right)\downarrow'\right\rangle,$$
$$|u_5\rangle' = \left|\frac{1}{2}, \frac{1}{2}\right\rangle' = \frac{1}{\sqrt{3}}\left|\left(X' + iY'\right)\downarrow' + Z'\uparrow'\right\rangle, \tag{11.85}$$
$$|u_6\rangle' = \left|\frac{1}{2}, -\frac{1}{2}\right\rangle' = \frac{1}{\sqrt{3}}\left|\left(X' - iY'\right)\uparrow' - Z'\downarrow'\right\rangle.$$

11.5.2 Wurtzite Structure

Analytical expressions for strain components for general crystal orientation can be obtained from the condition that the layer is grown pseudomorphically and the condition that strain energy density should be minimal. Figure 11.2 shows (a) configuration of the coordinate system (x', y', z') in $(hkil)$-oriented crystals and (b) a wurtzite primitive cell. The z-axis corresponds to the c-axis [0001] and the growth axis or the z'-axis is normal to the QW plane $(hkil)$. We define the unit vectors $\hat{x}', \hat{y}'$ and $\hat{z}'$ along the $\vec{x}'$-, $\vec{y}'$-, and $\vec{z}'$-axes and they are related to unit vectors $\hat{x}$, $\hat{y}$, and $\hat{z}\hat{z}'$ along the $\vec{x}$-, $\vec{y}$-, and $\vec{z}$-axes of the original crystal orientation (0001) by the rotation matrix Equation 11.69. We define the hexagonal primitive translational vectors as

$$\begin{aligned}
\vec{\alpha}_i &= a_i\hat{x}, \\
\vec{\beta}_i &= -\frac{a_i}{2}\hat{x} + \frac{\sqrt{3}a_i}{2}\hat{y}, \\
\vec{\gamma}_i &= c_i\hat{z},
\end{aligned} \tag{11.86}$$

where a_i and c_i are lattice constants of the hexagonal structure with the subscript i denoting the epilayer (e) and substrate (s), respectively. When the crystal is strained the primitive translation vectors become

$$\begin{aligned}
\vec{\alpha}''_i &= a_i\hat{x}'', \\
\vec{\beta}''_i &= -\frac{a_i}{2}\hat{x}'' + \frac{\sqrt{3}a_i}{2}\hat{y}'', \\
\vec{\gamma}''_i &= c_i\hat{z}'',
\end{aligned} \tag{11.87}$$

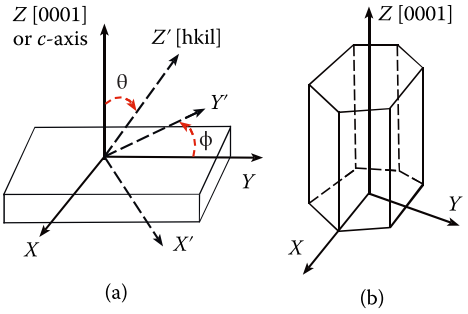

FIGURE 11.2 (a) Configuration of the coordinate system (x', y', z') in $(hkil)$-oriented crystals and (b) a wurtzite primitive cell.

with

$$\begin{aligned}\hat{x}'' &= (1+\varepsilon_{xx})\hat{x} + \varepsilon_{xy}\hat{y} + \varepsilon_{xz}\hat{z},\\ \hat{y}'' &= \varepsilon_{yx}\hat{x} + (1+\varepsilon_{yy})\hat{y} + \varepsilon_{yz}\hat{z},\\ \hat{z}'' &= \varepsilon_{zx}\hat{x} + \varepsilon_{zy}\hat{y} + (1+\varepsilon_{zz})\hat{z}.\end{aligned} \tag{11.88}$$

When a pseudomorphic interface is formed during the epitaxial growth, the translation vectors of each strained layer must have the same projection onto the growth plane, which is so-called the lattice commensurability constraint. By applying the lattice commensurability constraint on the QW and after some mathematical manipulations, we obtain the following analytical expressions for the strain tensors:

$$\begin{aligned}\varepsilon_{xx} &= \varepsilon_{xx}^{(0)} + \frac{\sin\theta\cos\phi}{\cos\theta}\varepsilon_{xz},\\ \varepsilon_{xy} &= \frac{\sin\theta\sin\phi}{\cos\theta}\varepsilon_{xz},\\ \varepsilon_{yy} &= \varepsilon_{xx}^{(0)} + \frac{\sin\theta\sin^2\phi}{\cos\theta\cos\phi}\varepsilon_{xz},\\ \varepsilon_{yz} &= \frac{\sin\phi}{\cos\phi}\varepsilon_{xz},\\ \varepsilon_{zz} &= \varepsilon_{zz}^{(0)} + \frac{\cos\theta}{\sin\theta\cos\phi}\varepsilon_{xz},\end{aligned} \tag{11.89}$$

where $\varepsilon_{xx}^{(0)} = \frac{a_s - a_e}{a_e}$ and $\varepsilon_{zz}^{(0)} = \frac{c_s - c_e}{c_e}$. The strain tensor ε_{xz} can be evaluated by minimizing the strain energy density which is given by

$$W = \frac{1}{2}\left[C_{11}\varepsilon_{xx}^2 + C_{11}\varepsilon_{yy}^2 + C_{33}\varepsilon_{12}^2 + 2C_{12}\varepsilon_{xx}\varepsilon_{yy} + 2C_{13}(\varepsilon_{xx}\varepsilon_{zz} + \varepsilon_{yy}\varepsilon_{zz}) + 4C_{44}\varepsilon_{zz}^2\right]. \tag{11.90}$$

By minimizing the strain energy density with respect to ε_{xz}, we get

$$\varepsilon_{xz} = -\frac{N(\theta,\phi)\sin\theta\cos\theta}{D(\theta,\phi)\cos\phi}\varepsilon_{xx}^{(0)}, \tag{11.91}$$

where

$$N(\theta,\phi) = \sin^2\theta\left(C_{11} + C_{12} + C_{13}\frac{\varepsilon_{zz}^{(0)}}{\varepsilon_{xx}^{(0)}}\right) + \cos^2\theta\left(2C_{13} + C_{33}\frac{\varepsilon_{zz}^{(0)}}{\varepsilon_{xx}^{(0)}}\right), \tag{11.92}$$

and

$$D(\theta,\phi) = \sin^4\theta\left(C_{11}\frac{\sin^4\phi + \cos^4\phi}{\cos^2\phi} + 2C_{12}\sin^2\phi\right) + C_{33}\frac{\cos^4\theta}{\cos^2\phi} + 2\left(\frac{C_{13}}{\cos^2\phi} + 2C_{44}\right)\sin^2\theta\cos^2\theta. \tag{11.93}$$

Using Equations 11.70 and 11.89, general expressions for the Hamiltonian for the valence-band structure in (x', y', z') coordinates can be written as

$$\tilde{H}'(\vec{k}', \bar{\bar{\varepsilon}}') = \begin{pmatrix} F' & -K'^* & -H'^* & 0 & 0 & 0 \\ -K' & G' & H' & 0 & 0 & \Delta \\ -H' & H'^* & \lambda' & 0 & \Delta & 0 \\ 0 & 0 & 0 & F' & -K' & H' \\ 0 & 0 & \Delta & -K'^* & G' & -H'^* \\ 0 & \Delta & 0 & H'^* & -H' & \lambda' \end{pmatrix}, \tag{11.94}$$

where

$$\begin{aligned}
F' &= \Delta_1 + \Delta_2 + \lambda' + \vartheta', \\
G' &= \Delta_1 - \Delta_2 + \lambda' + \vartheta', \\
\lambda' &= \frac{\hbar^2}{2m_0} A_1 \left(\sin^2\theta k_x'^2 - 2\sin\theta\cos\theta k_x' k_z' + \cos^2\theta k_z'^2 \right) \\
&\quad + \frac{\hbar^2}{2m_0} A_2 \left(\cos^2\theta k_x'^2 + s\cos\theta\sin\theta k_x' k_z' + \sin^2\theta k_z'^2 + k_y'^2 \right) \\
&\quad + D_1 \left(\varepsilon_{zz}^{(0)} - \frac{N(\theta,\phi)\cos^2\theta}{D(\theta,\phi)\cos^2\phi} \varepsilon_{xx}^{(0)} \right) + D_2 \left(2 - \frac{N(\theta,\phi)\sin^2\theta}{D(\theta,\phi)\cos^2\phi} \right) \varepsilon_{xx}^{(0)}, \\
\vartheta' &= \frac{\hbar^2}{2m_0} A_3 \left(\sin^2\theta k_x'^2 - 2\sin\theta\cos\theta k_x' k_z' + \cos^2\theta k_z'^2 \right) \\
&\quad + \frac{\hbar^2}{2m_0} A_4 \left(\cos^2\theta k_x'^2 + s\cos\theta\sin\theta k_x' k_z' + \sin^2\theta k_z'^2 + k_y'^2 \right) \\
&\quad + D_3 \left(\varepsilon_{zz}^{(0)} - \frac{N(\theta,\phi)\cos^2\theta}{D(\theta,\phi)\cos^2\phi} \varepsilon_{xx}^{(0)} \right) + D_4 \left(2 - \frac{N(\theta,\phi)\sin^2\theta}{D(\theta,\phi)\cos^2\phi} \right) \varepsilon_{xx}^{(0)}, \\
K' &= \frac{\hbar^2}{2m_0} A_5 e^{2i\phi} \left(\cos^2\theta k_x'^2 - k_y'^2 + \sin^2\theta k_z'^2 + 2\sin\theta\cos\theta k_x' k_y' + 2i\cos\theta k_x' k_y' + 2i\sin\theta k_y' k_z' \right) \\
&\quad - D_5 \frac{N(\theta,\phi)\left(\cos^2\phi - \sin^2\phi - 2i\sin\phi\cos\phi\right)\sin^2\theta}{D(\theta,\phi)\cos^2\phi} \varepsilon_{xx}^{(0)}, \\
H' &= \frac{\hbar^2}{2m_0} A_6 e^{i\phi} \left\{ \sin\theta\cos\theta \left(k_z'^2 - k_x'^2 \right) + \cos^2\theta k_x' k_z' - \sin^2\theta k_x' k_z' + i \left(\cos\theta k_y' k_z' - \sin\theta k_x' k_y' \right) \right\} \\
&\quad - D_6 e^{i\phi} \frac{N(\theta,\phi)\sin\theta\cos\theta}{D(\theta,\phi)\cos^2\phi} \varepsilon_{xx}^{(0)}.
\end{aligned} \tag{11.95}$$

We note that the old bases $\{|1\rangle, |2\rangle, |3\rangle, |4\rangle, |5\rangle, |6\rangle\}$ are used in the matrix representation of the Hamiltonian in Equation 11.94 and we assumed the hexagonal symmetry in the calculation. The Hamiltonian for nonpolar wurtzite QW structures can be obtained by substituting $\theta = \pi/2$ into the above equations.

11.5.3 Interband Optical-Matrix Elements for QW with an Arbitrary Crystal Orientation

11.5.3.1 (11*n*)-Oriented Zinc-Blende QW

The hole wave function in a QW can be written as

$$\Psi_m(z';k_t') = \frac{e^{i\vec{k}_t\cdot\vec{r}_t}}{\sqrt{A}}\sum_{\nu=1}^{6} g_m^{(\nu)}(z';k_t')\,|u_\nu\rangle\,, \tag{11.96}$$

where $g_m^{(\nu)}(z';k_t')(\nu = 1,2,3,4,5,$ and 6) is the wave function for the mth sub-band in (x',y',z') coordinates, $|\rangle\, u_\nu\rangle'$ is given by Equation 11.85, and $g_m'^{(i)}$ follows the normalization rules

$$\sum_{\nu=1}^{6}\int \mathrm{d}z'\,|g_m^{(\nu)}(z';\ k_t')|^2 = 1. \tag{11.97}$$

Since we solve one Schrödinger equation for electrons and 6×6 Hamiltonian for holes, we need only the optical matrix elements for two cases: spin up and spin down. The polarization-dependent interband momentum-matrix elements are written as follows.

TE polarization ($\hat{e}' = \cos\varphi'\hat{x}' + \sin\varphi'\hat{y}'$):

$$\begin{aligned}
\left|\hat{e}'\cdot\vec{M}'^{\uparrow}\right|^2 &= \left|\cos\varphi'\left\{-\frac{1}{\sqrt{2}}P_x\left\langle g_m'^{(1)}|\phi_l\right\rangle + \frac{1}{\sqrt{6}}P_x\left\langle g_m'^{(3)}|\phi_l\right\rangle + \frac{1}{\sqrt{3}}P_x\left\langle g_m'^{(6)}|\phi_l\right\rangle\right\}\right.\\
&\quad\left. + \sin\varphi'\left\{-i\frac{1}{\sqrt{2}}P_x\left\langle g_m'^{(1)}|\phi_l\right\rangle - i\frac{1}{\sqrt{6}}P_x\left\langle g_m'^{(3)}|\phi_l\right\rangle - i\frac{1}{\sqrt{3}}P_x\left\langle g_m'^{(6)}|\phi_l\right\rangle\right\}\right|^2\\
\left|\hat{e}'\cdot\vec{M}'^{\downarrow}\right|^2 &= \left|\cos\varphi'\left\{-\frac{1}{\sqrt{6}}P_x\left\langle g_m'^{(2)}|\phi_l\right\rangle + \frac{1}{\sqrt{2}}P_x\left\langle g_m'^{(4)}|\phi_l\right\rangle + \frac{1}{\sqrt{3}}P_x\left\langle g_m'^{(5)}|\phi_l\right\rangle\right\}\right.\\
&\quad\left. + \sin\varphi'\left\{-i\frac{1}{\sqrt{6}}P_x\left\langle g_m'^{(2)}|\phi_l\right\rangle - i\frac{1}{\sqrt{2}}P_x\left\langle g_m'^{(4)}|\phi_l\right\rangle + i\frac{1}{\sqrt{3}}P_x\left\langle g_m'^{(5)}|\phi_l\right\rangle\right\}\right|^2.
\end{aligned} \tag{11.98}$$

TE polarization ($\hat{e}' = \hat{z}'$):

$$\begin{aligned}
\left|\hat{e}'\vec{M}'^{\uparrow}\right|^2 &= \left|-\frac{2}{\sqrt{6}}P_z\left\langle g_m'^{(2)}|\phi_l\right\rangle + \frac{1}{\sqrt{3}}P_z\left\langle g_m'^{(5)}|\phi_l\right\rangle\right|^2\\
\left|\hat{e}'\cdot\vec{M}'^{\downarrow}\right|^2 &= \left|\frac{2}{\sqrt{6}}P_x\left\langle g_m'^{(3)}|\phi_l\right\rangle - \frac{1}{\sqrt{3}}P_x\left\langle g_m'^{(6)}|\phi_l\right\rangle\right|^2.
\end{aligned} \tag{11.99}$$

11.5.3.2 Wurtzite QW

The optical matrix elements general crystal orientation are given by

$$\left|\hat{e}'\cdot\vec{M}'^{\eta}\right|^2 = \left|\left\langle\Phi_l'^{\eta}\right|\hat{e}'\cdot\vec{p}'\,|\rangle\,\Psi_m'\right\rangle\right|^2. \tag{11.100}$$

The polarization-dependent interband momentum-matrix elements can be written as follows. TE polarization $\left(\hat{e}' = \cos\varphi'\hat{x} + \sin\varphi'\hat{y}\right)$:

$$\begin{aligned}
\hat{e}'\cdot\vec{M}'^{\uparrow} &= -\cos\varphi'\sin\theta P_z\left\langle\varphi_l'\,\middle|\,g_m'^{(3)}\right\rangle \\
&\quad -\frac{1}{\sqrt{2}}\cos\varphi'\cos\theta\cos\phi P_x\left\{(1+i)\left\langle\varphi_l'\,\middle|\,g_m'^{(1)}\right\rangle-(1-i)\left\langle\varphi_l'\,\middle|\,g_m'^{(2)}\right\rangle\right\} \\
&\quad +\frac{1}{\sqrt{2}}\sin\varphi' P_x\left\{\sin\phi\left(\left\langle\varphi_l'\,\middle|\,g_m'^{(1)}\right\rangle-\left\langle\varphi_l'\,\middle|\,g_m'^{(2)}\right\rangle\right)-i\cos\phi\left(\left\langle\varphi_l'\,\middle|\,g_m'^{(1)}\right\rangle+\left\langle\varphi_l'\,\middle|\,g_m'^{(2)}\right\rangle\right)\right\}, \\
\hat{e}'\cdot\vec{M}'^{\downarrow} &= -\cos\varphi'\sin\theta P_z\left\langle\varphi_l'\,\middle|\,g_m'^{(6)}\right\rangle \\
&\quad +\frac{1}{\sqrt{2}}\cos\varphi'\cos\theta\cos\phi P_x\left\{(1-i)\left\langle\varphi_l'\,\middle|\,g_m'^{(4)}\right\rangle-(1+i)\left\langle\varphi_l'\,\middle|\,g_m'^{(5)}\right\rangle\right\} \\
&\quad -\frac{1}{\sqrt{2}}\sin\varphi' P_x\left\{\sin\phi\left(\left\langle\varphi_l'\,\middle|\,g_m'^{(4)}\right\rangle-\left\langle\varphi_l'\,\middle|\,g_m'^{(5)}\right\rangle\right)+i\cos\phi\left(\left\langle\varphi_l'\,\middle|\,g_m'^{(4)}\right\rangle+\left\langle\varphi_l'\,\middle|\,g_m'^{(5)}\right\rangle\right)\right\}.
\end{aligned} \tag{11.101}$$

TM polarization $\left(\hat{e}' = \hat{z}'\right)$:

$$\begin{aligned}
\hat{e}'\cdot\vec{M}'^{\uparrow} &= \cos\theta P_z\left\langle\varphi_l'\,\middle|\,g_m'^{(3)}\right\rangle+\frac{1}{\sqrt{2}}\sin\theta\cos\phi\left\{-\left\langle\varphi\prime_l\,\middle|\,g_m'^{(1)}\right\rangle+\left\langle\varphi_l'\,\middle|\,g_m'^{(2)}\right\rangle\right\} \\
&\quad -\frac{i}{\sqrt{2}}\sin\theta\sin\phi\left\{\left\langle\varphi_l'\,\middle|\,g_m'^{(1)}\right\rangle+\left\langle\varphi_l'\,\middle|\,g_m'^{(2)}\right\rangle\right\}, \\
\hat{e}'\cdot\vec{M}'^{\downarrow} &= \cos\theta P_z\left\langle\varphi_l'\,\middle|\,g_m'^{(6)}\right\rangle+\frac{1}{\sqrt{2}}\sin\theta\cos\phi\left\{\left\langle\varphi_l'\,\middle|\,g_m'^{(4)}\right\rangle-\left\langle\varphi_l'\,\middle|\,g_m'^{(5)}\right\rangle\right\} \\
&\quad -\frac{i}{\sqrt{2}}\sin\theta\sin\phi\left\{\left\langle\varphi_l'\,\middle|\,g_m'^{(4)}\right\rangle+\left\langle\varphi_l'\,\middle|\,g_m'^{(5)}\right\rangle\right\},
\end{aligned} \tag{11.102}$$

11.6 Optical Gain Model with Many-Body Effects

The optical gain spectra are calculated using the non-Markovian gain model with many-body effects [11,13]. The many-body effects include the plasma screening, band-gap renormalization (BGR), and the excitonic or the Coulomb enhancement (CE) of the interband transition probability [14]. The simplest non-Markovian quantum kinetics is Gaussian line-shape function, which is connected with memory effects in the system–reservoir interaction [13]. The optical gain with many-body effects including the effects of anisotropy on the valence-band dispersion is given by

$$g(\omega) = \sqrt{\frac{\mu_o}{\varepsilon}}2\left(\frac{e^2}{m_o^2\omega}\right)\int_0^{2\pi}\mathrm{d}\varphi_0\int_0^{\infty}\mathrm{d}k_{||}\frac{2k_{||}}{(2\pi)^2 L_w}\left|M_{nm}(k_{||},\varphi_0)\right|^2\left[f_l^c(k_{||},\varphi_0)-f_m^v(k_{||},\varphi_0)\right]L(\omega,k_{||},\varphi_0), \tag{11.103}$$

where ω is the angular frequency, μ_o is the vacuum permeability, ε is the dielectric constant, e is the charge of an electron, m_o is the free electron mass, $k_{||}$ is the magnitude of the in-plane wave vector in the QW

plane, L_w is the well width, $\left|M_{nm}(k_{||}, \varphi_0)\right|^2$ is the momentum matrix element in the strained QW, f_l^c and f_m^v are the Fermi functions for occupation probability by the electrons in the conduction sub-band states and the valence sub-band states, respectively, and the indices l and m denote the electron states in conduction sub-band and heavy hole (light hole) sub-band states, respectively. Also, $E_{lm}(\hbar\omega, k_{||}, \varphi_0) = E_l^c(k_{||}, \varphi_0) - E_m^v(k_{||}, \varphi_0) + E_g + \Delta E_{SX} + \Delta E_{CH} - \hbar\omega$ is the renormalized transition energy between electrons and holes, where E_g is the band gap of the material, and ΔE_{SX} and ΔE_{CH} are the screened exchange and the Coulomb-hole contributions [14] to the BGR, respectively. The factor $Q(\hbar\omega, k_{||}, \varphi_0)$ accounts for the excitonic or CE of the interband transition probability [14,15]. The Gaussian line-shape function $L(E_{lm}(\hbar\omega, k_{||}, \varphi_0)$ renormalized with many-body effects is given by

$$\mathrm{Re}[L(E_{lm}(\hbar\omega, k_{||}, \varphi_0))] = \sqrt{\frac{\pi\tau_{in}(\hbar\omega, k_{||}, \varphi_0)\tau_c}{2\hbar^2}} \exp\left(-\frac{\tau_{in}(\hbar\omega, k_{||}, \varphi_0)\tau_c}{2\hbar^2} E_{lm}^2(\hbar\omega, k_{||}, \varphi_0)\right)$$

$$\mathrm{Im}[L(E_{lm}(\hbar\omega, k_{||}, \varphi_0))] = \frac{\tau_c}{\hbar}\int_0^\infty \mathrm{d}t \exp\left(-\frac{\tau_c}{2\tau_{in}(\hbar\omega, k_{||}, \varphi_0)} t^2\right) \sin\left(\frac{\tau_c E_{lm}(\hbar\omega, k_{||}, \varphi_0)}{\hbar} t\right). \quad (11.104)$$

The correlation time τ_c is related to the non-Markovian enhancement of optical gain [13] and is assumed to be constant. The τ_{in} and the τ_c used in the calculation are 25 and 10 fs, respectively. The BGR is given as a summation of the Coulomb-hole self-energy and the screened-exchange shift [14]. The φ_0 dependence of the BGR is very small and neglected for simplicity. The Coulomb-hole contribution to the BGR is written as

$$\Delta E_{CH} = -2E_R a_o \lambda_s \ln\left(1 + \sqrt{\frac{32\pi N L_w}{C\lambda_s^3 a_o}}\right), \quad (11.105)$$

where N is the carrier density, λ_s is the inverse screening length, and C is a constant usually taken between 1 and 4. The Rydberg constant E_R and the exciton Bohr radius a_o are given by

$$\Delta E_R(eV) = 13.6\frac{m_o/m_r}{(\varepsilon/\varepsilon_0)^2} \quad (11.106)$$

and

$$a_o(\text{Å}) = 0.53\frac{\varepsilon/\varepsilon_0}{m_o/m_r}, \quad (11.107)$$

where m_r is the reduced electron–hole mass defined by $1/m_r = 1/m_e + 1/m_h$. The exchange contribution to the BGR is given by

$$\Delta E_{SX} = -\frac{2E_R a_o}{\lambda_s}\int_0^\infty \mathrm{d}k_{||} k_{||} \frac{1 + \frac{C\lambda_s a_o k_{||}^2}{32\pi N L_w}}{1 + \frac{k_{||}}{\lambda_s} + \frac{C a_o k_{||}^3}{32\pi N L_w}} \left[f_n^c(k_{||}) + 1 - f_m^v(k_{||})\right]. \quad (11.108)$$

The factor $1/(1 - Q(k_{||}, \hbar\omega))$ represents the CE in the Padé approximation. Here, the factor $Q(k_{||}, \hbar\omega)$ is given by [11,14]

$$Q(k_{||}, \hbar\omega) = i\frac{E_R a_o}{\pi\lambda_s \left|M_{nm}(k_{||})\right|^2} \int_0^\infty dk'_{||} k'_{||} \left|M_{nm}(k_{||})\right| [f_n^c(k'_{||}) - f_m^v(k'_{||})]\Xi(E_{lm}(\hbar\omega, k_{||}))\Theta(k_{||}, k'_{||}), \quad (11.109)$$

where

$$\Theta(k_{||}, k'_{||}) = \int_0^\infty d\theta \frac{1 + \frac{C\lambda_s a_o q_{||}^2}{32\pi N L_w}}{1 + \frac{q_{||}}{\lambda_s} + \frac{C a_o q_{||}^3}{32\pi N L_w}} \quad (11.110)$$

and

$$q = \left|k_{||} - k'_{||}\right|. \quad (11.111)$$

The spontaneous emission coefficient $g_{SP}(\omega)$ can be obtained by replacing $[f_l^c(k_{||}, \varphi_0) - f_m^v(k_{||}, \varphi_0)]$ in Equation 11.103 by $f_l^c(k_{||}, \varphi_0)(1 - f_m^v(k_{||}, \varphi_0))$.

$$g_{SP}(\omega) = \sqrt{\frac{\mu_o}{\varepsilon}} 2 \left(\frac{e^2}{m_o^2 \omega}\right) \int_0^{2\pi} d\varphi_0 \int_0^\infty dk_{||} \frac{2k_{||}}{(2\pi)^2 L_w} \left|M_{nm}(k_{||}, \varphi_0)\right|^2 f_l^c(k_{||}, \varphi_0)(1 - f_m^v(k_{||}, \varphi_0))L(\omega, k_{||}, \varphi_0). \quad (11.112)$$

The spontaneous emission rate $r_{spon}(E)$ can be obtained from our calculated spontaneous emission spectrum $g_{SP}(E)$ by using

$$r_{spon}(E) = \left(\frac{4n^2}{\hbar\lambda^2}\right) g_{SP}(E),$$

where n is the refractive index of the QW and $E = 2\pi c\hbar/\lambda$ with c being the speed of light.

11.7 Numerical Example

As a numerical example, we calculate valence-band structure, optical matrix element, and optical gain as a function of crystal angle for zinc-blende and wurtzite GaN/AlGaN QW structures. For zinc-blende case, we consider the compressively strained GaN well by assuming that QW structures are grown on AlGaN substrate. On the other hand, in the case of the wurtzite semiconductor, we assume that QW structures are grown on GaN substrate. The material parameters for zinc-blende GaN and AlN used in the calculation were taken from Refs. [9,26–32] and references therein. Also, the material parameters for wurtzite GaN and AlN used in the calculation were taken from Refs. [33–35] and references therein. All parameters used in the calculation are summarized in Table 11.1.

11.7.1 Zinc-Blende Structure

Figure 11.3 shows the valence-band structures along $k_{x'}$ and $k_{y'}$ of (a) (001)-, (b) (111)-, and (c) (110)-oriented zinc-blende GaN/$Al_xGa_{1-x}N$ QWs (L_w = 2.5 nm). The sub-bands are labeled HH_i and LS_i, where i denotes the sub-band level. Here SL is the acronym for "splitoff-hole-light-hole," and the first letter denotes the dominant component for the wave function. The crystal angles for QWs with (11n) orientations

TABLE 11.1 Parameters Used in the Calculation

	Wurtzite			Zinc-Blende	
Parameters	GaN	AlN		GaN	AlN
Lattice constant (Å)					
a	3.1892	3.112	a	4.460	4.342
c	5.185	4.982			
Energy parameter					
E_g(eV)	3.44	6.16	E_g(eV)	3.1	4.9
$\Delta_1 = \Delta_{cr}$(meV)	22.0	−58.5	Δ (meV)	11.0	11.0
$\Delta_2 = \Delta/3$ (meV)	15.0	20.4			
$\Delta_3 = \Delta_2$(meV)	5.0	6.8			
Conduction-band effective masses					
m^w_{ez}/m_o	0.2	0.3	m^c_e/m_o	0.13	0.21
$m^w_{et} = m^w_{ez}$					
Valence-band effective-mass parameters					
A_1	−6.56	−3.95	γ_1	3.06	2.42
A_2	−0.91	−0.27	γ_2	0.91	0.58
A_3	−3.13	−1.95	γ_3	1.03	0.71
Deformation potentials					
a_c	−4.6	−4.5	a_c	−2.77	−6.8
D_1	−1.7	−2.89	a_v	3.63	2.3
D_2	6.3	4.89	b	−2.67	−1.5
D_3	−4.0	−3.34	d	−4.62	−4.5
Dielectric constant					
ε	10.0	8.5	ε	10.69	8.5
Elastic stiffness constant					
C_{11}	39.0	39.8	C_{11}	29.6	30.4
C_{12}	14.5	14.0	C_{12}	15.4	15.2
C_{13}	10.6	12.7			
C_{33}	39.8	38.2			
C_{44}	10.5	9.6	C_{44}	20.0	19.9
C_{66}	12.3	12.9			
Piezoelectric constant					
d_{31}(×10^{-12} m/V)	−1.7	−2.0	e_{14}(C/m^2)	−1.11	−0.526
Spontaneous polarization (C/m^2)					
P_{SP}	−0.029	−0.081			

($\varphi = \pi/4$) are obtained by using the relation $\theta = \tan^{-1}(\sqrt{2}/n)$. For example, crystal angles for (111) and (110) are $\theta \approx 55°$ and $\theta \approx 90°$, respectively. The valence-band structures of the self-consistent model are calculated at a carrier density of $N_{2D} = 10 \times 10^2$ cm^{-2}. The strain-induced piezoelectric field in the GaN well for the (111) orientation is about 1.33 MV/cm. The sub-band structures for different orientations are apparently different. For the (001) and the (111) orientations, the sub-band structures are nearly spherically symmetric. On the other hand, the valence-band structure of the (110)-oriented QW shows anisotropy. In particular, the effective mass of the first sub-band along $k_{y'}$ is observed to be greatly reduced for the (110)-oriented structure. A larger energy spacing between the first two sub-bands (HH1 and SL1) and higher sub-bands are observed for the (001)-oriented structure. The increase in the sub-band energy spacing will reduce the carrier population in the higher sub-bands. However, the energy spacing is gradually reduced with increasing polar angle θ.

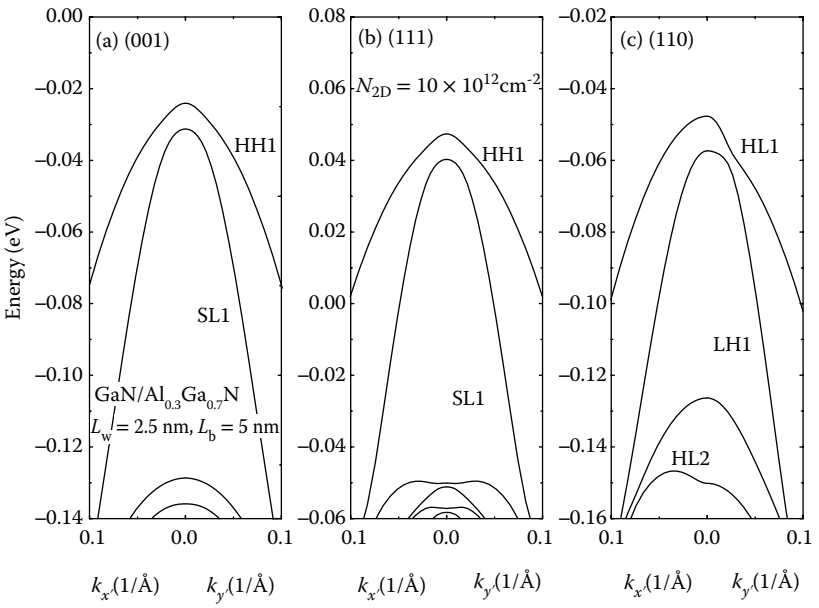

FIGURE 11.3 Valence-band structures along $k_{x'}$ and $k_{y'}$ of (a) (001)-, (b) (111)-, and (c) (110)-oriented zinc-blende GaN/Al$_x$Ga$_{1-x}$N QWs (L_w = 2.5 nm). The sub-bands are labeled HH$_i$ and LS$_i$, where i denotes the sub-band level.

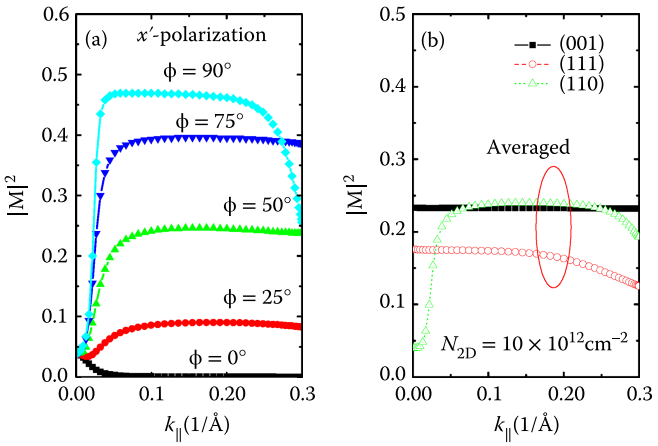

FIGURE 11.4 (a) x'-Polarized normalized optical matrix element as a function of $k_{||}$ for various values of φ for (110)-oriented zinc-blende GaN/Al$_x$Ga$_{1-x}$N quantum wells (QWs) (L_w = 2.5 nm) and (b) averaged optical matrix element as a function of $k_{||}$ for (001)-, (111)-, and (110)-oriented zinc-blende GaN/Al$_x$ Ga$_{1-x}$N QWs (L_w = 2.5 nm).

Figure 11.4 shows the x'-polarized normalized optical matrix element as a function of $k_{||}$ for various values of φ for (110)-oriented zinc-blende GaN/Al$_x$Ga$_{1-x}$N QWs (L_w = 2.5 nm) and (b) the averaged optical matrix element as a function of $k_{||}$ for (001)-, (111)-, and (110)-oriented zinc-blende GaN/Al$_x$Ga$_{1-x}$N QWs (L_w = 2.5 nm). The angle φ is defined by $\tan\phi = k_{y'}/k_{x'}$ and $k_{||} = \sqrt{k_{x'}^2 + k_{y'}^2}$. Here, φ = 0° and 90° mean that optical matrix elements are plotted along $k_{x'}$ and $k_{y'}$. The optical matrix elements of the self-consistent model are calculated at a carrier density of $N_{2D} = 10 \times 10^{12}$ cm^{-2}. The optical matrix element

increases with increasing angle between $k_{x'}$ and $k_{y'}$ wave vectors, that is, with changing from the x' to y' direction. The optical matrix elements averaged in the $k_{x'} - k_{y'}$ plane show that they significantly depend on the crystal angle. The optical matrix elements of the (001) orientation are nearly independent of the wave vectors. On the other hand, in the case of the (110) orientation, the optical matrix element rapidly increases with increasing wave vectors $k_{||}$ and begins to decrease in a range of large $k_{||}$. The optical matrix elements of the (111)-oriented QW structure slowly decreases with increasing wave vectors $k_{||}$. The (111)-oriented QW structure has a smaller matrix element than the (001)- or the (110)-oriented QW because the spatial separation between the electron and the hole wave functions increases due to the piezoelectric field, which results in a reduction in the transition probability between an electron and a hole.

Figure 11.5 shows x'- and y'-polarized many-body optical gain spectra of (a) (001)-, (b) (111)-, and (c) (110)-oriented zinc-blende GaN/$Al_xGa_{1-x}N$ QWs ($L_w = 2.5$ nm). Optical gains of the self-consistent model are calculated at a carrier density of $N_{2D} = 10 \times 10^{12}$ cm^{-2}. They are also averaged in the $k_{x'} - k_{y'}$ plane because of their anisotropy in the QW plane. The optical gain spectra have peaks corresponding to C1-HH1 and C1-HL1 transitions. The (111)-oriented QW structure shows that y'-polarized optical gain is larger than x'-polarized optical gain. The in-plane optical anisotropy ρ is about −0.17. Here, ρ is defined as $(I_{x'} - I_{y'})/(I_{x'} + I_{y'})$, where $I_i (i = x', y')$ means an intensity of peak optical gain. On the other hand, in the case of the (111)-oriented QW structure, y'-polarized optical gain is shown to be similar to x'-polarized optical gain with $\rho \approx 0$.

For a given carrier density, the (111)-oriented QW structure has a smaller optical gain than the (001)- or the (110)-oriented QW. This can be explained by the fact that the (111)-oriented QW has much smaller matrix elements than the (001)- or the (110)-oriented QW structures, as shown in Figure 11.4b. This is because the spatial separation between the electron and the hole wave functions is increased due to the piezoelectric field, which results in a reduction in the transition probability between electrons and holes.

11.7.2 Wurtzite Structure

Figure 11.6 shows the valence-band structures of wurtzite GaN/$Al_xGa_{1-x}N$ QWs ($L_w = 2.5$ nm) with (a) $\theta = 0°$ (c-plane), (b) 30°, and (c) 90° (a-plane). Here, the naming of the sub-bands follows the dominant composition of the wave function at the Γ point in terms of the $|X'\rangle$, $|Y'\rangle$, and $|Z'\rangle$ bases. The components

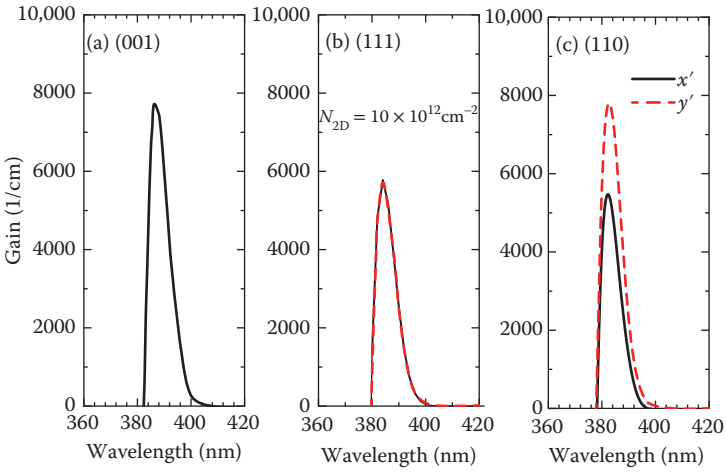

FIGURE 11.5 x'- and y'-polarized many-body optical gain spectra of (a) (001)-, (b) (111)-, and (c) (110)-oriented zinc-blende GaN/$Al_xGa_{1-x}N$ quantum wells (QWs) ($L_w = 2.5$ nm). Optical gains of the self-consistent model are calculated at a carrier density of N_{2D} 10×10^{12} cm^{-2}.

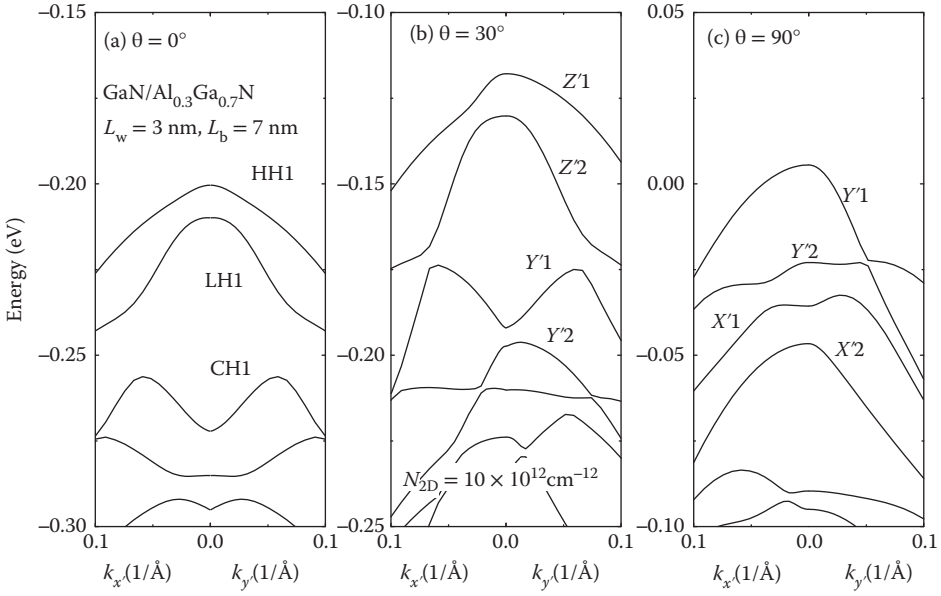

FIGURE 11.6 Valence-band structures of wurtzite GaN/$Al_xGa_{1-x}N$ quantum wells (QWs) (L_w = 2.5 nm) with (a) θ = 0° (*c*-plane), (b) 30°, and (c) 90° (*a*-plane).

$P_m^{i(=X',Y',Z')}$ of each wave function are given by

$$P_m^{X'} = \left\langle g_m'^{(3)} \mid g_m'^{(3)} \right\rangle + \left\langle g_m'^{(6)} \mid g_m'^{(6)} \right\rangle$$

$$P_m^{Y'} = \frac{\left\langle g_m'^{(1)} + g_m'^{(2)} \mid g_m'^{(1)} + g_m'^{(2)} \right\rangle + \left\langle g_m'^{(4)} + g_m'^{(5)} \mid g_m'^{(4)} + g_m'^{(5)} \right\rangle}{2}$$

$$P_m^{Z'} = \frac{\left\langle g_m'^{(2)} - g_m'^{(1)} \mid g_m'^{(2)} - g_m'^{(1)} \right\rangle + \left\langle g_m'^{(4)} - g_m'^{(5)} \mid g_m'^{(4)} - g_m'^{(5)} \right\rangle}{2}. \tag{11.113}$$

The valence-band structures of QW structures with θ = 30° and 90° show anisotropy in the QW plane, unlike the (0001)-oriented structure (*c*-plane). The effective mass of the topmost valence band along $k_{y'}$ is smaller than that along $k_{x'}$. Also, the hole effective mass (1.24 m_o for *a*-plane) of the topmost valence band along $k_{y'}$ is lower than that (1.56 m_o) of the *c*-plane. However, the hole effective mass (1.39 m_o) averaged in the $k_{x'} - k_{y'}$ QW plane is slightly larger than that along $k_{y'}$. Here, to estimate the magnitude of the hole effective mass, we considered a parabolic band fitted to the lowest sub-band of the exact band structure. The effective mass is determined so that, for a given carrier density and the quasi-Fermi level for holes, the carrier density and the quasi-Fermi level agree with those of the exact band structure. Hence, the effective mass of the fitted parabolic band reflects an averaged density of states.

Figure 11.7 shows the y'-polarized normalized optical matrix element as a function of $k_{||}$ for various values of φ for wurtzite *a*-plane GaN/$Al_xGa_{1-x}N$ QWs (L_w = 2.5 nm) and (b) averaged optical matrix element as a function of $k_{||}$ for QW structures with crystal angles of (a) θ = 0° (*c*-plane), (b) 30°, and (c) 90° (*a*-plane). The optical matrix elements of the self-consistent model are calculated at a carrier density of $N_{2D} = 10 \times 10^{12}$ cm^{-2}. The optical matrix element rapidly decreases with increasing $k_{||}$ and greatly depends on angle φ between $k_{x'}$ and $k_{y'}$ wave vectors. In particular, in the case with φ = 90°, the decrease in the optical matrix element is observed to be much more rapid, compared to the other case. The optical

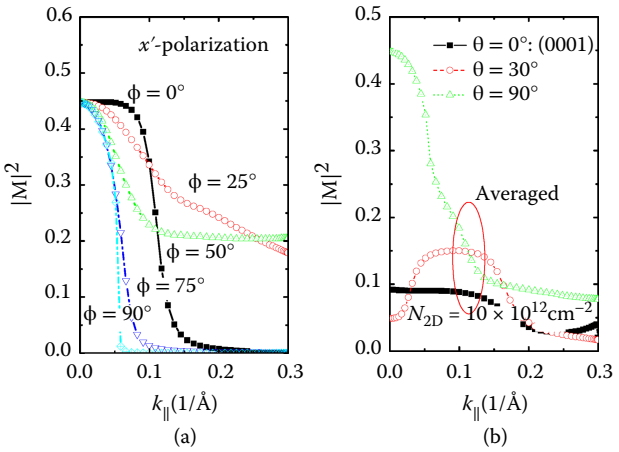

FIGURE 11.7 (a) y'-Polarized normalized optical matrix element as a function of $k_{||}$ for various values of φ for wurtzite a-plane GaN/$Al_xGa_{1-x}N$ quantum wells (QWs) (L_w = 2.5 nm) and (b) averaged optical matrix element as a function of $k_{||}$ for QW structures with crystal angles of θ = 0° (c-plane), 30°, and 90° (a-plane).

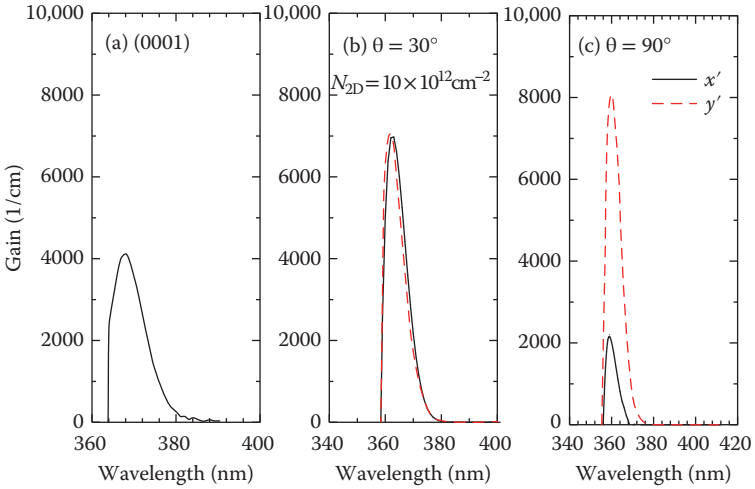

FIGURE 11.8 x'- and y'-polarized, many-body optical gain spectra for quantum well (QW) structures with crystal angles of θ = 0° (c-plane), 30°, and 90° (a-plane).

matrix elements averaged in the $k_{x'} - k_{y'}$ plane show that they significantly depend on the crystal angle, as observed for the zinc-blende case of Figure 11.3. The optical matrix element rapidly increases with increasing crystal angle. This can be explained by the fact that the internal field is reduced for the QW structure with larger crystal angle. In particular, in the QW structure with θ = 90°, the internal field is disappeared and the transition probability between an electron and a hole is significantly enhanced due to the reduction in the spatial separation between the electron and the hole wave functions.

Figure 11.8 shows x'- and y'-polarized many-body optical gain spectra for QW structures with crystal angles of (a) θ = 0° (c-plane), (b) 30°, and (c) 90° (a-plane). Optical gains of the self-consistent model are calculated at a carrier density of $N_{2D} = 10 \times 10^{12}$ cm^{-2}. They are also averaged in the $k_{x'} - k_{y'}$ plane because of their anisotropy in the QW plane. The (0001)-oriented QW structure shows that y'-polarized

optical gain is the same as x'-polarized optical gain because of its isotropy in the QW plane. The QW structure with $\theta = 30°$ also shows that x'- and y'-polarized optical gains are similar to each other. On the other hand, the *a*-plane QW structure shows that y'-polarized optical gain is much larger than x'-polarized optical gain. The in-plane optical anisotropy ρ is about -0.58, which is much larger than that observed for the zinc-blende QW structures.

11.8 Summary

In summary, we reviewed theoretical formalism to obtain eigenvalues and wave functions of (001)-oriented zinc-blende and (0001)-oriented wurtzite QW structures. In addition, we reviewed crystal orientation effects on electronic and optical properties of strained zinc-blende and wurtzite QW structures. As a numerical example, we calculated valence-band structures, optical matrix elements, and optical gains as a function of crystal orientation for zinc-blende and wurtzite AlGaN/GaN QW structures. These results can be used for a design of QW-based optoelectronic devices.

References

1. Corzine, S. W., R. H. Yan, and L. A. Coldren. 1993. Optical gain in III-V bulk and quantum well semiconductors. In *Quantum Well Lasers*, edited by P. S. Zory, Jr. San Diego, CA: Academic Press.
2. Chuang, S. L. 1995. *Physics of Optoelectronic Devices*. New York, NY: Wiley.
3. Piprek, J. 2003. *Semiconductor Optoelectronic Devices: Introduction to Physics and Simulation*. San Diego, CA: Academic Press.
4. Ahn, D. and S.-H. Park. 2011. *Engineering Quantum Mechanics*. Hoboken, NJ: John Wiley & Sons.
5. Chao, C. Y.-P. and S. L. Chuang. 1992. Spin–orbit-coupling effects on the valence-band structure of strained semiconductor quantum wells. *Phys. Rev. B* **46**: 4110–4122.
6. Chuang, S. L. and C. S. Chang. 1996. The k.p method for strained wurtzite semiconductors. *Phys. Rev. B* **54**: 2491–2504.
7. Park, S.-H. and S. L. Chuang. 1999. Crystal-orientation effects on the piezoelectric field and electronic properties of strained wurtzite semiconductors. *Phys. Rev. B* **59**: 4725–4737.
8. Takeuchi, T., H. Amano, and I. Akasaki. 2000. Theoretical study of orientation dependence of piezoelectric effects in wurtzite strained GaInN/GaN heterostructures and quantum wells. *Jpn. J. Appl. Phys.* **39**: 413–416.
9. Park, S.-H. and S. L. Chuang. 2000. Comparison of zinc-blende and wurtzite GaN semiconductors with spontaneous polarization and piezoelectric field effects. *J. Appl. Phys.* **87**: 353–364.
10. Bir, G. L. and G. E. Pikus, 1974. *Symmetry and Strain-Induced Effects in Semiconductors*. New York, NY: Wiley.
11. Park, S.-H., S. L. Chuang, and D. Ahn. 2000. Intraband relaxation time effects on non-Markovian gain with many-body effects and comparison with experiment. *Semicond. Sci. Technol.* **15**: 203–208.
12. Chuang, S. L. 1996. Optical gain of strained wurtzite GaN quantum-well lasers. *IEEE J. Quan. Electron.* **32**: 1791–1800.
13. Ahn, D. 1997. Theory of non-Markovian optical gain in quantum-well lasers. *Prog. Quant. Electron.* **21**: 249–287.
14. Chow, W. W., S. W. Koch, and M. Sergent III. 1994. *Semiconductor-Laser Physics*. Berlin: Springer.
15. Haug, H. and S. W. Koch. 1993. *Quantum Theory of the Optical and Electronic Properties of Semiconductors*. Singapore: World Scientific.
16. Nakamura, S. and G. Fasol. 1997. *The Blue Laser Diode*. Berlin: Springer.
17. Martin, G., A. Botchkarev, A. Rockett, and H. Morkoç. 1996. Valence-band discontinuities of wurtzite GaN, AlN, and InN heterojunctions measured by x-ray photoemission spectroscopy. *Appl. Phys. Lett.* **68**: 2541–2543.

18. Bernardini, F., V. Fiorentini, and D. Vanderbilt. 1997. Spontaneous polarization and piezoelectric constants of III–V nitrides. *Phys. Rev. B*. **56**: 10024–10027.
19. Park, S.-H. and S. L. Chuang. 1998. Piezoelectric effects on electrical and optical properties of wurtzite GaN/AlGaN quantum well lasers. *Appl. Phys. Lett.* **72**: 3103–3105.
20. Ambacher, O., J. Smart, J. R. Shealy, N. G. Weimann, K. Chu, M. Murphy, W. J. Schaff, L. F. Eastman, R. Dimitrov, L. Wittmer, M. Stutzmann, W. Rieger, and J. Hilsenbeck. 1999. Two-dimensional electron gases induced by spontaneous and piezoelectric polarization charges in N- and Ga-face AlGaN/GaN heterostructures. *J. Appl. Phys*. **85**: 3222–3233.
21. Park, S.-H. 2002. Crystal orientation effects on electronic properties of wurtzite InGaN/GaN quantum wells. *J. Appl. Phys.* **91**: 9904–9908.
22. Henderson, R. H. and E. Towe. 1995. Strain and crystallographic orientation effects on interband optical matrix elements and band gaps of [II/]-oriented III–V epilayers. *J. Appl. Phys.* **78**: 2447–2455.
23. Henderson, R. H. and E. Towe. 1996. Effective mass theory for III–V semiconductors on arbitrary (hkl) surfaces. *J. Appl. Phys.* **79**: 2029–2037.
24. Bykhovski, A., B. Gelmont, and S. Shur. 1993. Strain and charge distribution in GaN-AIN-GaN semiconductor–insulator–semiconductor structure for arbitrary growth orientation. *Appl. Phys. Lett.* **63**: 2243–2245.
25. Smith, D. L. and C. Mailhiot. 1998. Piezoelectric effects in strained-layer superlattices. *J. Appl. Phys.* **63**: 2717–2719.
26. Wright, A. F. and J. S. Nelson. 1995. Consistent structural properties for AlN, GaN, and InN. *Phys. Rev. B* **51**: 7866.
27. Rubio, A., J. L. Corkill, M. L. Cohen, E. L. Shirley, and S. G. Louie. 1993. Quasiparticle band structure of AlN and GaN. *Phys. Rev. B.* **48**: 11810.
28. Hellwege, K. H. and O. Madelung (Eds.). 1982. *Physics of Group IV Elements and III–V Compounds*, Landolt-Börnstein, New Series, Group III, Vol. 17, Part A. Berlin: Springer-Verlag.
29. Fan, W. J., M. F. Li, T. C. Chong, and J. B. Xia. 1996. Electronic properties of zinc-blende GaN, AlN, and their alloys $Ga_{1-x}Al_xN$. *J. Appl. Phys.* **79**: 18896.
30. Meney, A. T. and E. P. O'Reilly. 1995. Theory of optical gain in ideal GaN heterostructure lasers. *Appl. Phys. Lett.* **67**: 3013.
31. Kim, K., W. R. L. Lambrecht, and B. Segall. 1996. Elastic constants and related properties of tetrahedrally bonded BN, AlN, GaN, and InN. *Phys. Rev. B* **53**: 16310.
32. Van de Walle, C. G. and J. Neugebauer. 1997. Small valence-band offsets at GaN/InGaN heterojunctions. *Appl. Phys. Lett.* **70**: 2577.
33. Vurgaftman, I. and J. R. Meyer. 2003. Band parameters for nitrogen-containing semiconductors. *J. Appl. Phys.* **94**: 3675–3696.
34. Park, S.-H. 2011. Optical gain characteristics of non-polar Al-rich AlGaN/AlN quantum well structures. *J. Appl. Phys.* **110**: 063105.
35. Park, S.-H., Y.-T. Moon, D.-S. Han, J. S. Park, M.-S. Oh, and D. Ahn. 2011. Light emission enhancement in blue InGaAlN/InGaN quantum well structures. *Appl. Phys. Lett.* **99**: 181101.

12

Nanowires

Oliver Marquardt

Vladimir M. Kaganer

and

Pierre Corfdir

12.1 Introduction: Semiconductor Nanowires

During the past years, much research effort has been dedicated to explore the properties of semiconductor nanowires (NWs). Correspondingly, remarkable progress has been achieved in design, growth, characterization, and theoretical description of these structures. Many different applications have been suggested, ranging from energy harvesting and storage (Wallentin et al. 2013) to single-photon emitters (Heiss et al. 2013) and general lighting, with some of them poised to enter the commercial market in the near future.

Free-standing semiconductor NWs, sometimes also referred to as *nanorods*, are grown using molecular-beam epitaxy (MBE) or metal-organic chemical vapor deposition (MOCVD) or produced from planar layers by lithography. They have diameters of a few tens to hundreds of nanometers and lengths of a few hundred nanometers to a few micrometers and exhibit some unique properties in comparison to planar semiconductor quantum wells (QWs) or embedded quantum dots (QDs). If the diameter of a NW is small enough such that quantum confinement along the transverse direction is ensured, the NW is called a *quantum wire*. A particular advantage of the NW geometry is that it facilitates a much better strain relaxation, resulting in a better overall material quality. Additionally, the large surface-to-volume ratio of NWs represents an advantage for light emitting and detecting devices.

If heterostructures such as QWs or QDs are incorporated in NWs, the free side facets additionally facilitate elastic relaxation, allowing one to design heterostructures from materials that have a large lattice mismatch with respect to the NW material, which could not be produced in a planar structure. Typical heterostructures in NWs are of either axial or radial character, as depicted in Figure 12.1. The active layer of the heterostructure is then either perpendicular to the growth direction (axial heterostructure) or parallel to it (radial heterostructure). Thickness and material composition of axial and radial layers can be controlled during the growth process, so that the electronic properties of the NW can be tailored to suit specific applications.

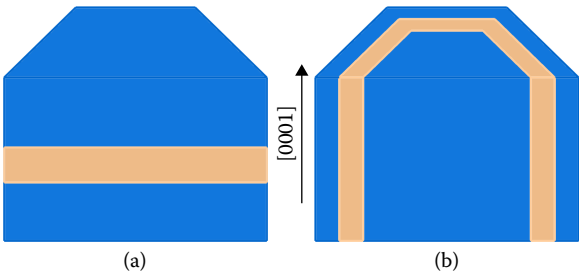

FIGURE 12.1 Cross-section of an axial (a) and a radial (b) heterostructure in a hexagonal nanowire (NW). The active layer is light gray.

Moreover, surface states at the side facets of NWs give rise to additional electrostatic potentials whose impact on the electronic properties needs to be understood. The simulation of semiconductor surfaces is a very challenging field and very few sophisticated approaches exist to model a semiconductor–vacuum interface.

Most III–V semiconductor NWs also show a high density of stacking faults. In some cases, the switching in the crystal structure along the length of an NW can be controlled. This allows one to investigate material properties of semiconductors in crystal structures that do not occur otherwise in the bulk. For example, whereas bulk GaAs exhibits a zinc-blende (ZB) crystal structure, GaAs NWs with the hexagonal wurtzite (WZ) structure have often been reported (Hoang et al. 2009; Heiss et al. 2011; Ahtapodov et al. 2012). Since the electronic properties of ZB and WZ GaAs are different, both phases have to be addressed in a simulation of an NW, with a particular focus on the treatment of their interface.

Finally, the photonic properties of NWs are altered by their peculiar geometry. In particular, the light absorption and extraction efficiencies depend strongly on the polarization of light with respect to the NW axis and it is mandatory to take this effect into account when correlating the degree of polarization of optical transitions with possible valence-band mixing in the NW.

12.2 Basic Charge Confining Mechanisms in NWs

In the following, we provide a brief overview of mechanisms that influence the localization and the energy levels of carriers in NWs. We compare these mechanisms to those that occur in planar layers or QDs and discuss specific features of NWs.

12.2.1 Bulk Electronic Properties: Band Offsets and Work Functions

Similar to planar heterostructures and QDs (see Chapters 11 and 13), the electronic properties of axial or radial semiconductor NW heterostructures are determined by the band offsets between the conduction and valence bands, E_c and E_v, of the materials involved. If we neglect surface states in a first step, the potential barrier arising at the interface between the NW and the surrounding vacuum is best described by the work function of the NW material. For the hole state, the following consideration yields a good description of the confining potential: The minimum energy to create a hole in the valence band is the sum of bandgap (to excite an electron from the valence to the conduction band) and the electron affinity. Therefore, we consider the sum of bandgap and electron affinity as a potential barrier for the valence band at the side facet of the NW. However, as the work function and the electron affinity of most semiconductors is on the order of a few eV and typically much larger than bandgaps and band offsets within an NW heterostructure, a good approximation can already be achieved by assuming an infinite potential at the semiconductor–vacuum interface. This simplification is in particular reasonable for the electron and hole states closest to the conduction and valence-band edges, respectively (Levine 1965).

12.2.2 Approaches to Compute Elastic Properties of Nanostructures

The electronic properties of a semiconductor nanostructure are significantly modified if the system is subject to elastic relaxation. The impact of strain on the electronic properties is commonly taken into account via deformation potentials. For example, within the ZB structure, the lowest conduction band is modified via the deformation potential a_c by the additional contribution $a_c \cdot \mathrm{Tr}(\varepsilon)$. In a WZ system, two deformation potentials a_1 and a_2 are employed such that the conduction band is modified by the term $a_2 \cdot (\varepsilon_{xx} + \varepsilon_{yy}) + a_1 \cdot \varepsilon_{zz}$. The strain-induced modifications of the valence bands are more complex. As an example, the impact of strain on the three highest valence bands can be computed using a six-band Pikus–Bir Hamiltonian (Ghosh et al. 2002). Chapter 1 provides a more detailed discussion on how strain is considered in the simulation of electronic properties of semiconductor nanostructures.

NWs exhibit quite specific strain profiles as the free side facets facilitate elastic relaxation. This feature allows not only the NW itself to relax toward the bulk lattice constants but also a much better elastic relaxation of axial and radial heterostructures incorporated within the wire. The possibility to relax even a large lattice mismatch is a particular strength of NWs in comparison to planar heterostructures, where large differences in the lattice constants lead to defects and crack formation.

To shed some light on the specific elastic properties of semiconductor NWs and their impact on the electronic properties, we discuss different approaches to compute the elastic properties of an NW in the following.

12.2.2.1 Valence Force Field Model

The most complete description of the elastic properties of an NW is achieved using an atomistic model. However, accurate models rely on *ab initio* methods, which are computationally highly expensive thus limiting the simulation domain to a small number of atoms (Yang et al. 2008; Xiang et al. 2008). Given the typical dimensions of semiconductor NWs in the order of tens to hundreds of nanometers, it is clear that such models are not well suited to describe NWs of the commonly observed dimensions.

An alternative approach is the computationally less expensive valence force field (VFF) model. Within the VFF model (Keating 1966), the atomic coordinates $\boldsymbol{r}_i$ of the individual atoms of the NW are the degrees of freedom. For an NW with a ZB crystal structure, the potential energy of the crystal to be minimized can be written as (Keating 1966):

$$E = \sum \sum \frac{3\alpha}{8r_{ij0}^2} \left(r_{ij}^2 - r_{ij0}^2 \right)^2 + \sum \sum \frac{3\beta}{8r_{ij0} \cdot r_{ik0}} \left(r_{ij} \cdot r_{ik} - r_{ij0} \cdot r_{ik0} \right)^2, \tag{12.1}$$

where the equilibrium bond angles satisfy $\boldsymbol{r}_{ij0} \cdot \boldsymbol{r}_{ik0} = -|\boldsymbol{r}_{ij0}||\boldsymbol{r}_{ik0}|/3$ and N indicates summation over the nearest neighbors. $\boldsymbol{r}_{ij} = \boldsymbol{r}_i - \boldsymbol{r}_j$ is the distance between atoms i and j and $\boldsymbol{r}_{ij0}$ denotes the value of $\boldsymbol{r}_{ij}$ in a bulk, strain-free crystal. The parameters α and β are coupling constants between neighboring atoms.

However, the basic VFF model is inaccurate for hexagonal crystal structures and thus requires improvements (Rücker and Methfessel 1995; Große and Neugebauer 2001), and it is questionable whether this model is an appropriate choice for the description of semiconductor NWs (Singh et al. 2011). Moreover, the computational effort of the VFF scales nearly linearly with the number of atoms involved. For typical NWs, even the VFF becomes computationally highly expensive if the super cell spans over a large segment of the NW.

12.2.2.2 Linear Continuum Elasticity Theory

As a computationally inexpensive alternative to the atomistic VFF model, the elastic properties of an NW can also be computed using a continuum-based approach. Here, the underlying atomistic lattice is ignored

and the strain tensor $\varepsilon_{ij}(\boldsymbol{r})$ is expressed in a continuous manner. The elastic energy to be minimized is then (Landau and Lifshitz 1959) given as follows:

$$E = \frac{1}{2}\int \sum C_{ijkl}(r)\varepsilon_{ij}(r)\varepsilon_{kl}(r)\mathrm{d}V, \tag{12.2}$$

where $C_{ijkl}(\boldsymbol{r})$ denotes the elastic constants of the materials involved and $i, j, k, l = x, y, z$. Note that all information on the underlying crystal structure is contained in the elastic constants. Within the ZB lattice, only the constants $C_{11} = C_{iiii} = C_{xxxx} = C_{yyyy} = C_{zzzz}$, $C_{12} = C_{iijj}$ $(i \neq j)$, and $C_{44} = C_{ijij}$ $(i \neq j)$ are nonzero. The advantage of the continuum elasticity model is of course its computational efficiency, as the discretization of the super cell is not bound to atomistic coordinates. However, the atomistic nature of the system under consideration is fully neglected.

12.2.3 Elastic Properties of NW Heterostructures

In this section, we discuss strain distributions in axial and radial NW heterostructures, as well as their impact on the electronic properties. We also discuss the specific advantages of axial and radial NW heterostructures in terms of elastic relaxation.

12.2.3.1 Radial NWs Heterostructures

In radial semiconductor NW heterostructures—typically referred to as core–shell NW—a core material is surrounded by one or more shells consisting of different materials or having different doping densities. The elastic properties of core and shells are controlled by their respective thicknesses, which finally allows the control of the electronic properties. In a simplified picture, a core–shell NW can be approximated as a cylinder with a lattice mismatched shell. The strain inside such a cylindrical core–shell NW is best described within cylindrical coordinates and the nonzero components of the strain tensor are (Menendez et al. 2011)

$$\begin{aligned}
\varepsilon_{rr} = \varepsilon_{\theta\theta} &= -\varepsilon_0\left(R_\mathrm{s}^2 - R_\mathrm{c}^2\right)\left[\frac{(\nu+1)(1-2\nu)}{(1-\gamma)(1-2\nu)R_\mathrm{c}^2 - (1-2\nu+\gamma)R_\mathrm{s}^2} + \frac{\nu}{\left(R_\mathrm{s}^2 - R_\mathrm{c}^2\right) + \gamma R_\mathrm{c}^2}\right] \\
\varepsilon_{zz} &= \varepsilon_0\left[\frac{\left(R_\mathrm{s}^2 - R_\mathrm{c}^2\right)}{\left(R_\mathrm{s}^2 - R_\mathrm{c}^2\right) + \gamma R_\mathrm{c}^2}\right]
\end{aligned} \tag{12.3}$$

in the core and

$$\begin{aligned}
\varepsilon_{rr}(r) &= \frac{\varepsilon_0(\nu+1)\gamma R_\mathrm{c}^2}{(1-\gamma)(1-2\nu)R_\mathrm{c}^2 - (1-2\nu+\gamma)R_\mathrm{s}^2}\left[(1-2\nu) - \frac{R_\mathrm{s}^2}{r^2}\right] + \frac{\nu\gamma R_\mathrm{c}^2\varepsilon_0}{\left(R_\mathrm{s}^2 - R_\mathrm{c}^2\right) + \gamma R_\mathrm{c}^2} \\
\varepsilon_{\theta\theta}(r) &= \frac{\varepsilon_0(\nu+1)\gamma R_\mathrm{c}^2}{(1-\gamma)(1-2\nu)R_\mathrm{c}^2 - (1-2\nu+\gamma)R_\mathrm{s}^2}\left[(1-2\nu) + \frac{R_\mathrm{s}^2}{r^2}\right] + \frac{\nu\gamma R_\mathrm{c}^2\varepsilon_0}{\left(R_\mathrm{s}^2 - R_\mathrm{c}^2\right) + \gamma R_\mathrm{c}^2} \\
\varepsilon_{zz} &= \frac{\gamma R_\mathrm{c}^2\varepsilon_0}{\left(R_\mathrm{s}^2 - R_\mathrm{c}^2\right) + \gamma R_\mathrm{c}^2}
\end{aligned} \tag{12.4}$$

in the shell, where r is the radial position. $\varepsilon_0 = \left(a_\mathrm{s} - a_\mathrm{c}\right)/a_\mathrm{c}$ is the initial misfit for core and shell materials with lattice constants a_c and a_s, respectively. R_c and R_s denote core and shell radii. For the Poisson ratio ν, we use the Voigt average (Hirth and Lothe 1968) and consider this ratio as being constant inside both the core and the shell. Within a cubic crystal, this ratio reads

$\nu^{\mathrm{c,s}} = C_{11}^{\mathrm{c,s}} + 4C_{12}^{\mathrm{c,s}} - 2C_{44}^{\mathrm{c,s}}/4C_{11}^{\mathrm{c,s}} + 6C_{12}^{\mathrm{c,s}} + 2C_{44}^{\mathrm{c,s}}$ where the $C_{ij}^{c,s}$ are the elastic constants of the core and the shell.

$\gamma = E_{\text{core}}/E_{\text{shell}}$ with E denoting Young's modulus. In the case of an NW grown along the [111] axis, this modulus can be determined from the stress–strain relation as (Hirth and Lothe 1968)

$$E_{111} = 3\frac{C_{44}\left(C_{11} + 2C_{12}\right)}{C_{44} + C_{11} + 2C_{12}}.$$

All components of the strain tensor are constants inside the core and depend only on the elastic constants of the materials involved as well as on the ratio between core and shell radii. Within the shell, ε_{zz} is constant. In contrast, $\varepsilon_{rr}(r)$ and $\varepsilon_{\theta\theta}(r)$ are functions of the radial distance from the NW core, and converge for large radii, as shown in Figure 12.2 for a Ge–Si core–shell NW with a core radius of 10 nm and a shell radius of 20 nm. The ratio between the radii of the core and the shell determines the strain state of the NW and is a property that can be easily controlled during the growth process.

For more realistic NW geometries that cannot be described using analytical models, the elastic properties can be computed only numerically using continuum-based or atomistic approaches. A comparison between the analytic solution for the cylindrical NW and a numerical solution using the finite-elements method for hexagonal, [011]-grown Ge–Si NWs can be found in Singh et al. (2011), where the authors report a very good agreement with the analytic model for the ε_{zz} component.

Discrepancies are observed in ε_{rr} and $\varepsilon_{\theta\theta}$ inside the shell toward the corners of the realistic, hexagonal NW, whereas a very reasonable agreement is reported in the vicinity of the core. Note, however, that ε_{rr} and $\varepsilon_{\theta\theta}$ are not uniform within the hexagonal core anymore, as derived analytically for a cylindrical NW. The agreement in ε_{zz} is very good with uniform values in both core and shell (Singh et al. 2011).

A comparative study between a continuum elasticity model and the atomistic VFF model is presented by Grönquist et al. (2009) for GaAs/GaP core–shell NWs with a total diameter of 27.8 nm. The agreement between VFF and continuum elasticity model is very good throughout the NW cross-section and deviations occur only at the free ends of the finite wires. For typical semiconductor NWs, elasticity theory within a continuum picture can thus be considered as a reasonable and computationally inexpensive model.

The elastic relaxation within the model NW has a significant influence on its electronic properties. $\text{Tr}(\varepsilon) > 0$ within the shell leads to a reduction of both the conduction and the valence-band energies. Within the core, the hydrostatic strain is negative and the energies of the conduction band minimum and valence band maximum increase. The impact of strain on the conduction band is larger, such that the bandgap of the core material (GaAs) increases and a red shift of luminescence from the core is predicted. As the magnitude of the strain depends on the thicknesses of the core and the shell, bandgap engineering can be applied in core–shell NWs by a variation of the core and shell thicknesses. An example for such bandgap engineering is presented in Liu et al. (2014), where the bandgap of an InAs shell around an InP core is tuned by up to 210 meV solely by a variation of the core diameter.

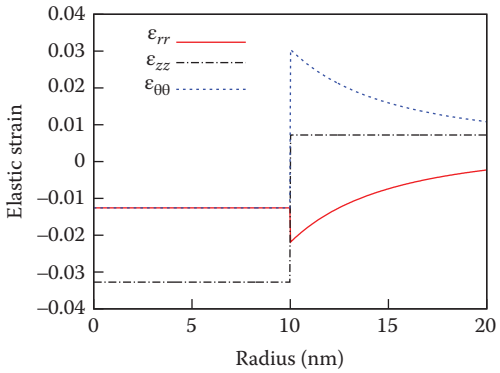

FIGURE 12.2 Nonzero strain components of a Ge–Si core–shell nanowire (NW) in cylindrical coordinates. The radius of the Ge core is 10 nm.

12.2.3.2 Axial NW Heterostructures

Axial NW heterostructures consist, similar to planar heterostructures, of thin layers separated by the matrix material (barriers). In contrast to their planar counterparts, the elastic strain caused by the mismatch between the insertion and the matrix crystal lattices relaxes at the side facets. This relaxation only partially eliminates the strain in the wells and gives rise to a complicated inhomogeneous strain distribution that depends on the relation between the NW diameter, the thickness of the insertion, the barrier thickness, and also the thickness of the whole stack of the insertions. The elastic relaxation facilitated by the free side facets is of particular importance for $In_xGa_{1-x}N$ disks in GaN NWs, as InN and GaN exhibit a lattice mismatch of approximately 10%, making it difficult to produce planar $In_xGa_{1-x}N$/GaN heterostructures with large In content and high crystal quality.

The difference in lattice parameters between the matrix and the inserted layer gives rise to the strain $\varepsilon^0_{zz} = (c - c_0)/c_0$, $\varepsilon^0_{rr} = \varepsilon^0_{\theta\theta} = (a - a_0)/a_0$, which itself does not cause stress and is called the intrinsic strain (the terms self-strain and eigenstrain are also used). a (a_0) and c (c_0) denote the layer (matrix) lattice constants. The elastic strain ε and the stress σ result from the coherency of the crystal lattices of the layer and the matrix. In a planar heterostructure, the layer is clamped by the matrix. Therefore, the components of the total strain in the plane $\varepsilon^T_{rr} = \varepsilon^0_{rr} + \varepsilon_{rr}$ and $\varepsilon^T_{\theta\theta} = \varepsilon^0_{\theta\theta} + \varepsilon_{\theta\theta}$ are equal to zero. The whole structure can freely expand in the direction normal to the layer, so that the stress $\sigma_{zz} = 0$. Using Hooke's law, these conditions lead to the stress in the layer plane $\sigma_{rr} = \sigma_{\theta\theta} = -\sigma^0$, where $\sigma^0 = (C_{11} + C_{12} - 2C^2_{13}/C_{33})\varepsilon^0_{rr}$. In NWs, this stress is relaxed at the free side facets. The solution of the above problem for the NW is, therefore, the sum of two terms: The solution in a planar structure described earlier and the image field due to stress $\sigma^{im}_{rr} = \sigma^0$ applied at the interface of the layer to satisfy the boundary conditions $\sigma_{rr} = 0$, $\sigma_{rz} = 0$ at the side facets. The elastic problem can be solved analytically for a circular cylinder shape of the NW and an arbitrary intrinsic strain distribution $\varepsilon^0(z)$ (Kaganer and Belov 2012). In this case, the image displacements are written as

$$u^{im}_r(x,\varsigma) = -\frac{R\sigma^0}{c_{44}}\int_0^\infty \left[\frac{1}{1+k_1}\frac{I_1(qx/\nu_1)}{I_1(q/\nu_1)} - \frac{1}{1+k_2}\frac{I_1(qx/\nu_2)}{I_1(q/\nu_2)}\right]\frac{g(q)}{D(q)}\cos q\varsigma \mathrm{d}q,$$

$$u^{im}_z(x,\varsigma) = -\frac{R\sigma^0}{c_{44}}\int_0^\infty \left[\frac{k_1\nu_1}{1+k_1}\frac{I_0(qx/\nu_1)}{I_1(q/\nu_1)} - \frac{k_2\nu_2}{1+k_2}\frac{I_0(qx/\nu_2)}{I_1(q/\nu_2)}\right]\frac{g(q)}{D(q)}\sin q\varsigma \mathrm{d}q,$$

$$D(q) = \frac{c_{11}-c_{12}}{c_{44}}\frac{k_2-k_1}{(1+k_1)(1+k_2)} - q\left(\nu_1\frac{I_0(q/\nu_1)}{I_1(q/\nu_1)} - \nu_2\frac{I_0(q/\nu_2)}{I_1(q/\nu_2)}\right). \tag{12.5}$$

I_0 and I_1 are modified Bessel functions, $\zeta = z/R$, and $x = r/R$ with R being the radius of the cylinder. The constants k_1, k_2, ν_1, ν_2 are defined by the elastic moduli of the material, which are $k_1 = 0.234$, $k_2 = 4.264$, $\nu_1 = 1.589$, $\nu_2 = 0.623$ for GaN. The function $g(z)$ describes the strain distribution and $g(q)$ its Fourier transform. For a homogeneous disk of thickness $2d$, it is given by $g(q) = 2sin(qd/R)/(\pi q)$.

Figure 12.3 illustrates the strain state in the disk center of a cylindrical NW with an inserted disk. In the limit of large diameters, the strain state approaches to the one of a planar structure. When the diameter is much smaller than the disk thickness, the inserted disk is strain free. As shown in Figure 12.3, the strain changes sign when going from very thin to thick NWs since both the insertion and the matrix are distorted. The insert in Figure 12.3 visualizes the displacements calculated by Equation 12.5 for the cylinder diameter corresponding to the minimum of ε_{zz} (exaggerated for better visibility).

The strain in the inserted disk is compensated by an opposite strain in the matrix. The displacements well above and well below the disk obtained using Equation 12.5 show that the total displacement due to transverse intrinsic strain is zero. In other words, the strain averaged over the whole cylinder is zero. This result has important consequences for multi-QW NW heterostructures since each QW experiences

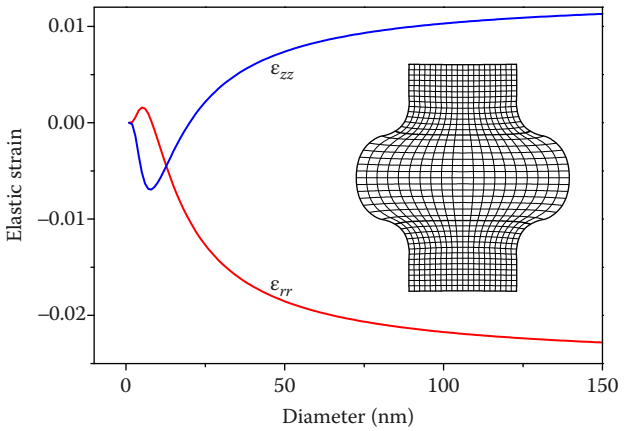

FIGURE 12.3 Elastic strain at the center of a 5-nm-thick $In_{0.25}Ga_{0.75}N$ disk inserted in a GaN cylinder as a function of the diameter of the cylinder. The inset shows the displacements (magnified for better visibility) for a cylinder diameter of 8 nm corresponding to the minimum of ε_{zz}.

strain produced by the other QWs. Variation of the barrier thickness and the NW diameter are therefore additional degrees of freedom for strain engineering of NW heterostructures (Wölz et al. 2013).

Figure 12.4 shows three strain states of lattice-mismatched heterostructures. In the case of a planar heterostructure (cf. Figure 12.4a), the Poisson effect gives rise to a homogeneous strain in the layer. In a multilayer NW heterostructure, the lateral elastic relaxation occurs in both materials. Figure 12.4b shows the displacement (center) and the out-of-plane strain (right) that were obtained for an axial superlattice using Equation 12.5. The strain is nonuniformly distributed. In the center of the cylinder, shown by the red line, the barrier crystal is laterally expanded. The QW is still under lateral compression, but to a smaller extent than in the planar case of Figure 12.4a. Going outwards, the strain distribution remains qualitatively similar to the one in the center up to about $0.8R$, shown by the green line. Toward the surface, the strain profile is more complex as shown by the blue line at $0.95R$.

This complicated strain state can be approximated by the strain state shown in Figure 12.4c, where it is assumed that the in-plane lattice spacing is constant everywhere in the superlattice. The strain can be determined by the condition that the average strain on the side surface is zero (Wölz et al. 2011, 2012). This relaxation reduces the strain in the QW by the factor $d_{\text{barrier}}/(d_{\text{barrier}} + d_{\text{well}})$ compared to the case of a planar heterostructure in Figure 12.4a. This lateral relaxation approximation is a much simpler description of the dominant strain states in the superlattice compared to the full consideration of the inhomogeneous strain resulting in a configuration as shown in Figure 12.4b. A systematic comparison of the full calculation based on Equation 12.5 with this approximation allows one to establish criteria for its validity (Wölz et al. 2013). First, the height of the whole stack of QWs should be larger than the NW diameter. Second, to achieve a large strain in QWs, the NWs should be sufficiently thick.

For more complicated geometries or material compositions, the strain can be determined by finite element calculations. For example, Figure 12.5 shows the strain distribution in an axial $In_xGa_{1-x}N$/GaN NW heterostructure containing two $In_xGa_{1-x}N$ insertions (Knelangen et al. 2011). Each insertion consists of a 2-nm-thick $In_xGa_{1-x}N$ layer with an In content of 19% and a 9-nm-thick layer with an In content of 39%. They are embedded in a GaN NW with a diameter of 30 nm and separated from the side facets by a 7-nm-thick GaN shell. The distance between the two disks is 3 nm. The magnitude of the strain increases toward the side facets, whereas elastic relaxation occurs near the center of the NW. Nevertheless, significant strain persists at the side facets of the NW.

A study on single $In_xGa_{1-x}N$ disks in a GaN NW that extend throughout the whole NW section without a GaN shell also reports persisting strains (Böcklin et al. 2010). An $In_{0.4}Ga_{0.6}N$ layer with a thickness

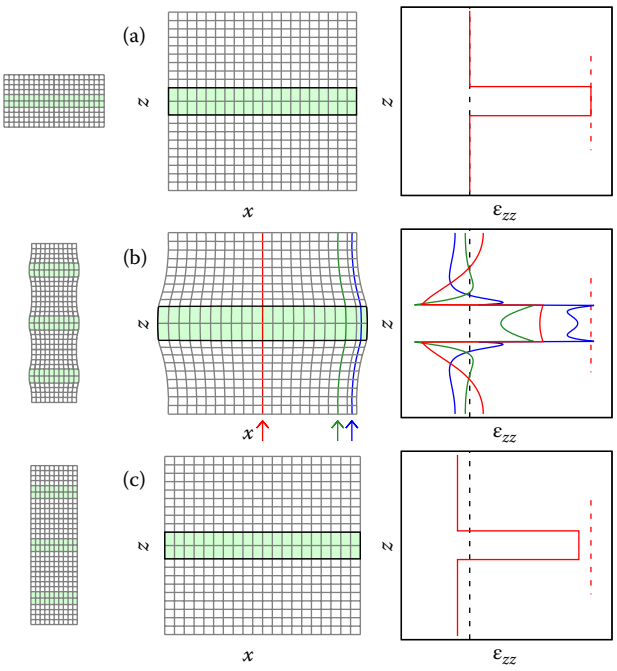

FIGURE 12.4 Strain states of lattice-mismatched heterostructures: sketch of the entire structure (left), atomic displacement (center), and out-of-plane strain ε_{zz} (right) near one quantum well (QW). (a) Planar pseudomorphic epitaxial QW growth, the substrate remains strain free. (b) Axial nanowire (NW) superlattice, nonuniform strain profile calculated with a cylindrical model. The red, green, and blue curves represent the displacement and strain at the radial positions indicated by the arrows (NW center, 80% and 95% of NW radius). (c) Axial NW superlattice, lateral relaxation estimate, QW and barrier assume a common average in-plane lattice parameter. (Reprinted with permission from Wölz et al. (2013). Copyright [2013] **American Chemical Society**.)

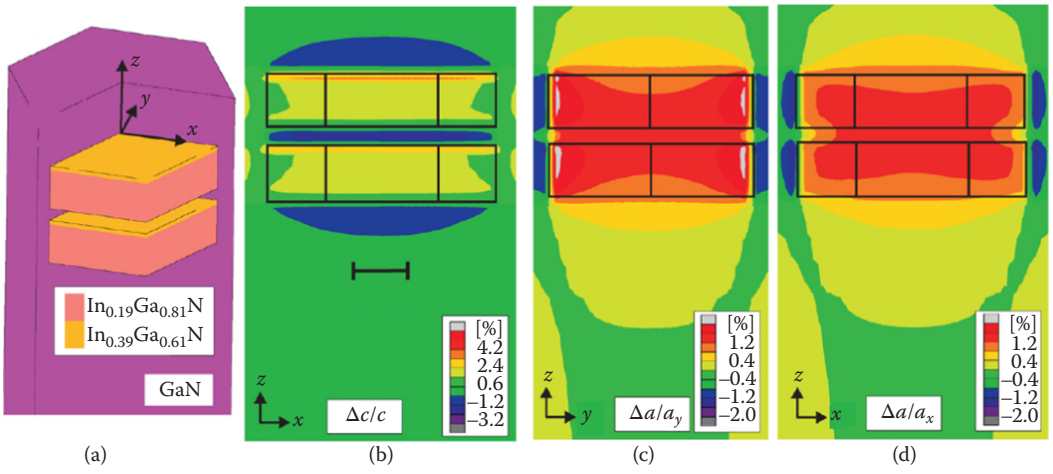

FIGURE 12.5 (a) Perspective view of the GaN nanowire (NW) with two $In_xGa_{1-x}N$ disks. The diagonal components of the strain tensor, ε_{zz}, ε_{yy}, and ε_{xx}, are shown in (b), (c), and (d), respectively. (Reproduced with permission from Knelangen, M. et al., *Nanotechnology*, 2011. © IOP Publishing. All rights reserved.)

of 3 nm within a hexagonal GaN NW with a diameter of 50 nm is highly strained due to the large In content and experiences a significant lateral extension. In the center of the layer, the magnitude of the strain approaches the value predicted in the biaxial limit. As a consequence, the bandgap is significantly lowered at the NW surface, inducing a localization of electrons and holes at the corners and at the side facets, respectively.

12.2.4 Piezoelectric and Spontaneous Polarization

Both GaN and InN exhibit pyroelectric and piezoelectric polarization. Planar heterostructures are commonly grown along the polar [0001] axis, resulting in large internal electrostatic fields in the $In_xGa_{1-x}N$ QWs. These fields lead to a spatial separation of electrons and holes and thus to a considerable reduction of the radiative recombination rate. The use of axial GaN NW heterostructures promises to resolve some of these problems (Kikuchi et al. 2004; Kim et al. 2004; Li and Waag 2012). As discussed in the previous section, the free side facets facilitate the elastic relaxation of strain in the $In_xGa_{1-x}N$ disk, so that higher In contents can be achieved (Björk et al. 2002; Ertekin et al. 2005; Kaganer and Belov 2012). Along with the elastic relaxation, a significant reduction of the built-in piezoelectric potential and hence the quantum confined Stark effect (QCSE) (Miller et al. 1984) is expected. In fact, some researchers report evidence of a vanishing QCSE (Lin et al. 2010; Armitage and Tsubaki 2010; Nguyen et al. 2011; Bardoux et al. 2009), whereas others observe a significant QCSE (Wölz et al. 2013). Typically, the polarization $\boldsymbol{P}$ is computed directly using the strain tensor. For WZ crystals, it reads (Nye 1985):

$$P(r) = \begin{pmatrix} 2e_{15}\varepsilon_{13} \\ 2e_{15}\varepsilon_{23} \\ e_{31}\left(\varepsilon_{11}+\varepsilon_{22}\right) + e_{33}\varepsilon_{33} + P_{\mathrm{sp}} \end{pmatrix}, \tag{12.6}$$

where $e_{ij} = e_{ij}(\boldsymbol{r})$ are the piezoelectric constants and $P_{\mathrm{sp}}(\boldsymbol{r})$ is the spontaneous polarization is the spontaneous polarization: WZ crystals exhibit a charge separation even in the absence of an electric field. The polarization potential $V(\boldsymbol{r})$ is computed solving the Poisson equation.

$$\rho_{\mathrm{p}}(r) = \varepsilon_0 \nabla \cdot \left[\varepsilon_r(r)\nabla V(r)\right] \quad \text{with} \quad \rho_{\mathrm{p}}(r) = -\nabla \cdot P(r). \tag{12.7}$$

ε_0 and ε_r denote the vacuum and the relative dielectric constants, respectively.

The resulting electrostatic potential $V(\boldsymbol{r})$ is used in the model to compute the electronic properties as a potential contribution added to both the conduction and the valence bands.

For the system considered by Böcklin et al. (2010), accounting for the polarization potential leads to a confinement of the hole ground state at the center of the NW, whereas the electron remains at the edges. Depending on size and material composition of NW and disk, both strain and the polarization potential can therefore dominate the localization of carriers in a NW.

For an $In_xGa_{1-x}N$ disk in a circular cylinder GaN NW, the Poisson equation can be solved using Equation 12.5 and the electrostatic potential is expressed in similar integrals (Kaganer et al. 2016). Figure 12.6 displays the electrostatic potential V along the axis of the cylinder ($r = 0$) for an $In_{0.25}Ga_{0.75}N$ disk of thickness 5 nm inserted into GaN cylinders of different diameters. The potential for a film (i.e., a cylinder of an infinite diameter) is included for comparison. Already for very thick NWs with a diameter of 150 nm, the electrostatic potential of the disk along the central NW axis is visibly reduced and this reduction increases for decreasing diameters. However, even for a very thin NW with a diameter of 15 nm, a potential difference of approximately 300 meV persists.

Although the polarization potential is significantly reduced compared to the one of planar heterostructures of similar growth direction, thickness, and In content, a very small NW diameter ($d < 30$ nm) is required to achieve a substantial reduction of the piezoelectric potential for the typical disk thicknesses of only a few nanometers. Even in such systems, a nonvanishing built-in potential remains. The magnitude of

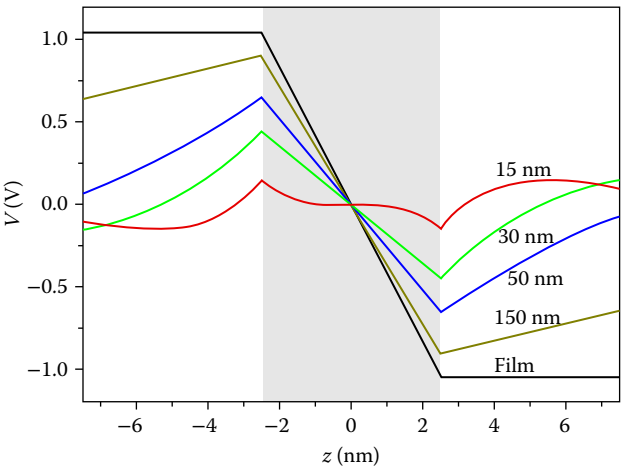

FIGURE 12.6 The electrostatic potential V along the axis of the cylinder ($r = 0$) for a 5-nm-thick $In_{0.25}Ga_{0.75}N$ disk inserted into GaN cylinders of different diameters, as indicated in the figure. The interior of the disk is highlighted. (Reproduced with permission from Kaganer, V. M. et al., *Nanotechnology*, 2016. ©IOP Publishing. All rights reserved.)

the polarization potential decreases with larger aspect ratio and the extrema of the potential shift from the central axis of the NW toward the side facets (Kaganer et al. 2016). The polarization potential therefore not only affects the transition energies but also the confinement of electrons and holes and thus the recombination rates. Strain relaxation in axial NW heterostructures represents an additional degree of freedom to control the emission wavelength and the electron–hole overlap in comparison to their planar counterparts. However, electrostatic potentials remain and exert an important influence on the electronic properties of the system.

12.3 Doping and the Influence of Surfaces

The electrical conductivity of semiconductor materials can be tuned over many orders of magnitude by the intentional incorporation of impurities (Schubert 2005). The ability to dope semiconductors in a controlled fashion is a key feature of all semiconductor-based technology (Sze 1985; Seeger 1991). For a theoretical description, doping poses a formidable problem, as dopants occupy random positions in the host lattice. Established solid-state simulation methods exploiting the periodicity of the system cannot be employed (Shklovskii and Efros 1984). For device simulations, the discrete nature of the dopant positions is therefore commonly ignored altogether, and a homogeneous, continuous distribution of charge is assumed (Keyes 1975).

In semiconductor NWs, intentional or unintentional doping is of particular importance due to the vicinity of free surfaces. In the case of an $In_xGa_{1-x}N$ based NW, the residual doping concentration of O and Si atoms is $\rho_D = 10^{16} - 10^{17}$ cm^{-3}. These donors transfer their extra electrons to surface states, leaving the donor ionized and the NW depleted at all temperatures, as long as the radius of the NW does not exceed the typical thickness of the depletion region in the bulk material. In the following, we discuss concepts to simulate the impact of doping on the electronic structure of semiconductor NWs.

12.3.1 Continuous, Homogeneous Doping-Related Background Charge

In a simple picture, the potential arising from ionized donors is treated as an electrostatic potential obtained from a homogeneous background charge density by solving the Poisson equation. In the case of a

semiconductor NW with doping density ρ_D and with a cylindrical shape, the surface potential V arising from a homogeneous background charge density is

$$V(r) = E_F - \frac{e \cdot \rho_D}{4\varepsilon_0 \cdot \varepsilon_r} R^2 \left[1 - \left(\frac{r}{R}\right)^2\right], \tag{12.8}$$

with R being the radius of the NW and E_F being the surface Fermi level pinning. For a cylindrical GaN NW with $R = 40$ nm and a doping density of 10^{17} cm^{-3}, the potential difference between the center and the side facets of the NW is 80 mV. If the residual doping consists of Si and O, which is typical for GaN, the potential is positive at the side facets, that is, attractive for holes.

Figure 12.7 shows the electron (red) and hole (blue) ground-state-charge density in an axial $In_xGa_{1-x}N$/GaN NW heterostructure for various values of x and accounting for the influence of a surface potential, computed using an eight-band **k·p** model (see Chapter 1 for details). It can be seen that the surface potential gives rise to a behavior specific to NWs. As long as the thickness of the $In_xGa_{1-x}N$ disk is small compared to the NW diameter, the strain-induced piezoelectric potential exhibits extrema at the top and bottom surface of the $In_xGa_{1-x}N$ disk along the central axis of the NW (see discussion in Section 12.2.4). For the hole state, this potential is attractive at the bottom center of the disk. A surface potential arising from an unintentional background doping, on the contrary, localizes hole states at the side facets of the NW. The localization of the hole state therefore depends on the magnitude of surface and polarization potentials: For large In contents, the polarization potential increases and induces a localization of holes along the central axis of the NW. For small In contents, the localization of the hole state is dominated by the surface potential and the hole is confined near the side facets of the NW (Marquardt et al. 2013).

12.3.2 Random Dopant Fluctuations: The Discrete Nature of Dopants in an NW

In a typical axial $In_xGa_{1-x}N$/GaN NW heterostructure, the number of dopants is commonly not large enough to justify the picture of a homogeneous doping-related background charge (Corfdir et al. 2014). With a diameter of 80 nm, a doping concentration of 10^{17} cm^{-3} corresponds on average to only 8.3 charges within a segment of 20 nm length. Therefore, individual, randomly distributed dopants have to be considered in NWs. Of course, the ideal approach to deal with influences of individual atoms is an atomistic picture, such as empirical tight binding or pseudopotential methods or density functional theory (for more details on these models see Chapter 1). At the same time, experimentally relevant semiconductor

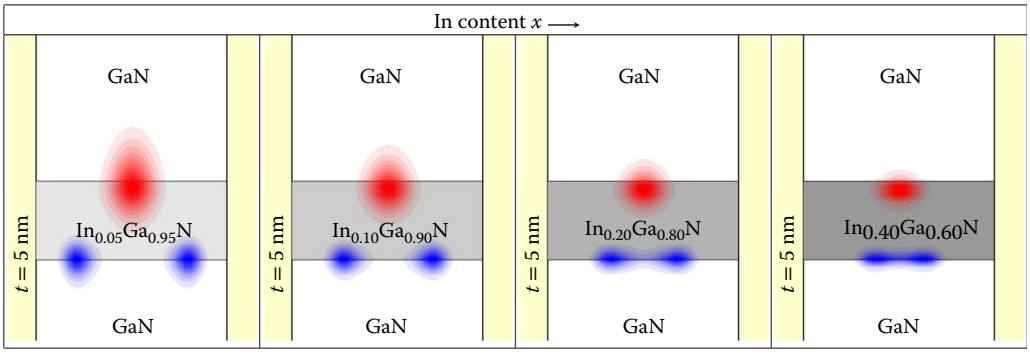

FIGURE 12.7 Side view of the electron (red) and hole (blue) ground-state charge density in an axial $In_xGa_{1-x}N$/GaN nanowire (NW) heterostructure for $x = 0.05, 0.1, 0.2$, and 0.4 (left to right). The NW diameter and disk thickness are not to scale. (Reprinted with permission from Marquardt et al. (2013). Copyright 2013 American Chemical Society.)

NWs exhibit typical diameters of a few tens to hundreds of nanometers. Therefore, a reasonable simulation domain commonly contains millions of atoms. Random dopant fluctuations (RDFs) induce variations of the electronic potential landscape between individual NWs. A meaningful study of the impact of RDFs on the electronic structure of NWs will thus require a statistical evaluation of these variations so that a correspondingly large number of simulations needs to be performed. Within an atomistic picture, this represents a massive computational effort. However, individual dopants can also be included within the computationally inexpensive continuum approaches, for example, effective mass or multiband **k·p** models (see Chapter 1 for more details) under certain conditions.

If an impurity can be considered a hydrogenic dopant, its main influence on the electronic structure is a long-range Coulomb potential. These shallow donors and acceptors induce only a small deformation to the host lattice and require only small energies (less than 3 k_BT ~75 meV at room temperature) to be ionized. Some examples for shallow donors and acceptors in different host materials are listed in Table 12.1.

A shallow impurity can be modeled as a point charge. The respective Coulomb potential obtained from the Poisson equation is then simply added to the potential landscape within the simulation cell. Note that a shallow impurity in a binary material is not necessarily a shallow impurity in a ternary system. For example, O is a shallow donor in GaN but its behavior in $Al_xGa_{1-x}N$ alloys is still a matter of controversy. Some studies report a shallow–deep transition (Park and Chadi 1997; McCluskey et al. 1998), whereas others claim O to remain a shallow donor in $Al_xGa_{1-x}N$ even for large Al contents (Kakanakova-Georgieva et al. 2013). For deep dopants, the model of a simple point charge for an ionized dopant becomes inaccurate.

Figure 12.8a shows an example of a Coulomb potential resulting from RDFs in an axial $In_xGa_{1-x}N$/GaN NW heterostructure. The respective electron (red) and hole (blue) ground-state charge densities are shown in the center of Figure 12.8, computed using an eight-band **k·p** model (Chuang and Chang 1996). To investigate the variations of electron and hole wave functions as well as binding energies in an ensemble of otherwise identical NWs, one can perform a statistical study by employing a larger number of random donor distributions. The electron–hole ground-state recombination energy resulting from such a study is shown in Figure 12.8c. Additionally, the energy values obtained by assuming a continuous, homogeneous

TABLE 12.1 Activation Energies of Shallow Impurities in Different Host Materials

Host Material	Donor	$E_c - E_d$ (meV)	Acceptor	$E_a - E_v$ (meV)
GaAs	S	6	Mg	28
	Se	6	Zn	31
	Te	30	Cd	35
	Si	58	Si	26
GaN	O	33.21 (Freitas et al. 2003)		
	Si	30.19 (Freitas et al. 2003)		
	C	34.0 (Wang and Chen 2000)		
	Ge	31.1 (Wang and Chen 2000)		
	S	29.5 (Wang and Chen 2000)		
	Se	29.5 (Wang and Chen 2000)		
AlN	Si	60 (Zeisel et al. 2000)		
ZnO	Al	53 (Meyer et al. 2005)		
	Ga	54.5 (Meyer et al. 2005)		
	In	63.2 (Meyer et al. 2005)		
Si	Sb	39	B	45
	P	45	Al	67
	As	54	Ga	73

Source: If not indicated differently, taken from Whitaker (2005). Freitas, J. A. et al., *Phys. Stat. Solidi B,* 240, 330, 2003; Wangm, H. and A.-B. Chen, *J. Appl. Phys.,* 87, 7859, 2000; Zeisel, R. et al. *Phys. Rev. B*, 61, R16283, 2000; Meyer, B. K. et al., *Semicond. Sci. Technol.*, 20, S62, 2005.

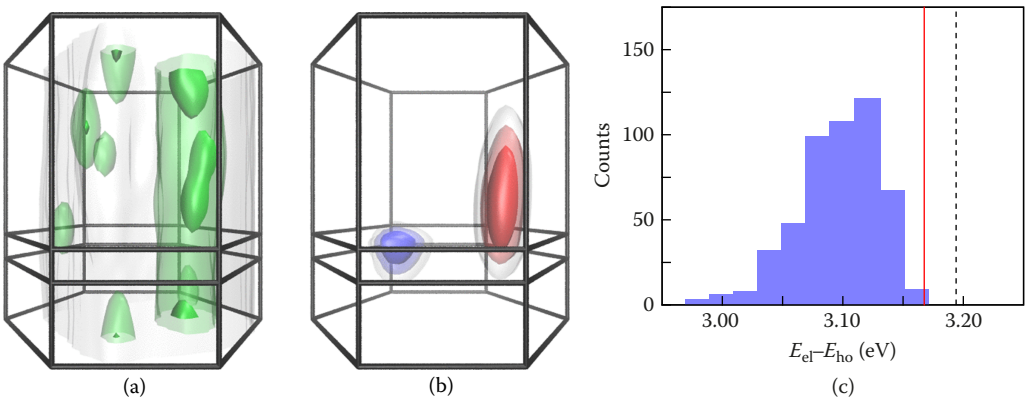

FIGURE 12.8 (a) Coulomb potential isosurfaces for a random distribution of nine individual dopants in an axial $In_xGa_{1-x}N/GaN$ nanowire (NW) heterostructure. The active layer is marked in the lower part of the NW. (b) Electron (red) and hole (blue) ground-state charge density. (c) Histogram of the electron–hole ground-state recombination energy for an ensemble of 500 different random dopant distributions with an average of 8.3 charges in the vicinity of the active layer. For the dimensions of the NW segment under consideration, this number corresponds to a doping density of 10^{17} cm^{-3}.

doping-related background charge (solid red line) as well as the one of an ideal, undoped NW (dashed black line) are shown. It is seen that both of these energies are larger than the ones obtained assuming RDFs, which arises from the fact that RDFs induce much deeper local potentials than the surface potential arising from a homogeneous doping-related background charge. The ground-state recombination energy of the undoped NW represents an upper limit for the other two models (Marquardt et al. 2015).

In contrast to shallow impurities, an atomistic treatment as described in Chapter 1 is necessary for deep donors or acceptors, that is, impurities that exhibit a large chemical shift or induce a large deformation to the host lattice.

12.3.3 Dielectrically Enhanced Ionization Energies of Dopants

In the previous sections, we have neglected the mismatch in dielectric constants at the surface of the NWs. Still, the discontinuity of the dielectric constant at the surface of NWs strongly affects the properties of dopants in these nanostructures. In particular, it has been demonstrated that the Coulomb interaction is enhanced in nanostructures surrounded by a barrier material with a lower dielectric constant (Keldysh 1979). This so-called dielectric confinement is a consequence of the electrostatic potential that is set up by the presence of a charged particle in a nanostructure and that can be calculated using the image charge method (Kumagai and Takagahara 1989).

Whereas the mismatch in dielectric constants at the interface of heterostructures such as GaN (Al,Ga)N or GaAs/(Al,Ga)As is small and induces only minor changes in the binding energy of excitons and in the ionization energy of dopants (Andreani and Pasquarello 1990), the mismatch is maximum for NWs in air or vacuum, and the increase in the dopant ionization energies due to the dielectric confinement is considerable (Shik 1993; Diarra et al. 2007). Although quantum confinement leads as well to increased ionization energies for impurities in NWs with a radius smaller than the Bohr radius of donors in the bulk, a_B (Bryant 1984), the dielectric confinement in semiconductor NWs in vacuum sets in already for NW radii smaller than $5a_B$ (Diarra et al. 2007). This finding has been verified experimentally by Björk et al. (2009) who measured a significant increase in resistivity in *n*-doped Si NWs with a diameter smaller than 15 nm.

The impact of image charges on donors or excitons in low dimensional systems has been studied using envelope function calculations (Kumagai and Takagahara 1989), tight binding methods (Allan et al. 1995; Diarra et al. 2007), or density functional calculations (Chan et al. 2008). All of these simulations predict an

increase in ionization energy for donors located at the center of a nanostructure as a result of the dielectric confinement (Shik 1993; Diarra et al. 2007; Chan et al. 2008; Corfdir and Lefebvre 2012a). The situation is more complicated for donors located off the axis of NWs. Early effective-mass simulations neglecting the dielectric confinement have predicted that the infinite potential barrier at the surface leads to a change in the symmetry of the wave function of the donor ground-state electron and to a monotonic decrease in donor ionization energy when getting closer to the surface (Levine 1965; Satpathy 1983). In particular, the binding energy of an electron to a donor located at the surface of a semi-infinite layer is one quarter of that of a bulk donor (Levine 1965). However, accounting for the combination of the surface potential barrier with the image charges, one obtains that the ionization energy of a donor in a thin slab bounded by vacuum exhibits a nonmonotonic dependence on the donor site: It first increases and then decreases when the donor moves from the center to the surface of the slab (Corfdir and Lefebvre 2012a). This position-dependent binding energy of electrons to donors has been used to explain the inhomogeneous broadening of the photoluminescence lines related to donor bound excitons in unintentionally doped and strain free GaN NWs (Corfdir et al. 2014). However, using scanning tunneling microscopy, Wijnheimer et al. measured a gradual increase in the ionization of Si over the last 1.2 nm below the (110) surface of a thick GaAs layer (Wijnheijmer et al. 2009). This increase, not reproduced by the effective mass theory, was tentatively explained by the authors in terms of broken symmetry at the surface.

The increase in dopant ionization energies in thin NWs due to the dielectric confinement leads to a decrease in conductivity, which may hinder the realization of efficient electronic and optoelectronic devices based on these structures (Björk et al. 2009). In addition, theoretical and experimental studies have shown that the deactivation of dopants is not homogeneous across the NW section (Wijnheijmer et al. 2009). This finding, together with the nonuniform distribution of dopants in NWs discussed above and the modification of the dielectric environment by electrical contacts or gates (Diarra et al. 2008), renders the simulation of NW-based devices challenging.

12.4 Crystal-Phase Bandgap Engineering

An additional peculiarity of III–V NWs with respect to their planar counterpart is their pronounced polytypism, that is, the crystal structure along the NW length alternates between WZ and ZB. Although III-nitride NWs usually grow within the WZ structure with only a few stacking faults, III-arsenides and III-phosphides NWs show frequent alternation between ZB and WZ segments of various lengths along their axis (Figure 12.9a; Algra et al. 2008; Bao et al. 2008; Caroff et al. 2009; Jacopin et al. 2011; Heiss et al. 2011; Graham et al. 2013). This polytypism is usually regarded as detrimental for the electronic properties of NWs: Alternation between WZ and ZB crystal structures efficiently scatters electrons (Konar et al. 2011) and leads to a significant increase in the resistivity of NWs (Thelander et al. 2011). In addition, the presence of stacking faults along the length of NWs leads to localization of electrons and holes, increasing the ionization energies of dopants in these nanostructures (Corfdir et al. 2009).

Although considerable effort has been devoted to the elimination of stacking defects in NWs, some degree of control on the switching between WZ and ZB during the NW growth has been demonstrated (Algra et al. 2008; Caroff et al. 2009). This control opens the possibility of realizing crystal-phase bandgap engineering in NWs, where carrier confinement is obtained by alternation of the crystal structure and not by the composition (Algra et al. 2008; Caroff et al. 2009; Akopian et al. 2010; Corfdir et al. 2013).

Such crystal-phase quantum structures exhibit atomically flat interfaces (Bolinsson et al. 2011) and are free of alloy disorder. Thanks to the high quality of their interfaces, crystal-phase quantum structures exhibit a sub-meV photoluminescence line width at cryogenic temperatures (Figure 12.9b; Akopian et al. 2010; Jacopin et al. 2011; Graham et al. 2013) and they can be of interest for quantum optics applications (Akopian et al. 2010; Castelletto et al. 2014).

The band alignment between the WZ and the ZB phases of a III–V semiconductor is usually considered to be of type II, with the electron and the hole localized in the ZB and WZ phases, respectively

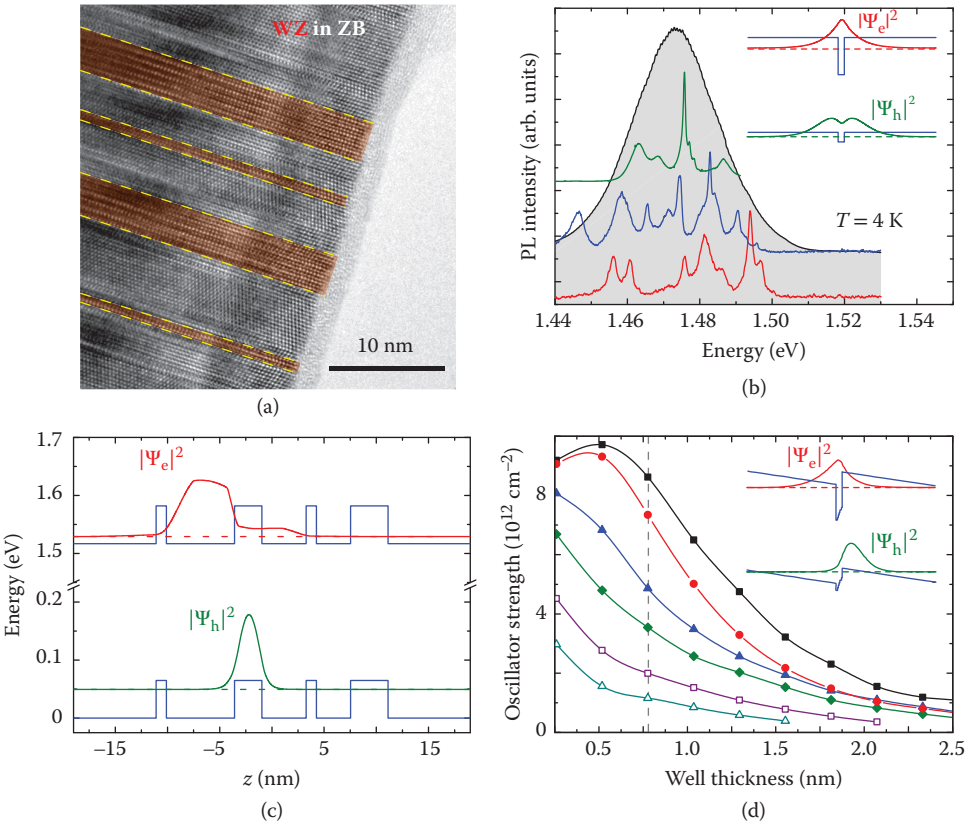

FIGURE 12.9 (a) High-resolution transmission electron micrograph of a polytype GaAs nanowire (NW). Wurtzite (WZ) segments are highlighted in red. (Adapted with permission from Corfdir et al. (2013). Copyright 2013 American Chemical Society.) (b) Photoluminescence spectra at 4 K from single NWs (red, blue, green) and from an ensemble (gray) of polytype GaAs NWs. The signal between 1.46 and 1.50 eV arises from recombination at crystal-phase quantum disks. The inset depicts schematically the band profile for a zinc-blende (ZB) inclusion in WZ GaAs (blue). The energy and the wave function of the electron (red) and the hole (green) are shown with solid and dashed lines, respectively. (Adapted with permission from Graham et al. (2013). Copyright 2013 by the American Physical Society.) (c) Band profiles and electron and hole energy and wave function for the NW segment shown in (a). (Adapted with permission from Corfdir et al. (2013). Copyright 2013 American Chemical Society.) (d) Oscillator strength for an exciton in a QW of ZB GaN in WZ GaN as a function of the well thickness and in the presence of built-in electric fields. The lines from top to bottom correspond to built-in electric fields of 0, 0.1, 0.5, 1, 2, and 3 MV/cm, respectively. The built-in electric fields along the confinement axis lead to a spatial separation of the electron and hole wave functions (inset). (Adapted with permission from Corfdir and Lefebvre (2012a). Copyright 2013, American Institute of Physics.)

(Figure 12.9b; Murayama and Nakayama 1994). *Ab initio* calculations have predicted small band offsets and small bandgap differences between the WZ and the ZB phases (Bechstedt and Belabbes 2013). Therefore, although long radiative lifetimes have been calculated for thick crystal-phase QWs in NWs (Zhang et al. 2010), the electron and hole wave functions for thin ZB insertions in the WZ show significant spreading into the WZ and the ZB, respectively. Their squared overlap is large (Corfdir and Lefebvre 2012b; Corfdir et al. 2013) and thin crystal-phase QWs and QDs in NWs act as efficient light emitters, despite their type-II band alignment.

Effective-mass calculations have been used to calculate the photoluminescence energy associated to single crystal-phase QWs and to polytype semiconductor NW segments (Bao et al. 2008; Jacopin et al. 2011; Heiss et al. 2011; Corfdir and Lefebvre 2012b; Lähnemann et al. 2012; Corfdir et al. 2013). For instance,

such calculations have been combined with photoluminescence and transmission electron microscopy experiments to obtain information on the band structure of WZ GaAs (Figure 12.9c; Corfdir et al. 2013). However, the conclusions obtained from those studies remain only qualitative since many parameters required for the calculations (e.g., the conduction and valence-band offsets) are poorly known. The situation is particularly complex for GaAs NWs. Despite a large number of experimental and theoretical studies, there is still no agreement on the value of the bandgap of WZ GaAs; values larger and smaller than the bandgap of ZB GaAs have been reported, and the symmetry of the lowest energy conduction band of this material remains under debate (Martelli et al. 2015). Note also that, due to the small conduction and valence-band offsets between the WZ and the ZB phases of a semiconductor, the energy computed for an exciton in crystal-phase quantum structures of a thickness of a few monolayers is not sensitive to the exact values of the band offsets. As shown in Lähnemann et al. (2012), the photoluminescence energy for three-to-five-monolayers-thick crystal-phase QWs in GaN NWs can in fact be reproduced using either a type-I or a type-II band alignment between WZ and ZB GaN.

The connection of bands at the interface of different crystal phases is of particular importance for the modeling of polytype NWs. In the case of GaAs, the lowest conduction band in the ZB phase has a Γ_6 symmetry and is well separated in energy from any remote conduction bands (Belabbes et al. 2012). The equivalent of the Γ_6 band in the WZ structure is the Γ_7 band. However, compounds with a WZ structure show an additional band with a Γ_8 symmetry that results from the zone folding due to doubling of the WZ unit cell along the [0001] direction when compared to the ZB unit cell. For WZ GaAs, recent studies (De and Pryor 2010; Heiss et al. 2011; Belabbes et al. 2012; Graham et al. 2013) suggest that this Γ_8 band is very close in energy to the Γ_7 band. It is, however, not clear yet which of these two bands exhibits the lowest energy. This point is, however, essential since (1) these two bands are expected to exhibit very different effective masses and a weak coupling with the fundamental Γ_7 and Γ_9 valence bands, and (2) confinement in crystal-phase quantum disks could give rise to some conduction band mixing.

Finally, although strain-free compounds with a ZB structure do not exhibit any spontaneous polarization, WZ III–V semiconductors are polar materials. Differential phase contrast microscopy and photoluminescence spectroscopy on polytype NWs have shown that the spontaneous polarization in WZ GaAs and GaN is aligned along the *c*-axis of the WZ and has a strength equal to 2.7×10^{-3} and -2.2×10^{-2} C/m^2, respectively (Bauer et al. 2014; Lähnemann et al. 2012). Since the axis of most WZ NWs is parallel to the *c* direction, the spontaneous polarization field in polytype NWs shows discontinuities at the interfaces between WZ and ZB segments. These discontinuities give rise to built-in electric fields parallel to the axis of polytype NWs with an intensity on the order of MV/cm in WZ/ZB GaN quantum structures (Corfdir and Lefebvre 2012b; Lähnemann et al. 2012). The resulting quantum confined Stark effect strongly affects the optical properties of thick crystal-phase QWs. The photoluminescence transitions are redshifted and may be centered even at energies below the band gap of ZB GaN, as observed for ZB insertions thicker than six monolayers (Jacopin et al. 2011), and the oscillator strength of the exciton is decreased (Figure 12.9d; Corfdir and Lefebvre 2012b).

12.5 Optical Anisotropy of NWs—The Antenna Effect

Semiconductors in the form of NWs not only show altered electrical characteristics but also strongly modified optical properties. As shown by Wang et al. (2001), the electric field at the sidewalls of NWs with a diameter much smaller than the wavelength is discontinuous (continuous) for light polarized perpendicular (parallel) to the NW axis. For an NW considered as an infinite dielectric cylinder, the electric field is $E_i = 2\kappa_0/\left(\kappa_r + \kappa_0\right) E_0$ for light polarized perpendicular to the NW axis, where E_0 is the electric field outside the NW and κ_0 and κ_r are the dielectric constants of vacuum and of the NW, respectively. Since $\kappa_r \approx 10$ in III–V semiconductors, this *antenna effect* leads to significant polarization anisotropies of interband transitions in NWs. In particular, the antenna effect may affect the emission intensity ratio between different interband transitions. For instance, while heavy holes have dipoles oriented perpendicular to the

axis of GaAs NWs, light holes exhibit a large on-axis dipole. Optical transitions related to light–hole excitons are thus promoted by the antenna effect (Spirkoska et al. 2012). Complete reversal of the polarization anisotropy with respect to the bulk has even been predicted for WZ NWs with crystal field and spin–orbit splittings such that the A and B valence bands are close in energy (Efros and Lambrecht 2014).

The situation is more complicated for NWs with diameters on the order of the wavelength, as they support guided modes (Ruppin 2002). The latter gives rise to oscillations in the degree of polarization of the NW luminescence with varying wavelength or NW diameter (Giblin et al. 2009; Corfdir et al. 2015). NW ensembles usually exhibit a broad diameter distribution, and the polarization response may vary from one NW to another (Corfdir et al. 2015).

The polarization properties of interband transitions have been commonly used to extract the strain state of planar structures. Comparison between the measured optical anisotropy and the one computed using **k·p** calculations has made it possible to estimate the strain state of group-III-nitrides layers grown along a nonpolar or a semipolar axis (Ghosh et al. 2002; Funato et al. 2013). The analysis is more complicated for NWs since the combination of strain, dielectric contrast and possible surface-induced wave function distortion (see Section 12.3.3) may result in a complex polarization anisotropy of the interband transitions in NWs. For instance, the degree of polarization of single InP/InAs NWs exhibiting valence-band mixing and dispersed on a carbon grid had to be calculated using detailed finite difference time domain calculations (Anufriev et al. 2015).

Acknowledgments

The authors would like to thank Oliver Brandt and Lutz Geelhaar as well as present and former colleagues at Paul-Drude-Institut für Festkörperelektronik, Berlin, Germany, and at the Cavendish Laboratory, Cambridge, UK, for stimulating and fruitful discussions. We also thank Lutz Schrottke for critical reading of the manuscript. Pierre Corfdir acknowledges funding from the Fonds National Suisse de la Recherche Scientifique through project 161032.

References

Ahtapodov, L., J. Todorovic, P. Olk, T. Mjåland, P. Slåttnes, D. L. Dheeraj, A. T. J. Van Helvoort, B. O. Fimland, and H. Weman. 2012. A story told by a single nanowire: Optical properties of wurtzite GaAs. *Nano Lett.* 12:6090.

Akopian, N., G. Patriarche, L. Liu, J.-C. Harmand, and V. Zwiller. 2010. Crystal phase quantum dots. *Nano Lett.* 10:1198.

Algra, R. E., M. A. Verheijen, M. T. Borgström, L.-F. Feiner, G. Immink, W. J. P. van Enckevort, E. Vlieg, and E. P. A. M. Bakkers. 2008. Twinning superlattices in indium phosphide nanowires. *Nature* 456:369.

Allan, G., C. Delerue, M. Lannoo, and E. Martin. 1995. Hydrogenic impurity levels, dielectric constant, and Coulomb charging effects in silicon crystallites. *Phys. Rev. B* 52:11982.

Andreani, L. C., and A. Pasquarello. 1990. Accurate theory of excitons in GaAs-$Ga_{1-x}Al_xAs$ quantum wells. *Phys. Rev. B* 42:8928.

Anufriev, R., J.-B. Barakat, G. Patriarche, X. Letartre, C. Bru-Chevallier, J.-C. Harmand, M. Gendry, and N. Chauvin. 2015. Optical polarization properties of InAs/InP quantum dot and quantum rod nanowires. *Nanotechnology* 26:395701.

Armitage, R., and K. Tsubaki. 2010. Multicolour luminescence from InGaN quantum wells grown over GaN nanowire arrays by molecular-beam epitaxy. *Nanotechnology* 21:195202.

Bao, J., D. C. Bell, F. Capasso, J. B. Wagner, T. Mårtensson, J. Trägårdh, and L. Samuelson. 2008. Optical properties of rotationally twinned InP nanowire heterostructures. *Nano Lett.* 8:836.

Bardoux, R., A. Kaneta, M. Funato, Y. Kawakami, A. Kikuchi, and K. Kishino. 2009. Positive binding energy of a biexciton confined in a localization center formed in a single $In_xGa_{1-x}N$/GaN quantum disk. *Phys. Rev. B* 79:155307.

Bauer, B., J. Hubmann, M. Lohr, E. Reiger, D. Bougeard, and J. Zweck. 2014. Direct detection of spontaneous polarization in wurtzite GaAs nanowires. *Appl. Phys. Lett*. 104:211902.

Bechstedt, F., and A. Belabbes. 2013. Structure, energetic, and electronic states of III–V compound types. *J. Phys. Condens. Matter* 25:273201.

Belabbes, A., C. Panse, J. Furthmüller, and F. Bechstedt. Electronic bands of III–V semiconductor polytypes and their alignment. 2012. *Phys. Rev. B* 86:075208.

Björk, M. T., B. J. Ohlsson, T. Sass, A. I. Persson, C. Thelander, M. H. Magnusson, K. Deppert, L. R. Wallenberg, and L. Samuelson. 2002. One-dimensional heterostructures in semiconductor whiskers. *Appl. Phys. Lett.* 80:1058.

Björk, M. T., H. Schmid, J. Knoch, H. Riel, and W. Riess. 2009. Donor deactivation in silicon nanostructures. *Nat. Nanotechnol*. 4:103.

Böcklin, C., R. G. Veprek, S. Steiger, and B. Witzigmann. 2010. Computational study of an InGaN/GaN nanocolumn light-emitting diode. *Phys. Rev. B* 81:155306.

Bolinsson, J., P. Caroff, B. Mandl, and K. A. Dick. 2011. Wurtzite-zincblende superlattices in InAs nanowires using a supply interruption method. *Nanotechnology* 22:265606.

Bryant, G. W. 1984. Hydrogenic impurity states in quantum-well wires. *Phys. Rev. B* 29:6632.

Caroff, P. K., A. Dick, J. Johansson, M. E. Messing, K. Deppert, and L. Samuelson. 2009. Controlled polytypic and twin-plane superlattices in III–V nanowires. *Nat. Nanotechnol*. 4:50.

Castelletto, S., Z. Bodrog, A. P. Magyar, A. Gentle, A. Gali, and I. Aharonovich. 2014. Quantum-confined single photon emission at room temperature from SiC tetrapods. *Nanoscale* 6:10027.

Chan, T. L., M. L. Tiago, E. Kaxiras, and J. R. Chelikowsky. 2008. Size limits on doping phosphorus into silicon nanocrystals. *Nano Lett*. 8:596.

Chuang, S. L., and C. S. Chang. 1996. **k·p** method for strained wurtzite semiconductors. *Phys. Rev. B* 54, 2491.

Corfdir, P., P. Lefebvre, J. Ristić, J.-D. Ganière, and B. Deveaud-Plédran. 2009. Electron localization by a donor in the vicinity of a basal stacking fault in GaN. *Phys. Rev. B* 80:153309.

Corfdir, P., and P. Lefebvre. 2012a. Role of the dielectric mismatch on the properties of donors in semiconductor nanostructures bounded by air. *J. Appl. Phys*. 112:106104.

Corfdir, P., and P. Lefebvre. 2012b. Importance of excitonic effects and the question of internal electric fields in stacking faults and crystal phase quantum discs: The model-case of GaN. *J. Appl. Phys.* 112:053512.

Corfdir, P., B. Van Hattem, E. Uccelli, S. Conesa-Boj, P. Lefebvre, A. Fontcuberta i Morral, and R. T. Phillips. 2013. Three-dimensional magneto-photoluminescence as a probe of the electronic properties of crystal-phase quantum disks in GaAs nanowires. *Nano Lett*. 13:5303.

Corfdir, P., J. K. Zettler, C. Hauswald, S. Férnandez-Garrido, O. Brandt, and P. Lefebvre. 2014. Sub-meV linewidth in GaN nanowire ensembles: Absence of surface excitons due to the field ionization of donors. *Phys. Rev. B* 90:205301.

Corfdir, P., F. Feix, J. K. Zettler, S. Fernández-Garrido, and O. Brandt. 2015. Importance of the dielectric contrast for the polarization of excitonic transitions in single GaN nanowires. *New J. Phys.* 17:033040.

De, A., and C. Pryor. 2010. Predicted band structures of III–V semiconductors in the wurtzite phase. *Phys. Rev. B* 81:155210.

Diarra, M., Y.-M. Niquet, C. Delerue, and G. Allan. 2007. Ionization energy of donor and acceptor impurities in semiconductor nanowires: Importance of dielectric confinement. *Phys. Rev. B* 75:045301.

Diarra, M., C. Delerue, Y.-M. Niquet, and G. Allan. 2008. Screening and polaronic effects induced by a metallic gate and a surrounding gate oxide on donor and acceptor impurities in silicon nanowires. *J. Appl. Phys.* 103:073703.

Efros, A. L., and W. R. L. Lambrecht. 2014. Theory of light emission polarization reversal in zinc-blende and wurtzite nanowires. *Phys. Rev. B* 89:035304.

Ertekin, E., P. A. Greaney, D. C. Chrzan, and T. D. Sands. 2005. Equilibrium limits of coherency in strained nanowire heterostructures. *J. Appl. Phys*. 97:114325.

Freitas, J. A., W. J. Moore, B. V, Shanabrook, G. C. B. Braga, D. D. Koleske, S. K. Lee, S. S. Park, and J. Y. Han. 2003. Shallow donors in GaN. *Phys. Stat. Solidi B* 240:330.

Funato, M., K. Matsuda, R. G. Banal, R. Ishii, and Y. Kawakami. 2013. Strong optical polarization in nonpolar (1–100) $Al_xGa_{1-x}N$/AlN quantum wells. *Phys. Rev. B* 87:041306(R).

Ghosh, S., P. Waltereit, O. Brandt, H. T. Grahn, and K. H. Ploog. 2002. Electronic band structure of wurtzite GaN under biaxial strain in the M plane investigated with photoreflectance spectroscopy. *Phys. Rev. B* 65: 075202.

Giblin, J., V. Protasenko, and M. Kuno. 2009. Wavelength sensitivity of single nanowire excitation polarization anisotropies explained through a generalized treatment of their linear absorption. *ACS Nano* 3:1979.

Graham, A. M., P. Corfdir, M. Heiss, S. Conesa-Boj, E. Uccelli, A. Fontcuberta, I. Morral, and R. T. Phillips. 2013. Exciton localization mechanisms in wurtzite/zinc-blende GaAs nanowires. *Phys. Rev. B* 87:125304.

Grönquist, J., N. Søndergaard, F. Boxberg, T. Guhr, S. Åberg, and H. Q. Xu. 2009. Strain in semiconductor core–shell nanowires. *J. Appl. Phys.* 106:053508.

Große, F., and J. Neugebauer. 2001. Limits and accuracy of valence force field models for $In_xGa_{1-x}N$ alloys. *Phys. Rev. B* 63:085207.

Heiss, M., S. Conesa-Boj, J. Ren, H. H. Tseng, A. Gali, A. Rudolph, E. Uccelli, F. Peiró, J. R. Morante, D. Schuh, E. Reiger, E. Kaxiras, J. Arbiol, and A. Fontcuberta i Morral. 2011. Direct correlation of crystal structure and optical properties in wurtzite/zinc-blende GaAs nanowire heterostructures. *Phys. Rev. B* 83:045303.

Heiss, M., Y. Fontana, A. Gustafsson, G. Wüst, C. Magen, D. D. O'Regan, J. W. Luo, B. Ketterer, S. Conesa-Boj, A. V. Kuhlmann, J. Houel, E. Russo-Averchi, J. R. Morante, M. Cantoni, N. Marzari, J. Arbiol, A. Zunger, R. J. Warburton, and A. Fontcuberta i Morral. 2013. Self-assembled quantum dots in a nanowire system for quantum photonics. *Nat. Materials* 12:439.

Hirth, J.P., and J. Lothe. 1968. *Theory of Dislocations*. New York: McGraw-Hill.

Hoang, T. B., A. F. Moses, H. L. Zhou, D. L. Dheeraj, B. O. Fimland, and H. Weman. 2009. Observation of free exciton photoluminescence emission from single wurtzite GaAs nanowires. *Appl. Phys. Lett.* 94:133105.

Jacopin, G., L. Rigutti, L. Largeau, F. Fortuna, F. Furtmayr, F. H. Julien, M. Eickhoff, and M. Tchernycheva. 2011. Optical properties of wurtzite/zincblende heterostructures in GaN nanowires. *J. Appl. Phys*. 110:064313.

Kaganer, V. M., and A. Yu. Belov. 2012. Strain and x-ray diffraction from axial nanowire heterostructures. *Phys. Rev. B* 85:125402.

Kaganer, V. M., O. Marquardt, and O. Brandt. 2016. Piezoelectric potential in axial (In,Ga)N/GaN nanowire heterostructures. *Nanotechnology* 27(16):165201.

Kakanakova-Georgieva, A., D. Nilsson, X. T. Trinh, U. Forsberg, N. T. Son, and E. Janzén. 2013. The complex impact of silicon and oxygen on the n-type conductivity of high-Al-content AlGaN. *Appl. Phys. Lett.* 102:132113.

Keating, P. N. 1966. Effect of invariance requirements on the elastic strain energy of crystals with application to the diamond structure. *Phys. Rev.* 145:637.

Keldysh, L. 1979. Coulomb interaction in thin semiconductor and semimetal films. *JETP Lett*. 29:658.

Keyes, W. R. 1975. The effect of randomness in the distribution of impurity atoms on FET thresholds. *Appl. Phys.* 8:251.

Kikuchi, A., M. Kawai, M. Tada, and K. Kishino. 2004. InGaN/GaN multiple quantum disk nanocolumn light-emitting diodes grown on (111) Si substrate. *Jpn. J. Appl. Phys.* 43:1524.

Kim, H., Y. Cho, H. Lee, S. Kim, S. Ryu, D. Kim, T. Kang, and K. Chung. 2004. High-brightness light emitting diodes using dislocation-free indium gallium nitride/gallium nitride multiquantum-well nanorod arrays. *Nano Lett*. 4:1059.

Knelangen, M., M. Hanke, E. Luna, L. Schrottke, O. Brandt, and A. Trampert. 2011. Monodisperse (In,Ga)N insertions in catalyst-free-grown GaN(0001) nanowires. *Nanotechnology* 22:365703.

Konar, A., T. Fang, N. Sun, and D. Jena. 2011. Charged basal stacking fault scattering in nitride semiconductors. *Appl. Phys. Lett*. 98:022109.

Kumagai, M., and T. Takagahara. 1989. Excitonic and nonlinear-optical properties of dielectric quantum-well structures. *Phys. Rev. B* 40:12359.

Lähnemann, J., O. Brandt, U. Jahn, C. Pfüller, C. Roder, P. Dogan, F. Grosse, A. Belabbes, F. Bechstedt, A. Trampert, and L. Geelhaar. 2012. Direct experimental determination of the spontaneous polarization of GaN. *Phys. Rev. B* 86:081302(R).

Landau, L. D., and E. M. Lifshitz. 1959. *Theory of Elasticity*. London: Pergamon.

Levine, J. 1965. Nodal hydrogenic wave functions of donors on semiconductor surfaces. *Phys. Rev.* 140:A586.

Li, S., and A. Waag. 2012. GaN-based nanorods for solid state lighting. *J. Appl. Phys*. 111:071101.

Lin, H. W., Y. J. Lu, H. Y. Chen, H. M. Lee, and S. Gwo. 2010. InGaN/GaN nanorod array white light-emitting diode. *Appl. Phys. Lett.* 97:073101.

Liu, P., H. Huang, X. Liu, M. Bai, D. Zhao, Z. Tang, X. Huang, J.-Y. Kim, and J. Guo. 2014. Core–shell nanowire diode based on strain-engineered bandgap. *Phys. Stat. Solidi A* 212:617.

Marquardt, O., C. Hauswald, M. Wölz, L. Geelhaar, and O. Brandt. 2013. Luminous efficiency of axial $In_xGa_{1-x}N$/GaN nanowire heterostructures: Interplay of polarization and surface potentials. *Nano Lett*. 13:3298.

Marquardt, O., L. Geelhaar, and O. Brandt. 2015. Impact of random dopant fluctuations on the electronic properties of $In_xGa_{1-x}N$/GaN axial nanowire heterostructures. *Nano Lett.* 15:4289.

Martelli, F., G. Priante, and S. Rubini. 2015. Photoluminescence of GaAs nanowires at an energy larger than the ZB band-gap: Dependence on growth parameters. *Semicond. Sci. Tech*. 30:055020.

McCluskey, M. D., N. M. Johnson, C. G. Van de Walle, D. P. Bour, M. Kneissl, and W. Walukiewicz. 1998. Metastability of oxygen donors in AlGaN. *Phys. Rev. Lett.* 80:4008.

Menéndez, J., R. Singh, and J. Drucker. 2011. Theory of strain effects on the Raman spectrum of Si–Ge core–shell nanowires. *Ann. Phys.* 523:145.

Meyer, B. K., J. Sann, D. M. Hofmann, C. Neumann, and A. Zeuner. 2005. Shallow donors and acceptors in ZnO. *Semicond. Sci. Technol.* 20:S62.

Miller, D. A. B., D. S. Chemla, T. C. Damen, A. C. Gossard, W. Wiegmann, T. H. Wood, and C. A. Burrus. 1984. Band-edge electroabsorption in quantum well structures: The quantum-confined stark effect. *Phys. Rev. Lett*. 53:2173.

Murayama, M., and T. Nakayama. 1994. Chemical trend of band offsets at wurtzite/zinc-blende heterocrystalline semiconductor interfaces. *Phys. Rev. B* 49:4710–4724.

Nguyen, H. P. T., S. Zhang, K. Cui, X. Han, S. Fathololoumi, M. Couillard, G. A. Botto, and Z. Mi. 2011. p-Type modulation doped InGaN/GaN dot-in-a-wire white-light-emitting diodes monolithically grown on Si(111). *Nano Lett*. 11:1919.

Nye, J. F. 1985. *Physical Properties of Crystals—Their Representation by Tensors and Matrices*. Clarendon: Oxford.

Park, C. H., and D. J. Chadi. 1997. Stability of deep donor and acceptor centers in GaN, AlN, and BN. *Phys. Rev. B* 55:12995.

Rücker, H., and M. Methfessel. 1995. Anharmonic Keating model for group-IV semiconductors with application to the lattice dynamics in alloys of Si, Ge, and C. *Phys. Rev. B* 52:11059.

Ruppin, R. 2002. Extinction by a circular cylinder in an absorbing medium. *Optics Commun.* 211:335.

Satpathy, S. 1983. Eigenstates of Wannier excitons near a semiconductor surface. *Phys. Rev. B* 28:4585.

Schubert, E. F. 2005. *Doping in III–V Semiconductors*. Cambridge: Cambridge University Press.

Seeger, K. 1991. *Semiconductor Physics*. Berlin: Springer.

Shik, A. 1993. Excitons and impurity centers in thin wires and in porous silicon. *J. Appl. Phys.* 74:2951.

Shklovskii, B. I., and A. L. Efros. 1984. *Electronic Properties of Doped Semiconductors*. Berlin: Springer.

Singh, R., E. J. Dailey, J. Drucker, and J. Menéndez. 2011. Raman scattering from Ge–Si core–shell nanowires: Validity of analytical strain models. *J. Appl. Phys.* 110:124305.

Spirkoska, D., A. L. Efros, W. R. L. Lambrecht, T. Cheiwchanchamnangij, A. Fontcuberta i Morral, and G. Abstreiter. 2012. Valence band structure of polytypic zinc-blende/wurtzite GaAs nanowires probed by polarization-dependent photoluminescence. *Phys. Rev. B* 85:045309.

Sze, S. M. 1985. *Semiconductor Devices: Physics and Technology*. New York: Wiley.

Thelander, C., P. Caroff, S. Plissard, A. W. Dey, and K. A. Dick. 2011. Effects of crystal phase mixing on the electrical properties of InAs nanowires. *Nano Lett*. 11:2424.

Wallentin, J., N. Anttu, D. Asoli, M. Huffman, I. Åberg, M. H. Magnusson, G. Siefer, P. Fuss-Kailuweit, F. Dimroth, B. Witzigmann, H. Q. Xu, L. Samuelson, K. Deppert, and M. T. Borgström. 2013. InP nanowire array solar cells achieving 13.8% efficiency by exceeding the ray optics limit. *Science* 339:1057.

Wang, H., and A.-B. Chen. 2000. Calculation of shallow donor levels in GaN. *J. Appl. Phys.* 87:7859.

Wang, J., M. S. Gudiksen, X. Duan, Y. Cui, and C. M. Lieber. 2001. Highly polarized photoluminescence and photodetection from single indium phosphide nanowires. *Science* 293:1455.

Whitaker, J. C. 2005. *Microelectronics*. Boca Raton, FL: Taylor & Francis.

Wijnheijmer, A. P., J. K. Garleff, K. Teichmann, M. Wenderoth, S. Loth, R. G. Ulbrich, P. A. Maksym, M. Roy, and P. M. Koenraad. 2009. Enhanced donor binding energy close to a semiconductor surface. *Phys. Rev. Lett.* 102:166101.

Wölz, M., V. M. Kaganer, O. Brandt, L. Geelhaar, and H. Riechert. 2011. Analyzing the growth of $In_xGa_{1-x}N$/GaN superlattices in self-induced GaN nanowires by x-ray diffraction. *Appl. Phys. Lett.* 98:261907.

Wölz, M., V. M. Kaganer, O. Brandt, L. Geelhaar, and H. Riechert. 2012. Erratum: Analyzing the growth of $In_xGa_{1-x}N$/GaN superlattices in self-induced GaN nanowires by x-ray diffraction. *Appl. Phys. Lett.* 100:179902.

Wölz, M., M. Ramsteiner, V. M. Kaganer, O. Brandt, L. Geelhaar, and H. Riechert. 2013. Strain engineering of nanowire multi-quantum well demonstrated by Raman spectroscopy. *Nano Lett.* 13:4053.

Xiang, H. J., S.-H. Wei, J. L. F. Da Silva, and J. Li. 2008. Strain relaxation and band-gap tunability in ternary $In_xGa_{1-x}N$ nanowires. *Phys. Rev. B* 77:193301.

Yang, L., R. N. Musin, X.-Q. Wang, and M. Y. Chou. 2008. Quantum confinement effect in Si/Ge core–shell nanowires: First-principles calculations. *Phys. Rev. B* 77:195325.

Zeisel, R., M. W. Bayerl, S. T. B. Goennenwein, R. Dimitrov, O. Ambacher, M. S. Brandt, and M. Stutzmann. 2000. DX-behavior of Si in AlN. *Phys. Rev. B* 61:R16283.

Zhang, L., J.-W. Luo, A. Zunger, N. Akopian, V. Zwiller, and J.-C. Harmand. 2010. Wide InP nanowires with wurtzite/zinc-blende superlattice segments are type-II whereas narrower nanowires become type-I: An atomistic pseudopotential calculation. *Nano Lett*. 10:4055.

13

Quantum Dots

Stanko Tomić
and
Nenad Vukmirović

13.1 Introduction

Semiconductor nanocrystals or quantum dots (QDs) are the subject of intensive research due to a number of novel properties, which make them attractive for both fundamental studies and technological applications [1–6]. QDs are of particular interest for solar cell applications due to their ability to increase efficiency via the generation of multiexcitons from a single photon [7–9]. QDs can be synthesized with a high degree of control using colloidal chemistry [10,11]. Much research effort has been directed toward studying QDs grown from more than one semiconductor, e.g., core/shell heterostructures [12–14]. Such core/shell nanostructures provide means to control the optical properties by tuning the electron–hole wave function overlap, which is affected by the alignment of the conduction band (CB) and valence band (VB) edges, as well as the QD shape and size [15–17]. In addition, such core/shell structures can provide for type-II alignments with staggered CB and VB edges so the lowest energy states for electrons and holes lie in different spatial regions, leading to charge separation between the carriers. Such staggered band alignments have several useful physical consequences, including longer radiative recombination times for more efficient charge extraction in photovoltaic applications [18,19], optical gaps that can be made smaller than the bulk values of constituent materials [12,20,21] and control of the electron–hole wave function overlap that determines the exchange interaction energy [22]. Charge separation in type-II QDs can also be used to increase the repulsion between like-sign charges in biexciton states [23,24], leading to the possibility of lasing in the single exciton regime [6,25,26].

To determine the energetics of many-body states in QDs both the confinement potential and many-body interactions between the carriers need to be taken into account. Many-body interactions lead to Coulomb (charge) and Fermi (spin) correlation. Coulomb correlation arises from the electrostatic interaction of charge carriers in the many-body complex, while spin correlation occurs due to the fermionic character

of charge carriers (i.e., the Pauli exclusion principle) [27]. Correlated many-body states may be calculated with the configuration interaction (CI) method, which can be used in the framework of continuum or atomistic descriptions of single-particle (SP) states. [28–38].

Colloidal QDs are usually embedded or dispersed in media [39] of lower dielectric constant than the semiconductor itself—this dielectric confinement leads to a modification of the Coulomb interaction, which can be described using classical image charge theory. While atomistic calculations [40] showed that dielectric confinement significantly affects the charging energies of QDs, in single-material spherical QDs the similar electron and hole charge distributions lead to a weakened dielectric confinement effect [41] on exciton states, which mainly increases the binding energy [28,42]. It is, therefore, natural to wonder if the optical properties of spherical type-II core/shell QDs can be significantly affected by the dielectric environment.

The effect of dielectric confinement and many-electron correlation means that the single-particle picture is not good enough to faithfully predict exciton energetics or wave functions in colloidal QDs. The proper treatment of charges requires a many-electron description that goes beyond mean-field theories, and the CI method is the most appropriate approach. However, the full configuration method becomes progressively computationally expensive as the number of states increases. Luckily, however, the interpretation of physical experiments often requires detailed knowledge of just a few excitons of particular symmetry. This allows us to evaluate the effect of correlation on exciton energies and dipole matrix elements for arbitrary QD geometries.

In this chapter, we review the theoretical methods for the description of the electronic structure and transport properties of colloidal nanocrystal quantum dots (NQD). The chapter is organized as follows: after introducing the core/shell morphology in Section 13.2, in Section 13.3, we review the widely used **k**·**p** theory for electronic structure of semiconductor nanostructures, both the 8-band and 14-band version, taking into account effects like band-mixing, strain, piezoelectric field, and spin–orbit interaction. In Section 13.4, we outline the plane wave (PW) basis set implementation of the **k**·**p** Hamiltonian and necessary modifications for QDs, i.e., three-dimensional (3D) confined systems. In Section 13.5, we discuss the effect of dielectric confinement, polarization, and self-polarization due to large contrast and spatial variation of the dielectric constant in colloidal NQD. In Section 13.6, we derive the CI Hamiltonian for the description of many-electron states in colloidal NQD. In Section 13.7, we present the basic theory of the electron transport through NQD networks. In Section 13.8, we present our results on the ground state exciton absorption edge, radiative times of the ground state exciton, and correlation energies of the ground state exciton, as well as electronic charge transport in NQD solids.

13.2 The Core/Shell QD Structure

The core/shell QD structures are comprised of a single-crystal "core" region enveloped by a second "shell" region of a different semiconductor material. The shell region usually possesses a larger band gap than the core region. In such a configuration, with appropriate materials combinations and alloys, such QD can exhibit a type-II band alignment, thus offering an extra degree of freedom in *quantum engineering* of nanostructures for specific purposes. The addition of the shell has two main effects. First, the shell serves to passivate dangling bonds at the surface of the dot (surface ligands), which arise from the lack of coordination in atoms at the dot surface, vastly reducing the number of trap states near the surface. This reduces the effect of the nonradiative processes on device operations. Second, the energy offsets forming the type-II heterostructure result in confinement of the charge carriers into separate regions of the QD: electrons are confined to the core, while holes are confined to the shell region, or vice versa, depending on the relative band alignment between the two regions. Such staggered band alignments have several useful physical consequences, including longer radiative recombination times for more efficient charge extraction in photovoltaic applications and optical gaps that can be made smaller than the bulk values of the constituent materials. To illustrate the problem and introduce basic parameters, in Figure 13.1 we show

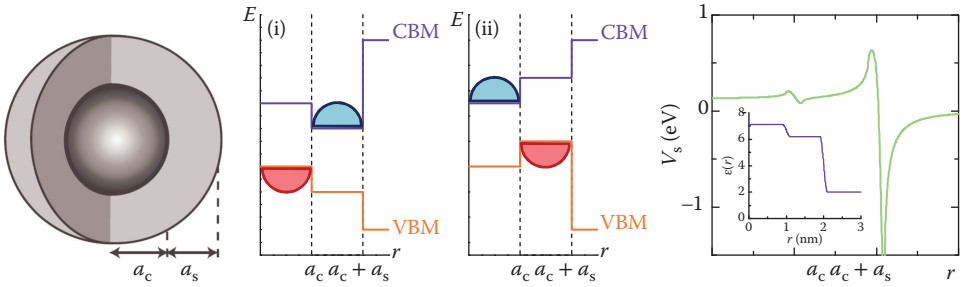

FIGURE 13.1 (a) Cutaway view of a spherical core/shell quantum dot (QD) characterized by core radius a_c and shell thickness a_s; (b) conduction band minimum (CBM) and the valence band maximum (VBM) corresponding to type-II alignment, in (i) an h/e CdTe/CdSe QD and (ii) an e/h CdSe/CdTe QD.

(a) a schematic of a spherical core/shell QD characterized by core radius a_c and shell thickness a_s, the staggered alignments of the CB minimum (CBM) and VB maximum (VBM) in type-II structures where (b,i) holes (h) are confined in the core and electrons (e) are confined in the shell, as in CdTe/CdSe QDs and (b,ii) electrons are confined in the core and holes are confined in the shell, as in CdSe/CdTe QDs.

13.3 The k•p Method for SP Electronic Structure

The quantum mechanical description of electrons in any material requires detailed knowledge of their wave functions, $\psi_{n,\mathbf{k}}(\mathbf{r})$, which are found by solving the Schrödinger equation (in the single-electron approximation):

$$H_0\psi_{n,\mathbf{k}}(\mathbf{r}) = E_{n,\mathbf{k}}\psi_{n,\mathbf{k}}(\mathbf{r}). \tag{13.1}$$

The Hamiltonian in Equation 13.1, $H_0 = p^2/2m_0 + V(\mathbf{r})$, is the function of the quantum mechanical momentum operator, $\mathbf{p} = -i\hbar\nabla$, and the crystal potential experienced by electrons, $V(\mathbf{r}) = V(\mathbf{r}+\mathbf{R})$, which is a periodic function, with the periodicity of the crystal lattice, $\mathbf{R}$, and m_0 is the free electron mass. According to Bloch's theorem, the solutions to this Schrödinger equation can be written as $\psi_{n,\mathbf{k}}(\mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{r}}u_{n,\mathbf{k}}(\mathbf{r})$ where $\mathbf{k}$ is the electron wave vector, n is the band index, and $u_{n,\mathbf{k}}$ is the *cell-periodic* function, with the same periodicity as the crystal lattice. The cell-periodic function, $u_{n,\mathbf{k}}$, satisfies the equation

$$H_{\mathbf{k}}u_{n,\mathbf{k}} = E_{n,\mathbf{k}}u_{n,\mathbf{k}}, \tag{13.2}$$

where the Hamiltonian

$$H_{\mathbf{k}} = H_0 + H'_{\mathbf{k}} = \frac{p^2}{2m_0} + V + \frac{\hbar^2k^2}{2m_0} + \frac{\hbar\mathbf{k}\cdot\mathbf{p}}{m_0} \tag{13.3}$$

is given as a sum of two terms: the unperturbed, H_0, which in fact equals the exact Hamiltonian at $\mathbf{k} = 0$ (i.e., at the Γ point in the Brillouin zone) and the "perturbation," $H'_{\mathbf{k}}$. The Hamiltonian given by Equation 13.3 is called the $\mathbf{k}\cdot\mathbf{p}$ Hamiltonian [43–45]. If the eigenvalues are nondegenerate, the first-order energy correction is given by $E'_{n,\mathbf{k}} \approx \langle u_{n,0}|H'_{\mathbf{k}}|u_{n,0}\rangle$ and there is no correction (to first order, in the absence of nondiagonal matrix elements) in the eigenfunctions. The second-order correction arises from nondiagonal terms; the energy correction is given by $E''_{n,\mathbf{k}} \approx \sum_{n'\neq n}|\langle u_{n,0}|H'_{\mathbf{k}}|u_{n',0}\rangle|^2/(E_{n,0} - E_{n',0})$. The "perturbation" term $H'_{\mathbf{k}}$ gets progressively smaller as $\mathbf{k}$ approaches zero. Therefore, the $\mathbf{k}\cdot\mathbf{p}$ perturbation theory is most

accurate for small values of **k**. However, if enough terms are included in the perturbative expansion, then the theory can in fact be reasonably accurate for any value of **k** in the entire Brillouin zone [46,47]. For a band n, with an extremum at $\mathbf{k} = 0$, and with no spin–orbit coupling, the result of **k·p** perturbation theory is (to the lowest nontrivial order):

$$u_{n,\mathbf{k}} = u_{n,0} + \frac{\hbar}{m_0} \sum_{n' \neq n} \frac{\langle u_{n,0} | \mathbf{k} \cdot \mathbf{p} | u_{n',0} \rangle}{E_{n,0} - E_{n',0}} u_{n',0} \tag{13.4}$$

$$E_{n,\mathbf{k}} = E_{n,0} + \frac{\hbar^2 k^2}{2m_0} + \frac{\hbar^2}{m_0^2} \sum_{n' \neq n} \frac{|\langle u_{n,0} | \mathbf{k} \cdot \mathbf{p} | u_{n',0} \rangle|^2}{E_{n,0} - E_{n',0}}. \tag{13.5}$$

The parameters required to do these calculations, the band edge energies, $E_{n,0}$, and the optical matrix elements, $\langle u_{n,0} | \mathbf{p} | u_{n',0} \rangle$, are typically inferred from experimental data or detailed atomistic-based theories.

A particular strength of the **k·p** theory is a straightforward inclusion of the spin–orbit interaction and of the strain effects on the band structure via deformation potential theory [48].

Relativistic effects in the **k·p** method are included also perturbatively via the spin–orbit (SO) interaction Hamiltonian, $H_{\mathrm{SO}} = \frac{\hbar}{4m_0^2 c^2} \left[\nabla V(\mathbf{r}) \times \mathbf{p} \right] \cdot \vec{\sigma}$, where $\vec{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$ is a vector consisting of the three Pauli spin matrices. Finally, the **k·p** Hamiltonian becomes

$$H_{\mathbf{k}} = \frac{p^2}{2m_0} + V + \frac{\hbar^2 k^2}{2m_0} + \frac{\hbar \mathbf{k} \cdot \mathbf{p}}{m_0} + \frac{\hbar}{4m_0^2 c^2} (\vec{\sigma} \times \nabla V) \cdot (\hbar \mathbf{k} + \mathbf{p}). \tag{13.6}$$

In a strained system, the lattice constants are stretched or compressed [48]. This affects the SP potential experienced by electrons and consequently yields an additional term in the **k·p** Hamiltonian. Taking into account only the effects up to first order in strain tensor ϵ, the strained **k·p** Hamiltonian becomes

$$H_{\mathbf{k}} = \frac{p^2}{2m_0} + V + \frac{\hbar^2 k^2}{2m_0} + \frac{\hbar \mathbf{k} \cdot \mathbf{p}}{m_0} + \frac{1}{4m_0^2 c^2} (\vec{\sigma} \times \nabla V) \cdot (\mathbf{k} + \mathbf{p}) + \mathbf{D}^{\boldsymbol{\epsilon}} \cdot \boldsymbol{\epsilon}, \tag{13.7}$$

where $\mathbf{D}^{\boldsymbol{\epsilon}}$ is the deformation potential operator. This Hamiltonian can be subjected to the same sort of perturbation theory analysis as above [49].

The most widely used version of the **k·p** Hamiltonian is the 8-band model [50], in which the operator given in Equation 13.7 is represented in the basis of s-bonding and p-antibonding states around the energy gap:

$\lvert u_1 \rangle = \lvert \frac{1}{2}, +\frac{1}{2} \rangle = \lvert s; \uparrow \rangle$	$\lvert u_5 \rangle = \lvert \frac{1}{2}, -\frac{1}{2} \rangle = -\lvert s; \downarrow \rangle$
$\lvert u_2 \rangle = \lvert \frac{3}{2}, +\frac{3}{2} \rangle = \frac{i}{\sqrt{2}} [\lvert x; \uparrow \rangle + i \lvert y; \uparrow \rangle]$	$\lvert u_6 \rangle = \lvert \frac{3}{2}, -\frac{3}{2} \rangle = -\frac{i}{\sqrt{2}} [\lvert x; \downarrow \rangle - i \lvert y; \downarrow \rangle]$
$\lvert u_3 \rangle = \lvert \frac{3}{2}, +\frac{1}{2} \rangle = \frac{i}{\sqrt{6}} [\lvert x; \downarrow \rangle + i \lvert y; \downarrow \rangle - 2 \lvert z; \uparrow \rangle]$	$\lvert u_7 \rangle = \lvert \frac{3}{2}, -\frac{1}{2} \rangle = +\frac{i}{\sqrt{6}} [\lvert x; \uparrow \rangle - i \lvert y; \uparrow \rangle + 2 \lvert z; \downarrow \rangle]$
$\lvert u_4 \rangle = \lvert \frac{1}{2}, +\frac{1}{2} \rangle = \frac{i}{\sqrt{3}} [\lvert x; \downarrow \rangle + i \lvert y; \downarrow \rangle + \lvert z; \uparrow \rangle]$	$\lvert u_8 \rangle = \lvert \frac{1}{2}, -\frac{1}{2} \rangle = +\frac{i}{\sqrt{3}} [\lvert x; \uparrow \rangle - i \lvert y; \uparrow \rangle - \lvert z; \downarrow \rangle]$

Occasionally, the 14-band **k·p** Hamiltonian is also used, in which one exploits the basis of p-bonding, s-bonding, and p-antibonding states around the energy gap [51]:

$\lvert u_1\rangle = \lvert\frac{3}{2},+\frac{3}{2}\rangle = \frac{i}{\sqrt{2}}\lvert(x_a+iy_a)\uparrow\rangle$	$\lvert u_8\rangle = \lvert\frac{3}{2},-\frac{3}{2}\rangle = -\frac{i}{\sqrt{2}}\lvert(x_a-iy_a)\downarrow\rangle$
$\lvert u_2\rangle = \lvert\frac{3}{2},+\frac{1}{2}\rangle = \frac{i}{\sqrt{6}}[\lvert(x_a+iy_a)\downarrow\rangle - 2\lvert z_a\uparrow\rangle]$	$\lvert u_9\rangle = \lvert\frac{3}{2},-\frac{1}{2}\rangle = +\frac{i}{\sqrt{6}}[\lvert(x_a-iy_a)\uparrow\rangle + 2\lvert z_a\downarrow\rangle]$
$\lvert u_3\rangle = \lvert\frac{1}{2},+\frac{1}{2}\rangle = \frac{i}{\sqrt{3}}[\lvert(x_a+iy_a)\downarrow\rangle + \lvert z_a\uparrow\rangle]$	$\lvert u_{10}\rangle = \lvert\frac{1}{2},-\frac{1}{2}\rangle = +\frac{i}{\sqrt{3}}[\lvert(x_a-iy_a)\uparrow\rangle - \lvert z_a\downarrow\rangle]$
$\lvert u_4\rangle = \lvert\frac{1}{2},+\frac{1}{2}\rangle = \lvert s_a\uparrow\rangle$	$\lvert u_{11}\rangle = \lvert\frac{1}{2},-\frac{1}{2}\rangle = -\lvert s_a\downarrow\rangle$
$\lvert u_5\rangle = \lvert\frac{3}{2},+\frac{3}{2}\rangle = \frac{i}{\sqrt{2}}\lvert(x_b+iy_b)\uparrow\rangle$	$\lvert u_{12}\rangle = \lvert\frac{3}{2},-\frac{3}{2}\rangle = -\frac{i}{\sqrt{2}}\lvert(x_b-iy_b)\downarrow\rangle$
$\lvert u_6\rangle = \lvert\frac{3}{2},+\frac{1}{2}\rangle = \frac{i}{\sqrt{6}}[\lvert(x_b+iy_b)\downarrow\rangle - 2\lvert z_b\uparrow\rangle]$	$\lvert u_{13}\rangle = \lvert\frac{3}{2},-\frac{1}{2}\rangle = +\frac{i}{\sqrt{6}}[\lvert(x_b-iy_b)\uparrow\rangle + 2\lvert z_b\downarrow\rangle]$
$\lvert u_7\rangle = \lvert\frac{1}{2},+\frac{1}{2}\rangle = \frac{i}{\sqrt{3}}[\lvert(x_b+iy_b)\downarrow\rangle + \lvert z_b\uparrow\rangle]$	$\lvert u_{14}\rangle = \lvert\frac{1}{2},-\frac{1}{2}\rangle = +\frac{i}{\sqrt{3}}[\lvert(x_b-iy_b)\uparrow\rangle - \lvert z_b\downarrow\rangle]$

13.4 PW Implementation of the k•p Hamiltonian and Coulomb Integrals for QD Structures

The QD as a three-dimensional (3D) object breaks the translational symmetry of the bulk material along all three Cartesian directions implying operator replacement $k_\nu \to -i\partial/\partial\nu$ in Equation 13.6, where $\nu = (x, y, z)$. To solve the multiband system of Schrödinger equations, Equation 13.6, the PW methodology is employed as an expansion method [33,52–56]. In the PW representation, the eigenvalues (E_i) and the coefficients $[A^{(i)}_{n,\mathbf{k}}]$ of the ith eigenvector $[\psi^{(i)}_n(\mathbf{r}) = \sum_\mathbf{k} A^{(i)}_{n,\mathbf{k}} e^{i\mathbf{kr}}]$, are linked by the relation

$$\sum_{n,\mathbf{k}'} h_{m,n}(\mathbf{k},\mathbf{k}') A^{(i)}_{n,\mathbf{k}'} = E_\mathrm{i} A^{(i)}_{m,\mathbf{k}}, \tag{13.8}$$

where $h_{m,n}(\mathbf{k},\mathbf{k}')$ are the Hamiltonian matrix elements in the PW basis, and $m, n \in \{1, ..., 8\}$ or $m, n \in \{1, ..., 14\}$ are the band indices in the $\mathbf{k}\cdot\mathbf{p}$ Hamiltonian. All the elements in the Hamiltonian matrix, Equation 13.8, can be expressed as a linear combination of different kinetic and strain related terms and their convolution with the characteristic function of the QD shape, $\chi_\mathrm{qd}(\mathbf{k})$ [33,57]. The whole $\mathbf{k}$-space is discretized by embedding the QD in a rectangular box of dimensions L_x, L_y, and L_z and volume $\Omega = L_x \times L_y \times L_z$ and choosing the $\mathbf{k}$-vectors in the form $\mathbf{k} = 2\pi(n_x/L_x, n_y/L_y, n_z/L_z)$, where n_x, n_y, and n_z are integers. Maximal absolute values of these integers control the accuracy of the method.

Having determined PW expansion coefficients, $A^{(i)}_{n,\mathbf{k}}$, from Equation 13.8, we can use the single-particle wave functions, $\psi^{(i)}_n(\mathbf{r})$, to calculate the Coulomb integrals, V_{ijkl}, relevant for many-electron processes that occur in QDs. The Coulomb integral among states i, j, k, and l is defined as

$$V_{ijkl} = \sum_{b=1}^{N_b}\sum_{b'=1}^{N_b}\int_\Omega d^3\mathbf{r}\int_\Omega d^3\mathbf{r}'\psi^{(i)}_b(\mathbf{r})^*\psi^{(j)}_b(\mathbf{r})V(|\mathbf{r}-\mathbf{r}'|)\psi^{(k)}_{b'}(\mathbf{r}')^*\psi^{(l)}_{b'}(\mathbf{r}'), \tag{13.9}$$

where $V(\mathbf{u}) = e^2/4\pi\epsilon\mathbf{u}$, with ϵ being the static dielectric constant, b is the band index, and N_b is the number of bands, i.e., 8 or 14 in the $\mathbf{k}\cdot\mathbf{p}$ Hamiltonian. The integral 13.9 can be rewritten as

$$V_{ijkl} = \int_\Omega d^3\mathbf{r}\int_\Omega d^3\mathbf{r}' B_{ij}(\mathbf{r})V(|\mathbf{r}-\mathbf{r}'|)B_{kl}(\mathbf{r}'), \tag{13.10}$$

where we introduce

$$B_{ij}(\mathbf{r}) = \sum_{b=1}^{N_b}\psi^{(i)}_b(\mathbf{r})^*\psi^{(j)}_b(\mathbf{r}). \tag{13.11}$$

The integral 13.10 can be calculated by performing a six-dimensional integration in real space, which is, however, numerically very demanding, especially when a large number of integrals has to be calculated, as in the CI approach [33,58]. One of the advantages of the PW method is that relevant physical quantities can be expressed in terms of the coefficients in the PW expansion, i.e., by the Fourier transform (FT). The last statement holds in principle when the evaluation of Coulomb integrals is concerned, since the FT of the Coulomb potential is fully analytical. This can be shown by the following set of transformations leading to Equation 13.16. Defining the PW expansion of $B_{ij}(\mathbf{r})$

$$B_{ij}(\mathbf{r}) = \sum_{\mathbf{k}} B_{ij}(\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{r}} \tag{13.12}$$

and putting the last expression into Equation 13.10 one obtains

$$V_{ijkl} = \sum_{\mathbf{k}} B_{ij}(\mathbf{k}) \sum_{\mathbf{k}'} B_{kl}(\mathbf{k}') \int_{\Omega} d^3\mathbf{r} \int_{\Omega} d^3\mathbf{r}' e^{i\mathbf{k}\cdot\mathbf{r}} V(|\mathbf{r}-\mathbf{r}'|) e^{i\mathbf{k}'\cdot\mathbf{r}'}. \tag{13.13}$$

Using the properties of convolution and Parseval's theorem in inverse space, the $B_{ij}(\mathbf{k})$ can be expressed in terms in the coefficients in the PW expansion of the envelope functions as

$$B_{ij}(\mathbf{k}) = \sum_{b=1}^{N_b} \sum_{\mathbf{q}} A^{(i)*}_{b,\mathbf{q}} A^{(j)}_{b,\mathbf{q}+\mathbf{k}}. \tag{13.14}$$

By introducing the approximation changing the domain of integration in one of the integrals in Equation 13.13 from Ω to the whole space (which is valid when Ω is large enough) one gets after the replacement of variables from $\mathbf{r}$ and $\mathbf{r}'$, to $\mathbf{r}$ and $\mathbf{u} = \mathbf{r} - \mathbf{r}'$

$$V_{ijkl} = \sum_{\mathbf{k}} B_{ij}(\mathbf{k}) \sum_{\mathbf{k}'} B_{kl}(\mathbf{k}') \left[\int_{\Omega} d^3\mathbf{r} e^{i\mathbf{k}\cdot\mathbf{r}} e^{i\mathbf{k}'\cdot\mathbf{r}}\right]\left[\int d^3\mathbf{u} V(|\mathbf{u}|) e^{-i\mathbf{k}'\cdot\mathbf{u}}\right]. \tag{13.15}$$

Using the relations

$$\int d^3\mathbf{u} e^{-i\mathbf{k}'\mathbf{u}} V(|\mathbf{u}|) = \frac{e^2}{\epsilon k'^2}$$

and

$$\frac{1}{\Omega}\int_{\Omega} d^3\mathbf{r} e^{i(\mathbf{k}+\mathbf{k}')\mathbf{r}} = \delta_{\mathbf{k}+\mathbf{k}',0}$$

one obtains

$$V_{ijkl} = \Omega \sum_{\substack{\mathbf{k} \\ \mathbf{k}\neq 0}} B_{ij}(\mathbf{k}) B_{kl}(-\mathbf{k}) \frac{e^2}{\epsilon k^2}. \tag{13.16}$$

13.5 The Effect of Dielectric Confinement

For colloidal QDs, the dielectric constant ϵ of the QD material is typically much larger than that of the surrounding medium, and the spatial variation $\epsilon = \epsilon(\mathbf{r})$ cannot be ignored. Such dielectric contrast means that

any free charge in the QD induces polarization charge in the QD and its surroundings. As a consequence, the Coulomb energy for the electron–hole pair for a system with a spatially varying dielectric constant is

$$V(\mathbf{r}_e, \mathbf{r}_h) = V_{\mathrm{c}}(\mathbf{r}_e, \mathbf{r}_h) + V_{\mathrm{s}}(\mathbf{r}_e) + V_{\mathrm{s}}(\mathbf{r}_h), \tag{13.17}$$

where V_{c} is the interparticle Coulomb potential and V_{s} is the self-polarization potential. The potential, $V_{\mathrm{c}}(\mathbf{r}_e, \mathbf{r}_h)$, encompasses both the "direct" interparticle Coulomb interaction and the interface polarization potential due to the interaction between a real particle and the induced charge of the other particle. The self-polarization potential $V_{\mathrm{s}}(\mathbf{r})$ arises from the interaction of a particle and its own induced charge.

The interparticle Coulomb potential in real space:

$$V_{\mathrm{c}}(\mathbf{r}_e, \mathbf{r}_h) = \frac{1}{\epsilon(\mathbf{r}_e, \mathbf{r}_h)} \frac{e^2}{4\pi\epsilon_0 |\mathbf{r}_e - \mathbf{r}_h|}, \tag{13.18}$$

under the assumption, $\epsilon(\mathbf{r}_e, \mathbf{r}_h) \approx \epsilon(\mathbf{r}_e - \mathbf{r}_h) = \epsilon(\mathbf{r})$ [59], can be recast in the form of product of two functions

$$V_{\mathrm{c}}(\mathbf{r}) = \frac{1}{\epsilon(\mathbf{r})} \frac{e^2}{4\pi\epsilon_0 |\mathbf{r}|} = \frac{1}{\epsilon(\mathbf{r})} U(\mathbf{r}), \tag{13.19}$$

where $U(\mathbf{r}) = e^2/4\pi\epsilon_0\mathbf{r}$ is the bare Coulomb potential due to charge e, with the well-known FT in the inverse space: $\mathcal{F}[U(\mathbf{r})] = \tilde{U}(\mathbf{k}) = e^2/\epsilon_0 k^2$. Considering stepwise uniform regions where $\epsilon(r) = \epsilon_m = \text{const.}$, and m is the index of the region, and spherical QD shape, the general expression, Equation 13.18, acquires the simple analytical form:

$$\tilde{V}_{\mathrm{c}}(\mathbf{k}) = \frac{e^2}{\epsilon_0 k^2} \left[\frac{1}{\epsilon_{\mathrm{c}}} + \left(\frac{1}{\epsilon_{\mathrm{s}}} - \frac{1}{\epsilon_{\mathrm{c}}} \right) \cos(kR_{\mathrm{c}}) + \left(\frac{1}{\epsilon_{\mathrm{coll.}}} - \frac{1}{\epsilon_{\mathrm{s}}} \right) \cos(kR_{\mathrm{s}}) \right], \tag{13.20}$$

where ϵ_{c}, ϵ_{s}, and $\epsilon_{\mathrm{coll.}}$ are the dielectric constants of the QD's core, shell, and surrounding colloid respectively, and R_{c} and R_{s} are the radius of core and outer radius of the shell, respectively.

Calculation of the self-polarization potential for the arbitrary QD shape in the inverse space is a bit involved. Alternatively, for several characteristic QD shapes, like spherical QD [60], spherical core/shell QD [61,62], elliptical QD [63], elliptical core/shell QD [64–66], and cuboidal QD [67], there are available expressions for $V_{\mathrm{s}}(\mathbf{r})$ in real space. Here, we list a general expression for the self-polarization potential of spherical QD with an arbitrary number, m, of core/shells:

$$V_{\mathrm{s}}^{m}(r) = \frac{e^2}{8\pi\epsilon_0\epsilon_m} \sum_{l=0}^{\infty} \frac{1}{1 - p_{m,l}q_{m,l}} \left[p_{m,l} r^{2l} + p_{m,l} q_{m,l} r^{-1} + q_{m,l} r^{-2(l+1)} \right], \tag{13.21}$$

where $p_{m,l}$ and $q_{m,l}$ are recursive coefficients that depend on the dielectric constants, ϵ_m, and radii, R_m, in the region m, of the core/shell structures (see Ref. [61]). Such expressions are then simple to FT into $\tilde{V}_{\mathrm{s}}(\mathbf{k})$ and are used when calculating the Coulomb matrix elements, V_{ijkl} for the CI Hamiltonian; see Section 13.6. The expression for the contribution of the direct Coulomb, Equation 13.20, and self-polarization energy, Equation 13.21, to the total Coulomb integral reads as

$$V_{ijkl}^{(c)} = \Omega \sum_{\substack{\mathbf{k} \\ \mathbf{k} \neq 0}} B_{ij}(\mathbf{k}) B_{kl}(-\mathbf{k}) \tilde{V}_{\mathrm{c}}(\mathbf{k}) \tag{13.22}$$

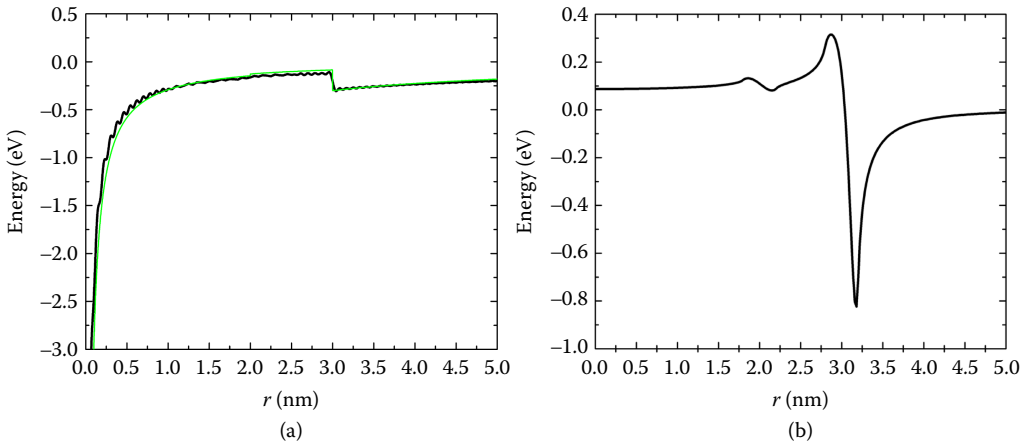

FIGURE 13.2 Variation of the interparticle Coulomb potential (a), $V_c(\mathbf{r})$, and self-polarization potential (b), $V_s(\mathbf{r})$, of CdSe/CdTe core shell quantum dot (QD) with $\epsilon_c = 7.1$, $\epsilon_s = 6.2$, $\epsilon_{coll.} = 2$, and $R_c = 2$ nm and $R_s = 3$ nm. In (a), the tiny bright line represents analytic $V_c(r)$ obtained by Equation 13.19, while solid thick line represents the same $V_c(r)$ recovered from the analytical expression in the inverse space, Equation 13.20.

and

$$V_{ijkl}^{(s)} = \Omega \sum_{\mathbf{k}} [B_{ij}(\mathbf{k})\delta_{kl} + \delta_{ij}B_{kl}(\mathbf{k})]\tilde{V}_s(\mathbf{k}). \tag{13.23}$$

In colloidal core/shell QDs, the self-polarization potential is characterized by a small peak and well near the core/shell interface due to the small dielectric mismatch between the core and shell materials. However, a much larger peak just inside $r = R_c + R_s$ and a deep well slightly outside the QD are due to the far greater dielectric mismatch of the shell and matrix material. In Figure 13.2, we present the radial variation of $V_c(\mathbf{r})$ and $V_s(\mathbf{r})$ of CdSe/CdTe core shell QD with $\epsilon_c = 7.1$, $\epsilon_s = 6.2$, $\epsilon_{coll.} = 2$, and $R_c = 2$ nm and $R_s = 3$ nm.

13.6 The Configuration Interaction Hamiltonian for NQD

In this section, we derive the Hamiltonian describing interacting electrons and holes in QDs. This Hamiltonian can be used to calculate neutral and charged excitonic and multiexcitonic states that are populated upon photoexcitation of the QD.

We start with a general Hamiltonian describing interacting electrons in the solid

$$H = H_s + H_{int}, \tag{13.24}$$

where the noninteracting part of the Hamiltonian is

$$H_s = \int d^3x\, \psi^\dagger(x) \left[-\frac{\hbar^2}{2m_0}\nabla^2 + V(x)\right] \psi(x)\, d^3x \tag{13.25}$$

and the interaction part is

$$H_{int} = \frac{1}{2}\int d^3x\, d^3y\, \psi^\dagger(x)\, \psi^\dagger(y)\, V_{int}(x,y)\, \psi(y)\, \psi(x). \tag{13.26}$$

In previous equations, $\psi(x)$ is the fermionic field operator, $\psi^{\dagger}(x)$ is its adjoint operator, V is the potential of atomic cores in the solid, and $V_{\text{int}}(x,y)$ is the interaction potential between the particles at x and y, which for the bare Coulomb interaction reads $V_{\text{int}}(x,y) = \frac{e^2}{4\pi\epsilon_0|x-y|}$.

Next, we express the field operator in terms of annihilation operators as

$$\psi(x) = \sum_{v\in \text{VB}} a_v \phi_v(x) + \sum_{c\in \text{CB}} a_c \phi_c(x), \tag{13.27}$$

where $\phi_c(x)$ and $\phi_v(x)$ are the SP wave function for states in the CB and VB, while a_c and a_v are the corresponding electron annihilation operators. The operator a_c acts on arbitrary many-body state of the system by destroying an electron in state c and a_v acts by destroying an electron in state v. Their adjoint operators $a_c^{\dagger}$ and $a_v^{\dagger}$ act by creating an electron in the corresponding state. After substitution of Equation 13.27 into Equations 13.25 and 13.26, one arrives at the expressions

$$H_s = \sum_{j_1 j_2} a_{j_1}^{\dagger} a_{j_2} \int d^3x\, \phi_{j_1}^{*}(x) \left[-\frac{\hbar^2}{2m_0}\nabla^2 + V(x)\right] \phi_{j_2}(x) \tag{13.28}$$

$$H_{\text{int}} = \frac{1}{2} \sum_{j_1 j_2 j_3 j_4} a_{j_1}^{\dagger} a_{j_2}^{\dagger} a_{j_3} a_{j_4} \int d^3x\, d^3y\, \phi_{j_1}^{*}(x)\, \phi_{j_4}(x)\, V_{\text{int}}(x,y)\, \phi_{j_2}^{*}(y)\, \phi_{j_3}(y), \tag{13.29}$$

i.e.,

$$H = \sum_{j_1 j_2} v_{j_1 j_2} a_{j_1}^{\dagger} a_{j_2} + \frac{1}{2} \sum_{j_1 j_2 j_3 j_4} V_{j_1 j_4 j_2 j_3} a_{j_1}^{\dagger} a_{j_2}^{\dagger} a_{j_3} a_{j_4}, \tag{13.30}$$

where

$$v_{j_1 j_2} = \int d^3x\, \phi_{j_1}^{*}(x) \left[-\frac{\hbar^2}{2m_0}\nabla^2 + V(x)\right] \phi_{j_2}(x) \tag{13.31}$$

$$V_{j_1 j_4 j_2 j_3} = \int d^3x\, d^3y\, \phi_{j_1}^{*}(x)\, \phi_{j_4}(x)\, V_{\text{int}}(x,y)\, \phi_{j_2}^{*}(y)\, \phi_{j_3}(y). \tag{13.32}$$

The summations in Equation 13.30 should in principle include all SP states in the valence and the CB. However, we are interested mainly in the lowest energy excitations of the system where electrons at the bottom of the CB and holes at the top of the VB are created. To study these excitations, it is more convenient to use electron and hole creation and annihilation operators $e_j(e_j^{\dagger})$ and $h_j(h_j^{\dagger})$ defined as $e_j = a_j$, when $j \in$ CB and $h_j = a_j^{\dagger}$ when $j \in$ VB. These operators satisfy the following fermionic commutation rules: $\{e_i, e_j^{\dagger}\} = \delta_{ij}$, $\{h_i, h_j^{\dagger}\} = \delta_{ij}$, $\{e_i, e_j\} = 0$, $\{h_i, h_j\} = 0$, $\{e_i^{\dagger}, h_j\} = 0$, $\{e_i, h_j^{\dagger}\} = 0$, where $\{a,b\} = ab + ba$ denotes the anticommutator of two operators. The H_s term then reads

$$H_s = \sum_{j_1 j_2 \in \text{CB}} v_{j_1 j_2} e_{j_1}^{\dagger} e_{j_2} - \sum_{j_1 j_2 \in \text{VB}} v_{j_2 j_1} h_{j_1}^{\dagger} h_{j_2} + \sum_{j\in \text{VB}} v_{jj}, \tag{13.33}$$

where the last term contributes an irrelevant constant shift and can be omitted from further consideration. To obtain the H_{int} term, we first introduce an approximation to consider only the terms in H_{int} that contain the same number of electron creation and annihilation operators, as well as the same number of hole

creation and annihilation operators. The other terms that do not conserve the number of pair excitations give rise to the polarization of the electronic orbitals, and their main effects can be taken into account by a dielectric constant in the Coulomb potential, which can be inserted into the final result [68].

Under such an assumption the H_{int} term consists of three sets of terms $H_{\text{int}} = H_{\text{int}}^{(1)} + H_{\text{int}}^{(2)} + H_{\text{int}}^{(3)}$. The first set of these terms (when $j_1, j_2, j_3, j_4 \in \text{CB}$) reads:

$$H_{\text{int}}^{(1)} = \frac{1}{2} \sum_{j_1 j_2 j_3 j_4} V_{j_1 j_4 j_2 j_3} e_{j_1}^{\dagger} e_{j_2}^{\dagger} e_{j_3} e_{j_4}. \tag{13.34}$$

The second set of these terms $H_{\text{int}}^{(2)}$ that is obtained when $j_1, j_2, j_3, j_4 \in \text{VB}$ contains the term $h_{j_1} h_{j_2} h_{j_3}^{\dagger} h_{j_4}^{\dagger}$. By exploiting the fermionic commutation relations, this term can be rewritten as

$$h_{j_1} h_{j_2} h_{j_3}^{\dagger} h_{j_4}^{\dagger} = \delta_{j_2 j_3} \delta_{j_1 j_4} - \delta_{j_1 j_3} \delta_{j_2 j_4} - \delta_{j_2 j_3} h_{j_4}^{\dagger} h_{j_1} - \delta_{j_1 j_4} h_{j_3}^{\dagger} h_{j_2} + \delta_{j_1 j_3} h_{j_4}^{\dagger} h_{j_2} + \delta_{j_2 j_4} h_{j_3}^{\dagger} h_{j_1} + h_{j_3}^{\dagger} h_{j_4}^{\dagger} h_{j_1} h_{j_2}. \tag{13.35}$$

The $H_{\text{int}}^{(2)}$ term is then given as

$$\begin{aligned} H_{\text{int}}^{(2)} = {} & \frac{1}{2} \sum_{j_1 j_2} V_{j_1 j_1 j_2 j_2} - \frac{1}{2} \sum_{j_1 j_2} V_{j_1 j_2 j_2 j_1} - \frac{1}{2} \sum_{j_1 j_2 j_4} V_{j_1 j_4 j_2 j_2} h_{j_4}^{\dagger} h_{j_1} - \frac{1}{2} \sum_{j_1 j_2 j_3} V_{j_1 j_1 j_2 j_3} h_{j_3}^{\dagger} h_{j_2} \\ & + \frac{1}{2} \sum_{j_1 j_2 j_4} V_{j_1 j_4 j_2 j_1} h_{j_4}^{\dagger} h_{j_2} + \frac{1}{2} \sum_{j_1 j_2 j_3} V_{j_1 j_2 j_2 j_3} h_{j_3}^{\dagger} h_{j_1} + \frac{1}{2} \sum_{j_1 j_2 j_3 j_4} V_{j_1 j_4 j_2 j_3} h_{j_3}^{\dagger} h_{j_4}^{\dagger} h_{j_1} h_{j_2}. \end{aligned} \tag{13.36}$$

In Equation 13.36, the first and the second terms describe the Coulomb direct and exchange interactions between the carriers in fully occupied VB. These terms lead to an irrelevant shift of the total energy and can be omitted from the Hamiltonian. The third and the fourth terms describe the direct Coulomb interaction of holes with fully occupied VB, while the fifth and the sixth terms describe the exchange interaction of holes with fully occupied VB. The third, fourth, fifth, and sixth terms are of the same mathematical form as the second term in H_{s} [Equation 13.33] and can be considered to originate from the effective potential that describes the interaction of holes with occupied VB. Consequently, these terms can be incorporated into H_{s}. Finally, the last term in Equation 13.36 describes the hole–hole interaction.

The third set of terms, $H_{\text{int}}^{(3)}$, includes the terms that contain two electron and two hole operators. Among these terms, the terms that contain $h_{j_1} e_{j_2}^{\dagger} e_{j_3} h_{j_4}^{\dagger}$ and $e_{j_1}^{\dagger} h_{j_2} h_{j_3}^{\dagger} e_{j_4}$ contribute equally to $H_{\text{int}}^{(3)}$. The same holds for the terms that contain $h_{j_1} e_{j_2}^{\dagger} h_{j_3}^{\dagger} e_{j_4}$ and $e_{j_1}^{\dagger} h_{j_2} e_{j_3} h_{j_4}^{\dagger}$. By exploiting the identities

$$h_{j_1} e_{j_2}^{\dagger} e_{j_3} h_{j_4}^{\dagger} = e_{j_2}^{\dagger} e_{j_3} h_{j_1} h_{j_4}^{\dagger} = \delta_{j_1 j_4} e_{j_2}^{\dagger} e_{j_3} + e_{j_2}^{\dagger} h_{j_4}^{\dagger} e_{j_3} h_{j_1} \tag{13.37}$$

and

$$h_{j_1} e_{j_2}^{\dagger} h_{j_3}^{\dagger} e_{j_4} = -e_{j_2}^{\dagger} e_{j_4} h_{j_1} h_{j_3}^{\dagger} = -\delta_{j_1 j_3} e_{j_2}^{\dagger} e_{j_4} - e_{j_2}^{\dagger} h_{j_3}^{\dagger} e_{j_4} h_{j_1} \tag{13.38}$$

one obtains

$$\begin{aligned} H_{\text{int}}^{(3)} = {} & \sum_{j_1 j_2 j_3 j_4} V_{j_1 j_4 j_2 j_3} e_{j_2}^{\dagger} h_{j_4}^{\dagger} e_{j_3} h_{j_1} - \sum_{j_1 j_2 j_3 j_4} V_{j_1 j_4 j_2 j_3} e_{j_2}^{\dagger} h_{j_4}^{\dagger} e_{j_3} h_{j_1} \\ & + \sum_{j_1 j_2 j_3} V_{j_1 j_1 j_2 j_3} e_{j_2}^{\dagger} e_{j_3} - \sum_{j_1 j_2 j_4} V_{j_1 j_4 j_2 j_1} e_{j_2}^{\dagger} e_{j_4}. \end{aligned} \tag{13.39}$$

The last two terms in Equation 13.39 describe direct and exchange Coulomb interaction of electrons with fully occupied VB. In the same way as similar terms for holes these can be incorporated into H_s. One thus obtains the Hamiltonian

$$\begin{aligned} H = & \sum_{j_1 j_2 \in \mathrm{CB}} \tilde{v}_{j_1 j_2} e^{\dagger}_{j_1} e_{j_2} - \sum_{j_1 j_2 \in \mathrm{VB}} \tilde{v}_{j_2 j_1} h^{\dagger}_{j_1} h_{j_2} \\ & + \frac{1}{2} \sum_{ijkl} V_{iljk} e^{\dagger}_{\mathrm{i}} e^{\dagger}_{j} e_k e_l + \frac{1}{2} \sum_{ijkl} V_{likj} h^{\dagger}_{\mathrm{i}} h^{\dagger}_{j} h_k h_l \\ & - \sum_{ijkl} e^{\dagger}_{\mathrm{i}} h^{\dagger}_{j} h_k e_l \left(V_{ilkj} - V_{ijkl} \right). \end{aligned} \tag{13.40}$$

So far, we have not specified the SP states $\phi_{\mathrm{c}}(x)$ and $\phi_v(x)$. If we choose these states to be eigenstates of the effective SP Hamiltonian $\tilde{v}_{ij}$, Equation 13.40 reduces to

$$\begin{aligned} H = & \sum_{j \in \mathrm{CB}} \varepsilon_j e^{\dagger}_{j} e_j - \sum_{j \in \mathrm{VB}} \varepsilon_j h^{\dagger}_{j} h_j \\ & + \frac{1}{2} \sum_{ijkl} V_{iljk} e^{\dagger}_{\mathrm{i}} e^{\dagger}_{j} e_k e_l + \frac{1}{2} \sum_{ijkl} V_{likj} h^{\dagger}_{\mathrm{i}} h^{\dagger}_{j} h_k h_l \\ & - \sum_{ijkl} e^{\dagger}_{\mathrm{i}} h^{\dagger}_{j} h_k e_l \left(V_{ilkj} - V_{ijkl} \right). \end{aligned} \tag{13.41}$$

To obtain the energies of exciton states, one needs to diagonalize the Hamiltonian given in Equation 13.41. Since all terms in the Hamiltonian contain only the operators that conserve the number of electrons and the number of holes, one can separately find exciton states (one electron and one hole), biexciton states (two electrons and two holes), negative trions (two electrons and one hole), etc. For example, to find the exciton states one has to diagonalize the Hamiltonian in the Hilbert space spanned by the vectors $e^{\dagger}_{a} h^{\dagger}_{b} |G\rangle$, where $|G\rangle$ is the ground state of the system (fully occupied VB and empty CB), $a \in$ CB and $b \in$ VB. Each element of the Hamiltonian matrix depends on Coulomb integrals defined in Equation 13.32. When the **k**·**p** method is used to evaluate the SP wave functions, these are expressed as

$$\phi_j(x) = \sum_b \psi^{(b)}_j(x)\, u_b(x), \tag{13.42}$$

where b is the band index, $u_b(x)$ is the Bloch function of underlying crystal, and $\psi^{(b)}_j$ is the slowly varying envelope function. The Coulomb integral (Equation 13.32) then reads

$$\begin{aligned} V_{j_1 j_4 j_2 j_3} = \sum_{b_1 b_2 b_3 b_4} \int d^3x\, d^3y\, \psi^{(j_1)}_{b_1}(x)^*\, \psi^{(j_4)}_{b_4}(x)\, V_{\mathrm{int}}(x,y)\, \psi^{(j_2)}_{b_2}(y)^*\, \psi^{(j_3)}_{b_3}(y) \times \\ \times u^*_{b_1}(x)\, u_{b_4}(x)\, u^*_{b_2}(y)\, u_{b_3}(y). \end{aligned} \tag{13.43}$$

Next, we exploit the fact that for the slowly varying function $f(\mathbf{u})$ and rapidly varying periodic function $g(\mathbf{u})$, the following approximation holds

$$\int d\mathbf{u}\, f(\mathbf{u})\, g(\mathbf{u}) \approx \int d\mathbf{u}\, f(\mathbf{u}) \langle g(\mathbf{u}) \rangle_T, \tag{13.44}$$

where $\langle g(\mathbf{u})\rangle_T$ is the average of the function $g(\mathbf{u})$ over its period. In Equation 13.43, the terms in the first line define a slowly varying function, while the terms in the second line define a rapidly varying function whose average over the period is

$$\left\langle u^*_{b_1}(x)\, u_{b_4}(x)\, u^*_{b_2}(y)\, u_{b_3}(y)\right\rangle_T = \delta_{b_1 b_4}\delta_{b_2 b_3}, \tag{13.45}$$

due to orthonormality of the Bloch functions. Consequently, the Coulomb integral 13.43 can be expressed only in terms of envelope functions as

$$V_{j_1 j_4 j_2 j_3} = \sum_{b_1 b_2}\int d^3x\, d^3y\, \psi^{(j_1)}_{b_1}(x)^*\, \psi^{(j_4)}_{b_1}(x)\, V_{\text{int}}(x,y)\, \psi^{(j_2)}_{b_2}(y)^*\, \psi^{(j_3)}_{b_2}(y). \tag{13.46}$$

These integrals can be evaluated in the PW representation using the expression given by Equation 13.16 in Section 13.4. More details on the implementation of a numerically efficient scheme for evaluation of Coulomb integrals when the PW basis is used to represent the envelope functions is given in Ref. [33].

13.7 Transport Properties of QD Nanocrystal Arrays

In this section, we discuss electrical transport properties of colloidal QD nanocrystal arrays and we derive the formula for the charge carrier hopping probability from one dot to another.

QDs obtained by colloidal chemistry typically contain ligand molecules that passivate the surface and eliminate unwanted surface trap states. However, the ligands also increase the separation between the neighboring dots and consequently decrease the wave function overlap between the electron wave functions in neighboring dots (and their electronic coupling). The presence of ligands, therefore, significantly inhibits charge transport between neighboring dots. The approaches where short ligands are used enable the improvement of electrical transport between the dots.

To engineer the transport in QD arrays, it is highly desirable to have a theoretical approach for calculating the electrical transport properties of these systems. Since electronic coupling between the dots is weak, it can be considered a perturbation that leads to occasional carrier hopping from one dot to another. To evaluate the hopping rate of charge carrier between the dots, we model the system using the following Hamiltonian

$$H = H_1 + H_2 + V. \tag{13.47}$$

In the previous equation, H_1 is the Hamiltonian of the first dot, H_2 is the Hamiltonian of the second dot, and V is the part of the Hamiltonian that describes the electronic coupling between two dots. For notational simplicity, we will assume that each of the dots accommodates only one electronic energy level that is coupled to all vibrational modes of that dot. The Hamiltonian H_1 is then given as

$$H_1 = \varepsilon_1 a_1^\dagger a_1 + \sum_s \hbar\omega_{1s} b_{1s}^\dagger b_{1s} + \sum_s g_{1s}\hbar\omega_{1s} a_1^\dagger a_1\left(b_{1s}^\dagger + b_{1s}\right). \tag{13.48}$$

In the previous equation, a_1 is the electron annihilation operator that annihilates an electron in state of energy ε_1, g_{1s} is the dimensionless coupling constant describing the coupling of electron in dot 1 to the phonon mode s of dot 1, and $\hbar\omega_{1s}$ is the energy of that phonon mode, while b_{1s} is the corresponding phonon annihilation operator. Electronic operators satisfy fermionic commutation relations, while phonon

operators satisfy bosonic commutation relations. Analogously, the Hamiltonian H_2 reads

$$H_2 = \varepsilon_2 a_2^\dagger a_2 + \sum_s \hbar\omega_{2s} b_{2s}^\dagger b_{2s} + \sum_s g_{2s}\hbar\omega_{2s} a_2^\dagger a_2 \left(b_{2s}^\dagger + b_{2s}\right), \tag{13.49}$$

where the notation is analogous to that in the Hamiltonian H_1. The last term in Equation 13.47 reads

$$V = -t\left(a_1^\dagger a_2 + a_2^\dagger a_1\right), \tag{13.50}$$

where t is the electronic coupling parameter between the two electronic states in dots 1 and 2.

We are interested in determining the transition rate for the process in which an electron that is initially in dot 1 makes a transition to dot 2. Possible initial states for this process are $|i\rangle = a_1^\dagger \left|i_{\text{ph}}\right\rangle$, where $\left|i_{\text{ph}}\right\rangle$ denotes a state of the phonon subsystem. The final state is given as $\left|f\right\rangle = a_2^\dagger \left|f_{\text{ph}}\right\rangle$, where $\left|f_{\text{ph}}\right\rangle$ is the final state of the phonon subsystem. The transition rate is then given by Fermi's Golden Rule rule expression

$$W = \frac{2\pi}{\hbar}\sum_i p_i \sum_f \left|V_{if}\right|^2 \delta\left(E_{\text{i}} - E_{\text{f}}\right), \tag{13.51}$$

where p_{i} is the probability that the initial state is $|i\rangle$ $V_{if} = \langle i\,|V|\,f\rangle$ is the matrix element of the operator V, while E_{i} and E_{f} are energies of the unperturbed system in the initial and final states. We will transform this conventional expression to a form that is more convenient for our analysis. Using the identities

$$\delta\left(E_{\text{i}} - E_{\text{f}}\right) = \frac{1}{2\pi\hbar}\int_{-\infty}^{\infty} dt \exp\left[\frac{i}{\hbar}\left(E_{\text{i}} - E_{\text{f}}\right)t\right], \tag{13.52}$$

$$\left|V_{if}\right|^2 = \langle i\,|V|\,f\rangle\,\langle f\,|V|\,i\rangle, \tag{13.53}$$

and

$$\langle i\,|V|\,f\rangle \exp\left[\frac{i}{\hbar}\left(E_{\text{i}} - E_{\text{f}}\right)t\right] = \left\langle i\left|\exp\left[\frac{i}{\hbar}H_0 t\right] V \exp\left[-\frac{i}{\hbar}H_0 t\right]\right|f\right\rangle \tag{13.54}$$

where $H_0 = H_1 + H_2$, we arrive at the expression for transition rates that is convenient for our analysis

$$W = \frac{1}{\hbar^2}\int_{-\infty}^{\infty} dt \sum_i p_i \sum_f \left\langle i\left|V_I(t)\right|f\right\rangle \left\langle f\left|V_I(0)\right|i\right\rangle, \tag{13.55}$$

where $V_I(t) = \exp\left[\frac{i}{\hbar}H_0 t\right] V \exp\left[-\frac{i}{\hbar}H_0 t\right]$.

The Hamiltonians H_1 and H_2 can be diagonalized exactly using the unitary transformation

$$U = \exp\left[a_1^\dagger a_1 \sum_s g_{1s}\left(b_{1s}^\dagger - b_{1s}\right)\right] \exp\left[a_2^\dagger a_2 \sum_s g_{2s}\left(b_{2s}^\dagger - b_{2s}\right)\right]. \tag{13.56}$$

This transformation acts on electron and phonon operators as

$$Ua_{\mathrm{i}}U^{\dagger} = a_{\mathrm{i}} \exp\left[\sum_{\mathrm{s}} g_{is}\left(b_{is}^{\dagger} - b_{is}\right)\right] \tag{13.57}$$

and

$$Ub_{is}U^{\dagger} = b_{is} - g_{is}a_{\mathrm{i}}^{\dagger}a_{\mathrm{i}}. \tag{13.58}$$

The transformed Hamiltonian then reads

$$\tilde{H} = UHU^{\dagger} = \tilde{H}_1 + \tilde{H}_2 + \tilde{V}, \tag{13.59}$$

with

$$\tilde{H}_1 = \left(\varepsilon_1 - \sum_{\mathrm{s}} g_{1s}^2 \hbar\omega_{\mathrm{s}}\right) a_1^{\dagger}a_1 + \sum_{\mathrm{s}} \hbar\omega_{1s} b_{1s}^{\dagger} b_{1s}, \tag{13.60}$$

$$\tilde{H}_2 = \left(\varepsilon_2 - \sum_{\mathrm{s}} g_{2s}^2 \hbar\omega_{\mathrm{s}}\right) a_2^{\dagger}a_2 + \sum_{\mathrm{s}} \hbar\omega_{2s} b_{2s}^{\dagger} b_{2s}, \tag{13.61}$$

and

$$\tilde{V} = -t\left[a_1^{\dagger}a_2 \exp\left[-\sum_{\mathrm{s}} g_{1s}\left(b_{1s} - b_{1s}^{\dagger}\right)\right] \exp\left[\sum_{\mathrm{s}} g_{2s}\left(b_{2s} - b_{2s}^{\dagger}\right)\right] + c.c.\right]. \tag{13.62}$$

To obtain the expression for transition probability given by Equation 13.55, we first express the matrix elements as

$$\left\langle i\left|V_I(t)\right|f\right\rangle = \langle\Psi_{\mathrm{i}}| \exp\left(\frac{i}{\hbar}\tilde{H}_0 t\right)\tilde{V}\exp\left(-\frac{i}{\hbar}\tilde{H}_0 t\right)|\Psi_{\mathrm{f}}\rangle. \tag{13.63}$$

In the last expression $|\Psi_{\mathrm{i}}\rangle = U|i\rangle$ and $|\Psi_{\mathrm{f}}\rangle = U|f\rangle$. By noting that $\tilde{H}_0 = \tilde{H}_0^{\mathrm{e}} + \tilde{H}_0^{\mathrm{ph}}$ and $\exp\left[-\frac{i}{\hbar}\tilde{H}_0^{\mathrm{e}}t\right]|\Psi_{\mathrm{f}}\rangle = \exp\left[-\frac{i}{\hbar}\varepsilon_{\mathrm{f}}t\right]|\Psi_{\mathrm{f}}\rangle$, using Equations 13.62 and 13.55 one arrives at

$$\begin{aligned} W = {} & \frac{t^2}{\hbar^2}\int_{-\infty}^{\infty} du \exp\left[\frac{i}{\hbar}\left(\varepsilon_{\mathrm{i}} - \varepsilon_{\mathrm{f}}\right)u\right] \sum_{i_{\mathrm{ph}}} p_{i_{\mathrm{ph}}} \\ & \times \sum_{f_{\mathrm{ph}}} \langle i_{\mathrm{ph}}| \exp\left(\frac{i}{\hbar}\tilde{H}_0^{\mathrm{ph}}u\right) \exp\left[-\sum_{\mathrm{s}} g_{1s}\left(b_{1s} - b_{1s}^{\dagger}\right)\right] \exp\left[\sum_{\mathrm{s}} g_{2s}\left(b_{2s} - b_{2s}^{\dagger}\right)\right] \exp\left(-\frac{i}{\hbar}\tilde{H}_0^{\mathrm{ph}}u\right)|f_{\mathrm{ph}}\rangle \\ & \times \langle f_{\mathrm{ph}}| \exp\left[\sum_{\mathrm{s}} g_{1s}\left(b_{1s} - b_{1s}^{\dagger}\right)\right] \exp\left[-\sum_{\mathrm{s}} g_{2s}\left(b_{2s} - b_{2s}^{\dagger}\right)\right] |i_{\mathrm{ph}}\rangle. \end{aligned} \tag{13.64}$$

Taking into account that $\sum_{f_{\rm ph}} |f_{\rm ph}\rangle\langle f_{\rm ph}| = 1$ and $\sum_{i_{\rm ph}} p_{i_{\rm ph}} \langle i_{\rm ph}|X|i_{\rm ph}\rangle = \langle X\rangle_{\rm ph}$, where $\langle X\rangle_{\rm ph}$ denotes the expectation value of the operator X, we obtain

$$\begin{aligned} W = &\frac{t^2}{\hbar^2}\int_{-\infty}^{\infty} du \exp\left[\frac{i}{\hbar}\left(\varepsilon_{\rm i}-\varepsilon_{\rm f}\right)u\right] \\ &\times\left\langle \exp\left\{-\sum_{\rm s} g_{1s}\left[b_{1s}(u)-b_{1s}^{\dagger}(u)\right]\right\}\exp\left\{\sum_{\rm s} g_{1s}\left[b_{1s}-b_{1s}^{\dagger}\right]\right\}\right\rangle \\ &\times\left\langle \exp\left\{\sum_{\rm s} g_{2s}\left[b_{2s}(u)-b_{2s}^{\dagger}(u)\right]\right\}\exp\left\{-\sum_{\rm s} g_{2s}\left[b_{2s}-b_{2s}^{\dagger}\right]\right\}\right\rangle. \end{aligned} \tag{13.65}$$

Next, we exploit the identity

$$\begin{aligned} &\left\langle \exp\left\{-g_{is}\left[b_{is}(u)-b_{is}^{\dagger}(u)\right]\right\}\exp\left\{g_{is}\left[b_{is}-b_{is}^{\dagger}\right]\right\}\right\rangle = \\ &\exp\left\{-g_{is}^2\left[2n_{is}+1-(n_{is}+1)e^{-i\omega_{is}u}-n_{is}e^{i\omega_{is}u}\right]\right\}, \end{aligned} \tag{13.66}$$

where g_{is}, n_{is}, and $\hbar\omega_{is}$ are, respectively, the electron–phonon coupling constant, the phonon occupation number and the phonon energy for the sth phonon mode at site i. We finally obtain the transition rate as [69–71]

$$\begin{aligned} W = &\frac{t^2}{\hbar^2}\int_{-\infty}^{\infty} du \exp\left[\frac{i}{\hbar}\left(\varepsilon_{\rm i}-\varepsilon_{\rm f}\right)u\right]\times \\ &\prod_{\rm s}\exp\left\{-g_{1s}^2\left[2n_{1s}+1-(n_{1s}+1)e^{-i\omega_{1s}u}-n_{1s}e^{i\omega_{1s}u}\right]\right\}\times \\ &\prod_{\rm s}\exp\left\{-g_{2s}^2\left[2n_{2s}+1-(n_{2s}+1)e^{-i\omega_{2s}u}-n_{2s}e^{i\omega_{2s}u}\right]\right\}. \end{aligned} \tag{13.67}$$

Equation 13.67 can be used directly to calculate the transition rate from one dot to another. In some cases, a simpler Marcus formula is often used instead of Equation 13.67. In what follows, we introduce additional approximations that lead to the reduction of Equation 13.67 to the Marcus formula. However, we alert the reader that the conditions for the validity of the Marcus formula should be carefully checked before that formula is used.

A detailed analysis of the exponential terms in Equation 13.67 leads to the conclusion that when electron–phonon interaction is strong ($\sum_s g_{is}^2 \gg 1$) exponential terms decay very quickly when $|\omega_{is}u| \gg 1$. Therefore, only the terms with $|\omega_{is}u| \ll 1$ contribute significantly to the integral. By keeping terms up to the quadratic in the Taylor expansion $e^{i\omega_{is}t} = 1 + i\omega_{is}t - \frac{1}{2}\omega_{is}^2 t^2$, one obtains

$$\begin{aligned} &\prod_{\rm s}\exp\left\{-g_{1s}^2\left[2n_{1s}+1-(n_{1s}+1)e^{-i\omega_{1s}u}-n_{1s}e^{i\omega_{1s}u}\right]\right\}\approx \\ &\prod_{\rm s}\exp\left(i\omega_{1s}g_{1s}^2 u\right)\exp\left[-\frac{1}{2}g_{1s}^2\omega_{1s}^2 u^2\left(2n_{1s}+1\right)\right]. \end{aligned} \tag{13.68}$$

When the temperature is high, $k_B T \gg \hbar\omega_{is}$, the $\left(2n_{is}+1\right)$ term reduces to $\frac{2k_B T}{\hbar\omega_{is}}$. Using that approximation one obtains

$$W \approx \frac{t^2}{\hbar^2}\int_{-\infty}^{\infty} du \exp\left[\frac{i}{\hbar}\left(\varepsilon_{\rm f}-\varepsilon_{\rm i}+\lambda_1+\lambda_2\right)u\right]\exp\left[-\frac{k_B T}{\hbar^2}\left(\lambda_1+\lambda_2\right)u^2\right], \tag{13.69}$$

where $\lambda_i = \sum_s \hbar\omega_{is}\, g_{is}^2$. The solution of the last integral gives the Marcus formula

$$W \approx \frac{t^2}{\hbar}\sqrt{\frac{\pi}{k_B T\lambda}}\exp\left[-\frac{\left(\lambda+\varepsilon_f-\varepsilon_i\right)^2}{4k_B T\lambda}\right], \tag{13.70}$$

where $\lambda = \lambda_1 + \lambda_2$.

13.8 Practical Examples: Cd-Based Chalcogenide QDs

13.8.1 Comparison with Experiment: Absorption Edge Wavelength of CdSe/CdTe Type II CQDs

In this section, we apply the theoretical approach described in previous sections to calculate the absorption wavelength in CdSe/CdTe core/shell QDs. We analyze the influence of CdSe/CdTe core/shell colloidal QD (CQD) morphology on the variation of the first exciton peak in the absorption spectra, which corresponds to excitation of the $1S^e_{1/2}1S^h_{3/2}$ state. As a first step, we validate our theoretical methodology, Sections 13.3–13.6, based on the combination of $\mathbf{k}\cdot\mathbf{p}$ and CI Hamiltonians, against available experimental and theoretical results on CdSe and CdTe core-only CQD, as they are the constituent materials of our core/shell structure. Figure 13.3 shows the variation of the $1S^e_{1/2}1S^h_{3/2}$ exciton energy as a function of CdSe and CdTe QD radius, a_c, respectively. Solid symbols represent experimental results from several independent measurements from different groups and the line is an empirical inverse polynomial fitting curve, $E_X^{1S_{1/2}1S_{3/2}}(a_c) = E_g^{\text{bulk}} + (Aa_c^2 + Ba_c + C)^{-1}$, to this experimental data proposed by de Mello Donega in Ref. [72], where A, B, and C are fitting parameters. Results of $1S^e_{1/2}1S^h_{3/2}$ exciton energies predicted by our calculation are shown on the same figures and exhibit excellent agreement with the experimental measurements [73–82].

Having validated the methodology for core-only CQD structures, we now employ our method to describe CdSe/CdTe core/shell type-II CQD structures. The remaining question when modeling the system of core/shell structures is the parametrization of the valence band offset (VBO) between core and shell materials, in our case CdSe and CdTe, respectively. From our previous analysis [58] of the variation of $1S^e_{1/2}1S^h_{3/2}$ and $1S^e_{1/2}2S^h_{3/2}$ excitonic energies with the shell thickness in CdTe/CdSe core/shell CQD, we have estimated the $\text{VBO}_{\text{CdTe/CdSe}} = 0.4$ eV. We have assumed the same VBO for inverse structures, i.e., for CdSe/CdTe core/shell CQD considered here.

The $1S^{(e)}_{1/2}nS^{(h)}_{3/2}$ $(n = 1, 2)$ states are the two lowest-energy excitons observed in the absorption spectra of colloidal CdTe/CdSe NCs [23,83,84], making them the most important for understanding the near band-edge absorption characteristics of such nanoparticles. Figure 13.3 shows the $1S^{(e)}_{1/2}1S^{(h)}_{3/2}$ and $1S^{(e)}_{1/2}2S^{(h)}_{3/2}$ exciton energies (solid lines) calculated by the CI as a function of shell thickness for CdTe/CdSe QDs with (a) $a_c = 1.7$ nm, (b) $a_c = 1.72$ nm, (c) $a_c = 1.75$ nm, and (d) $a_c = 1.95$ nm. Dashed lines show upper and lower limits on the exciton energies resulting from an uncertainty of 1 monolayer (ML) in the displayed core radii ($\sim \pm 0.3$ nm). Filled circles show exciton energies taken from the first and second absorption peak positions in absorption spectra measured by (a) Gong et al., [84], (b) Ma et al., [85], (c) Cai et al. [86], and (d) Oron et al. [23]. We see good quantitative agreement between the calculated exciton energies and the experimental data, with the data lying in the channels defined by an uncertainty of ± 1 ML width in the core size. It should be noted that the results of Oron et al. [23] were obtained on zinc-blende NC structures, in addition to those of Cai et al. [86] The papers by Gong et al. [84] and Ma et al. [85] do not explicitly state the crystal structures of the core/shell nanoparticles, although Ma et al. [85] note that their core/shell NCs gave very similar absorption and photoluminescence spectra to those of Cai and coworkers [86]. Our calculations accurately reproduce the 0.25 eV energy separation between the $1S^{(e)}_{1/2}1S^{(h)}_{3/2}$ and $1S^{(e)}_{1/2}2S^{(h)}_{3/2}$ excitons that is nearly independent of shell thickness [84]. This constant energy separation is

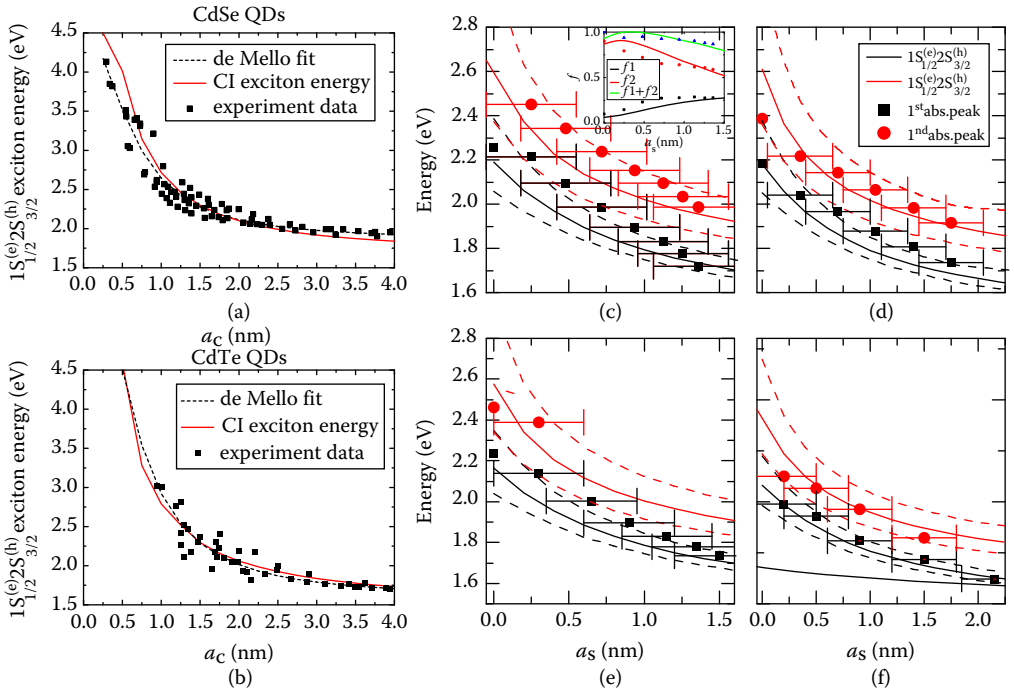

FIGURE 13.3 Variation of the $1S^{e}_{1/2}1S^{h}_{3/2}$ exciton energy as a function of colloquial quantum dot (CQD) core radius, a_c, in (a) CdSe and (b) CdTe CQDs: experimental results (square symbols), de Mello Donega's empirical fit to the experimental data (dotted line) [72], and theoretical model (solid line). Calculated (lines) energies of the $1S^{(e)}_{1/2}nS^{(h)}_{3/2}$ ($n = 1, 2$) excitons for CdTe/CdSe QDs with shell thickness a_s and core radius of (c) $a_c = 1.7$ nm, (d) $a_c = 1.72$ nm, (e) $a_c = 1.75$ nm, and (f) $a_c = 1.95$ nm; experimental data taken from Refs. [84], [85],[86], and [23] are shown as filled symbols. Error bars represent an uncertainty of 1 mono-layer (ML) ($\approx \pm 0.3$ nm) in the shell thickness. Dashed lines show upper and lower limits on the exciton energies resulting from an uncertainty of 1 ML in the nominal core radii.

characteristic of changing electron confinement but approximately constant hole confinement in the h/e heterostructure. We also find good agreement between the oscillator strength obtained by Gong et al. [84] from the absorption spectra and our calculations. Calculating the oscillator strength f_n of the $1S^{(e)}_{1/2}nS^{(h)}_{3/2}$ ($n = 1, 2$) excitons as $f_n = 2P^2_X/m_0E_X$ we find that $f_1 + f_2 \sim$ constant (inset Figure 13.3), confirming the validity of relevant excitonic wave functions too.

In contrast to core-only CQDs, the size dependence of $1S^{e}_{1/2}1S^{h}_{3/2}$ exciton energies in core/shell CQDs is more complex and it is not possible to capture its trend by a simple polynomial fitting curve. Figure 13.5 shows the shell thickness, a_s, dependence of the $1S^{e}_{1/2}1S^{h}_{3/2}$ absorption wavelengths for CdSe/CdTe core/shell CQDs for different core radii, a_c, ranging from a_c=1.5 nm to 2.5 nm. Also shown in Figure 13.5 are experimental data for a CdSe core of 1.7-nm radius, as measured by transmission electron microscopy, and different CdTe shell thicknesses [87,88]. Good agreement is seen between the calculated energy of the $1S^{e}_{1/2}1S^{h}_{3/2}$ exciton and the spectral position observed for the first absorption peak, further validating our theoretical approach.

To illustrate our theoretical model outlined in Section 13.6, in Figure 13.4, we show the excitonic spectra of an $a_c = 2$ nm CdSe core-only QD ($a_s = 0$ case) and a set of CdSe/CdTe core/shell QDs with shell thicknesses varying from $a_s = 0.5$ to 3.0 nm. By changing the shell thickness from $a_s = 0$ (no shell) to $a_s = 3$ nm the excitonic ground state energy, E_{X0}, is changed from 2.30 eV to 1.49 eV. It is also interesting to note that for certain shell thicknesses, i.e., between $a_s = 0.5$ and 1 nm, the character of the ground state exciton changes from $1S_{1/2}1S_{3/2}$ to $1S_{1/2}1P_{3/2}$. It can be observed that the density of excitonic states

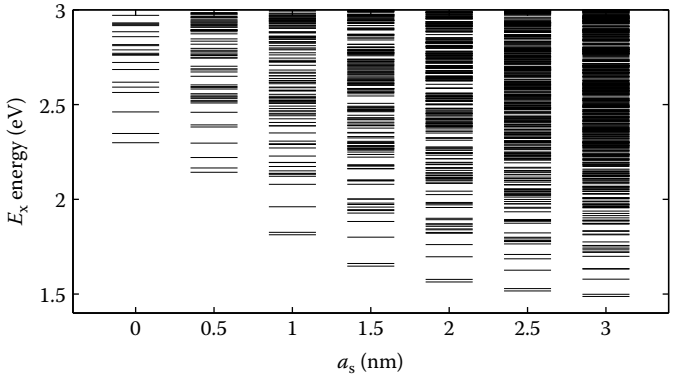

FIGURE 13.4 Excitonic spectra of CdSe and CdSe/CdTe quantum dots (QDs) as a function of the shell thickness, a_s.

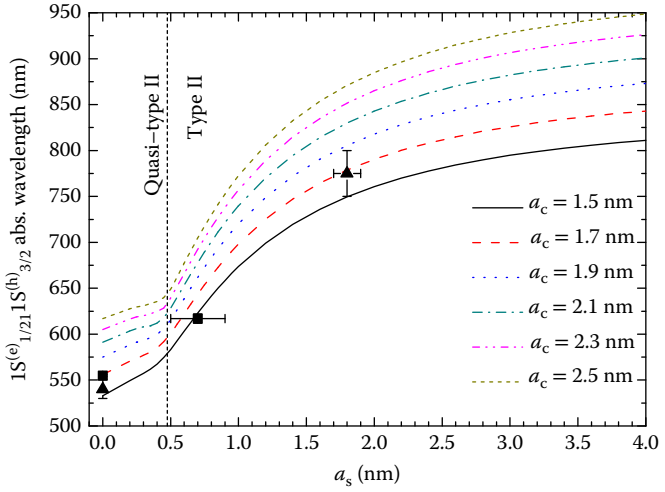

FIGURE 13.5 Core radius, a_c, and shell thickness, a_s, dependence of the $1S^e_{1/2}1S^h_{3/2}$ exciton peak in the absorption spectra of CdSe/CdTe core/shell quantum dots (QDs). Also shown for comparison are experimental data from Ref. [87] (squares) and Ref. [88] (triangles).

increases with shell thickness: this suggests that phononic gaps are more likely to be found in core-only or core/shell structures with thin shells [16].

In Figure 13.5, we can clearly distinguish between two trends in the $1S^e_{1/2}1S^h_{3/2}$ absorption wavelengths in the region of shell thicknesses <0.5 nm and for shell thicknesses >0.5 nm. For the QDs with shells of thickness approximately or less than 0.5 nm, we can see that (1) the $1S^e_{1/2}1S^h_{3/2}$ absorption wavelengths are approximately linearly dependent on a_s, suggesting strong influence of the dielectric confinement on $1S^e_{1/2}1S^h_{3/2}$ excitons and their wavelengths in this region and (2) as the core size, a_c, increases the gradient of the $1S^e_{1/2}1S^h_{3/2}$ absorption wavelength versus a_s curves decreases. We explain such nonmonotonic behavior in terms of the changing localization regime of the $1s^h_{3/2}$ SP hole, in the correlated $1S^e_{1/2}1S^h_{3/2}$ exciton, as the QD dimensions change [58]. The lower gradient of the curves for shell widths <0.5 nm is due to the fact that the $1s^h_{3/2}$ hole is in the delocalized regime, i.e., its probability density is spread over the whole heterostructure and its energy is mainly determined by the global confinement provided by the QD potential well of radius $a_c + a_s$. In this regime, the size dependence of the hole confinement is closer to that

of one confined in a core-only QD. For the QDs with shells of width greater than 0.5 nm, we can see that the $1S^e_{1/2}1S^h_{3/2}$ absorption wavelengths again acquire a trend that can be described by a quadratic polynomial function. As a_s increases, for a particular core size, the $1s^h_{3/2}$ SP hole localizes in the shell fully so the absorption wavelength is again more strongly affected by the effect of shell thickness on the confinement energy of $1s^h_{3/2}$. This behavior is closer to the SCR in a core/shell heterostructure. Maps of the absorption wavelength against QD core/shell dimensions as depicted in Figure 13.5 can only be obtained numerically and should be of use to experimentalists studying such systems.

13.8.2 Radiative Lifetimes in CdSe/CdTe CQDs

Next, we analyze the radiative lifetimes in core/shell CdSe/CdTe CQDs [17]. In order to assess the variation of the radiative times, $\tau_{\text{rad.}}$, with the core size, a_c, and the shell thickness, a_s, we use the following expression:

$$\frac{1}{\tau_{\text{rad.}}} = \frac{1}{3}\frac{F^2\bar{n}e^2}{\pi\epsilon_0 c^3\hbar^4}E_X\frac{1}{d_X}\sum_{i=1}^{d_X}|\mathbf{P}_{X_i}|^2, \tag{13.71}$$

where $\bar{n}$ is the refractive index of the colloidal material, e is the electron charge, ϵ_0 is the permittivity of free space, m_0 is the electron rest mass, c is the speed of light, and $\hbar$ is the reduced Planck constant. In the expression above, $|\mathbf{P}_X|$ is the modulus of excitonic dipole matrix element, given in (eVÅ), and E_X is the excitonic energy, both obtained from the CI calculation, and d_X is the integer number representing the degree of degeneracy of a particular excitonic state. For the $1S^e_{1/2}1S^h_{3/2}$ exciton considered below the degeneracy is $d_X = 8$. The expression for dielectric screening of spherical core/shell QDs is given as $F = 9\epsilon_s\epsilon_{\text{coll.}}/(\epsilon_s\epsilon_a + 2\epsilon_{\text{coll.}}\epsilon_b)$, where $\epsilon_a = \epsilon_c(3 - 2\Omega_s/\Omega_{qd}) + 2\epsilon_s\Omega_s/\Omega_{qd}$, $\epsilon_b = \epsilon_c\Omega_s/\Omega_{qd} + \epsilon_s(3 - \Omega_s/\Omega_{qd})$, and Ω_{qd} and Ω_s are the volume of QD and shell, respectively [89].

Figure 13.6 shows the mean radiative lifetime of the $1S^e_{1/2}1S^h_{3/2}$ exciton in CdSe/CdTe CQDs as a function of shell thickness a_s for several different core radii. The most noticeable feature is the sudden increase in $\tau_{\text{rad.}}$ at around $a_s \sim 0.5$ nm. In comparison to core-only CQDs, in which the value of $1S^e_{1/2}1S^h_{3/2}$ excitonic radiative time τ_{rad} changes relatively weakly with the CQD size [35,90], in core/shell structures it is

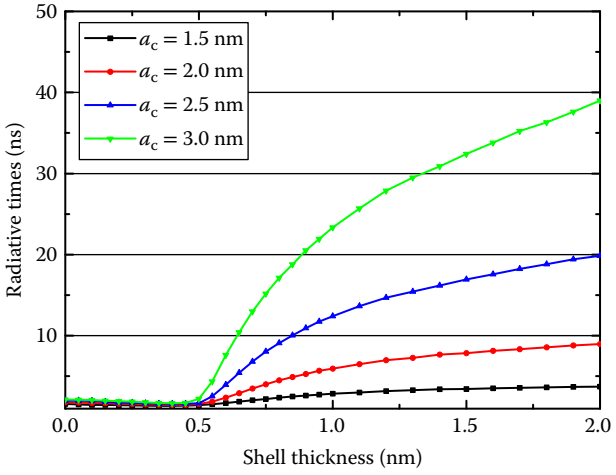

FIGURE 13.6 Plot of the $1S^e_{1/2}1S^h_{3/2}$ exciton radiative lifetime as a function of the shell thickness for different cores: $a_c = 1.5, 2.0, 2.5$, and 3.0 nm in CdSe/CdTe colloquial quantum dots (CQDs).

possible to increase this radiative time over one order of magnitude with suitable change in the core and shell sizes. Several observations regarding the trend in τ_{rad} in CdSe/CdTe core/shell type II CQDs can be made here: (1) again as in the case of the absorption wavelength, shown in Figure 13.5, we can distinguish clearly between two different regions of the shell thicknesses $a_s < 0.5$ nm and $a_s > 0.5$ nm. In the region of $a_s < 0.5$ nm, the whole CQD system behaves similar to core-only CQDs, since both electron and hole correlated charge densities are either largely confined in the core region or just starting to delocalize over the whole CQD structure. In this region, the optical dipole matrix element, $|\mathbf{P}_X|$, is strong while the magnitude of both $|\mathbf{P}_X|$ and E_X changes very little with the overall sizes of QDs, which explains the almost constant τ_{rad}. (2) As shell thickness increases, and in particular beyond $a_s > 1$ nm, the hole states become strongly confined in the shell region. Here, dielectric confinement does not have enough strength to overcome the confinement imposed by the type II aligned VB edge of CdTe, so that the exciton reaches the SCR. Consequently, the electron–hole wave function overlap is dramatically reduced and E_X is mainly now determined by the variation of the $1s_{3/2}^{(h)}$ hole confinement with a_s. For $a_s > 1$ nm, $\tau_{rad.}$ continues to increase, but more slowly, reaching values that are about one order of magnitude greater than those for core-only CQDs. For all shell thicknesses, a monotonic trend to larger $\tau_{rad.}$ with increasing core size is also observed. It is interesting to observe that for CdSe/CdTe type II CQDs, the value of $\tau_{rad.}$ increases over "only" one order of magnitude, while in other systems, such as epitaxially grown type II InAs/GaAs/GaAsSb structure, the radiative times can increase over three orders of magnitude [91,92]. This difference in behavior is attributed to the absence of significant dielectric confinement in epitaxially grown QDs, compared to CQDs.

13.8.3 Correlation Energy

The correlation energy of the exciton can be defined as

$$E_{corr} = E_{X,CI} - E_X, \tag{13.72}$$

where E_X is the exciton energy calculated according to first-order perturbation theory inside the strong confinement approximation for the exciton wave function [42,61,83,93], and $E_{X,CI}$ is the excitonic energy calculated using the CI method.

13.8.3.1 CdTe/CdSe QD: Effect of Electron Shell Localization

In Figure 13.7a, we present E_{corr} for the $1S_{1/2}^{(e)}1S_{3/2}^{(h)}$ exciton as a function of core radius for fixed shell $a_s = 2$ nm CdTe/CdSe QDs. We see that in the presence of dielectric confinement $|E_{corr}| \lesssim 20$ meV and that E_{corr} exhibits at least one minimum as a function of a_c in the $\epsilon = \text{const.}$ and $\epsilon = \epsilon(r)$ case. $|E_{corr}|$ is up to four times greater in the presence of dielectric confinement, $\epsilon = \epsilon(r)$, compared to the $\epsilon = \text{const.}$ case. This result highlights the importance of a proper treatment of the dielectric environment in such nanostructures. The minimum in E_{corr} for the $1S_{1/2}^{(e)}1S_{3/2}^{(h)}$ exciton, Figure 13.7a, is a consequence of two competing effects: proximity of the self-polarization potential peak, which tends to reduce the electron–hole separation and the effect of the type-II confinement profile, which tends to separate the carriers as a_c increases.

13.8.3.2 CdSe/CdTe QD: Effect of Hole Shell Localization

In Figure 13.7b, we present E_{corr} for the $1S_{1/2}^{(e)}1S_{3/2}^{(h)}$ exciton as a function of shell thickness for an $a_c = 3.5$ nm fixed core in CdSe/CdTe QDs. We observe the largest size correlation energy of the considered excitons for the $1S_{1/2}^{(e)}1S_{3/2}^{(h)}$ state in CdSe/CdTe QDs when $\epsilon = \epsilon(r)$, with E_{corr} reaching -62 meV for an $a_c = 3.5$ nm, $a_s = 0.6$ nm QD, Figure 13.7b. This value is more than six times larger than the corresponding value for the $\epsilon = \text{const.}$ case, highlighting the particularly strong effect of the dielectric environment on this exciton. It can be observed from Figure 13.7b that the effect of dielectric mismatch on E_{corr} is strongest

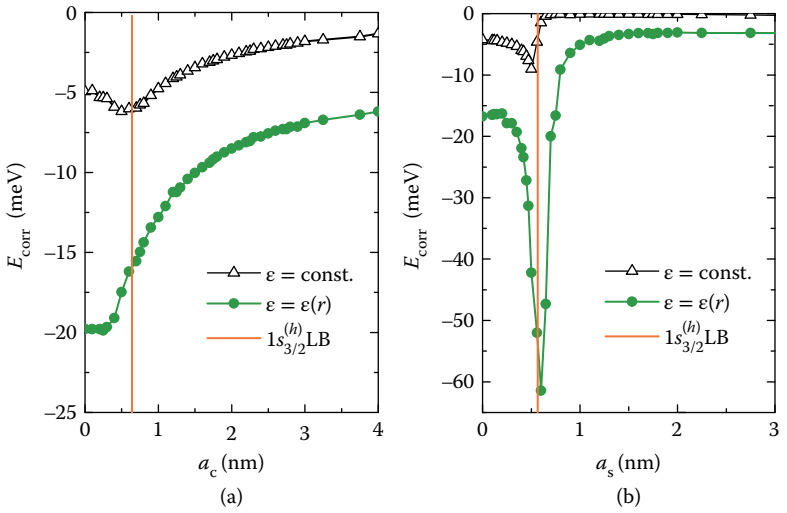

FIGURE 13.7 Correlation energy, E_{corr}, of $1S^{(e)}_{1/2}1S^{(h)}_{3/2}$ excitons in: (a) $a_s = 2.0$ nm CdTe/CdSe quantum dots (QDs), and (b) $a_c = 3.5$ nm CdSe/CdTe QDs. Cases of $\epsilon = \text{const.}$ and $\epsilon = \epsilon(r)$ are represented by open and solid symbols, respectively. Vertical line points to the localization/delocalization boundary (LB) of $1s_{3/2}$ hole state.

in the vicinity of the $1s^{(h)}_{3/2}$ localization boundary (LB), i.e., once the SP hole starts to move significantly toward the shell region and becomes delocalized over the whole QD.

$1S^{(e)}_{1/2}1S^{(h)}_{3/2}$ *exciton:* In the case of a core-only CdSe QD correlation causes both carriers, correlated electron and correlated hole, to move closest toward the center of the QD compared with their strongest SP character. This is purely a result of the direct interparticle Coulomb interaction. Introduction of the self-polarization potential, i.e., $\epsilon = \epsilon(r)$, further exaggerates this move of the radial probability densities (RPD) of both carriers away from the QD surface in the core-only CdSe QD, Figure 13.8a. This effect increases localization of both carriers near the center of QD, increasing overlap between them and giving correlation energies $E_{corr} = -18$ meV. For the $1S^{(e)}_{1/2}1S^{(h)}_{3/2}$ exciton in the CdSe QD, the shift in RPD is mainly due to an increase in $2s^{(e)}_{1/2}1s^{(h)}_{3/2}$ character.

To assess the effect of dielectric confinement on the correlated carriers in the CdSe QD, we consider the expectation value of the $1s$ electron (hole) radial coordinate, denoted $\langle r_{e(h)}\rangle$. When $\epsilon = \text{const.}$ (no self-polarization) we find $\langle r_h\rangle = 1.55$ nm compared to $\langle r_h\rangle = 1.44$ nm when $\epsilon = \epsilon(r)$ for the $1s^{(h)}_{3/2}$ state. In contrast, the effect of dielectric confinement moves the $1s^{(e)}_{1/2}$ electron from $\langle r_e\rangle = 2.01$ nm to $\langle r_e\rangle = 1.89$. Although the SP $1s^{(e)}_{1/2}$ RPD has significantly greater overlap with the repulsive peak in self-polarization potential near the QD surface than the SP $1s^{(h)}_{3/2}$ hole RPD, the correlated electron is shifted by dielectric confinement by almost the same distance as the correlated hole. These results reflect the larger sensitivity of the correlated hole wave function to the dielectric environment compared to the electron in the CdSe core-only QD.

In Figure 13.8b, we see that the introduction of a thin CdTe shell allows the uncorrelated hole to start to localize nearer the QD surface (at $r = a_c + a_s$), dramatically reducing its overlap with the uncorrelated electron. However, the effect of correlation is strong enough to pull the hole back toward the center, mainly due to the addition of the $1s^{(e)}_{1/2}2s^{(h)}_{3/2}$ electron–hole pair (EHP) character to the exciton wave function. We see that the introduction of dielectric confinement exaggerates this move of the carriers further away from the QD surface compared to the $\epsilon = \text{const.}$ case [58]. The effect of dielectric confinement is particularly strong in this case because $\langle r_h\rangle$ for the $1s^{(h)}_{3/2}$ state is close to the value of QD's outermost radius, $a_c + a_s$.

The close proximity of the hole to the QD surface reduces the distance $\xi = \langle r_h^{\mathrm{QD}} \rangle - \langle r_h^{\mathrm{ind.}} \rangle$ between the hole in the QD and its induced charge in the colloid, increasing the Coulombic repulsion between them which scales as $\propto 1/\xi$. Such repulsion, in addition to the strong effect of correlations as discussed above, causes the hole to be pushed back toward the center of the QD, thereby dramatically increasing overlap with the correlated electron wave function. The presence of dielectric confinement means the exciton wave function is an almost equal superposition of the $1s_{1/2}^{(e)} ns_{3/2}^{(h)}$ ($n = 1, 2$) states, with characters $|c_1|^2 = 0.449$ and $|c_2|^2 = 0.458$. For comparison, when $\epsilon = \text{const.}$ the $1s_{1/2}^{(e)} 2s_{3/2}^{(h)}$ character amounts to only $|c_2|^2 = 0.019$. The much stronger configuration mixing in the dielectric confinement case allows E_{corr} to reach ~ -62 meV, compared to -9 meV without dielectric confinement.

Further increase of the CdSe/CdTe QD shell thickness to $a_s = 1$ nm, Figure 13.8c, allows the SP hole to fully localize in the shell, while the SP electron stays in the core, reaching the type-II localization limit. The carriers effectively enter the strong confinement regime (SCR) in which the Coulomb effects are overridden

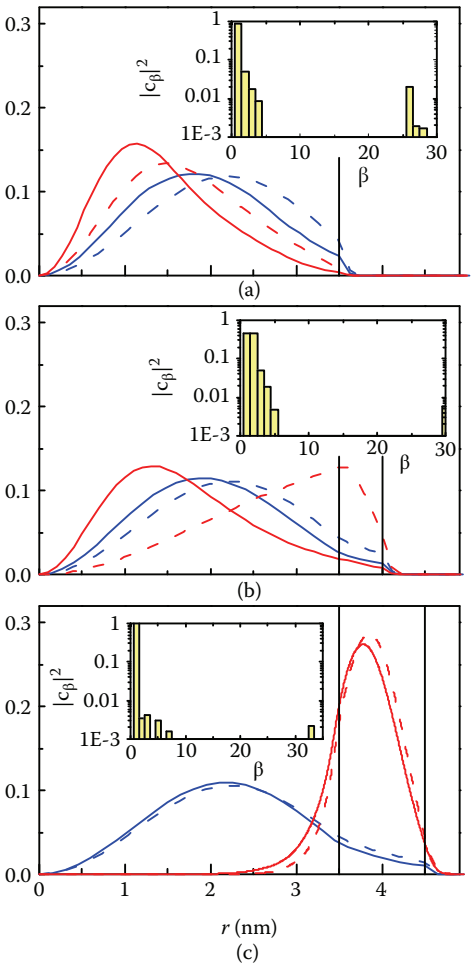

FIGURE 13.8 Solid (dashed) lines represent radial probability densities (RPDs) of the $1S_{1/2}^{(e)} 1S_{3/2}^{(h)}$ exciton (single particle $1s_{1/2}^{(e)} 1s_{3/2}^{(h)}$ electron–hole pair [EHP]) of CdSe/CdTe quantum dots (QDs) with $a_c = 3.5$ nm and (a) $a_s = 0$, (b) $a_s = 0.5$ nm, and (c) $a_s = 1$ nm. Vertical lines denote the boundaries between the core, shell, and external medium. Insets are bar charts of single EHP characters, $|c_\beta|^2$, in the exciton.

by the effects of the type-II spatial confinement. In the SCR the charge density of a correlated excitonic state is very similar to the charge density of the uncorrelated product of SP states, $\rho_X^{e(h)} \simeq \rho_{SP}^{e(h)}$, and the effect of correlations is lost. Again, E_{corr} is only nonzero when the hole is delocalized; once it localizes in the shell, the effect of VBM confinement overrides the interparticle Coulomb attraction.

Overall, we find that dielectric confinement affects the correlated hole density more than the correlated electron density for two reasons. First, the larger effective mass and deeper potential well experienced by SP hole states compared to electron states allows the former to localize more fully in the shell, closer to the peak in $V_s(r)$ at $r \simeq a_c + a_s$. Second, the smaller energy spacing between the hole SP basis states (i.e., the larger density of hole states) compared to electron SP basis states means that the resulting correlated hole density has more "degrees of freedom" to adjust to the effects of dielectric confinement.

13.8.4 Absorption Spectra of Equivalent CdTe and CdTe/CdSe QD

In Figure 13.9, we compare the exciton dipole spectrum of (a) an $a_c = 2$ nm, $a_s = 1$ nm CdTe/CdSe QD and (b) an "equivalent" $a_c = 3.8$ nm CdTe QD calculated using the CI Hamiltonian. The radius of the CdTe core-only QD is chosen such that the absorption wavelength of its ground-state exciton is the same as that of the CdTe/CdSe core/shell QD. The energy gap between the $1S_{1/2}^{(e)}1S_{3/2}^{(h)}$ and $1S_{1/2}^{(e)}2S_{3/2}^{(h)}$ excitons is increased from $\sim$ 0.1 eV in the core-only QD to $\sim$ 0.18 eV in the core/shell structure. The size of the ground-state exciton optical dipole matrix element is reduced by about 30%, from $0.15P_0^2$ (where P_0 is the bulk optical dipole matrix element) in the core-only QD to $0.11P_0^2$ in the core/shell QD due to reduced electron–hole overlap and electron delocalization (the $a_c = 2$ nm, $a_s = 1$ nm QD lies in the quasi-type-II regime). As expected, the CdTe QD shows slightly stronger absorption than its CdTe/CdSe QD counterpart (see Figure 13.9c), mainly due to the better overall overlap between electron and hole states in core-only QDs compared to type-II structures. However, we note that type-II QDs overall have superior absorption

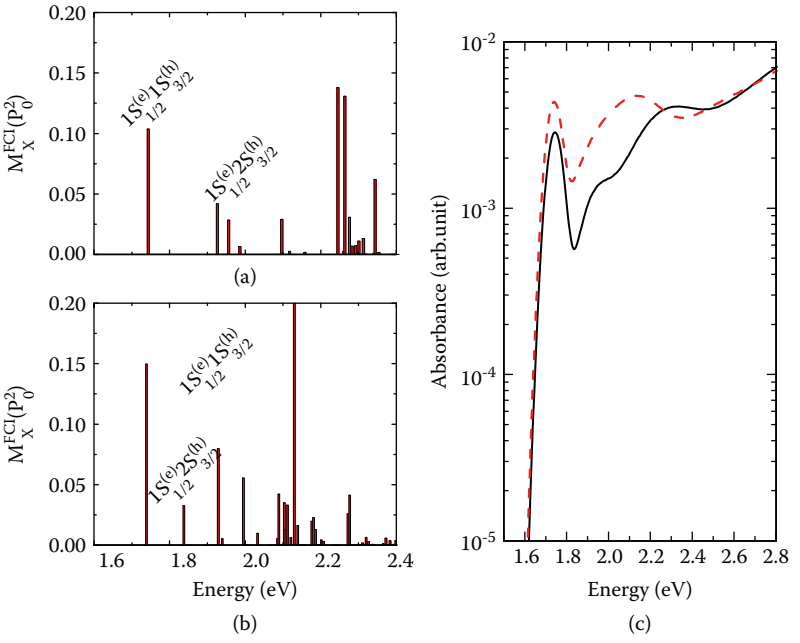

FIGURE 13.9 The exciton dipole spectrum of (a) an $a_c = 2$ nm, $a_s = 1$ nm CdTe/CdSe quantum dot (QD) and (b) an $a_c = 3.8$ nm CdTe QD. (c) The absorption spectra corresponding to the dipole spectra shown in (a) and (b) are shown as solid and dashed lines, respectively.

properties compared with core-only QDs for the important application area of QD-sensitized solar cells. For example the type-II band alignment allows the band-edge absorption to be red-shifted compared to the core-only QD—this is often desirable because the optimum energy for exploitation of the solar spectrum is ~1.35 eV. Type-II QDs also allow greater absorption ranges to be achieved compared to core-only QDs since the absorption edges are not limited by the energy gap of the underlying bulk materials [94].

13.8.5 Charge Carrier Mobility in CdSe QD Arrays

In this section, we evaluate the charge carrier mobility in QD arrays. Good charge carrier mobility is essential for applications of QDs in solar cells since photogenerated carriers have to reach the contacts to form the current in external circuit.

We consider a three-dimensional lattice of CdSe QDs and calculate the temperature dependence of electron mobility in such a system. For hopping transport in a cubic lattice of QDs, the mobility is related to hopping rate between two neighboring dots via $\mu = \frac{ea^2}{k_B T} W$, where a is the distance between centers of two neighboring dots (lattice constant of a QD supercrystal) and W is the transition rate between the dots that can be evaluated using either Equation 13.67 or 13.70 from Section 13.7. If all dots are identical the mobility reads

$$\mu = \int_{-\infty}^{\infty} dt \frac{ea^2 t^2}{\hbar^2 k_B T} \prod_s \exp\left\{-2g_s^2 \left[2n_s + 1 - (n_s + 1)e^{-i\omega_s t} - n_s e^{i\omega_s t}\right]\right\} \tag{13.73}$$

and takes the form $\mu = \int_{-\infty}^{\infty} dt\, h(t)$. By exploiting the fact that $h(t) = h(-t)^*$, the last integral reduces to $\mu = \int_0^{\infty} dt\, f(t)$, where $f(t) = 2\,\mathrm{Re}\, h(t)$.

We consider two different dots. For dot 1, the diameter is $D = 3.6$ nm, relevant phonon energies and electron–phonon coupling constants are $\hbar\omega_s(1-6) = \{3.78, 8.10, 1.46, 2.87, 24.0, 24.0\}$ meV, $g_{1-6}\hbar\omega_s(1-6) = \{2.82, 0.14, 0.78, 0.52, 20.33, 3.52\}$ meV, the supercrystal lattice constant is $a = 5.0$ nm, and the electronic coupling parameter is $t = 4.0$ meV. For dot 2, the relevant parameters are $D = 7.0$ nm, $\hbar\omega_s(1-6) = \{1.95, 4.16, 0.75, 1.48, 24.0, 24.0\}$ meV, $g_{1-6}\hbar\omega_s(1-6) = \{1.06, 0.14, 0.40, 0.27, 14.58, 2.52\}$ meV, the supercrystal lattice constant is $a = 8.0$nm, the electronic coupling parameter is $t = 2.5$meV.

One might expect that it is difficult to numerically integrate the function $f(t)$ because of oscillatory terms of the form $\exp(i\omega_s t)$. However, in the strong electron–phonon coupling regime this function is smooth and decays with t very quickly. Its form for dot 1 at temperatures of $T = 200$ K and $T = 300$ K is presented in Figure 13.10.

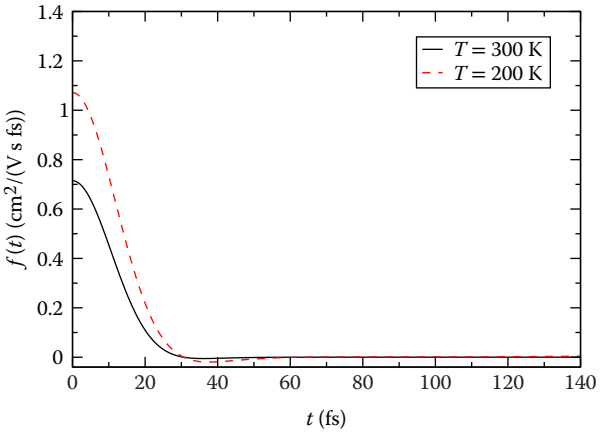

FIGURE 13.10 The function $f(t)$ for dot 1 at temperatures of $T = 200$ K and $T = 300$ K.

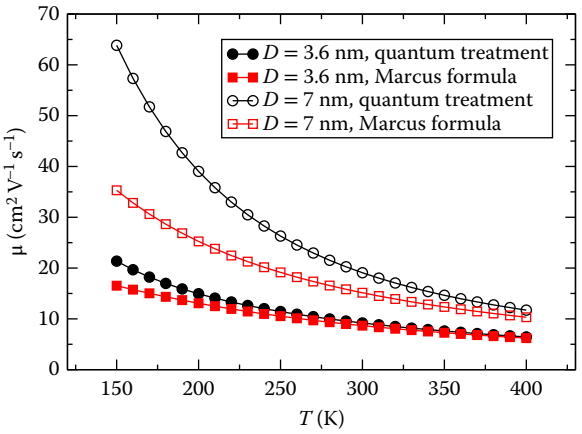

FIGURE 13.11 The dependence of mobility, μ, on temperature, T, for dots 1 and 2 with diameters $D = 3.6$ and 7 nm, respectively, calculated using Equations 13.70 and 13.67.

Temperature dependence of the mobility for dots 1 and 2 is presented in Figure 13.11. It has been checked that the assumption of strong electron–phonon coupling is valid in this temperature range [95]. The mobility was calculated using both the fully quantum treatment from Equation 13.67 and the Marcus formula given by Equation 13.70. We find that for this system in the relevant temperature range the mobility decreases with increasing temperature. This temperature dependence is often observed in the case of band transport. However, one should note that the charge transport in our case takes place by carrier hopping between neighboring dots. This suggests that the measurement of temperature dependence of mobility is not sufficient to identify the charge carrier transport regime. By comparing the results obtained using fully quantum treatment and Marcus formula, we find that, as expected, the agreement between these approaches is the best at high temperatures.

13.9 Conclusion

In conclusion, we have reviewed the methodology for calculation of SP, (multi)excitonic states, and charge transport in colloidal QDs. The methodology was then applied to analyze the optical absorption wavelength, radiative lifetimes, and correlation energy in CdSe/CdTe core/shell QDs, as well as charge carrier mobility in CdSe QD arrays.

Acknowledgments

Stanko Tomić acknowledges the EPSRC UK grant "Enhanced multiple exciton generation in colloidal quantum dots" (EP/K008587/1) for financial support. Nenad Vukmirović is supported by the Ministry of Education, Science, and Technological Development of the Republic of Serbia under project ON171017, and by the European Commission under H2020 project VI-SEEM, Grant No. 675121. We would also like to acknowledge the contribution of the COST Action MP1406 "MultiscaleSolar," and we wish to thank David J. Binks, Jacek M. Miloszewski, Nikola Prodanović, and Edward J. Tyrrell for useful discussions.

References

1. B. N. Pal, Y. Ghosh, S. Brovelli, R. Laocharoensuk, V. I. Klimov, J. A. Hollingsworth, and H. Htoon. Giant CdSe/CdS core/shell nanocrystal quantum dots as efficient electroluminescent materials: Strong influence of shell thickness on light-emitting diode performance. *Nano Lett.* **12**, 331 (2012).

2. P. V. Kamat. Quantum dot solar cells. Semiconductor nanocrystals as light harvesters. *J. Phys. Chem. C* **112**, 18737 (2008).
3. V. Sukhovatkin, S. Hinds, L. Brzozowski, and E. H. Sargent. Colloidal quantum-dot photodetectors exploiting multiexciton generation. *Science* **324**, 1542 (2009).
4. G. Konstantatos, and E. H. Sargent. Nanostructured materials for photon detection *Nat. Nanotechnol.* **5**, 391 (2010).
5. J.-M. Caruge, Y. Chan, V. Sundar, H. J. Eisler, and M. G. Bawendi. Transient photoluminescence and simultaneous amplified spontaneous emission from multiexciton states in CdSe quantum dots. *Phys. Rev. B* **70**, 085316 (2004).
6. V. I. Klimov, S. A. Ivanov, J. Nanda, M. Achermannn, I. Bezel, J. A. McGuire, and A. Piryatinski. Single-exciton optical gain in semiconductor nanocrystals. *Nature* **447**, 441 (2007).
7. A. Shabaev, A. L. Efros, and A. J. Nozik. Multiexciton generation by a single photon in nanocrystals. *Nano Lett.* **6**, 2856 (2006).
8. O. E. Semonin, J. M. Luther, S. Choi, H.-Y. Chen, J. Gao, A. J. Nozik, and M. C. Beard. Peak external photocurrent quantum efficiency exceeding 100 in a quantum dot solar cell. *Science* **334**, 1530 (2011).
9. M. T. Trinh, L. Polak, J. M. Schins, A. J. Houtepen, R. Vaxenburg, G. I. Maikov, G. Grinbom, A. G. Midgett, J. M. Luther, M. C. Beard, A. J. Nozik, M. Bonn, E. Lifshitz, and L. D. A. Siebbeles. Anomalous independence of multiple exciton generation on different group IV–VI quantum dot architectures. *Nano Lett.* **11**, 1623 (2011).
10. S. Ithurria, M. D. Tessier, B. Mahler, R. P. S. M. Lobo, B. Dubertret, and A. L. Efros. Colloidal nanoplatelets with two-dimensional electronic structure. *Nat. Mater.* **10**, 936 (2011).
11. J. Jasieniak, M. Califano, and S. E. Watkins. Size-dependent valence and conduction band-edge energies of semiconductor nanocrystals. *ACS Nano* **5**, 5888 (2011).
12. S. Kim, B. Fisher, H.-J. Eisler, and M. Bawendi. Type-II quantum dots: CdTe/CdSe(core/shell) and CdSe/ZnTe(core/shell) heterostructures. *J. Am. Chem. Soc.* **125**, 11466 (2003).
13. J. J. Li, J. M. Tsay, X. Michalet, and S. Weiss. Wavefunction engineering: From quantum wells to near-infrared type-II colloidal quantum dots synthesized by layer-by-layer colloidal epitaxy. *Chem. Phys.* **318**, 82 (2005).
14. R. Xie, U. Kolb, J. Li, T. Basch, and A. Mews. Synthesis and characterization of highly luminescent CdSe-Core CdS/$Zn_{0.5}Cd_{0.5}$S/ZnS multishell nanocrystals. *J. Am. Chem. Soc.* **127**, 7480 (2005).
15. U. Aeberhard, R. Vaxenburg, E. Lifshitz, and S. Tomić. Fluorescence of colloidal PbSe/PbS QDs in NIR luminescent solar concentrators. *Phys. Chem. Chem. Phys.* **14**, 16223 (2012).
16. C. T. Smith, E. J. Tyrrell, M. A. Leontiadou, J. Miloszewski, T. Walsh, M. Cadirci, R. Page, P. O'Brien, D. Binks, and S. Tomić. Energy structure of CdSe/CdTe type {II}colloidal quantum dots: Do phonon bottle necks remain for thick shells? *Sol. Energ. Mater.* **158**, 160–167 (2016),
17. M. A. Leontiadou, E. J. Tyrrell, C. T. Smith, D. Espinobarro-Velazquez, R. Page, P. O'Brien, J. Miloszewski, T. Walsh, D. Binks, and S. Tomić. Influence of elevated radiative lifetime on efficiency of CdSe/CdTe type {II}colloidal quantum dot based solar cells. *Sol. Energ. Mater.* **159**, 657–663 (2017).
18. S. Kumar, M. Jones, S. Lo, and G. D. Scholes. Nanorod heterostructures showing photoinduced charge separation. *Small* **3**, 1633 (2007).
19. H. Zhong, Y. Zhou, Y. Yang, C. Yang, and Y. Li. Synthesis of Type II CdTe–CdSe nanocrystal heterostructured multiple-branched rods and their photovoltaic applications. *J. Phys. Chem. C* **111**, 6538 (2007).
20. S. Itzhakov, H. Shen, S. Buhbut, H. Lin, and D. Oron. Type-II quantum-dot-sensitized solar cell spanning the visible and near-infrared spectrum. *J. Phys. Chem. C* **117**, 22203 (2013).
21. N. McElroy, R. Page, D. Espinbarro-Valazquez, E. Lewis, S. Haigh, P. O'Brien, and D. Binks. Comparison of solar cells sensitised by CdTe/CdSe and CdSe/CdTe core/shell colloidal quantum dots with and without a CdS outer layer. *Thin Solid Films* **560**, 65 (2014).

22. S. Brovelli, R. D. Schaller, S. A. Crooker, F. Garcia-Santamaria, Y. Chen, R. Viswanatha, J. A. Hollingsworth, H. Htoon, and V. I. Klimov. Nano-engineered electron–hole exchange interaction controls exciton dynamics in core–shell semiconductor nanocrystals. *Nat. Commun.* **2**, 280 (2011).
23. D. Oron, M. Kazes, and U. Banin. Multiexcitons in type-II colloidal semiconductor quantum dots. *Phys. Rev. B* **75**, 035330 (2007).
24. P. G. McDonald, E. J. Tyrrell, J. Shumway, J. M. Smith, and I. Galbraith. Tuning biexciton binding and antibinding in core/shell quantum dots. *Phys. Rev. B* **86**, 125310 (2012).
25. S. A. Ivanov, J. Nanda, A. Piryatinski, M. Achermann, L. P. Balet, I. V. Bezel, P. O. Anikeeva, S. Tretiak, and V. I. Klimov. Light amplification using inverted core/shell nanocrystals: Towards lasing in the single-exciton regime. *J. Phys. Chem. B* **108**, 10625 (2004).
26. J. Nanda, S. A. Ivanov, M. Achermann, I. Bezel, A. Piryatinski, and V. I. Klimov. Light amplification in the single-exciton regime using exciton–exciton repulsion in type-II nanocrystal quantum dots. *J. Phys. Chem. C* **111**, 15382 (2007).
27. A. Szabo, and N. S. Ostlund. *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory* (New York, NY: Dover Publications, 1982).
28. V. A. Fonoberov, E. P. Pokatilov, and A. A. Balandin. Exciton states and optical transitions in colloidal CdS quantum dots: Shape and dielectric mismatch effects. *Phys. Rev. B* **66**, 085310 (2002).
29. E. Menéndez-Proupin, and C. Trallero-Giner. Electric-field and exciton structure in CdSe nanocrystals. *Phys. Rev. B* **69**, 125336 (2004).
30. N. Vukmirović, Ž. Gačević, Z. Ikonić, D. Indjin, P. Harrison, and V. Milanović. Electric-field and exciton structure in CdSe nanocrystals. *Semicond. Sci. Technol.* **21**, 1098 (2006).
31. N. Vukmirović, Z. Ikonić, I. Savić, D. Indjin, and P. Harrison. A microscopic model of electron transport in quantum dot infrared photodetectors. *J. Appl. Phys.* **100**, 074502 (2006).
32. N. Vukmirović, D. Indjin, Z. Ikonić, and P. Harrison. Origin of detection wavelength tuning in quantum dots-in-a-well infrared photodetectors. *Appl. Phys. Lett.* **88**, 251107 (2006).
33. N. Vukmirović, and S. Tomić. Plane wave methodology for single quantum dot electronic structure calculations. *J. Appl. Phys*. **103**, 103718 (2008).
34. N. Vukmirović, Z. Ikonić, V. D. Jovanović, D. Indjin, and P. Harrison. Optically pumped intersublevel mid-infrared lasers based on InAs/GaAs quantum dots. *IEEE J. Quant. Electron.* **41**, 1361 (2005).
35. M. Califano, A. Franceschetti, and A. Zunger. Lifetime and polarization of the radiative decay of excitons, biexcitons, and trions in CdSe nanocrystal quantum dots. *Phys. Rev. B* **75**, 115401 (2007).
36. S. Tomić, and N. Vukmirović. Excitonic and biexcitonic properties of single GaN quantum dots modeled by 8-band $\mathbf{k}\cdot\mathbf{p}$ theory and configuration-interaction method. *Phys. Rev. B* **79**, 245330 (2009).
37. M. Korkusinski, O. Voznyy, and P. Hawrylak. Fine structure and size dependence of exciton and biexciton optical spectra in CdSe nanocrystals. *Phys. Rev. B* **82**, 245304 (2010).
38. G. Allan, and C. Delerue. Tight-binding calculations of the optical properties of HgTe nanocrystals. *Phys. Rev. B* **86**, 165437 (2012).
39. A. L. Efros, and M. Rosen. The electronic structure of semiconductor nanocrystals. *Ann. Rev. Mater. Sci.* **30**, 475 (2000).
40. A. Franceschetti, A. Williamson, and A. Zunger. Addition spectra of quantum dots: The role of dielectric mismatch. *J. Phys. Chem. B* **104**, 3398 (2000).
41. J. I. Climente, M. Royo, J. L. Movilla, and J. Planelles. Strong configuration mixing due to dielectric confinement in semiconductor nanorods. *Phys. Rev. B* **79**, 161301 (2009).
42. L. E. Brus. A simple model for the ionization potential, electron affinity, and aqueous redox potentials of small semiconductor crystallites. *J. Chem. Phys.* **80**, 4403 (1984).
43. J. M. Luttinger, and W. Kohn. Motion of electrons and holes in perturbed periodic fields. *Phys. Rev.* **97**, 869 (1955).
44. E. O. Kane. Band structure of indium antimonide. *J. Phys. Chem. Solids* **1**, 249 (1957).
45. E. O. Kane. The k · p method. *Semiconduct. Semimet.* **1**, 75 (1966).

46. M. Cardona, and F. H. Pollak. Energy-band structure of germanium and silicon: The k · p method. *Phys. Rev.* **142**, 530 (1966).
47. D. Rideau, M. Feraille, L. Ciampolini, M. Minondo, C. Tavernier, H. Jaouen, and A. Ghetti. Strained Si, Ge, and $Si1-xGex$ alloys modeled with a first-principles-optimized full-zone k · p method. *Phys. Rev. B* **74**, 195208 (2006).
48. G. L. Bir, and G. E. Pikus. *Symmetry and Strain-Induced Effects in Semiconductors* (New York, NY: Wiley, 1974).
49. J. P. Loehr. *Physics of Strained Quantum Well Lasers* (New York, NY: Kluwer Academic Publishers, 1998).
50. C. R. Pidgeon, and R. N. Brown. Interband magneto-absorption and Faraday rotation in InSb. *Phys. Rev.* **146**, 575 (1966).
51. S. Tomić, and N. Vukmirović. Symmetry reduction in multiband Hamiltonians for semiconductor quantum dots: The role of interfaces and higher energy bands. *J. Appl. Phys.* **110**, 053710 (2011).
52. M. A. Cusack, P. R. Briddon, and M. Jaros. Electronic structure of InAs/GaAs self-assembled quantum dots. *Phys. Rev. B* **54**, R2300 (1996).
53. A. D. Andreev, and E. P. O'Reilly. Strain distributions in quantum dots of arbitrary shape. *Phys. Rev. B* **62**, 15851 (2000).
54. N. Vukmirović, D. Indjin, V. D. Jovanović, Z. Ikonić, and P. Harrison. Symmetry of $k \cdot p$ Hamiltonian in pyramidal InAs/GaAs quantum dots: Application to the calculation of electronic structure. *Phys. Rev. B* **72**, 075356 (2005).
55. S. Tomić, A. G. Sunderland, and I. J. Bush. Parallel multi-band $k \cdot p$ code for electronic structure of zinc blend semiconductor quantum dots. *J. Mater. Chem.* **16**, 1963 (2006).
56. N. Vukmirović, Z. Ikonić, D. Indjin, and P. Harrison. Symmetry-based calculation of single-particle states and intraband absorption in hexagonal GaN/AlN quantum dot superlattices. *J. Phys. Condens. Matter* **18**, 6249 (2006).
57. A. D. Andreev, J. R. Downes, D. A. Faux, and E. P. O'Reilly. Strain distributions in quantum dots of arbitrary shape. *J. Appl. Phys.* **86**, 297 (1999).
58. E. J. Tyrrell, and S. Tomić. Effect of correlation and dielectric confinement on $1s^{(e)}_{1/2}ns^{(h)}_{3/2}$ excitons in CdTe/CdSe and CdSe/CdTe type-II quantum dots. *J. Phys. Chem. C* **119**, 12720 (2015).
59. A. Franceschetti, H. Fu, L. W. Wang, and A. Zunger. Many-body pseudopotential theory of excitons in InP and CdSe quantum dots. *Phys. Rev. B* **60**, 1819 (1999).
60. Y. Z. Hu, M. Lindberg, and S. W. Koch. Theory of optically excited intrinsic semiconductor quantum dots. *Phys. Rev. B* **42**, 1713 (1990).
61. P. G. Bolcatto, and C. R. Proetto. Partially confined excitons in semiconductor nanocrystals with a finite size dielectric interface. *J. Phys. Condens. Matter* **13**, 319 (2001).
62. J. Movilla, and J. Planelles. Image charges in spherical quantum dots with an off-centered impurity: Algorithm and numerical results. *Comput. Phys. Comm.* **170**, 144 (2005).
63. C. G. Koay, J. E. Sarlls, and E. zarslan. Three-dimensional analytical magnetic resonance imaging phantom in the Fourier domain. *Magn. Reson. Med.* **58**, 430 (2007).
64. C. Xue, and S. Deng. Three-layer dielectric models for generalized coulomb potential calculation in ellipsoidal geometry. *Phys. Rev. E* **83**, 056709 (2011).
65. S. Deng. A robust numerical method for self-polarization energy of spherical quantum dots with finite confinement barriers. *Comput. Phys. Commun.* **181**, 787 (2010).
66. A. Sihvola, and I. Lindell. Polarizability and effective permittivity of layered and continuously inhomogeneous dielectric ellipsoids. *J. Electromagnet. Wave.* **4**, 1 (1990).
67. T. Takagahara. Effects of dielectric confinement and electron–hole exchange interaction on excitonic states in semiconductor quantum dots. *Phys. Rev. B* **47**, 4569 (1993).
68. E. Hanamura, and H. Haug. Condensation effects of excitons. *Phys. Rep.* **33**, 209 (1977).

69. G. Nan, X. Yang, L. Wang, Z. Shuai, and Y. Zhao. Nuclear tunneling effects of charge transport in rubrene, tetracene, and pentacene. *Phys. Rev. B* **79**, 115203 (2009).
70. S. H. Lin, C. H. Chang, K. K. Liang, R. Chang, Y. J. Shiu, J. M. Zhang, T.-S. Yang, M. Hayashi, and F. C. Hsu. *Ultrafast Dynamics and Spectroscopy of Bacterial Photosynthetic Reaction Centers* (New York, NY: John Wiley & Sons, Inc., 2002), pp. 1–88.
71. I.-H. Chu, M. Radulaski, N. Vukmirovic, H.-P. Cheng, and L.-W. Wang. Charge transport in a quantum dot supercrystal. *J. Phys. Chem. C* **115**, 21409 (2011).
72. C. de Mello Donegá, and R. Koole. Size dependence of the spontaneous emission rate and absorption cross section of CdSe and CdTe quantum dots. *J. Phys. Chem. C* **113**, 6511 (2009).
73. W. W. Yu, L. Qu, W. Guo, and X. Peng. Experimental determination of the extinction coefficient of CdTe, CdSe, and CdS nanocrystals. *Chem. Mater.* **15**, 2854 (2003), doi:10.1021/cm034081k.
74. T. Rajh, O. I. Micic, and A. J. Nozik. Synthesis and characterization of surface-modified colloidal cadmium telluride quantum dots. *J. Phys. Chem.* **97**, 11999 (1993), doi:10.1021/j100148a026.
75. C. B. Murray, D. J. Norris, and M. G. Bawendi. Synthesis and characterization of nearly monodisperse CdE (e = sulfur, selenium, tellurium) semiconductor nanocrystallites. *J. Am. Chem. Soc.* **115**, 8706 (1993), doi:10.1021/ja00072a025.
76. X. Peng, J. Wickham, and A. P. Alivisatos. Kinetics of II–VI and III–V colloidal semiconductor nanocrystal growth: "Focusing" of size distributions. *J. Am. Chem. Soc.* **120**, 5343 (1998), doi:10.1021/ja9805425.
77. A. L. Rogach, A. Kornowski, M. Gao, A. Eychmuller, and H. Weller. Synthesis and characterization of a size series of extremely small thiol-stabilized CdSe nanocrystals. *J. Phys. Chem. B* **103**, 3065 (1999), doi:10.1021/jp984833b.
78. V. N. Soloviev, A. Eichhofer, D. Fenske, and U. Banin. Molecular limit of a bulk semiconductor: Size dependence of the band gap in CdSe cluster molecules. *J. Am. Chem. Soc.* **122**, 2673 (2000), doi:10.1021/ja9940367.
79. F. V. Mikulec, M. Kuno, M. Bennati, D. A. Hall, R. G. Griffin, and M. G. Bawendi. Organometallic synthesis and spectroscopic characterization of manganese-doped CdSe nanocrystals. *J. Am. Chem. Soc.* **122**, 2532 (2000), doi:10.1021/ja991249n.
80. Y. Masumoto, and K. Sonobe. Size-dependent energy levels of CdTe quantum dots. *Phys. Rev. B* **56**, 9734 (1997).
81. D. V. Talapin, S. Haubold, A. L. Rogach, A. Kornowski, M. Haase, and H. Weller. A novel organometallic synthesis of highly luminescent CdTe nanocrystals. *J. Phys. Chem. B* **105**, 2260 (2001), doi:10.1021/jp003177o.
82. P. Dagtepe, V. Chikan, J. Jasinski, and V. J. Leppert. Quantized growth of CdTe quantum dots; observation of magic-sized CdTe quantum dots. *J. Phys. Chem. C* **111**, 14977 (2007).
83. E. J. Tyrrell, and J. M. Smith. Effective mass modeling of excitons in type-II quantum dot heterostructures. *Phys. Rev. B* **84**, 165328 (2011).
84. K. Gong, Y. Zeng, and D. F. Kelley. Extinction coefficients, oscillator strengths, and radiative lifetimes of CdSe, CdTe, and CdTe/CdSe nanocrystals. *J. Phys. Chem. C* **117**, 20268 (2013).
85. X. Ma, A. Mews, and T. Kipp. Determination of electronic energy levels in type-II CdTe-core/CdSe-shell and CdSe-core/CdTe-shell nanocrystals by cyclic voltammetry and optical spectroscopy. *J. Phys. Chem. C* **117**, 16698 (2013).
86. X. Cai, H. Mirafzal, K. Nguyen, V. Leppert, and D. F. Kelley. Spectroscopy of CdTe/CdSe type-II nanostructures: Morphology, lattice mismatch, and band-bowing effects. *J. Phys. Chem. C* **116**, 8118 (2012).
87. E. A. Lewis, R. C. Page, D. J. Binks, T. J. Pennycook, P. O'Brien, and S. J. Haigh. Probing the core–shell–shell structure of CdSe/CdTe/CdS type II quantum dots for solar cell applications. *J. Phys.* **522**, 012069 (2014).

88. C.-Y. Chen, C.-T. Cheng, C.-W. Lai, Y.-H. Hu, P.-T. Chou, Y.-H. Chou, and H.-T. Chiu. Type-II CdSe/CdTe/ZnTe (core/shell/shell) quantum dots with cascade band edges: The separation of electron (at CdSe) and hole (at ZnTe) by the CdTe layer. *Small* **1**, 1215 (2005).
89. B. D. Geyter, Y. Justo, I. Moreels, K. Lambert, P. F. Smet, D. V. Thourhout, A. J. Houtepen, D. Grodzinska, C. de Mello Donega, A. Meijerink, D. Vanmaekelbergh, and Z. Hens. The different nature of band edge absorption and emission in colloidal PbSe /CdSe core/shell quantum dots. *ACS Nano* **5**, 58 (2011), PMID: 21189031.
90. D. Espinobarro-Velazquez, M. A. Leontiadou, R. C. Page, M. Califano, P. O'Brien, and D. J. Binks. Effect of chloride passivation on recombination dynamics in CdTe colloidal quantum dots. *Chem. Phys. Chem.* **16**, 1239 (2015).
91. K. Nishikawa, Y. Takeda, T. Motohiro, D. Sato, J. Ota, N. Miyashita, and Y. Okada. Extremely long carrier lifetime over 200ns in GaAs wall-inserted type II InAs quantum dots. *Appl. Phys. Lett.* **100**, 113105 (2012).
92. S. Tomić. Effect of Sb induced type II alignment on dynamical processes in InAs/GaAs/GaAsSb quantum dots: Implication to solar cell design. *Appl. Phys. Lett.* **103**, 072112 (2013).
93. L. E. Brus. A simple model for the ionization potential, electron affinity, and aqueous redox potentials of small semiconductor crystallites. *J. Chem. Phys.* **79**, 5566 (1983).
94. S. Tomić, J. M. Miloszewski, E. J. Tyrrell, and D. J. Binks. Design of core/shell colloidal quantum dots for meg solar cells. *IEEE J. Photovolt.* **6**, 179 (2016).
95. N. Prodanović, N. Vukmirović, Z. Ikonić, P. Harrison, and D. Indjin. Synthesis and characterization of surface-modified colloidal cadmium telluride quantum dots. *J. Phys. Chem. Lett.* **5**, 1335 (2014).

IV

Light-Emitting Diodes (LEDs)

14

Light-Emitting Diode Fundamentals

Sergey Yu. Karpov

14.1 Introduction

Due to the breakthroughs in technological advances in the development of III-nitride semiconductors made in the early 1990s, visible light-emitting diodes (LEDs) now occupy the dominant sector of the optoelectronics market. The emerging LED production is also stimulated by ever-increasing demands from solid-state lighting, which is expected to not only reduce world energy consumption but also change human conceptions of light quality and ways to utilize it. However, because of the constantly increasing capacity of production equipment and, in particular, of growth reactors, the cost of research and development in the LED industry continues to increase at a very high rate, making comprehensive investigations virtually impossible for all but the biggest universities and industrial players. In this situation, modeling and simulation are especially important, saving time, money, and manpower at the research and development stage.

Simulation of LEDs is essentially a multiscale and multidisciplinary problem. The relevant specific sizes range from a few nanometers for the widths of LED active regions to a few centimeters for the LED lamp dimensions, that is, about seven orders of magnitude. This means that, electromechanical, electrical, thermal, and optical phenomena involved in LED operations are often considered to be self-consistent and, in the case of state-of-the-art LEDs, in three-dimensional (3D) formulation. In practice, this makes straightforward LED simulations unacceptably time- and computer resource-consuming. The commonly utilized approach to resolving this problem and making simulation a helpful research and engineering tool is to split the task into independent but interrelated subtasks. This approach correlates three stages of LED fabrication: heterostructure growth, chip fabrication, and lamp assembly. We choose the LED structure, chip, and lamp as the major objects of investigation within the above-mentioned subtasks.

This chapter focuses on fundamental mechanisms underlying the LED operation that should be included in the simulation models. LED heterostructures, chips, and lamps are considered separately, with a discussion on links between them. Here, inorganic LEDs will be considered only when the organic LED models require special modifications.

14.2 LED Family

There is no wonder that most LED applications are related to various kinds of lighting, and thus utilize the visible optical spectrum. The visible spectral range is shared by phosphide LEDs with $(Al_xGa_{1-x})_{0.51}In_{0.49}P$ active regions (emission wavelength $\lambda = 555-625$ nm), usually lattice matched to the GaAs substrates and nitride LEDs with InGaN-based active regions ($\lambda = 400-600$ nm). Beyond the visible spectral range, the most important region is the ultraviolet (UV) region where UVA ($\lambda = 315-400$ nm), UVB ($\lambda = 280-315$ nm), and UVC ($\lambda = 100-280$ nm) diapasons are commonly distinguished. The active regions of UV LEDs are normally based on AlGaN compounds, with the emission wavelength decreasing with the Al content. These LEDs are used for resin curing, sterilization of water, air, and food, and in medical and biotechnology applications. Infrared (IR) LEDs are situated at the other siide of the visible spectral range, with the active region made up of AlGaInAs and GaInAsSb compounds. They are used in niche applications such as IR lighting, remote control, and gas sensing with mid-IR LEDs ($\lambda = 1.8-5.0\,\mu$m).

LED efficiency is the most important characteristic of the devices (see Section 14.3.4). It is also simultaneously a measure for the maturity of their fabrication technology. Figure 14.1 summarizes the maximum external quantum efficiency (EQE) achieved to date in different spectral ranges by LEDs from various industrial companies and leading research centers. There are two distinct EQE maxima: one of ~77%–82% has been reported for InGaN-based LEDs ($\lambda = 410-440$ nm), whereas another of ~71% has been attained by red AlGaInP ($\lambda = 625$ nm) and IR AlGaAs ($\lambda = 850$ nm) devices. Between them, that is, in the spectral range of 500–600 nm, a so-called "green gap" is seen where EQE drops to ~10%–30% in both nitride and phosphide LEDs. Electron leakage into p-type layers has been proved to be responsible for the efficiency

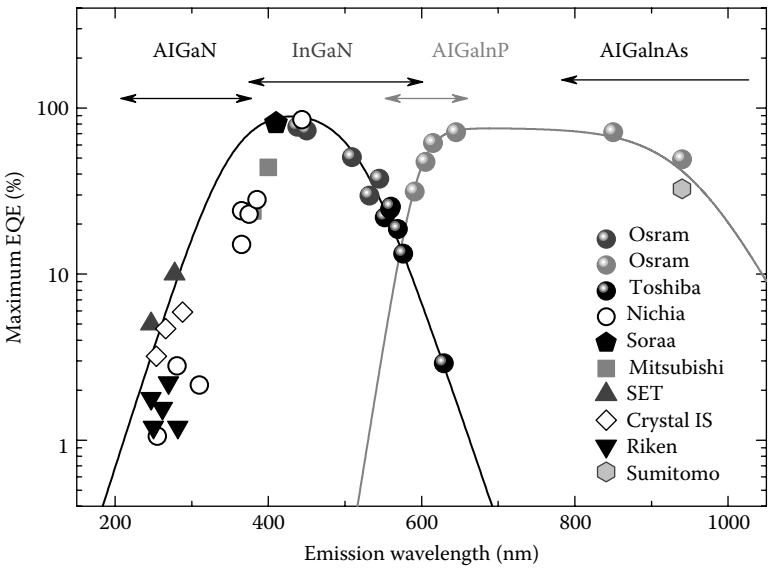

FIGURE 14.1 Maximum external quantum efficiencies (EQEs) of light-emitting diodes (LEDs) as a function of their emission wavelength (symbols are data compiled from numerous sources, lines are drawn for eye). Arrows indicate the ranges of utilizing various III–V compounds for the active region fabrication.

reduction in phosphide LED (Altieri et al., 2005), whereas the nature of the "green gap" in nitride LEDs is not yet completely understood.

Figure 14.1 shows that EQE of AlGaN-based ultraviolet (UV) LEDs decrease dramatically with the Al content in the alloy (at shorter λ), which is attributed to (1) insufficiently high light extraction efficiency partly related to strong transverse-magnetic polarization of the emitted light and absence of highly reflective metallic contacts in this spectral range, (2) problems of p-doping of AlGaN alloys with high Al content, and (3) higher defect density in the grown materials, as compared to Al-free nitride compounds.

Efficiency of IR LEDs suffers primarily from nonradiative Auger recombination inherent in narrow-bandgap semiconductors. In particular, EQE of mid-IR GaInAsSb LEDs emitting in the spectral range of 2.0–4.5 μm (not shown in Figure 14.1) does not exceed ~0.5%–3.0% at room temperature.

To date, as shown earlier, nitride and phosphide LEDs have attained the best performance in terms of emission efficiency. Therefore, simulation examples given in this chapter focus primarily on these two material systems.

14.3 Carrier Transport and Recombination in LED Heterostructures

Any LED structure consists of a stack of epitaxial layers grown in a certain sequence on a substrate of choice. At an arbitrary orientation of heterostructure growth surface, the epilayer interfaces do not always correspond to the main low-index facets of the grown crystals. Therefore, it is convenient to distinguish between the epitaxial coordinate system (ECS) $x'y'z'$, with the x' and y' axes lying in the interface plane and the z'-axis being normal to this plane, and the crystal coordinate system (CCS) xyz, with the axes corresponding to the symmetry axes of the crystal (see Figure 14.2). While ECS is suitable for analysis of carrier transport across the LED structure, CCS enables easy description of anisotropic properties of the crystalline materials (Nye, 1964).

Any crystal of wurtzite symmetry has elastic properties isotropic in the plane normal to its hexagonal *C*-axes. If the CCS *z*-axis corresponds to the *C*-axis of the crystal, the other *x*- and *y*-axes can be chosen rather arbitrarily. It is convenient to choose the *x*-axis of CCS and the x'-axis of ECS to coincide with each other and to be normal to both *z*- and z'-axes (Romanov et al., 2006). Then the y'- or *y*-axis orthogonal to both x'- and z'- or *x*- and *z*-axes can be unambiguously determined. In ECS defined in such a way, the in-plane lattice constant mismatch can be characterized by two parameters corresponding to the lattice

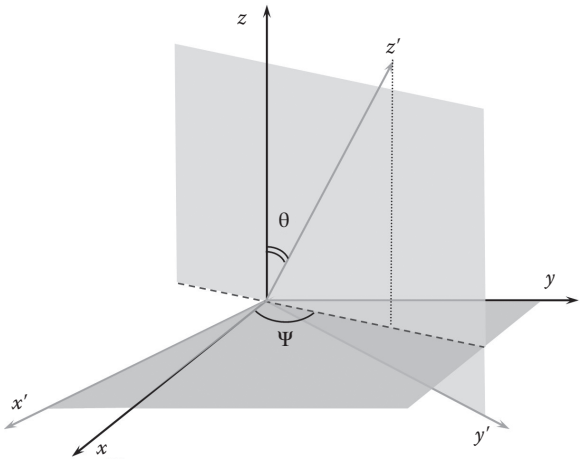

FIGURE 14.2 Crystal coordinate system (CCS) *xyz*, epitaxial coordinate system (ECS) $x'y'z'$, and specific angles between their axes.

constant variation along the x'- and y'-axes (Romanov et al., 2006):

$$\eta_x = 1 - a_S/a_E \quad \text{and} \quad \eta_y = 1 - [(a_S/a_E)^2 \cos^2\theta + (c_S/c_E)^2 \sin^2\theta]^{1/2}, \tag{14.1}$$

where a_E and c_E are the lattice constant of the epitaxial layer considered, a_S and c_S are the lattice constants of the substrate or a thick template layer on which the LED structure is coherently grown, and θ is the polar angle between the z'- and z-axes (see Figure 14.2). This angle can be expressed via four-digit Miller indices $[klmh]$ ($m = -k-l$) characterizing the orientation of the LED structure interfaces (Gil, 2009):

$$\cos\theta = \frac{h}{[h^2 + \frac{4}{3}(c_E/a_E)^2(k^2 + l^2 + kl)]^{1/2}} \tag{14.2}$$

In the case of zinc-blende materials having no in-plane isotropy of elastic properties, the x'-axis normal to both z'- and z-axes does not coincide with the x-axis of CCS but lies in the xy-plane (Figure 14.2). Then the angles θ and ψ (see Figure 14.2) can be expressed via three-digit Miller indexes $[klm]$ as follows:

$$\cos\theta = \frac{m}{(k^2 + l^2 + m^2)^{1/2}}, \quad \cos\psi = \frac{l}{(k^2 + l^2)^{1/2}} \tag{14.3}$$

The lattice mismatch in zinc-blende crystals is described by the only parameter η_x (see Equation 14.1), where a_E and a_S denote the lattice constants of the epilayer and substrate/template, respectively.

Generally, epitaxial layers may have the lattice constants highly mismatched with those of substrate/template layer. The mismatch produces elastic strain affecting the band structure of the material, its electric polarization, and the band offsets at the structure interfaces. Therefore, determination of strain distribution in an LED structure is the primary task in analysis of its operation.

14.3.1 Electromechanical Coupling

Under isothermal conditions, the equations for elastic strain, stress, and electric polarization in every epitaxial layer of the LED structure can be written in matrix form (Nye, 1964):

$$\boldsymbol{\sigma} = \hat{C}\boldsymbol{\varepsilon} - \hat{e}^T\mathbf{E} \text{ (a)}, \quad \boldsymbol{\varepsilon} = \hat{S}\boldsymbol{\sigma} + \hat{d}^T\mathbf{E} \text{ (b)}, \quad \mathbf{D} = \mathbf{P}_{sp} + \hat{d}\boldsymbol{\sigma} + \kappa_0\hat{\kappa}\mathbf{E} \text{ (c)}. \tag{14.4}$$

Here, $\boldsymbol{\sigma}$ is the vector built up of six components of the stress tensor (σ_{xx}, σ_{yy}, σ_{zz}, σ_{yz}, σ_{xz}, and σ_{xy}), $\boldsymbol{\varepsilon}$ is the vector consisting of the diagonal and doubled nondiagonal components of the strain tensor (u_{xx}, u_{yy}, u_{zz}, $2u_{yz}$, $2u_{xz}$, and $2u_{xy}$), $\mathbf{E}$ and $\mathbf{D}$ are the electric field and induction vectors, respectively, and $\mathbf{P}_{sp}$ is the spontaneous electric polarization that does not vanish in wurtzite crystals. The first and second of Equations 14.4 correspond to the direct and reverse Hook laws, where $\hat{C}$ and $\hat{S}$ are the 6×6 matrices of stiffness and compliance constants having a standard form in CCS (Nye, 1964); $\hat{C} \cdot \hat{S} = \hat{S} \cdot \hat{C} = \hat{I}$, where $\hat{I}$ is the 6×6 unity matrix. Here, $\hat{d}$ and $\hat{e}$ are the 3×6 matrices of piezoelectric coefficients ($\hat{e} = \hat{d} \cdot \hat{C}$ and $\hat{d} = \hat{e} \cdot \hat{S}$) having a standard form in CCS too (Nye, 1964), and the superscript "T" denotes transposition of a matrix. The third of Equations Equation 14.4 links the electric field, induction, and stress vectors; κ_0 is the electric constant; and $\hat{\kappa}$ is the 3×3 matrix of static dielectric constants, being diagonal in CCS for both zinc-blende and hexagonal crystals.

The strain components in every epitaxial layer can be determined as follows (Romanov et al., 2006). First, the stress components are assumed to be uniform within the layer, which is valid, if the lateral dimensions of an LED structure are much larger than the vertical ones and bending of the structure is negligible. Second, the components of the stress vector in ECS $\sigma_{z'z'}$, $\sigma_{y'z'}$, and $\sigma_{x'z'}$ are regarded as vanishing according to the absence of external mechanical loading of the top heterostructure surface. Third, the shear component $\sigma_{x'y'}$ is also equal to zero due to uniformity of the strain distribution in the x'- and y'-directions. Then the strain vector contains only two components, $\sigma_{x'x'}$ and $\sigma_{y'y'}$. To determine them, the respective components

of the strain vector $\varepsilon_{x'x'}$ and $\varepsilon_{y'y'}$ should be expressed through $\sigma_{x'x'}$ and $\sigma_{y'y'}$ using Equation 14.4 in ECS. These components are known from geometrical consideration, that is, $\varepsilon_{x'x'} = -\eta_x$ and $\varepsilon_{y'y'} = -\eta_y$, which enables the determination of two nonzero components $\sigma_{x'x'}$ and $\sigma_{y'y'}$. As soon as they are found other components of the strain vector $\boldsymbol{\varepsilon}$ in ECS can be calculated using Equation 14.4 and then transformed into CCS in a standard way.

In the simplest case of wurtzite crystals, the above procedure provides the analytical solution, which is formulated as follows, neglecting the electric field contribution to the stress components:

$$\sigma_{x'x'} = \sigma_1 = -\frac{S'_{22}\eta_x - S'_{12}\eta_y}{S'_{11}S'_{22} - (S'_{12})^2}, \quad \sigma_{y'y'} = \sigma_2 = -\frac{S'_{11}\eta_y - S'_{12}\eta_x}{S'_{11}S'_{22} - (S'_{12})^2}, \tag{14.5}$$

where $S'_{11} = S_{11}$, $S'_{12} = S_{12}\cos^2\theta + S_{13}\sin^2\theta$ and $S'_{22} = S_{11}\cos^2\theta + S_{33}\sin^2\theta + \frac{1}{4}(2S_{13} + S_{44} - S_{11} - S_{33})\sin^2 2\theta$.

Then the strain components in CCS can be obtained from Equation 14.5:

$$\begin{aligned}
&\varepsilon_{xx} = -\eta_a; \quad \varepsilon_{xz} = \varepsilon_{xy} = 0\\
&\varepsilon_{yy} = S_{12}\sigma_1 + (S_{11}\cos^2\theta + S_{13}\sin^2\theta)\,\sigma_2\\
&\varepsilon_{zz} = S_{13}\sigma_1 + (S_{13}\cos^2\theta + S_{33}\sin^2\theta)\,\sigma_2\\
&\varepsilon_{yz} = -\tfrac{1}{2}S_{44}\sigma_2\sin 2\theta.
\end{aligned} \tag{14.6}$$

Nonzero stress components in zinc-blende crystals depend generally on both angles θ and ψ. Using the above approach, one can obtain the strain components in CCS for three practically important orientations of the LED structure interfaces, that is,

- for (001) orientation:

$$\varepsilon_{xx} = \varepsilon_{yy} = -\eta_x, \quad \varepsilon_{zz} = -\frac{2S_{12}}{S_{11} + S_{12}}\eta_x, \quad \varepsilon_{xz} = \varepsilon_{xy} = \varepsilon_{yz} = 0; \tag{14.7}$$

- for (111) orientation:

$$\varepsilon_{xx} = \varepsilon_{yy} = \varepsilon_{zz} = -\eta_x\frac{S_{11} + 2S_{12}}{S_{11} + 2S_{12} + \frac{1}{4}S_{44}}, \quad \varepsilon_{xz} = \varepsilon_{xy} = \varepsilon_{yz} = \eta_x\frac{\frac{1}{2}S_{44}}{S_{11} + 2S_{12} + \frac{1}{4}S_{44}}; \tag{14.8}$$

- for (0*lm*) orientations: strain components can be found from Equation 14.6), assuming $S_{33} = S_{11}$, $S_{13} = S_{12}$, and finding the angle θ from Equation 14.3); in particular, $\theta = 45°$ for the (011) orientation.

The obtained strain components in CCS can then be used for calculating strain-dependent band structure and electric polarization of every epitaxial layer. In wurtzite crystals, the components of the total electric polarization vector **P** in ECS are

$$\begin{aligned}
&P_{x'} = 0\\
&P_{y'} = e_{15}\varepsilon_{yz}\cdot\cos\theta - [P_{\text{sp}} + e_{31}(\varepsilon_{xx} + \varepsilon_{yy}) + e_{33}\varepsilon_{zz}]\cdot\sin\theta\\
&P_{z'} = [P_{sp} + e_{31}(\varepsilon_{xx} + \varepsilon_{yy}) + e_{33}\varepsilon_{zz}]\cdot\cos\theta + e_{15}\varepsilon_{yz}\cdot\sin\theta.
\end{aligned} \tag{14.9}$$

In zinc-blende crystals,

$$\begin{aligned} &P_{x'} = P_{y'} = P_{z'} = 0, && \text{for (001) orientation} \\ &P_{x'} = P_{y'} = 0 \text{ and } P_{z'} = \sqrt{3}e_{14}\varepsilon_{yz}, && \text{for (111) orientation} \\ &P_{x'} = e_{14}\varepsilon_{yz} \text{ and } P_{y'} = P_{z'} = 0, && \text{for } (0lm) \text{ orientation.} \end{aligned} \tag{14.10}$$

The difference in the z'-component of the electric polarization vector $\Delta P_{z'}$ of neighboring layers determines the polarization charge accumulated at their interface. Figure 14.3a shows $\Delta P_{z'}$ at the interfaces between the $In_{0.2}Ga_{0.8}N$, $Al_{0.2}Ga_{0.8}N$, $In_{0.17}Al_{0.83}N$, and GaN layer as a function of the C-axis inclination angle θ (the sign of $\Delta P_{z'}$ corresponds to the sign of the interface charge). To calculate $\Delta P_{z'}$ for the nitride alloys, Vegard's rule was used for estimation of the compliance constants S_{ij}. The application of Vegard's rule to the stiffness constants C_{ij} had proved to provide the results graphically indistinguishable from those shown in Figure 14.3a. The $\Delta P_{z'}$ values plotted in Figure 14.3a refine the similar results reported by Romanov et al. 2006, as the nondiagonal component of the strain tensor u_{yz} was erroneously used in this study instead of ε_{yz}.

An alternative approach to analysis of electromechanical coupling in nitride compounds is based on minimization of the elastic energy in every epitaxial layer (Park and Chuang, 1999). This approach provides incorrect results as the predicted stress does not correspond to unloaded free surface of the LED structure and its interfaces.

Equation 14.10 relevant to zinc-blende semiconductors shows, in particular, that only the (111)-orientation possesses polarization charges at the LED structure interfaces.

14.3.2 Transport Equations

In the hierarchy of transport models, the drift-diffusion one (DDM) was first successfully applied to electron and hole transport in semiconductor devices. Introduced in the mid-twentieth century (Van Roosbroeck, 1950), DDM is commonly recognized as the best compromise between the predictability of the theory and the cost of its numerical implementation.

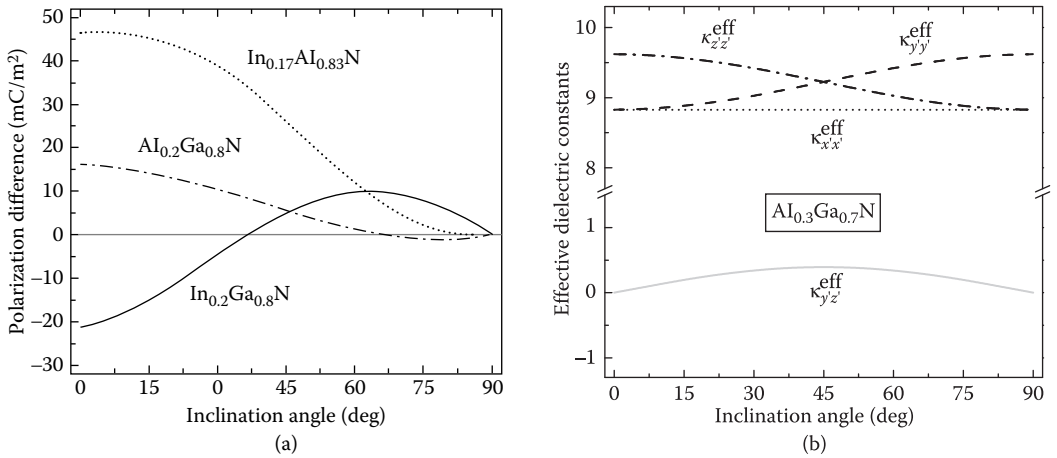

FIGURE 14.3 Difference between the electric polarization of III-nitride alloys and GaN (a) and components of the dielectric constant matrix of $Al_{0.3}Ga_{0.7}N$ (b) in the epitaxial coordinate system (ECS) as a function of inclination angle θ.

Within DDM, the electron n and hole p concentrations can be found from the continuity equations solved in ECS:

$$\frac{\partial n}{\partial t} + \nabla \cdot \mathbf{J_n} = G - R, \quad \frac{\partial p}{\partial t} + \nabla \cdot \mathbf{J_p} = G - R, \tag{14.11}$$

where G is the carrier generation rate, for example, due to light absorption, R is the electron/hole recombination rate, and $\mathbf{J_n}$ and $\mathbf{J_p}$ are the electron and hole fluxes, respectively, related to their partial current densities $\mathbf{j_n}$ and $\mathbf{j_p}$ as follows: $\mathbf{j_n} = -q\mathbf{J_n}, \mathbf{j_p} = q\mathbf{J_p}$; q is the elementary charge.

The most compact definition of the carrier fluxes $\mathbf{J_n}$ and $\mathbf{J_p}$ is based on linear nonequilibrium thermodynamics, involving electron (F_n) and hole (F_p) quasi-Fermi levels measured in electron volts (eV):

$$\mathbf{J_n} = -\mu_n n \nabla F_n, \quad \mathbf{J_p} = -\mu_p p \nabla F_p. \tag{14.12}$$

Here, scalar electron (μ_n) and hole (μ_p) mobilities are used as there is no distinct evidence in the literature for noticeable anisotropy of carrier electric conductivities in various LED structures. The expression 14.12 for the carrier fluxes account for all the forces driving the carrier movement if the electron and hole concentrations are defined as

$$n = N_C \cdot \mathcal{F}_n\left(\frac{F_n - E_C + q\phi}{kT}\right), \quad p = \sum_s N_s \cdot \mathcal{F}_s\left(\frac{E_s - F_p - q\phi}{kT}\right). \tag{14.13}$$

Here, E_C is the conduction band energy determined from the known electron affinity of every epitaxial layer in the LED structure, E_s is the energy of the sth valence sub-band obtained from E_C and the energy gap of the sub-band, ϕ is the electric potential, k is the Boltzmann constant, T is temperature, N_C is the effective density of states in the conduction band, and N_s is effective density of states in the s-th valence sub-band (heavy-hole, light-hole, or split-off hole one), that is,

$$N_C = 2m^n_{xx}(m^n_{zz})^{1/2}\left(\frac{kT}{2\pi\hbar^2}\right)^{3/2}, \quad N_s = 2m^s_{xx}(m^s_{zz})^{1/2}\left(\frac{kT}{2\pi\hbar^2}\right)^{3/2}, \tag{14.14}$$

where $\hbar$ is Planck's constant $(m^n_{xx})^{-1}$, $(m^n_{zz})^{-1}$, and $(m^s_{xx})^{-1}$, $(m^s_{zz})^{-1}$ are the components of the inverse effective mass tensor of electrons and holes in the s-th valence sub-band ($s = hh, lh$, or so) in CCS, respectively ($m^n_{xx} = m^n_{zz} = m_n$ and $m^s_{xx} = m^s_{zz} = m_s$ in cubic crystals within isotropic approximation). Integrals $\mathcal{F}_n$ and $\mathcal{F}_s$ in Equation 14.13 are

$$\mathcal{F}_\nu(\zeta) = \int_0^\infty \frac{g_\nu(x)}{1 + \exp(x - \zeta)}, \quad (\nu = n, s), \tag{14.15}$$

where $g_\nu(x)$ is the density of states in the ν-th band, which is expressed as a function of carrier kinetic energy normalized by kT. For a simple parabolic band, $g_\nu(x) = x^{1/2}$ and the integral 14.15 is reduced to the standard Fermi integral of the 1/2-order. In a general case, Equation 14.15 enables accounting for the band nonparabolicity or a more complex energy spectrum typical for holes.

Electric potential distribution in the LED structure can be found from the Poisson equation solved in ECS. In wurtzite heterostructures of arbitrary orientation, the Poisson equation should account for anisotropic properties of the materials:

$$\kappa_0 \nabla \cdot (\hat{\kappa}^{\text{eff}} \nabla\phi) = \nabla \cdot \mathbf{P} - \rho, \quad \hat{\kappa}^{\text{eff}} = \hat{\kappa} - \frac{(\hat{e}\hat{S}\hat{e}^T)}{\kappa_0}, \tag{14.16}$$

where $\rho = q(N_D^+ + p - N_A^- - n)$ is the charge density, and N_D^+ and N_A^- are the concentrations of ionized donors and acceptors, respectively. The effective dielectric constant $\hat{\kappa}^{\text{eff}}$ is a diagonal matrix in CCS. To

obtain its components in the ECS, the standard transformation should be applied. The second term in the expression for $\hat{\kappa}^{\text{eff}}$ originates from the elastic stress produced by the electric field (see Equation 14.4) and the back contribution of the stress to the electric polarization.

The dielectric constant of cubic crystals κ is a scalar as well as the effective dielectric constant $\kappa^{\text{eff}} = \kappa - S_{44}e_{14}^2/\kappa_0$. Therefore, the Poisson equation for the cubic heterostructures has the same form in both CCS and ECS. As a rule, $\kappa >> S_{44}e_{14}^2/\kappa_0$, which allows neglecting the stress contribution to the dielectric constants. In the case of wurtzite semiconductors, the stress contribution to κ_{zz}^{eff} of ~7%–8% is not negligible, changing remarkably the anisotropy degree of the dielectric constant. Variation of $\hat{\kappa}^{\text{eff}}$ with the *C*-axis inclination angle (Figure 14.3b) may exceed ~10% and should be accounted for in the analysis of nonpolar and semipolar LEDs; in the latter case, the nondiagonal $\kappa_{y'z'}^{\text{eff}}$ component of the effective dielectric constant should also be considered.

Concentrations of ionized donors and acceptors in Equation 14.16 depend locally on the quasi-Fermi level positions relative to the band edges, that is,

$$N_{\text{D}}^{+} = \frac{N_{\text{D}}}{1 + g_{\text{D}} \exp\left(\frac{F_n - E_{\text{C}} + E_{\text{D}} + q\phi}{kT}\right)}, \quad N_{\text{A}}^{-} = \frac{N_{\text{A}}}{1 + g_{\text{A}} \exp\left(\frac{E_{\text{V}} + E_{\text{A}} - F_p - q\phi}{kT}\right)}, \tag{14.17}$$

Here N_{D} and N_{A} are total donor and acceptor concentrations, respectively, E_{D} and E_{A} are the activation energies of the impurities, and g_{D} and g_{A} are their degeneracy factors. These equations assume free carriers and those trapped by donors and acceptors to be in thermodynamic equilibrium with each other.

Equations 14.11 through 14.17 form the simplest DDM for the carrier transport in LED structures. More elaborate models (Lundstrom, 2000; Jüngel, 2009; Jacoboni, 2010; Querlioz and Dollfus, 2010), both semiclassical and quantum mechanical, are not discussed here as most of the practically important results have been obtained with DDM allowing clear and straightforward interpretation.

Figure 14.4 presents band diagrams of single-quantum well (SQW) LED structures consisting of a thick n-GaN contact layer, an undoped 15 nm i-GaN interlayer, an undoped 3-nm $In_{0.17}Ga_{0.83}N$ SQW, a an 8-nm i-GaN spacer where Mg was assumed to diffuse to from the neighboring 20-nm p-$Al_{0.1}Ga_{0.9}N$ electron-blocking layer (EBL), and a 150 nm p-GaN contact layer. The above structures differ in crystal orientation only, demonstrating orientation-dependent band diagrams primarily controlled by the interface polarization charges. The Ga- and N-polar structures (Figure 14.4a and b) form QWs with strong built-in polarization fields and asymmetric barriers, whereas nonpolar (Figure 14.4c) and semipolar (Figure 14.4d) ones possess a rather weak electric field inside the well, and nearly symmetric barriers. The QW asymmetry varies with the forward bias applied and produces a strong, by tens of nanometers, bias-dependent blue shift of the emission wavelength of the Ga- and N-polar structures. In contrast, the emission wavelengths of the nonpolar and semipolar structures are predicted to be stable within ~3–5 nm under the bias/current variation. It is also interesting that GaN claddings of the InGaN SQW in N-polar structure form rather high natural barriers, preventing electrons and holes from leakage to the p- and n-layers of the structure, respectively (Figure 14.4b).

14.3.3 Recombination Models

Various recombination channels should be accounted for in LED simulations. The first one is the nonradiative Shockley–Read–Hall (SRH) recombination mediated by defects. The respective recombination rate is

$$R^{\text{SRH}} = \frac{np}{\tau_p n(1 + \xi_n) + \tau_n p(1 + \xi_p)} \cdot \left[1 - \exp\left(-\frac{F_n - F_p}{kT}\right)\right];$$
$$\xi_n = \exp\left(\frac{E_{\text{t}} - F_n}{kT}\right), \quad \xi_p = \exp\left(\frac{F_p - E_{\text{t}}}{kT}\right). \tag{14.18}$$

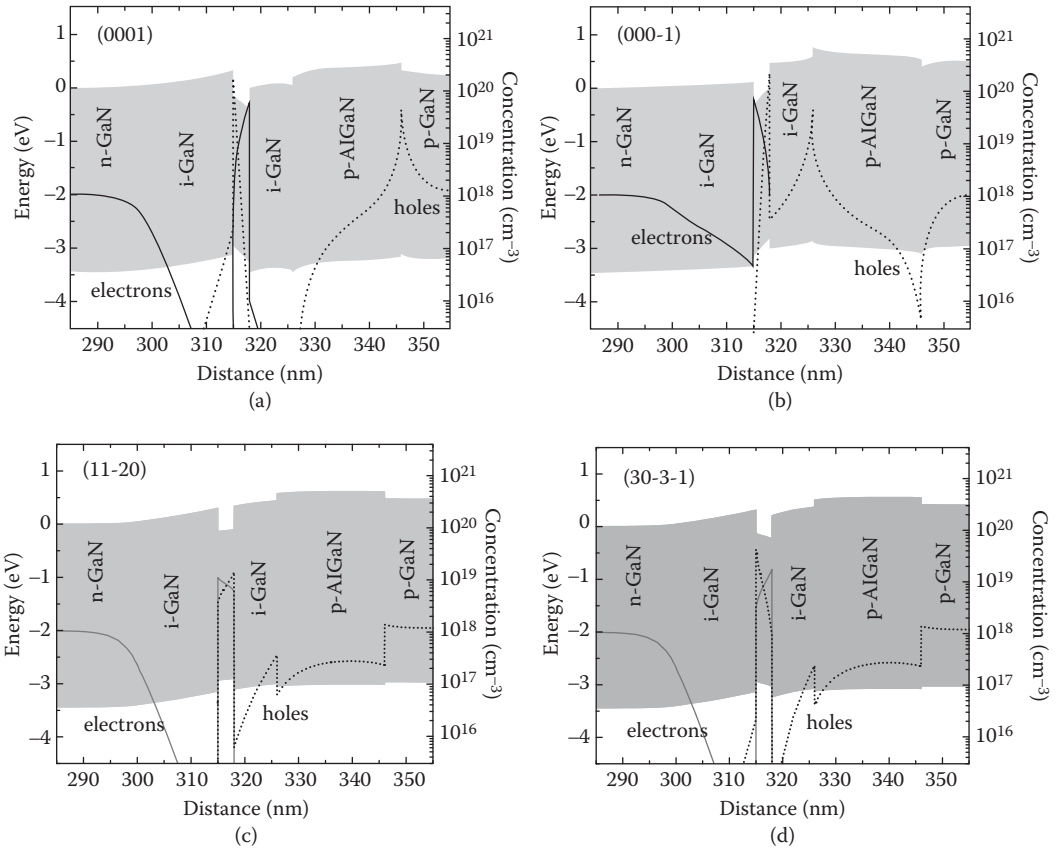

FIGURE 14.4 Band diagrams (gray shading indicates the bandgap) and distributions of electron (solid lines) and hole (dotted lines) concentrations in Ga-polar (a), N-polar (b), nonpolar (c), and semipolar (d) InGaN-based SQW light-emitting diode (LED) structures at the current density of 30 A/cm^2.

Here, E_t is the dominant energy level of a defect responsible for SRH recombination, τ_n and τ_p are the electron and hole life times. One can see from Equation 14.18 that ξ_n, $\xi_p << 1$, if the energy E_t is located deep in the materials bandgap.

In the case of recombination via point defects, $\tau_{n,p} = (c_{n,p} V_{n,p} N_t)^{-1}$, where $c_{n,p}$ is the electron/hole trapping cross-section by the defects, $V_{n,p}$ is the carrier thermal velocity, and N_t is the defect concentration in the material. Dislocations threading through the LED active region are inherent defects in III-nitride epitaxial structures because of lack of native substrates. Here, the electron/hole life times can be estimated by the expression (Karpov and Makarov, 2002)

$$\tau_{n,p} = \frac{1}{4\pi D_{n,p} N_d} \left\{ \ln\left(\frac{1}{\pi r_c^2 N_d}\right) + \frac{2D_{n,p}}{r_c V_{n,p}} - \frac{3}{2} \right\}, \tag{14.19}$$

where $D_{n,p}$ is the carrier diffusivity, N_d is the dislocation density, and r_c is the dislocation core radius equal to the lattice constant a in the order of magnitude.

Another recombination channel is the radiative one. The rate of radiative recombination between free carriers in the conduction and valence bands can be calculated as

$$R^{rad} = Bnp \cdot \left[1 - \exp\left(-\frac{F_n - F_p}{kT}\right)\right], \tag{14.20}$$

where the radiative recombination constant B depends essentially on the particular mechanism of light emission. In the case of bulk unstrained zinc-blende semiconductors, the B constant corresponding to optical transitions occurring with the **k**-selection rule at a rather low excitation level can be estimated as

$$B_0 = (N_{\rm cv}\tau_{\rm B})^{-1} \cdot \left(1 + \frac{3kT}{2E_{\rm g}}\right); \quad N_{\rm cv} = 2\left[\frac{(m_n + m_{\rm hh})kT}{2\pi\hbar^2}\right]^{3/2}, \quad \tau_{\rm B} = \frac{3\hbar m_0 c^2}{2\alpha\overline{n}_{\rm r}E_{\rm g}E_{\rm P}} \tag{14.21}$$

Here, m_n and $m_{\rm hh}$ are the effective masses of electrons and heavy holes, respectively, m_0 is the free electron mass, c is the light velocity in vacuum, $\alpha = q^2/(4\pi\varepsilon_0\hbar c) \approx 1/137$ is the fine structure constant, $\overline{n}_{\rm r}$ is the group refractive index at the wavelength corresponding to the optical transition, $E_{\rm g}$ is the bandgap of the semiconductor, and $E_{\rm P}$ is the energy related to the Kane's matrix element P (Kane, 1957): $E_{\rm P} = 2m_0P^2/\hbar^2$. At high excitation levels, the B constant starts to depend on the carrier concentrations. Numerical calculations of the radiative recombination rate under these conditions, accounting for the fact that $m_n << m_{\rm hh}$ have found the B constant to be well fitted by the expression: $B = B_0N_{\rm cv}/(N_{\rm cv} + p)$.

Many other processes may also contribute to radiative recombination, including excitonic optical transitions, band-to-impurity (donor, acceptor, or a deep center), donor-to-acceptor, phonon-assisted band-to-band recombination, and so on (Pankove, 1971), each providing its own specific radiative recombination constant. In addition, each of the above processes occurs differently in bulk materials, QWs, and quantum dots (QDs), multiplying a variety of B constants. Therefore, the choice of dominant mechanisms should be based on experimental data in every particular case and corresponding models should be applied in the framework of DDM.

The third important mechanism is nonradiative Auger recombination with the rate

$$R^{\rm A} = (C_n n + C_p p)\, np \cdot \left[1 - \exp\left(-\frac{F_n - F_p}{kT}\right)\right], \tag{14.22}$$

where C_n and C_p are the Auger recombination coefficients related to the microscopic processes involving two electrons and a hole and two holes and an electron, respectively. In bulk unstrained zinc-blende semiconductors, the C_n coefficient corresponding to the process involving two electrons and a heavy hole, that is, the CHCC† process, is given by the equation (Gelmont and Sokolova, 1982; Abakumov et al., 1991)

$$\begin{aligned} C_n &= \frac{16}{\sqrt{\pi}}\chi\frac{\nu_{\rm D}}{N_{\rm C}N_{\rm hh}}\left(\frac{kT}{E_{\rm g}}\right)^{3/2}\frac{\langle\varepsilon_{\rm c}\rangle}{E_{\rm g}}\exp\left(-E_{\rm th}/kT\right); \\ E_{\rm th} &= \frac{m_n}{m_{\rm hh}}E_{\rm g}\frac{(2E_{\rm g}+\Delta)(3E_{\rm g}+2\Delta)}{(E_{\rm g}+\Delta)(3E_{\rm g}+\Delta)}, \quad \nu_{\rm D} = \frac{E_{\rm D}}{\hbar}, \\ \chi &= \left(\frac{E_{\rm g}+\Delta}{3E_{\rm g}+2\Delta}\right)^{3/2}\left(\frac{3E_{\rm g}+\Delta}{2E_{\rm g}+\Delta}\right)^{1/2}, \quad E_{\rm D} = \frac{m_{\rm n}q^4}{32\pi^2(\kappa_0\kappa)^2\hbar^2}. \end{aligned} \tag{14.23}$$

† To distinguish between various microscopic Auger processes, the following four-letter notation is commonly accepted. First two letters indicate the initial and final electronic states of the recombining particle, e.g., an electron in the conduction band (C) and a hole in the heavy hole sub-band (H). Last two letters indicate the initial and final states of the third particle to which the excess energy released in course of recombination is transferred. Therefore, the CHCC notation means that an electron in the conduction band and a heavy hole recombine with each other, transferring the excess energy to one more electron (C), which is scattered within the same conduction band (C). In turn, the CHHS notation means that an electron in the conduction band and a heavy hole recombine, transferring the excess energy to another heavy hole (H), which is scattered into the split-off hole sub-band (S).

Here, N_C and N_{hh} are the effective densities of states in the conduction band and heavy-hole sub-band, respectively (see Equation 14.14), E_D is the ionization energy of a shallow donor, $\langle \varepsilon_c \rangle$ is the mean electron energy in the conduction band, and Δ is the spin–orbital splitting of the valence band. The energy $\langle \varepsilon_c \rangle = \frac{3}{2}kT$ in the case of nondegenerated electrons and $\langle \varepsilon_c \rangle = \frac{3}{5}\varepsilon_F \propto n^{2/3}$ for degenerated electrons, and ε_F is the Fermi energy. In the latter case, Auger recombination depends even more on the electron concentration than in the case of nondegenerated carriers.

The CHCC Auger process has a threshold accounted for by the exponential factor in Equation 14.23, with E_{th} being the threshold energy. The existence of the threshold originates from simultaneous conservation of the momentum and energy during recombination. As follows from Equation 14.23, the rate of the CHCC Auger recombination decreases dramatically with the material bandgap; therefore, this nonradiative process is especially important in narrow-bandgap semiconductors.

Among the processes involving two holes and an electron in bulk cubic semiconductors, the most important is the CHHS one, where the high-energy hole transfers from the heavy-hole to the split-off hole sub-band. At $E_g - \Delta >> kT$, the recombination coefficient associated with this process has an analytical representation (Abakumov et al., 1991):

$$C_p = \frac{216}{\sqrt{\pi}}\chi\frac{\nu_D}{N_{hh}N_{so}}\left(\frac{kT}{E_{th}}\right)^{1/2}\left(\frac{kT}{E_g}\right)^3 \exp\left(\frac{-E_{th}}{kT}\right);$$
$$E_{th} = \frac{m_{so}}{m_{hh}}E_g\frac{E_g(E_g-\Delta)}{(E_g+\Delta)(3E_g-2\Delta)}, \quad \chi = \left(\frac{m_{hh}}{m_{so}}\right)^{1/2}\left(\frac{E_g+\Delta}{E_g}\right)^2. \tag{14.24}$$

Equation 14.24 shows that the energy threshold of the CHHS process disappears under resonance conditions $\Delta \approx E_g$; near the resonance, the threshold energy is relatively small, resulting in a considerable increase in the recombination rate.

Similarly to the case of radiative recombination, there is a variety of microscopic Auger recombination processes, involving, in addition to free carriers, those trapped by shallow donors, acceptors, and deep centers (Abakumov et al., 1991), or injected into QWs and QDs (Zegrya and Kharchenko, 1992; Dyakonov and Kachorovskii, 1994; Zegrya and Samosvat, 2007) as well as phonons (Kioupakis et al., 2011). In all these cases, the momentum conservation is violated in one or more directions, leading to the disappearance of the Auger recombination threshold and, as a result, to an increase in the nonradiative recombination rate. Unfortunately, not all of the mechanisms allow even simplified analytical treatment to derive the recombination coefficient; moreover, the theory for some important mechanisms is not yet completely developed (Abakumov et al., 1991). Therefore, only experimental observations may provide guidelines for the choice of dominant Auger mechanisms, in particular materials and device structures, and fitting of the C_n and C_p coefficient to the available data instead of using a rigorous theory may be applied in some cases.

The final nonradiative mechanism of carrier losses to be considered is that their surface recombination is occurs mainly at free surfaces of the LED active region. The surface recombination is accounted for via the flux J_s of carriers arriving at the free surface to recombine nonradiatively. The flux may be calculated as

$$J_s = \frac{np}{n/S_p + p/S_n}\cdot\left[1 - \exp\left(-\frac{F_n - F_p}{kT}\right)\right], \tag{14.25}$$

where S_n and S_p are the surface recombination velocities associated with n-type and p-type materials, respectively. Frequently the carrier losses localized at the active region surface are accounted for by the effective recombination rate distributed over the whole active region with the area $\mathcal{A}$ and perimeter $\mathcal{P}$, which can be calculated by Equation 14.18 with $\xi_n = \xi_p = 0$ and $\tau_{n,p} = \mathcal{A}/(\mathcal{P}S_{n,p})$ (Royo et al., 2002). Such a simplified model is valid if the carrier concentration variation in the active region is much lower than their mean concentrations. In other cases, a more elaborate approach should be applied (see Section 14.4.2).

14.3.4 Efficiency of Light Emission

A number of parameters characterize the efficiency of LED structures. One is the injection efficiency η_{inj}, which accounts for the losses of minority carriers arriving at the contact electrodes without recombination inside the LED structure. The injection efficiency is defined as the ratio of the recombination current $I_{\rm rec} = q\int_V R(\mathbf{r})\,d^3\mathbf{r}$ to the total current I flowing through the LED (here, the integration is performed over the whole LED volume V). Another important parameter is the internal quantum efficiency (IQE) η_i, defined as the ratio of the radiative recombination current $I_{\rm rad} = q\int_V R^{\rm rad}(\mathbf{r})\,d^3\mathbf{r}$ to $I_{\rm rec}$. EQE $\eta_{\rm e}$ characterizes the number of photons outgoing from the LED per one electron–hole pair passing through the LED contact electrodes. Generally, $\eta_e = \eta_{\rm ext}\eta_{\rm inj}\eta_i$, where $\eta_{\rm ext}$ is the efficiency of light extraction (LEE) from the LED die. Finally, the wall–plug efficiency (WPE) $\eta_{\rm w}$ is the ratio of the output optical power to the electrical power consumed by LED during its operation. The ratio $\eta_{\rm w}/\eta_{\rm e}$ is sometimes regarded as the so-called electric efficiency. However, this ratio may be greater than unity at low currents and, therefore, use of this parameter appears to be somewhat misleading.

Both blue/green InGaN-based and red/amber AlGaInP-based LEDs exhibit rather strong EQE reduction at high operating currents but the dominant mechanisms responsible for the reduction are different in nitride and phosphide devices. In AlGaInP LEDs, the EQE reduction with current has been primarily attributed to electron leakage into the p-side of the LED structure because of insufficiently high barriers inherent in the AlGaInP materials system (Royo et al., 2002; Altieri et al., 2005). The leakage is found to be strongly increased by elevated temperature, with the activation energy nearly equal to the difference between the electron energy levels involved in optical transitions and the edge of the barrier conduction band (Altieri et al., 2005). This experimental fact was the major argument for attributing the EQE reduction with current to the electron leakage.

Blue/green nitride LEDs operate at the nonequilibrium carrier density in the active region, at least an order of magnitude higher than that in phosphide LEDs (compare Figures 14.4 and 14.5a). Therefore, the EQE reduction (droop) with current is associated here with Auger recombination (Shen et al., 2007; Bulashevich and Karpov, 2008; Laubsch et al., 2009); see also various reviews (Piprek, 2010; Avrutin et al., 2013; Verzellesi et al., 2013; Cho et al., 2013) for more detailed discussion on this and alternative mechanisms and experimental studies (Iveland et al., 2013; Galler et al., 2013; Binder et al., 2013) directly demonstrating the importance of Auger recombination. Although particular microscopic mechanisms of Auger recombination in nitride LEDs are not reliably identified, the concept of enlarged recombination volume used in the LED structure design (see, e.g., the review by Weisbuch et al. 2015) and aimed at suppression of the nonradiative carrier losses due to lowering the carrier concentration in the active region works well in practice.

Figure 14.5a shows the band diagrams and nonequilibrium carrier concentrations in a red ($\lambda = 622$ nm) LED structure grown on an n-GaAs substrate and consisting of an undoped $5 \times$ (3.5 nm $Ga_{0.51}In_{0.49}P$/7 nm $(Al_{0.8}Ga_{0.2})_{0.51}In_{0.49}P$) multiple-quantum well (MQW) active region sandwiched between n- and p-$Al_{0.51}In_{0.49}P$ confinement layers, both 100 nm thick; a thick p-GaP contact layer completes the structure. As one can see from a comparison of Figures 14.5a and 14.4, the electron/hole concentration in the active region of the red LED is ~1–2 orders of magnitude lower than in InGaN-based LEDs at the same operating current density of 30 A/cm^2. This originates from different properties of the phosphide and nitride semiconductors: conduction and valence band offsets, carrier mobilities, and recombination constants. The band alignment in the MQW active region is rather flat, providing a uniform distribution of the recombination rate among all the QWs.

The injection efficiency of the red LED structure plotted in Figure 14.5b is strongly dependent on temperature. It differs from unity rather weakly at 300 K, and is considerably reduced at 340 and 370 K, indicating increasing carrier losses via leakage. The electron leakage is predicted to dominate over the hole one at all current densities considered. IQE of the red LED structure varies weakly with temperature (Figure 14.5c) due to temperature-dependent recombination constants. The IQE dependence on current

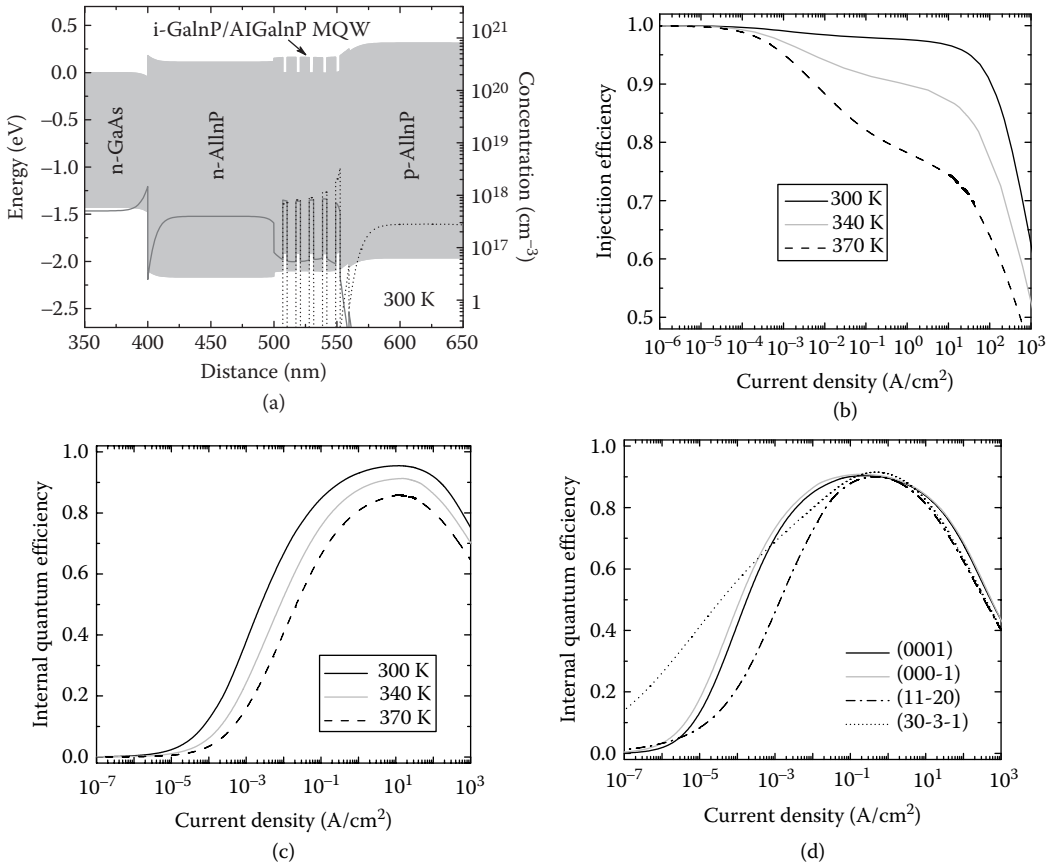

FIGURE 14.5 Band diagram (gray shading indicates the bandgap) and distributions of electron (solid lines) and hole (dotted lines) concentrations in a red MQW light-emitting diodes (LEDs) structure at the current density of 30 A/cm^2 (a). Injection efficiency (b) and IQE (c) of the structure as a function of current density at three different temperatures. IQEs of the InGaN-based SQW LED structures shown in Figure 14.4 are also given versus current density for comparison (d).

has a dome-like shape, reflecting competition between the radiative and nonradiative recombination channels. In particular, the IQE droop at high currents is related to nonradiative carrier losses caused by Auger recombination. Because IQE starts to decline at much higher current density than IE does, degradation of the overall red LED efficiency with temperature should be primarily attributed to the carrier leakage.

Figure 14.5d displays current-dependent IQEs of the nitride SQW LED structures discussed in Section 14.3.2. Simulation of the structures was intentionally carried out with the same recombination constants in order to emphasize the role of the crystal polarity. The computed injection efficiency is equal to unity for all the structures in the whole range of the current density variation. Therefore, the IQE droop clearly seen in Figure 14.5d should be attributed entirely to Auger recombination. As the carrier concentrations in the active region of InGaN-based LEDs considerably exceed those in the red LED structure, the IQE droop in nitride emitters starts at lower current densities (compare Figure 14.5c and d).

The efficiency droop is also observed in UV LEDs with AlGaN active regions (Mickevičius et al., 2015). The nature of the droop here, however, remains unclear and requires further investigation.

14.4 Current Spreading, Heat Transfer, and Light Extraction in LED Dice

State-of-the-art LEDs utilize rather complex essentially 3D chip designs aimed at achieving various goals: providing a uniform carrier injection into the active regions at minimum series resistance, efficient light extraction from the LED dice, and efficient heat removal from the active region. Most of the physical processes involved in the chip operation are coupled with one another so that their straightforward modeling requires a large amount of computational time and computer resources. Therefore, implementation of approximation approaches to create engineering simulations of the processes occurring in the LED chips is highly desirable.

A rather effective approach can be developed on the basis of the equation for the total current density $\mathbf{j} = \mathbf{j_n} + \mathbf{j_p}$; under steady-state conditions, this equation follows from Equation 14.11: $\nabla \cdot \mathbf{j} = 0$. It does not contain any recombination terms and enables much easier and faster numerical solution than full DDM equations, if the relationship between $\mathbf{j}$ and electric potential ϕ is specified. In the contact layers, this relationship can be approximated by a unipolar conductivity, that is, $\mathbf{j} = -\sigma_{n,p}\nabla\phi$, where σ_n and σ_p are the electric conductivities of the n- and p-contact layers, respectively. This approximation does not work, however, in the LED active region. On the other hand, electrons and holes are transferred here mainly in the vertical direction due to dominant vertical components of the electric field and gradients of carrier concentrations. Therefore, the carrier transport in the active region may be well approximated by a 1D model. The 1D solution to the DDM equations provides a nonlinear dependence of the vertical current density j_v on the p-n junction bias $U_b = F_n - F_p$, where the values of the quasi-Fermi levels F_n and F_p are taken at the n- and p-sides of the active region, respectively. Thus, the $j_v(U_b)$ dependence obtained may be used as a nonlinear boundary condition in 3D chip simulations, connecting the vertical component of the local current density $\mathbf{j}$ with the electric potential difference on the n-side (ϕ_n) and p-side (ϕ_p) borders of the active region: $\mathbf{n} \cdot \mathbf{j} = j_v(\phi_p - \phi_n)$, where $\mathbf{n}$ is the normal to the active region directed from the p- to n-contact layer. Solution to the above problem provides 3D distribution of the current density in an LED die.

This approximate approach has proved to be quite efficient for coupled modeling of LEDs with rather complex chip designs (Bogdanov et al., 2008); it can also be easily refined by using additional experimental input (López and Margalith, 2008). Comparison of theoretical results obtained by the above approach with available observations has demonstrated its good predictability (Chernyakov et al., 2013).

As it was already mentioned, the carrier transport in the LED active region occurs primarily in the vertical direction. Therefore, considering the lateral electron and hole transfer as a perturbation, one can also account for the lateral redistribution of carriers due to ambipolar diffusion and their surface recombination at the free surface of the active region.

To illustrate, Figure 14.6 shows simulated 2D current density distribution over the active region and some characteristics of a vertical red LED die shaped in such a way as to increase its light extraction efficiency (LEE) using inclined side walls of the die (see Figure 14.7a). The LED structure presented in Figure 14.5a has been chosen for the simulations, whereas modeling of current spreading, heat transfer, and light extraction was carried out self-consistently in 3D approximation. Due to the vertical chip geometry, current crowding (see Section 14.4.1 for a more detailed discussion) is rather weak in the chip (Figure 14.6a), resulting in a nearly linear relationship between the mean current density and current (Figure 14.6b). Simulated current–voltage and light–current characteristics of the red LED are shown in Figure 14.6c, whereas EQE and WPE variation with current is given in Figure 14.6d. At relatively low currents, WPE exceeds EQE (see Figure 14.6d), which is typical for LEDs operating at the forward voltage $V_f < E_{ph}/q$, where E_{ph} is the mean energy of emitted photons. The decrease of EQE and WPE at high currents is caused by the thermally enhanced electron leakage, as discussed in Section 14.3.4. One can see from Figure 14.6b that the mean temperature in the LED die increases superlinearly with the operating current.

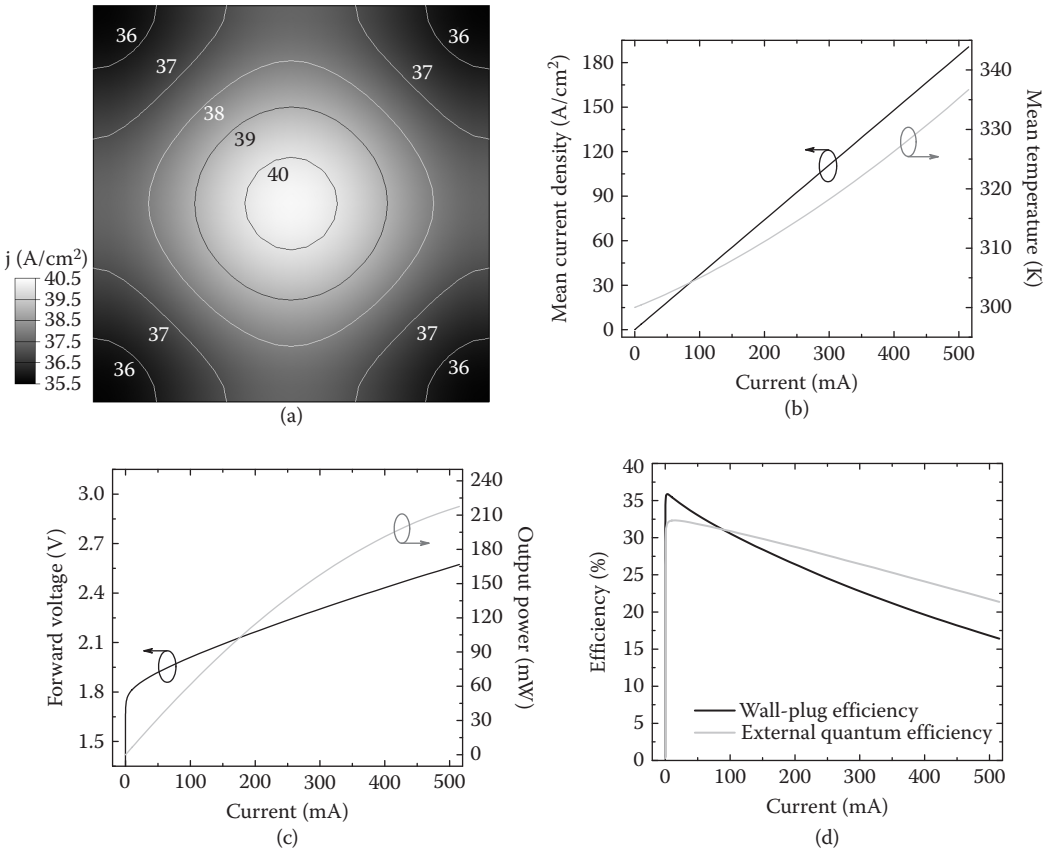

FIGURE 14.6 Two-dimensional (2D) vertical current density distribution in the active region at the current of 100 mA (a) and mean current density and temperature (b), forward voltage and output optical power (c), and external quantum efficiency (EQE) and wall–plug efficiency (WPE) (d) as a function of operating current of the red light-emitting diode (LED) chip shown in Figure 14.7a.

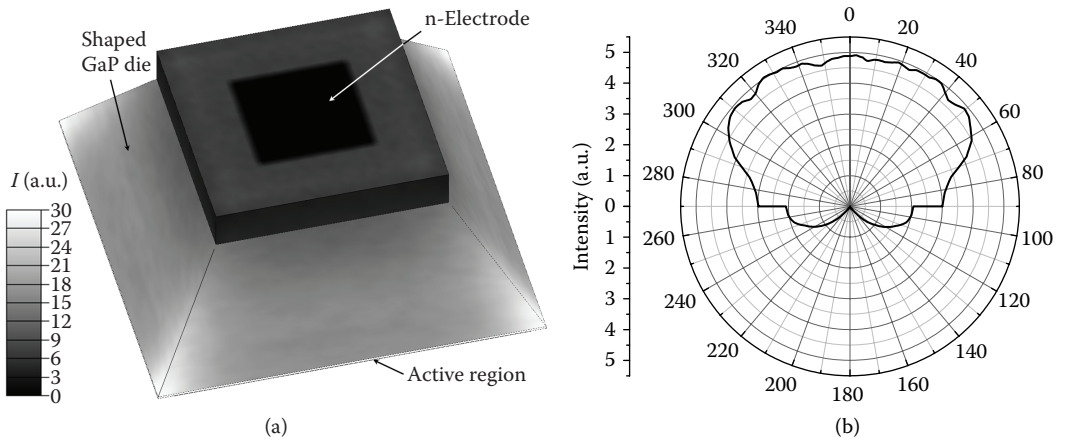

FIGURE 14.7 Near-field emission intensity distribution over the outer surfaces of red light-emitting diode (LED) die with shaped side walls (a) and far-field emission pattern of the LED (b).

As soon as the distribution of electron and hole concentrations in the active region is determined, the radiative recombination rate in each point can be obtained (Bogdanov et al., 2008). Using the radiation pattern specific for the active region band structure, one can use the ray tracing to simulate photon propagation and absorption in the LED die, as well as their extraction efficiency.

Figure 14.7 shows results of the ray-tracing modeling of the shaped red LED die considered previously. LEE has been found to be independent of operating current due to weak current crowding and equal to 35.2%. Near-field emission intensity distribution over the outer surfaces of the LED die (Figure 14.7a) demonstrates that most of the emitted photons are extracted through the inclined side walls of the die and primarily through the regions adjacent to the die edges. As a result, the far-field emission pattern differs substantially from the Lambertian one, providing nearly constant intensity in a wide observation angle of ~100° (Figure 14.7b).

Ray tracing is quite suitable for accounting properties of the photon ensemble in LED dice. However, its applicability is rather limited as the approach ignores the wave nature of photons, that is, the effects of their diffraction and interference. Ray tracing is warranted if specific LED chip dimensions are much larger than the photon wavelength. However, any elements of the chip design based on the photon interference or diffraction, such as photonic crystals or textured surface commonly used for better extraction of photons from LEDs, cannot be considered to be accurate in the ray tracing procedure.

An alternative approach frequently applied for estimation of LEE is the direct finite-difference time domain (FDTD) simulations, where the Maxwell's equations are solved for the electromagnetic field induced by a dipole imitating the act of radiative recombination of an electron and a hole. The FDTD method is quite suitable, in particular, for analysis of light diffraction at a textured surface of the LED dice. Despite the accounting for photon wave nature, the FDTD approach has also some limitations related to (1) ignoring the ensemble properties of electrons/holes and, hence, of emitted photons and (2) using some artificial boundary conditions for the Maxwell's equations, the impact of which on the accuracy of theoretical predictions is still uncertain (Zhmakin, 2011).

To summarize, ray tracing and FDTD simulation are complementary approaches to LEE calculation rather than competing ones. The choice between them depends on the LED chip geometry and parameters desirable to predict LEE value, emission pattern in the near- and far-field zones, light absorption in LED die, emission polarization, and so on.

14.4.1 Current Crowding

Current crowding is the effect of the current density localization inside the LED die around the streamlines with minimum path resistance. Generally, every LED chip can be represented via two distributed linear resistors corresponding to n- and p-contact layers and a nonlinear resistor corresponding to the LED active region, including p-n junction, all connected in series. The nonlinear p-n junction resistor has a nearly exponential resistance dependence on the internal p-n junction bias. At a low operating current, the p-n junction resistance is much higher than both linear ones and the lateral current flow in the contact layers dominates, resulting in a rather uniform distribution of the vertical current density over the LED active region. At higher currents, the increase in the p-n junction bias leads to a dramatic reduction of the p-n junction resistance, which becomes comparable to or lower than both contact layer resistances. As a result, the vertical current density redistributes significantly in such a way as to minimize the total path resistance along the current streamlines. Therefore, current crowding is essentially a nonlinear effect, which develops with the LED operating current I and depends on temperature primarily through the nonlinear p-n junction resistance.

Because of current crowding, the main LED characteristics should be obtained by averaging over the active region area $\mathcal{A}$ with the local current density j_v serving as the weight function. So, the mean current density in the active region $\bar{j} = I^{-1} \int_{\mathcal{A}} j_v^2(\mathbf{r})\, d^2\mathbf{r} > I/\mathcal{A}$ is no longer proportional to $I = \int_{\mathcal{A}} j_v(\mathbf{r})\, d^2\mathbf{r}$, if the current crowding is pronounced. Also the mean IQE of an LED becomes dependent on current crowding (see simulation results obtained by Bogdanov et al. (2010) and experimental data by Malyutenko et al.

(2010)): $\overline{\eta_i} = I^{-1} \int_{\mathcal{A}} \eta_i(j_v) \cdot j_v(\mathbf{r})\, d^2\mathbf{r}$, where $\eta_i(j_v)$ is the IQE of LED structure dependent on the local current density j_v.

Another important effect of current crowding found by simulations (Bogdanov et al., 2010) and reported experimentally by Laubsch et al. (2010) is the current localization under metallic electrodes deposited on the emitting surface of LEDs. Depending largely on the operating current, the localization results in LEE decreasing with current, enhancing the overall efficiency droop. The effect was found to be much stronger than the IQE reduction caused by current crowding (Bogdanov et al., 2010). Avoiding metallic electrodes on the emitting surface by the use of advance chip designs (see, for instance, Laubsch et al. 2010) improves LEE substantially.

Generally, current crowding is especially pronounced in lateral LED chips utilizing nonconductive sapphire substrate and having one-side electrode access to both n- and p-contact layers. This is because of a relatively thin n-contact layer providing dominant lateral current flow inside the chip. In a vertical LED die similar to that shown in Figure 14.7a, the current crowding is weaker but, nevertheless, still tangible (see Figure 14.6a) despite a relatively large distance, ~200 μm, between the n- and p-electrodes, which looks at first glance to be quite sufficient for the complete homogenization of the current density distribution inside the die.

14.4.2 Surface Recombination

There are three main parameters affecting the surface recombination impact on the LED efficiency: the recombination velocity S, the ambipolar carrier diffusion coefficient D_a in the active region, and the carrier differential life time τ_d generally dependent on the carrier concentration in the active region. The latter two parameters determine the carrier diffusion length $L_d = (D_a\tau_d)^{1/2}$, which characterizes the width of the region adjacent to the free surface of the active region where contribution of surface recombination to the carrier losses is significant.

Red/amber AlGaInP LEDs operate normally at a reduced, ~10^{17}–10^{18} cm^{-3}, carrier concentration in the active region (Figure 14.5a), providing a relatively long τ_d. The carrier diffusivity in the active region enhanced in undoped QWs may be rather large. As a result, L_d may approach ~10–25 μm in such devices. Therefore, surface recombination results in a remarkable rise to the carrier losses already at the LED chip dimensions as small as ~150–250 μm (Royo et al., 2002).

The carrier concentration in the active regions of blue and green InGaN-based LEDs is ~1–2 orders of magnitude higher (Figure 14.4), shortening significantly the lifetime τ_d; the carrier diffusivity here is also about an order of magnitude smaller. Therefore, L_d does not normally exceed a few micrometers in nitride LEDs. On the other hand, the free active region surface is normally located close to the regions of strong current crowding in the III-nitride LED chips. Estimates show that surface recombination may be valuable in blue/green LEDs at low currents corresponding to the maximum of their efficiency. Surface recombination is expected to be especially pronounced in deep-UV LEDs because of (1) a high Al content in the active region normally increasing the surface recombination velocity, (2) a weaker carrier localization resulting in their higher diffusivity in the active region, and (3) smaller LED die dimensions usually utilized in practice.

14.5 Phosphor-Converted White Light Emission

An important sector of LED industry is the production of white light sources for solid-state lighting. Such sources are based on emulation of sunlight by a number of LEDs emitting in different spectral ranges or by using a partial light conversion by one or more phosphors. As it has been shown by Bulashevich et al. (2015), the phosphor-converted LEDs (PC-LEDs) provide a higher efficacy at a better color rendition, as compared to white light sources utilizing LEDs only. In addition, fabrication of PC-LEDs represents the

main industrial approach to the white light production. Therefore, this section will focus on simulation models necessary for the analysis of light conversion by phosphors pumped with LEDs.

14.5.1 White Light Characteristics

The fundamentals of color perception and white light characterization are described in detail by Judd and Wyszecki (1975). Here, the main white light characteristics will be briefly discussed in view of discussion on their further modeling.

Starting with J.C. Maxwell's studies in the mid-nineteenth century, it became customary to decompose the light color into three components. According to CIE (1986), three color coordinates, X, Y, and Z, can be brought in by the convolution of a light source spectral power density emission spectrum $S_\lambda(\lambda)$ with three tabulated color matching functions, $\overline{x}(\lambda)$, $\overline{y}(\lambda)$, and $\overline{z}(\lambda)$. Then the relative chromatic coordinates $x = X/(X + Y + Z)$ and $y = Y/(X + Y + Z)$ project the emission spectrum into the CIE1931 color chart, which is represented by points in the (x, y) space, all the possible colors that can be encountered in practice.

A natural source producing white light is a black body heated up to a certain temperature. Under the temperature variation, chromatic coordinates of the black-body radiation (BBR) form in the CIE1931 diagram, a locus nearly corresponding to white light. In practice, some deviations of the white-light chromatic coordinates x_w and y_w from the BBR locus are admissible. In this case, a certain projection of the (x_w, y_w) point to the BBR locus produces the so-called correlated color temperature (CCT), which is one of the most important characteristics of white light.

Another important characteristic is the luminous efficacy of radiation (LER) obtained by convolution of the spectral power density spectrum $S_\lambda(\lambda)$ with the human eye sensitivity function $\Phi(\lambda)$ (Judd and Wyszecki, 1975), accounting statistically for the perception of human eye: $\mathrm{LER} = \int S_\lambda(\lambda)\,\Phi(\lambda)\,d\lambda$. The efficacy of a white light source is the product of LER and total WPE of the source, which depends on WPEs of individual LEDs and the power fractions of the emission spectra supplied by these LEDs into the total spectrum $S_\lambda(\lambda)$ (Bulashevich et al., 2015). If a phosphor emission is used for white light generation, its effective WPE depends on the WPE of the pumping LED, the Stokes shift caused by down conversion of the photon energy, and quantum yield of the phosphor emission (Bulashevich et al., 2015).

Ability of a light source to reproduce the colors of illuminated objects accurately is characterized by the color rendering index (CRI). The CRI accounts for the differences between the chromatic coordinates of light reflected from a number of standard color samples at their illuminating by the light source studied and by one or another standard illuminant. Detailed procedure of the CRI evaluation is recommended by CIE (1986). The CRI is reduced from the maximum value of 100 by every chromatic coordinate difference obtained from the selected color samples.

14.5.2 Light Conversion by Phosphor

Conversion medium in an LED lamp typically consists of a silicone layer with embedded particles of one or more phosphors deposited on the surface of pumping LED. Therefore, the light absorption and scattering by the phosphor particle, as well as emission of the down-converted light and its scattering by the particles, occurs in the medium. Here, spectral ray tracing is the approach most suitable for simulations of light conversion by phosphors. Assuming the phosphor particles to be randomly and uniformly distributed in silicone, one can derive the light conversion to depend on scattering properties of an individual particle and total volume fraction f_{ph} of the particles in the conversion medium.

Rigorous Mie theory (Van de Hulst, 1957) is currently the best model for analysis of light scattering by individual phosphor particles. Assuming spherical particle shape, the theory predicts the wavelength-dependent scattering and absorption cross-sections necessary for ray tracing simulations, as well as the scattering pattern in the far-field zone. For particles of realistic size and monochromatic light, the scattering patterns exhibit strong interference oscillations (black lines in Figure 14.8), which should be absent in

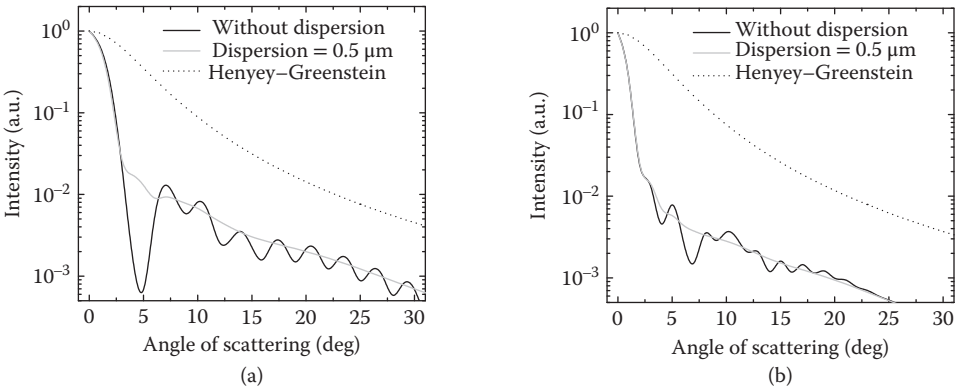

FIGURE 14.8 Scattering patterns of YAG:Ce^{3+} particle with the radius of 4 μm obtained by the Mie theory with and without account of particle size dispersion and by Henyey–Greenstein approximation for the wavelength of 600 nm (a) and 400 nm (b).

practice in view of (1) dispersion of the particle sizes, (2) deviation of the particle shapes from the spherical one, and (3) a nonmonochromatic spectra of the pumping LED and phosphor. Accounting for the size dispersion of phosphor particles is already sufficient to avoid the oscillations in the scattering pattern (gray lines in Figure 14.8). It is interesting that the Henyey–Greenstein scattering functions (Henyey and Greenstein, 1941) commonly used for empirical fitting of the scattering pattern do not work well in the case of realistic phosphor particles, if the asymmetry parameter g of the function is correctly estimated as the first moment of the scattering pattern predicted by the Mie theory (dotted lines in Figure 14.8). Being fitted to the computed pattern, the Henyey–Greenstein scattering function at g ~0.98–0.99 is capable of predicting the enhanced forward scattering seen in Figure 14.8 but fails to reproduce the wings of the scattering pattern predicted by the Mie theory. So, there is an unlikely alternative to the use of Mie theory for modeling the light interaction with the phosphor particles.

The absorption/scattering cross-sections and scattering pattern of individual phosphor particles involved in spectral ray tracing allow simulating the total emission spectrum in the far-field zone, estimating the color characteristics of white light, that is, chromatic coordinates x and y, CCT, and CRI, and their dependence on the observation angle. The latter dependence is quite important as the overall white light LED performance requires uniform color characteristics in a wide range of the observation of angle variation. As soon as the angle dependence of the total emission spectrum is obtained, integration over the angle enables calculation of LER and then of efficacy of the white light source.

One of the approaches aimed at improving angle uniformity of color characteristics is based on the fact that the scattering cross-section of phosphor particles depends very critically on the particle size. Figure 14.9 shows the simulated angle dependence of CCT and CRI of the white light source comprising of a blue (450 nm) LED and a yellow YAG:Ce^{3+} phosphor with the emission spectrum peaked at 560 nm and its full width at half maximum of 120 nm. Simulations were carried out for a square-shaped LED chip uniformly coated by a 230-nm silicone layer with 8% of its volume filled by the phosphor particles. One can see from Figure 14.9 that the angle uniformity of CCT and CRI is strongly dependent on the particle size. Among the chosen values of the mean particle radius, only that of 2 μm provides simultaneously the angle-uniform CCT and CRI.

The above results have been obtained by ray tracing performed in 50 spectral intervals from 380 to 730 nm, using 10^8 rays. Even at that, the simulated color characteristics in the far-field zone are rather noisy because of insufficient ray statistics. Therefore, parallel computing is normally required for the spectral ray tracing in order to accelerate simulations.

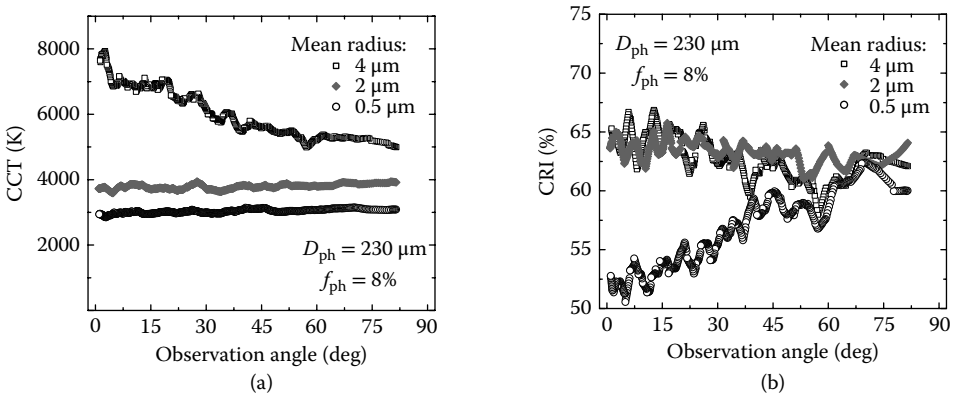

FIGURE 14.9 Angle-dependent correlated color temperature (CCT) (a) and color rendering index (CRI) (b) of a white light source consisting of a blue light-emitting diode (LED) and a YAG:Ce^{3+} yellow phosphor with the particles of various sizes. The conversion media thickness D_{ph} = 230 μm is assumed in the ray-tracing simulations.

14.6 Concluding Remarks

This chapter presents a general frame and a strategy for multiscale and multidisciplinary simulations of state-of-the-art LEDs from heterostructure to lamp. The basic models and the most important mechanisms underlying the simulations are reviewed with the focus on those that were not discussed in detail previously. It should be noted that many aspects of simulations are just outlined in the chapter and should be adopted to specific subjects of the study. In particular, the radiative and nonradiative recombination mechanisms require more detailed description that accounts for both direct and phonon-assisted processes, as well as the variety of active region designs (quasi-bulk, QWs, or QDs). Another issue beyond the scope of the chapter is the proper choice of materials properties, which are known to often be highly inaccurate, and its impact on the simulation results. Finally, the role of a number of material factors important for LED operation, such as high-density threading dislocations and compositional fluctuations in InGaN and other alloys, are still in the early stages of study. A better understanding of these factors may result in a reconsideration of the basic carrier transport and recombination models in the future.

References

Abakumov VN, Perel VI, Yassievich IN (1991) Nonradiative recombination in semiconductors (Ch. 11). In: *Modern Problems in Condensed Matter Science*, V. M. Agranovich and A. A. Maradudin (Eds.), vol. 33. Amsterdam: North-Holland.

Altieri P, Jaeger A, Windisch R, Linder N, Stauss P, Oberschmid R, Streubel K (2005) Internal quantum efficiency of high-brightness AlGaInP light-emitting devices. *J. Appl. Phys.* 98: 086101.

Avrutin V, Hafiz S-D-A, Zhang F, Özgür Ü, Morkoç H (2013) InGaN light-emitting diodes: Efficiency-limiting processes at high injection. *J. Vac. Sci. Technol. A* 31: 050809.

Binder M, Nirschl A, Zeisel R, Hager T, Lugauer H-J, Sabathil M, Bougeard D, Wagner J, Galler B (2013) Identification of nnp and npp Auger recombination as significant contributor to the efficiency droop in (GaIn)N quantum wells by visualization of hot carriers in photoluminescence. *Appl. Phys. Lett.* 103: 071108.

Bogdanov MV, Bulashevich KA, Evstratov IY, Zhmakin AI, Karpov SY (2008) Coupled modeling of current spreading, thermal effects, and light extraction in III-nitride light-emitting diodes. *Semicond. Sci. Technol.* 23: 125023.

Bogdanov MV, Bulashevich KA, Khokhlev OV, Evstratov IY, Ramm MS, Karpov SY (2010) Current crowding effect on light extraction efficiency of thin-film LEDs. *Phys. Status Solidi C* 7: 2124–2126.

Bulashevich KA, Karpov SY (2008) Is Auger recombination responsible for the efficiency rollover in III-nitride light-emitting diodes? *Phys. Status Solidi C* 5: 2066–2069.

Bulashevich KA, Kulik AV, Karpov SY (2015) Optimal ways of colour mixing for high-quality white-light LED sources. *Phys. Status Solidi A* 212: 914–919.

Chernyakov AE, Bulashevich KA, Karpov SY, Zakgeim AL (2013) Experimental and theoretical study of electrical, thermal, and optical characteristics of InGaN/GaN high-power flip-chip LEDs. *Phys. Status Solidi A* 210: 466–469.

Cho J, Schubert EF, Kim JK (2013) Efficiency droop in light-emitting diodes: Challenges and countermeasures. *Laser Photonics Rev.* 7: 408–421.

CIE Publ. No. 15.2 (1986) Colorimetry. *Official Recommendations of the International Commission on Illumination*. Vienna: Central Bureau of the CIE.

Dyakonov MI, Kachorovskii VY (1994) Nonthreshold Auger recombination in quantum wells. *Phys. Rev. B*, 49: 17130–17138.

Galler B, Lugauer H-J, Binder M, Hollweck R, Folwill Y, Nirschl A, Gomez-Iglesias A, Hahn B, Wagner J, Sabathil M (2013) Experimental determination of the dominant type of Auger recombination in InGaN quantum wells. *Appl. Phys. Express* 6: 112101.

Gelmont B, Sokolova Z (1982) Auger recombination in direct-gap N-type semiconductors. *Sov. Phys. Semicond.* 16: 1067–1068.

Gil B (2009) Wurtzitic semiconductors heterostructures grown on (hk.l) oriented substrates: The interplay between spontaneous and piezoelectric polarization fields, elastic energy and the modification of quantum confined Stark effect. *Proc. SPIE* 7216: 7216OE.

Henyey LG, Greenstein JL (1941) Diffuse radiation in the galaxy. *Astrophys. J.* 93: 70–83.

Iveland J, Martinelli L, Peretti J, Speck JS, Weisbuch C (2013) Direct measurement of Auger electrons emitted from a semiconductor light-emitting diode under electrical injection: Identification of the dominant mechanism for efficiency droop. *Phys. Rev. Lett.* 110: 177406.

Jacoboni C (2010) *Theory of Electron Transport in Semiconductors*. Berlin: Springer.

Judd DB, Wyszecki G (1975) *Color in Business, Science and Industry*. New York, NY: Wiley.

Jüngel A (2009) *Transport Equations for Semiconductors*. Berlin: Springer.

Kane EO (1957) Band structure of indium antimonide. *J. Phys. Chem Solids* 1: 249–261.

Karpov SY, Makarov YN (2002) Dislocation effect on light emission efficiency in gallium nitride. *Appl. Phys. Lett.* 81: 4721–4723.

Kioupakis E, Rinke P, Delaney KT, Van de Walle CG (2011) Indirect Auger recombination as a cause of efficiency droop in nitride light-emitting diodes. *Appl. Phys. Lett.* 98: 161107.

Laubsch A, Sabathil M, Bergbauer W, Strassburg M, Lugauer H, Peter M, Lutgen S, Linder N, Streubel K, Hader J, Moloney JV, Pasenow B, Koch SW (2009) On the origin of IQE-"droop" in InGaN LEDs. *Phys. Status Solidi C* 6: S913–S916.

Laubsch A, Sabathil M, Baur J, Peter M, Hahn B (2010) High-power and high-efficiency InGaN-based light emitters. *IEEE Trans. Electron Devices* 57: 79–87.

López T, Margalith T (2008) Electro-thermal modelling of high power light emitting diodes based on experimental device characterisation. www.comsol.de/paper/5004.

Lundstrom M (2000) *Fundamentals of Carrier Transport*. Cambridge: Cambridge University Press.

Malyutenko VK, Bolgov SS, Podoltsev AD (2010) Current crowding effect on the ideality factor and efficiency droop in blue lateral InGaN/GaN light emitting diodes. *Appl. Phys. Lett.* 97: 251110.

Mickevičius J, Tamulaitis G, Jurkevičius J, Shur MS, Shatalov M, Yang J, Gaska R (2015) Efficiency droop and carrier transport in AlGaN epilayers and heterostructures. *Phys. Status Solidi B* 252: 961–964.

Nye JF (1964) *Physical Properties of Crystals: Their Representation by Tensors and Matrices* (Ch. 10). Oxford: Clarendon Press.

Pankove JI (1971) *Optical Processes in Semiconductors*. New York, NY: Dover Publications.

Park S-H, Chuang S-L (1999) Crystal-orientation effects on the piezoelectric field and electronic properties of strained wurtzite semiconductors. *Phys. Rev. B* 59: 4725–4737.

Piprek J (2010) Efficiency droop in nitride-based light-emitting diodes. *Phys. Status Solidi A* 207: 2217–2225.

Querlioz D, Dollfus P (2010) *The Wigner Monte Carlo Method for Nanoelectronic Devices*. London: ISTE Ltd. & John Wiley & Sons.

Romanov AE, Baker TJ, Nakamura S, Speck JS (2006) Strain-induced polarization in wurtzite III-nitride semipolar layers. *J. Appl. Phys.* 100: 023522.

Royo P, Stanley RP, Ilegems M, Streubel K, Gulden KH (2002) Experimental determination of the internal quantum efficiency of AlGaInP microcavity light-emitting diodes *J. Appl. Phys.* 91: 2563–2568.

Shen YC, Müller GO, Watanabe S, Gardner NF, Munkholm A, Krames MR (2007) Auger recombination in InGaN measured by photoluminescence. *Appl. Phys. Lett.* 91: 141101.

Van de Hulst HC (1957) *Light Scattering by Small Particles*. New York, NY: Dover Publications.

Van Roosbroeck W (1950) Theory of flow of electrons and holes in germanium and other semiconductors. *Bell System Tech. J.* 29: 560–607.

Verzellesi G, Saguatti D, Meneghini M, Bertazzi F, Goano M, Meneghesso G, Zanoni E (2013) Efficiency droop in InGaN/GaN blue light-emitting diodes: Physical mechanisms and remedies. *J. Appl. Phys.* 114: 071101.

Weisbuch C, Piccardo M, Martinelli L, Iveland J, Peretti J, Speck JS (2015) The efficiency challenge of nitride light-emitting diodes for lighting. *Phys. Status Solidi A* 212: 899–913.

Zegrya GG, Kharchenko VA (1992) New mechanism of Auger recombination of nonequilibrium current carriers in semiconductor heterostructures. *Sov. Phys. JETP* 74: 173–181.

Zegrya GG, Samosvat DM (2007) Mechanisms of Auger recombination in semiconducting quantum dots. *J. Exp. Theor. Phys.* 104: 951–965.

Zhmakin AI (2011) Enhancement of light extraction from light emitting diodes. *Phys. Rep.* 498: 189–241.

15

Organic Light-Emitting Diodes

Pascal Kordt
Peter Bobbert
Reinder Coehoorn
Falk May
Christian Lennartz
and
Denis Andrienko

15.1 Introduction

The 2014 Nobel Prize in Physics was awarded "for the invention of efficient blue light-emitting diodes (LEDs) which has enabled bright and energy-saving white light sources" [1–3], setting a clear target to mankind: energy-efficient and environmental-friendly light sources [4]. The first *organic* LED (OLED) was reported in 1987 by a team at Kodak [5]. This publication, cited to date more than 10,000 times, stipulated the entire field of organic electronics. Shortly afterward, a polymer LED (PLED) was demonstrated [6],

paving the way for flexible lighting applications [7]. Nowadays, OLEDs are successfully used in displays of mobile phones and televisions: In 2008, Samsung announced a flexible display that was only 50 μm thick [8], about half the thickness of a sheet of paper. A prototype of an OLED display for the automotive market was presented recently by Continental [9]: In OLED displays, black pixels are completely switched off, allowing the driver's eye to adapt better to the darkness. Contrary to liquid crystal displays (LCDs), OLED screens do not require backlight illumination, yielding exceptionally good contrast ratios and reduced power consumption. OLED displays also provide viewing angles and response times superior to LCDs and are, in general, thinner and lighter. Last but not least, many organic materials can be printed from solution, enabling cost-effective large-scale manufacturing on mechanically flexible films.

The prime challenge in OLED development is an improvement of device operation lifetimes. In many cases, the steady decrease of luminescence efficiency of OLEDs under continuous operation is compensated by a gradual rise in bias voltages, or by using larger pixel sizes, which can then be operated at lower voltages and luminances. In particular, blue phosphorescent devices, with their high energy of emitted photons and long-lived excited states, are prone to rapid material degradation. The underlying mechanisms of degradation, discussed in Section 15.12.5, are unfortunately not fully understood.

Designing new organic materials is crucial for tuning OLED properties, performance, and stability. One, therefore, needs to provide rigorous links between device characteristics and the chemical composition of the layers, or structure–property relationships. This can be done with the help of cheminformatics tools [10], i.e., by correlating available experimental data to underlying chemical structures. However, the number of experimental samples is normally limited to a few hundred since synthesis, device optimization, and characterization are often costly and time consuming. A substantial extension of the training set is therefore experimentally not feasible to achieve, again motivating the development of computer simulations techniques capable of predicting device characteristics *in silico* [11]. In this chapter, we aim to provide an overview of theoretical models and simulation approaches developed to study charge and exciton transport in organic semiconductors, as well as to predict and eventually optimize current–voltage characteristics, electroluminescence efficiency, and lifetimes of OLEDs.

15.2 Working Principles of an OLED

We start by reviewing the elementary processes taking place in an OLED. The simplest two-layer OLED consists of hole and (luminescent) electron transporting layers (ETLs), which are sandwiched between two electrodes, as shown in Figure 15.1a. When a voltage is applied to the electrodes, holes are injected from the anode into the hole transporting layer (HTL) and electrons from the cathode into the ETL. The external field forces electrons and holes to drift toward the interface between these layers, where they recombine and emit light. Energetically, such a heterojunction is designed to facilitate the hole injection from HTL to ETL as well as to block electron diffusion into the opposite direction.

To fine-tune device properties such as luminescence, driving voltage, light outcoupling, and lifetime, more layers are added to a two-layer OLED: A typical phosphorescent OLED is shown in Figure 15.1b. Here, when a voltage is applied to the electrodes, electrons are injected from the reflective metal cathode (Al) and holes from the semitransparent anode (indium tin oxide or ITO). The cathode has a low work function, while the anode has a high one. This energy difference is compensated by the externally applied voltage, which forces charge carriers to drift into the emission layer. In the emission layer, holes and electrons form excitons, predominantly on the emitter (guest) molecules. Finally, radiative decay of an exciton leads to light emission. In the next sections, we briefly review the functionality of all layers.

15.2.1 Electrodes

The outermost layers, metallic or metal-oxide electrodes, inject electrons and holes into organic layers. A common material for the semitransparent anode is ITO [12]. Its low ionization potential (IP) ensures

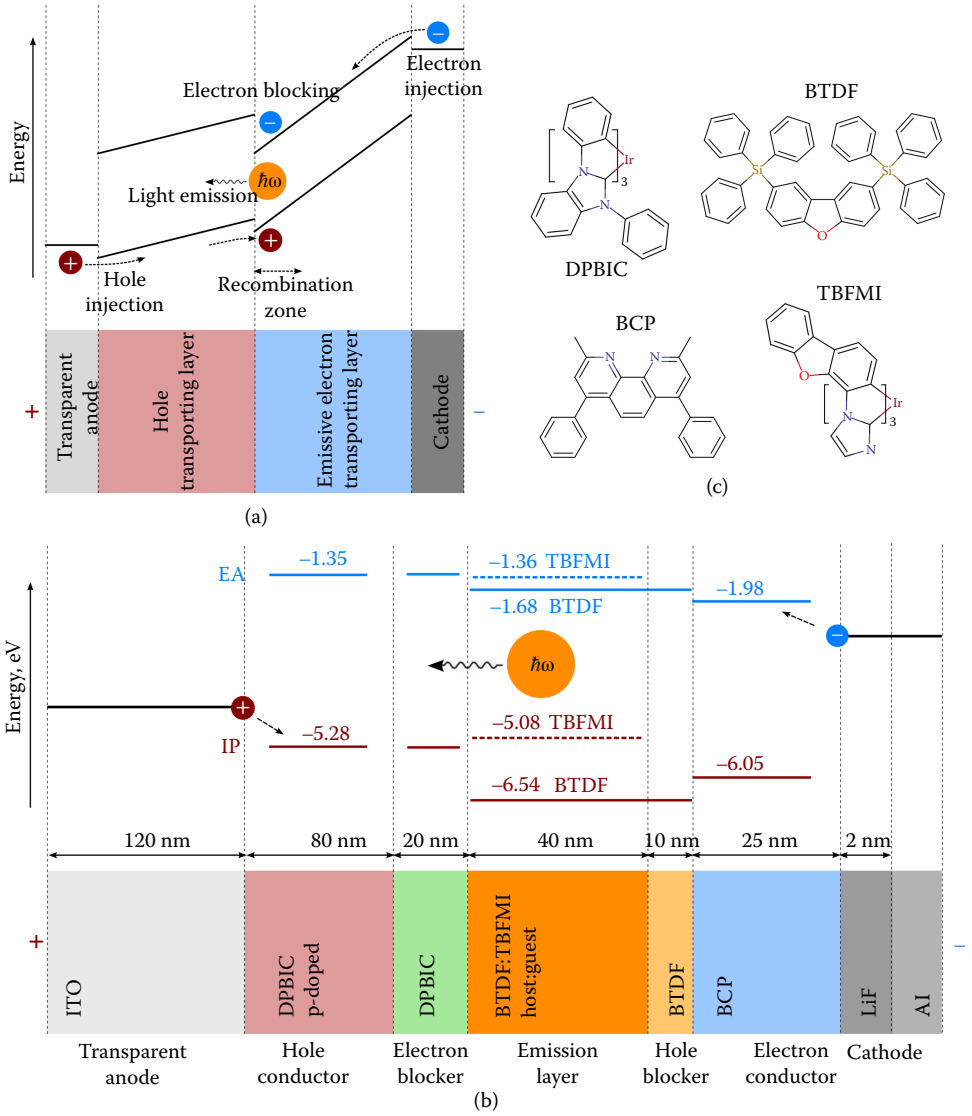

FIGURE 15.1 (a) Schematic of a two-layer electroluminescent organic light-emitting diode (OLED) and its energy level diagram under applied voltage. (b) Structure of a blue phosphorescent OLED with a transparent anode (indium tin oxide or ITO), a reflective cathode (Al), emission layer, and hole/electron conduction/blocking layers. Indicated energy levels are without the applied bias. (c) Compounds used in the OLED (b). (Adapted from P. Kordt et al., *Adv. Funct. Mater.*, 25, 1955–1971, 2015.)

efficient hole injection into the organic layer. For efficient electron injection a low work function cathode is required. Since such materials are often air-unstable, normally bilayered structures are used [13], e.g., Al/LiF [14,15] or Ag/MgAg [16,17].

15.2.2 Charge Transport Layers

The role of electron and hole transport layers (HTLs) is to provide an Ohmic contact to the electrodes (barrierless injection) and to help control light outcoupling. These layers are often doped. By inserting

electron donating or accepting impurities into a material, their intrinsic charge carrier density is increased, leading to higher conductivities and a reduction of the injection barrier. While it is a standard technique in inorganic semiconductors, it can be challenging in organic materials: Dopant molecules create energetic traps, hindering the formation of mobile charges [18].

Realizing *p*-type doping, which increases the hole density, requires a dopant material with a lowest unoccupied molecular orbital (LUMO) level close to the highest occupied molecular orbital (HOMO) of the host. This is feasible in most cases and, consequently, it is used in many OLEDs. For example, the OLED stack in Figure 15.1b has a *p*-doped Tris[(3-phenyl-1H-benzimidazol-1-yl-2(3H)-ylidene)-1,2-phenylene]Ir (DPBIC) HTL. On the contrary, *n*-type doping requires a dopant material with an HOMO level near or above the LUMO of the host. The necessity to find air-stable dopant materials with such a high HOMO level makes *n*-type doping challenging. Moreover, when using small molecules such as O_2, Br_2, and I_2 as dopant materials, these can diffuse into an organic host, leading to undesired doping profiles. This effect, however, can also be used to actually dope the adjacent layer with an interlayer of dopant molecules. In the stack shown in Figure 15.1b this technique is applied: Here an LiF interlayer is placed between the aluminum cathode and the electron transport layer (ETL) doping the latter [19]. Alq_3 (aluminum-tris(8-hydroxychinolin)) is a material that can be used for both the hole- and electron-conducting layers. In the stack in Figure 15.1b DPBIC is used for hole conduction and bathocuproine-4,7-diphenyl-2,9-dimethyl-1,10-phenanthroline (BCP) serves as an electron-conducting layer [14,16,17,20,21].

15.2.3 Blocking Layers

Hole/electron blocking layers suppress charge flow to the opposite electrodes, enhancing their recombination probability in the emission layer. The hole blocking layer has an IP lower than the IP of the emission layer. Similarly, an electron blocking layer has an electron affinity (EA) higher than the EA of the emission layer. In Figure 15.1, DPBIC is used as an electron blocker and BCP as a hole blocker [14,16,17,20,21].

15.2.4 Emission Layer

In the emission layer, electrons and holes recombine, leading to exciton formation. Upon radiative decay of the excitons, photons are emitted. To avoid exciton quenching, emitters (host molecules) are embedded into a matrix (guest molecules). In the stack shown in Figure 15.1b, Tris[(1,2-dibenzofurane-4-ylene)(3-methyl-1/1-imidazole-1-yl-2(3/1)-ylidene)]Ir(III) (TBFMI) is a blue phosphorescent emitter, while 8-bis(triphenylsilyl)-dibenzofuran (BTDF) is the host material.

Exciton generation and the emission of photons proceeds in several steps, which are illustrated in Figure 15.2. First, a carrier (in this case a hole) is trapped on an emitter molecule and an oxidized complex is formed. Driven by Coulomb forces, a carrier of opposite charge (electron) moves on host molecules toward the trapped carrier. When the electron reaches a host molecule' neighboring oxidized emitter, a charge transfer (CT) state is formed, as shown in Figure 15.2b. In a CT state, the short-range exchange interaction leads to an energy splitting of singlet (S) and triplet (T) states. The triplet consists of three triplet substates, which differ from one another by their relative spin orientations. Statistically, one obtains a population ratio of 1:3 between singlet and triplet substates. In a final step, the electron moves in a very rapid process directly to the emitter molecule, forming an excited emitter, see Figure 15.2c. This process may occur via a singlet or a triplet path, depending on the initial spin orientation of the electron–hole pair. The two spins of an electron and a hole are then coupled to four new combined states: one singlet state, with total spin momentum 0, and one triplet state, with the total spin momentum 1. In a statistical limit, all four substates are populated with an equal probability.

An exciton formed on an emitter molecule can decay radiatively by fluorescence or phosphorescence. In *fluorescence*, which is the most common radiative decay pathway for organic molecules, the radiative decay occurs from the excited singlet state, S_1, to the ground state. If triplet states are not harvested, the internal quantum efficiency of an OLED is limited to 25%. *Phosphorescence*, conversely, allows harvesting of

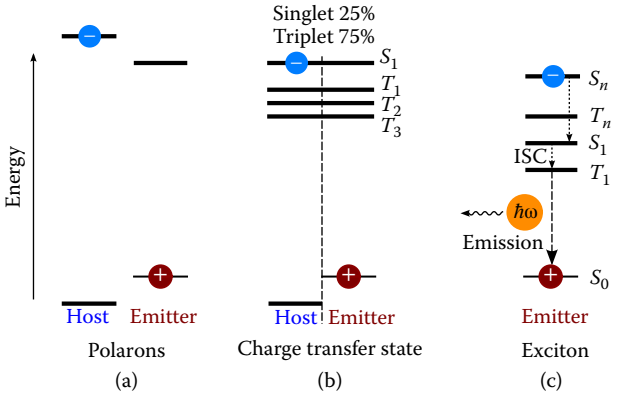

FIGURE 15.2 Exciton formation in the emission layer of a phosphorescent organic light-emitting diode (OLED). (a) Electron moves toward a hole which is trapped on the emitter molecule. (b) Charge transfer state is formed with the electron on a host and hole on a neighboring emitter molecule. Spins of hole and electron are already correlated to one singlet and three triplet states. (c) Exciton is formed on the emitter molecule with a statistical probability to be in a 25% singlet and 75% triplet state. Singlet state undergoes fast intersystem crossing (ISC) into a triplet state which then emits light.

triplet states. One can use compounds with large spin–orbit couplings, such as organometallic complexes, to achieve this. The large spin–orbit coupling capacitates a spin flip, (see Figure 15.2c), leading to ISC, i.e., a transition between singlet and triplet states. Thus, triplets can also decay radiatively ($T_1 \rightarrow S_0$), yielding an internal quantum efficiency of 100%.

Another way of harvesting triplet states is via a *thermally activated delayed fluorescence* (TADF) [22]. TADF can be realized in organic molecules even with small spin–orbit coupling. In TADF triplets are converted to singlets thermally (reverse ISC) and a radiative $S_1 \rightarrow S_0$ decay (fluorescence) takes place. TADF requires small energy differences between T_1 and S_1 levels $\Delta E_{ST} \sim k_B T$. To achieve this, a large spatial HOMO–LUMO separation is required. This, however, competes with the requirement of large oscillator strengths of fluorescence, i.e., a large transition dipole moment [23].

Several processes in an OLED can lead to efficiency losses. In a phosphorescent OLED, these are *triplet–triplet quenching*, where two triplets are annihilated as a result of the interaction of neighboring excited emitter molecules, and *triplet–polaron quenching*, i.e., an interaction between electronic and vibrational modes, energy transfer to polarons, and dissociation into free carriers [12,24]. These processes are discussed in more detail in Sections 15.12.3 and 15.12.5.

A comprehensive OLED model should incorporate all the aforementioned elementary processes: charge injection from electrodes, charge trapping, and transport to the emission layer, long-range electron–hole interactions, formation of CT and excited states, exciton–exciton and exciton–polaron interactions, and radiative decay of excitons. The ultimate goal of such a model is to predict the current–voltage–luminescence characteristics of the multilayered OLED structure. Since the typical thickness of an organic layer of an OLED stack ranges from 10 to 100 nm, it is computationally prohibitive to use only first-principles methods for OLED modeling. State-of-the-art OLED simulations employ either continuous models, such as drift–diffusion equations (see Section 15.3), or solve the master equation for charge/exciton occupation probabilities (see Section 15.4). Both approaches use models of different complexity. For example, if the target is to optimize the composition of the stack made of well-characterized organic layers, drift–diffusion equations and lattice models with phenomenologically fitted parameters are used. If one also needs to retain the link to the underlying chemical composition of layers, off-lattice models based on first-principles parametrizations are required. In all cases, we are dealing with a typical multiscale problem, which requires the development of scale-bridging techniques (see Section 15.11). In the next sections, we review these approaches, paying special attention to limitations, parametrizations, and scale-bridging issues.

15.3 Drift–Diffusion Equations

We start with the most coarse model: On a macroscopic level, the drift and diffusion of electrons, holes, and excitons in an OLED stack can be described by the corresponding densities, denoted here as n, p, and s, respectively. The drift–diffusion model assumes that local charge densities, charge mobilities, μ_n, μ_p, diffusion constants, D_e, D_n, and field strength, F, all vary continuously in space. The set of equations describing the time-dependent spatial distribution of charge and exciton densities is then based on the respective conservation laws. In what follows, we describe a model where charges drift-diffuse in a given density of states (DOS) $g(E)$ (see also Section 15.7), can be trapped, have density- and field-dependent mobilities, and recombine either radiatively or nonradiatively.

15.3.1 Charges

Charge conservation leads to two continuity equations, for electrons and holes

$$e\frac{\partial n}{\partial t} = \vec{\nabla}\cdot\vec{J}_n - eR(n,p) - e\frac{\partial n_\mathrm{t}}{\partial t}, \qquad e\frac{\partial p}{\partial t} = -\vec{\nabla}\cdot\vec{J}_p - eR(n,p) - e\frac{\partial p_\mathrm{t}}{\partial t}, \tag{15.1}$$

where $R(n,p)$ is the recombination rate (see Section 15.3.2), n_t and p_t are densities of trapped charges, and e is the electron charge. The current equations for electron and holes drift–diffusion in an electrostatic potential ψ read [25,26]

$$\vec{J}_n = -en\mu_n\vec{\nabla}\psi + eD_n\vec{\nabla}n, \qquad \vec{J}_p = -ep\mu_p\vec{\nabla}\psi - eD_p\vec{\nabla}p. \tag{15.2}$$

Note that only mobile charges contribute to the current. Summing up electron and hole currents yields the total current in the device, $\vec{J} = \vec{J}_n + \vec{J}_p$. The electrostatic potential ψ is related to the electron and hole densities via the Poisson equation

$$\varepsilon_0\varepsilon_r\Delta\psi = e\left(n - p + n_\mathrm{t} - p_\mathrm{t}\right), \tag{15.3}$$

where ε_0 is the vacuum and ε_r is the relative permittivity. The densities of trapped charges obey the phenomenological rate equations

$$\frac{\partial n_\mathrm{t}}{\partial t} = r_{\mathrm{c},n}n(N_t - n_\mathrm{t}) - r_{\mathrm{e},n}n_t, \qquad \frac{\partial p_\mathrm{t}}{\partial t} = r_{\mathrm{c},p}p(N_t - p_\mathrm{t}) - r_{\mathrm{e},p}p_t, \tag{15.4}$$

where it is assumed that traps occupy two levels (one for holes and one for electrons). Here N_t is the trap density, r_e is the escape rate, and r_c is the capture rates of electrons (n) and holes (p), respectively.

Since two electrons or holes cannot occupy the same energy level at the same time, the occupation probabilities of energy levels follow the Fermi–Dirac statistics [27]

$$f(E,E_\mathrm{F}) = \left[1 + \exp\left(\frac{E - E_\mathrm{F}}{k_\mathrm{B}T}\right)\right]^{-1}, \tag{15.5}$$

where $f(E,E_\mathrm{F})$ is the occupation probability of a level with energy E, k_B is the Boltzmann constant, and T is the temperature. The carrier density is then related to the quasi-Fermi level, E_F, as

$$p\left(E_\mathrm{F}\right) = \frac{N}{V}\int_{-\infty}^{\infty} g(E)f(E,E_\mathrm{F})\mathrm{d}E, \tag{15.6}$$

where $g(E)$ is the DOS (see Section 15.7), N is the number of holes, and V is the box volume. A similar relation holds also for electrons. Fermi–Dirac statistics implies that the diffusion coefficient and mobility

in Equation 15.1 are related via the generalized Einstein relation [28]

$$D = \frac{p\mu}{e}\left(\frac{\partial p}{\partial E_{\mathrm{F}}}\right)^{-1}. \tag{15.7}$$

In order to solve the drift–diffusion equations one needs to know the dependence of mobility on charge density, electric field, and temperature, $\mu(p, n, \vec{F}, T)$. These can be obtained either from experiments or Monte Carlo simulations, as discussed in Section 15.11.

15.3.2 Recombination

The recombination rate $R(n, p)$ includes two loss mechanisms: Shockley–Read–Hall (SRH), or trap-assisted recombination, and bimolecular recombination. In SRH recombination, electrons are trapped in low-energy states and recombine with free holes. The macroscopic rate for this process reads [29]

$$R_{\mathrm{SRH}} = \frac{C_n C_p N_{\mathrm{t}}(np - n_i p_i)}{C_n(n + n_i) + C_p(p + p_i)}. \tag{15.8}$$

Here C_n and C_p are the capture coefficients for electrons and holes, respectively, N_{t} is the trap density for electrons, n and p are the electron and hole densities, and n_i and p_i are the intrinsic electron and hole densities.

Bimolecular recombination is often modeled using the Langevin rate

$$R_{\mathrm{L}} = \gamma(np - n_i p_i), \qquad \gamma = \frac{q}{\varepsilon_0 \varepsilon_{\mathrm{r}}}\left(\mu_n + \mu_p\right). \tag{15.9}$$

Here γ is the recombination constant and μ_n and μ_p are electron and hole mobilities. The intrinsic electron and hole carrier densities are often neglected [29,30], simplifying the expression to

$$R_{\mathrm{L}} \approx \gamma np. \tag{15.10}$$

There are several assumptions on which the Langevin rate is based and which might not hold in the case of organic semiconductors. Its validity can be verified by performing kinetic Monte Carlo (KMC) simulations on lattices, as discussed in Section 15.6.2.

The overall recombination rate, $R_{\mathrm{SRH}} + R_{\mathrm{L}}$, enters the continuity Equation 15.1. Experimentally, the prevailing recombination mechanism can be probed by means of the classical Shockley diode equation with an ideality factor η that differs depending on the dominant mechanism [31,32].

15.3.3 Excitons

The generation, transport, and decay of excitons can be described with a phenomenological equation for the population of excitons S_i, where i denotes the exciton type (triplet, singlet) [26]

$$\frac{\partial S_i}{\partial \mathrm{t}} = G_i R(n, p) + \vec{\nabla} \cdot \vec{J}_i - \left(k_i^{(r)} + k_i^{(n)}\right) S_i - k_i^{(a)} f S_i^2 + \sum_{j=S,T} (k_{ji} S_j - k_{ij} S_i) - k_{\mathrm{TPQ}}(n + p). \tag{15.11}$$

Here G_i is the exciton generation efficiency: For singlet excitons $G_{\mathrm{S}} = 0.25$, while for triplet excitons $G_{\mathrm{T}} = 0.75$. $k^{(r)}$ and $k^{(n)}$ are the radiative and nonradiative (position-dependent) decay rates, and $k^{(a)}$ is the annihilation rate. The factor f is 0.5 if only one triplet is lost or 1 if both are lost. The radiative decay rate can be calculated using the dipole emission model [26,33]. The exciton energy transfer rate, k_{ij}, describes the conversion of triplet excitons to singlets and vice versa. k_{TPQ} is the triplet–polaron quenching rate [24] (assumed here to be the same for electrons and holes). Since excitons are charge-neutral, their transport

is purely diffusive, i.e., $\vec{J}_i = D_i \vec{\nabla} S_i$. Note that the exciton dissociation into an electron and a hole has been neglected in Equation 15.11.

15.3.4 Boundary Conditions and Numerical Solution

Equations 15.1 through 15.11 are complemented by the boundary conditions for the electrostatic potential, ψ, by setting the potential difference at the boundaries to $\psi_{\mathrm{eff}} = V_{\mathrm{app}} - V_{\mathrm{int}}$, where V_{app} is the applied potential and V_{int} is the built-in potential, defined as the difference of the materials' work functions.

The set of Equations 15.1 through 15.11 is normally solved using iterative schemes, until self-consistency for the electrostatic potential, density, and current is reached. The Gummel iteration method [34] with a discretization according to a scheme proposed by Scharfetter and Gummel [35] can be used to solve linearized equations. This method is less sensitive to the initial guess than a Newton algorithm and thus is the method of choice despite its slower convergence in terms of iteration steps [36].

15.3.5 Limitations

Field, temperature, and charge density dependencies of the mobility entering drift–diffusion equations are normally parametrized in equilibrium or under stationary conditions. These dependencies cannot in general be used to describe nonequilibrium, time-dependent processes. For example, charge-carrier relaxation makes the mobility effectively time dependent, as evident from, e.g., impedance and dark-injection studies (see Section 15.12.2). Moreover, it has been shown that the current density in the device can be spatially inhomogeneous in all three dimensions (filamentary structures) [37–39], which then questions the applicability of mean-field descriptions. Last, some processes, e.g., exciton–electron interactions or energetic barriers between organic layers, are difficult to incorporate into drift–diffusion equations. The master equation, which we review in the next section, allows us to consider the individual rates between CT and other events. In the subsequent section, we describe the Monte Carlo technique that also allows us to include excitonic processes.

15.4 Master Equation

In inorganic, crystalline semiconductors charges are delocalized and energy eigenvalues of the electronic Hamiltonian form smooth bands. Consequently, one deals with band-like charge transport, where mobility decreases with increasing temperature. In contrast, in amorphous organic semiconductors, molecules are weakly bound by van der Waals forces and, consequently, intermolecular electronic couplings are small. As a result, excess electrons are localized on single molecules or their fragments. Moreover, orientational and positional *energetic* disorder helps further localize charged excitations. In this situation, charge transport can be modeled as thermally activated charge *hopping* between neighboring molecules. Mathematically, this is described by a Poisson process: The probability of an event to happen depends only on the time interval (no memory) and events do not occur simultaneously. The corresponding master equation, which describes the time evolution of a system with a discrete set of states i, then reads

$$\frac{\mathrm{d}P_i(t)}{\mathrm{d}t} = \sum_{j \neq i} \Big[w_{ij} P_i(t) - w_{ji} P_j(t) \Big], \tag{15.12}$$

where $P_i(t)$ is the probability of finding the system in state i at time t and w_{ij} is the transition rate from state i to j. In most situations, the dimension of the state space is too large to use direct numerical differential equation solvers to solve for P_i. In special cases, e.g., for single charge carrier transport or transport of many carriers in a mean-field approximation, this equation can be rewritten in terms of site occupation probabilities, p_i.

15.4.1 Single-Carrier Charge Transport

If only one charge carrier is present in the system, the state of the system is fully determined by the position of the charge or the index of the molecule which this charge occupies. In this situation, the occupation probability of the system state, P_i, is equivalent to the site (molecule) occupation probability, p_i. The rates for transitions between states are then given by CT rates, ω_{ij} (see Section 15.6), and the master equation can be rewritten in terms of site occupation probabilities

$$\frac{\mathrm{d}p_i(t)}{\mathrm{d}t} = \sum_{j\neq i}\Big[\omega_{ij}p_i(t) - \omega_{ji}p_j(t)\Big]. \tag{15.13}$$

Equation 15.13 is a set of linear equations of size N, where N is the number of molecules in the system. It can be solved using standard numerical differential equation solvers [40]. In special cases, e.g., for one-dimensional charge transport, it is even possible to obtain an analytic solution [41–43]. Alternatively, one can use the KMC algorithm, which is discussed in Section 15.5.

15.4.2 Finite Charge-Carrier Densities

In experimentally relevant conditions, charge densities have finite values. The system state is now given by a vector of indices of occupied molecules and the number of states increases dramatically. It is, however, still possible to reformulate the master Equation 15.12 in terms of site occupation probabilities by using a mean field approximation [44,45]. The master equation then reads

$$\frac{\mathrm{d}p_i}{\mathrm{d}t} = \sum_{j\neq i}\Big[\omega_{ij}p_i\left(1-p_j\right) - \omega_{ji}p_j\left(1-p_i\right)\Big]. \tag{15.14}$$

The resulting equation is no longer linear in p_i and thus requires more involved numerical solvers. In most cases, where one deals with many carriers the KMC method (see Section 15.5) becomes more practical: It does not rely on the mean field approximation but solves the original master equation, and it can easily be extended to other processes, such as exciton transport, triplet–triplet annihilation, etc.

15.4.3 Mobility and Diffusion Constant

The stationary solution of the master equation for a system in an external field $\vec{F}$ allows to evaluate both charge carrier mobility and diffusion constant. For single-carrier transport, the mobility tensor $\hat{\mu} = \vec{v}\otimes\vec{F}/F^2$ reads

$$\mu_{\alpha\beta} = F^{-2}\sum_{ij}\omega_{ij}p_i r_{ij,\alpha}F_\beta, \tag{15.15}$$

where $r_{ij,\alpha} = r_{i,\alpha} - r_{j,\alpha}$ and $\vec{v}$ is the average velocity of a charge carrier [45]. For finite charge carrier density $\rho = N/V$, where N is the number of carriers and V is the volume of the box, Equation 15.15 takes the form

$$\mu_{\alpha\beta} = \frac{1}{\rho F^2 V}\sum_{ij}\omega_{ij}p_i\left(1-p_j\right)r_{ij,\alpha}F_\beta. \tag{15.16}$$

In a similar fashion, occupation probabilities can be used to calculate charge and current distributions in the system [45]. Alternatively, one can directly analyze charge trajectories of KMC simulations and evaluate mobility tensor as

$$\hat{\mu} = \frac{\langle\vec{v}\rangle\otimes\vec{F}}{F^2}, \qquad \text{with} \qquad \langle\vec{v}\rangle = \frac{\Delta\vec{r}}{\Delta t}. \tag{15.17}$$

Here $\vec{F}$ is the external electric field, $\vec{r}$ is the charge position, and Δt is the simulation time. The diffusion tensor can also be calculated directly from the charge trajectory generated without the applied external field [46]

$$D_{\alpha\beta} = \frac{\left\langle \Delta r_\alpha \Delta r_\beta \right\rangle}{2\tau}, \tag{15.18}$$

where $\Delta r_\alpha(t) = r_\alpha(t+\tau) - r_\alpha(\tau)$, or can be obtained from charge mobility with the help of the generalized Einstein relation, Equation 15.7.

15.5 Kinetic Monte Carlo

As we saw in the previous section, rewriting the master Equation 15.12 to a form tractable for large systems (i.e., in terms of site occupation probabilities) requires certain approximations. Moreover, this step becomes less and less straightforward if events other than CT are included, e.g., electron–hole recombination, exciton splitting, exciton decay, or transfer. For these reasons, a different way of solving the master equation becomes more practical. Here, system dynamics, or trajectories in the phase space, or Markov chains, are generated explicitly by using the so-called KMC method.

KMC, also known as dynamic Monte Carlo, Gillespie algorithm, residence-time algorithm, *n*-fold way, or the Botz–Kalos–Lebowitz algorithm, was initially developed by Doob [47,48]. The time update as it is used here was first introduced by Young and Elcock [49]. Bortz, Kalos, and Lebowitz developed the same algorithm independently and applied it to the Ising model [50]. Gillespie provided a physics-based derivation of the algorithm [51,52], which was then improved in terms of computational efficiency by Fichthorn [53] and Jansen [54]. The version we describe here is known as a variable step size method (VSSM) [54].

15.5.1 Variable Step Size Method

The VSSM allows us to group certain events and treat the groups and events in the group hierarchically. In the case of charge transport with multiple charge carriers, e.g., a two-level approach can be used. First, a charge is selected with the probability proportional to its escape rate ω_i, that is the sum of the rates of all possible moves of this charge away from the occupied site

$$\omega_i = \sum_{j=1}^{m} \omega_{ij}, \tag{15.19}$$

where m is the coordination number, i.e., the number of neighbors to which the site is connected, and ω_{ij} is the rate for all possible moves to connected sites. Afterward, the destination site is chosen with a probability ω_{ij}/ω_i. Since we are dealing with a Poisson process, waiting times are exponentially distributed with a parameter λ that is the inverse of the sum of all escape rates

$$\Delta t \sim \exp(\lambda), \ \lambda = \left(\sum_i \omega_i \right)^{-1}. \tag{15.20}$$

In order to reproduce this distribution, the time in the VSSM algorithm is updated after each move with Δt drawn from an exponential distribution. In practice, this is achieved by drawing a random number u from a uniform distribution on the interval $(0, 1]$ and setting Δt to

$$\Delta t = -\lambda \ln(u), \qquad u \sim \mathcal{U}((0,1]). \tag{15.21}$$

15.5.2 Forbidden Events

The VSSM algorithm can be adapted to efficiently treat forbidden events [54]. In the case of charge transport, e.g., each site (molecule) can be occupied by one charge carrier at a time (the Pauli exclusion principle). As a result, charge carriers obey Fermi–Dirac occupation statistics (see Section 15.3). Hence, if site j is already occupied, all incoming rates, $\omega_{i\to j}$ should be set to zero. This is computationally inefficient, since all rates (as well as escape rates) must be updated after every event. A much more efficient way is to keep the forbidden event in the event list. Once this event is attempted, i.e., a carrier attempts to hop to an occupied site, this destination site is marked as forbidden and another destination is selected from the remaining possibilities. If all surrounding sites are occupied, the algorithm switches to the level above, i.e., a different charge is selected and the previous one is added to a temporary list of forbidden events. The time is updated regardless of whether or not the event is forbidden (the charge moved or not). One can show that this strategy results in the same statistics as *a priori* removing all forbidden events from the event list [54].

15.5.3 Efficiency and Parallelization

Developing an efficient KMC code for charge/exciton transport is a rather challenging task. Indeed, if only the Pauli exclusion principle is taken into account, a single charge transfer event does not affect the rates, and only the (local) list of forbidden events should be updated. Hence, all rates can be precomputed before the KMC simulation is performed. If (long-range) Coulomb interactions are included, every CT event changes all rates. In other words, energy differences between two states of the system depend now on the relative positions of all charges in the system. To avoid updating rates at every KMC step, one can use scale-separation techniques: Since mesoscopic charge densities evolve on time scales much longer than typical CT times, electrostatic fields created by far-off charges can be treated in a mean-field way, e.g., by solving the Poisson equation. In this case, one can update local rates at every step, while all rates are updated only every 100–1000 steps [55].

Even the local rate updates can become computationally costly for systems with many charge carriers. In this situation, the binary tree search algorithm helps improve computational efficiency by providing a faster way of identifying which events and escape rates should be updated. By using this algorithm one can gain a factor of 10 speed-up for systems with 10% occupied sites [55]. Another alternative is to use the next reaction method [56], an improved version of the first reaction method, where a dependency graph is used to determine which rates to update.

Apart from trivial parallelization, where multiple copies of the system with different initial conditions are run in parallel, Monte Carlo schemes can be parallelized by dividing the simulation volume into subvolumes treated as individual simulations [57–60]. An efficient parallel implementation on graphics processing units (GPUs) has also been demonstrated [61].

15.6 Rates

To parametrize the master Equation 15.12, we need to evaluate the rates of all elementary processes of interest. For example, for charge transport, CT rates must be evaluated for all neighboring molecular pairs. To make this computationally feasible, we often rely on simplified theoretical treatments of transfer reactions. In this section, we review several of such theories as well as introduce the corresponding rate expressions.

15.6.1 Charge Transfer

The simplest expression for the CT rate, proposed by Miller and Abrahams [62], is the rate of thermally activated, barrierless tunneling between two localized electronic states. The corresponding rate is proportional to the Boltzmann prefactor of the free energy difference between the initial and final states, ΔE_{ij},

and the electronic overlap between the states, which decays exponentially with the molecular separation

$$\omega_{ij} = \begin{cases} \omega_0 \exp\left(-2\gamma r_{ij}\right) \exp\left(-\frac{\Delta E_{ij}}{k_B T}\right) & \Delta E_{ij} > 0, \\ 1, & \Delta E_{ij} \leq 0, \end{cases} \tag{15.22}$$

where $1/\gamma$ is the localization length of the charge, r_{ij} is the distance between the two molecules, and ω_0 is the attempt frequency. The simplicity and intuitiveness of this rate justified its use in the early stages of understanding charge transport in organic semiconductors [63].

Environmental effects, as well as molecular reorganization upon charging/discharging (i.e., the energetic barrier between the initial and final states), have been accounted for in the so-called Marcus CT rate [64,65]

$$\omega_{ij} = \frac{2\pi}{\hbar} \frac{J_{ij}^2}{\sqrt{4\pi\lambda_{ij} k_B T}} \exp\left[-\frac{\left(\Delta E_{ij} - \lambda_{ij}\right)^2}{4\lambda_{ij} k_B T} \right]. \tag{15.23}$$

Again, $\Delta E_{ij} = E_i - E_j$ is the driving force or the free energy difference of the final and initial states of the CT reaction [66]. An accurate evaluation of these energies using polarizable force-fields is discussed in Section 15.7. J_{ij} is the electronic coupling element. They are intimately related to the overlap of the diabatic electronic states, as discussed in Section 15.9. Finally, λ_{ij} is the sum of the external and internal reorganization energies, discussed in more detail in Section 15.8.

The main issue with the classical Marcus rate is that the intramolecular vibrational modes promoting the CT reaction are energetically comparable to the C–C bond stretching mode at room temperature, $\hbar\omega_{CC} \sim 0.2\,\text{eV} \gg k_B T \sim 0.025\,\text{eV}$. Therefore, these modes should be treated quantum mechanically. For a common set of intramolecular high-frequency (quantum-mechanical) and an outer sphere low-frequency (classical) vibrational coordinates, a mixed quantum-classical multichannel generalization of the Marcus formula is readily available [66,67]. A generalization for the bimolecular electron transfer rate with independent sets of coordinates for donor and acceptor has also been proposed [45].

Another rate expression has been proposed by Weiss and Dorsey [68–70]. In the low temperature limit, $\hbar\nu_c / k_B T \gg 1$, it reads

$$\omega_{ij}(\epsilon) = \frac{J_{ij}^2}{\hbar^2 \nu_c} \left(\frac{\hbar\nu_c}{2\pi k_B T} \right)^{1-2\alpha} \frac{\left| \Gamma\left(\alpha + i\frac{\Delta E_{ij}}{2\pi k_B T}\right) \right|^2}{\Gamma(2\alpha)} \exp\left(\frac{\Delta E_{ij}}{2k_B T} - \frac{\left|\Delta E_{ij}\right|}{\hbar\nu_c} \right). \tag{15.24}$$

Here $\Gamma(z)$ is the gamma function and ν_c is the characteristic frequency, or the largest frequency in the Ohmic bath, which is related to the reorganization energy, λ, by $\lambda = 2\alpha\hbar\nu_c$. The Kondo parameter, α, describes the coupling strength between the charge and the heat bath [71]. In the high temperature limit, the Weiss–Dorsey rate simplifies to the Marcus rate. Indications that the Weiss–Dorsey rate is better suited for describing charge transport, especially at low temperatures and high fields, have recently been reported [72].

15.6.2 Electron–Hole Recombination

The microscopic process of electron–hole recombination is straightforward: Whenever a hole and an electron occupy the same site (molecule), they recombine with a certain rate, w_r. Radiative recombination leads to a photon emission which is the basis of OLED functionality.

On a macroscopic level, electron–hole recombination is traditionally described by the Langevin equation (see Section 15.3.2), with a prefactor proportional to the sum of electron and hole mobilities. Extensive KMC simulations were performed to verify this dependency [73]. An excellent agreement was found if the

electron and hole mobilities are extracted from simulations with both carrier types, which was attributed to the change in mobilities upon inclusion of two carrier types due to their Coulomb attraction. Deviations at high charge concentrations ($>10^{-3}$ carriers per molecule) were also observed: Here, the average electron–hole distance becomes smaller than the thermal capture radius, $r_c = e^2/4\pi\epsilon\epsilon_0 k_B T$, violating assumptions used in the derivation of the Langevin formula.

15.6.3 Energy Transfer

We now briefly review rate expressions used to describe electronic excitation (energy) transfer (EET). EET is also known as a resonant energy transfer (RET).

A theory explaining the mechanism of EET was first proposed by Förster [74]. The electronic interaction promoting EET relies on a coupling of the donor and acceptor molecules via Coulombic interaction. Similar to CT, Förster theory relies on Fermi's Golden Rule with the electronic coupling between donor and acceptor treated perturbatively. Additional assumptions are that the system equilibrates after the electronic excitation of the donor on a time scale much faster than that of EET and that coupling to the bath (given by the absorption line shape) is much greater that the electronic coupling between donor and acceptor. Energy conservation in the weak coupling limit results in a coupling element, which is proportional to the overlap of the donor fluorescence spectrum with the acceptor absorption spectrum. The spectral overlap includes nuclear overlap (in the form of Franck–Condon factors), which depends on the spectral line shapes and thus provides the temperature dependence of the EET rate. Under these assumptions, the EET rate between emitters i (donor) and j (acceptor) takes a simple form

$$\omega_{ij}^{F} = \frac{1}{\tau_{r,i}} \left(\frac{R_{F,ij}}{r_{ij}} \right)^6, \tag{15.25}$$

where $\tau_{r,i}$ is the radiative life time of the emitter (donor), r_{ij} is the separation between donor and acceptor, and $R_{F,ij}$ is the Förster radius for transfer from donor to acceptor.

An EET process can occur even when Coulomb-term-mediated electronic transitions are forbidden. Dexter provided a derivation for the case when the Coulombic interaction is negligible and EET is due to the exchange part [75]. This rate decays exponentially with intermolecular separation

$$\omega_{ij}^{D} = k_D \exp\left(-2\gamma r_{ij}\right), \tag{15.26}$$

where $1/\gamma$ is the wavefunction decay length and k_D is the Dexter prefactor proportional to the exchange integral.

Note that since the exchange integral is a quantum mechanical correction to the Coulombic repulsion, the total EET rate is always a sum of the two rates. However, due to the exponential decay of the Dexter coupling, the electronic interaction that mediates EET at separations greater than 5 Å is invariably Coulombic.

The importance of exchange and other short-range interactions, as well as higher multipole contributions to the Coulombic interaction, has been examined fairly extensively, helping improve the accuracy of rate expressions 15.25 and 15.26. An overview of these extensions can be found in Refs. [76–80].

15.7 Density of States

In the previous section, we saw that CT rates depend on the energies of the initial and final states of the CT complex. In an amorphous organic semiconductor, every molecule is embedded in a unique environment and therefore every molecule has its unique set of energy levels, EA, and IP. The set of energy levels available for an excess charge is termed the DOS. In this section, we discuss several methods of evaluation of the DOS of an amorphous organic semiconductor.

15.7.1 Gaussian Disorder

The simplest approach to model the DOS of organic semiconductor is to assume a Gaussian distribution of energy levels

$$g(E) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{\left(E-\bar{E}\right)^2}{2\sigma^2}\right]. \tag{15.27}$$

This phenomenological expression, first proposed by Bässler [63] and termed the Gaussian disorder model (GDM), is motivated as follows: Randomly oriented dipole moments in an amorphous material interact with one another, thus influencing the energy levels of neighboring molecules. As a result of the central limit theorem, this leads to approximately Gaussian distributed energy levels. The width of the distribution, σ, is called the *energetic disorder* and can be extracted from temperature-dependent mobility measurements [81, 82] or from simulations, as discussed in Section 15.11.1.

The original model was extended to the correlated disorder model (CDM) by accounting for spatial correlations that arise from the dipole interaction of neighboring molecules [83]. For randomly oriented dipoles, $\vec{p}_j$, the the electrostatic energy of a site i is given by [84]

$$E_i = -\frac{q}{4\varepsilon} \sum_{j\neq i} \frac{q\vec{p}_j \cdot (\vec{r}_j - \vec{r}_i)}{|\vec{r}_j - \vec{r}_i|^3}, \tag{15.28}$$

where ε is the dielectric permittivity and q is the charge. In the case of equal absolute values, $|\vec{p}_j| = p$, and dipoles fixed on a cubic lattice with lattice constant a, the sum can be evaluated using Ewald summation [85,86], yielding a Gaussian DOS with energetic disorder $\sigma = 2.35\, q\, p/\varepsilon\, a^2$ [83,87]. The spatial energy correlation function

$$\kappa(r) = \frac{\mathbb{E}\left[\left(E(\vec{r}_i) - \bar{E}\right)\left(E(\vec{r}_j) - \bar{E}\right)\right]}{\sigma^2} \tag{15.29}$$

is then given by $\kappa(r) \approx 0.74\frac{a}{r}$ [88]. Here $r = \left|\vec{r}_i - \vec{r}_j\right|$ is the distance between two molecules, $\bar{E}$ is the mean of the energy distribution, and $\mathbb{E}[\cdot]$ is the expectation value. Note that the spatial correlation function of this simple model system depends only on the lattice spacing a, limiting its applicability to realistic morphologies [89], as discussed in Section 15.11.1.

The Gaussian DOS is at the heart of the family of GDM, where Miller–Abrahams rates, Equation 15.22, with the Gaussian DOS are assumed for a charge hopping on a cubic lattice. Extensive KMC simulations of these models helped parametrize the mobility as a function of temperature, field, and field carrier density [63]. It was extended by including the important influence of charge carrier density in the extended Gaussian disorder model (EGDM) [90] without spatial correlation of site energies. The mobility expression in the EGDM at finite can be written as a product of three functions,

$$\mu(T, F, \rho) = \mu_0(T) g(T, \rho) f(T, F), \tag{15.30}$$

where

$$\mu_0(T) = 1.8 \times 10^{-9} \mu_0 \exp\left[-C\hat{\sigma}^2\right], \qquad g(T, \rho) = \exp\left[\frac{1}{2}\left(\hat{\sigma}^2 - \hat{\sigma}\right)\left(2\rho a^3\right)^{\delta}\right], \tag{15.31}$$

$$f(T, F) = \exp\left[0.44\left(\hat{\sigma}^{3/2} - 2.2\right)\left(\sqrt{1 + 0.8\hat{F}^2} - 1\right)\right], \qquad \delta = 2\frac{\ln\left(\hat{\sigma}^2 - \hat{\sigma}\right) - \ln(\ln 4)}{\hat{\sigma}^2},$$

with $C = 0.42$, $\hat{F} = eaF/\sigma$ and $\hat{\sigma} = \sigma/k_B T$. A similar parametrization, termed the extended correlated disorder model (ECDM), also exists for the spatially correlated DOS [91].

15.7.2 Perturbative Approach

While Gaussian disorder models have been successful in describing many properties of organic semiconductors, they do not provide a direct link to the material morphology or chemical composition. A perturbative approach allows us to evaluate site energies in atomistically resolved (see Section 15.10), large-scale morphologies. Furthermore, it can also be used to parametrize Gaussian disorder models and perform large-scale simulations (see Section 15.11.2).

For every molecular pair, our quantity of interest is the energy difference $\Delta E_{ij} = E_i - E_j$, which is the energy separation between the minima of the diabatic potential energy surfaces (PES). In systems with weak intermolecular couplings, one can treat interactions with the environment perturbatively: The intermolecular electrostatic and induction contributions are then given by the first- and second-order terms in the expansion of the interaction energy [92].

The total energy energy of an ion embedded in a molecular environment thus includes an internal contribution E_i^{int}, i.e., the electron affinities for electrons and IPs for holes of *isolated* molecules. These can vary from one molecular pair to another because of different energy levels for different types of molecules, or different conformers of the same molecule. Correspondingly, the *external* contribution is due to the electrostatic and induction interactions, E_i^{el} and E_i^{ind}, of a charged molecule with the environment. These interaction energies are determined by the electrical charge distribution and the polarizability distribution, respectively, in the environment of the charged molecule. Overall, the total energy of molecule i is then given by

$$E_i = E_i^{\text{int}} + E_i^{\text{el}} + E_i^{\text{ind}}. \tag{15.32}$$

Most difficult to evaluate is the interaction with the environment. This is for two reasons: First, the underlying interactions are long-ranged and thus large system sizes are needed to converge the values of site energies. Second, special summation techniques are needed in order to evaluate interactions of an ion with a neutral *periodic* environment [93–96]. Also note that the perturbative evaluation of site energies relies on accurate molecular representations in terms of distributed multipoles and polarizabilities and is computationally demanding, in spite of being classical [97].

15.7.3 Hybrid Approaches

Using Equation 15.32, the site energies can be evaluated by solving a microscopic analogue of the Poisson equation. The self-consistent solution is normally achieved using iterative schemes for induced multipoles and Ewald summation techniques for static multipoles, which is computationally demanding. The role of hybrid schemes is to reduce computational cost. Here, one relies on the fact that the fields created by the far-off charges can be treated in a mean-field way, i.e., obtained by solving the Poisson equation. Nearest-neighbor interactions (i.e., interactions between molecular pairs within a certain cutoff distance) are still evaluated explicitly. This scheme also allows us to add metallic electrodes as image charges and has been used extensively to simulate multilayered OLED devices [98–100]. To improve computational efficiency, in this approach all induction interactions are taken into account effectively, by rescaling interaction energies by the (effective) dielectric constant of the medium.

15.8 Reorganization Energy

The reorganization energy, responsible for the energy barrier between diabatic states as well as for the broadening of energy levels of the electron detachment/attachment spectra, has two contributions. The *internal* reorganization energy is a measure for how much the geometry of the CT complex adapts while the charge is transferred. It can be estimated based on four points on the diabatic PES [45,101]

$$\lambda_{ij}^{\text{int}} = E_i^{nC} - E_i^{nN} + E_j^{cN} - E_j^{cC}. \tag{15.33}$$

Here, small letters denote the state and capital the geometry of a molecule, e.g., E_i^{nC} is the internal energy of the molecule *i* in the geometry of its charged state. Treatments that do not approximate the PES in terms of a single shared normal mode are also available [45,66,102].

An additional contribution to the overall reorganization energy results from the rearrangement of the environment in which the CT takes place. In a classical case, this outer-sphere reorganization energy, λ^{out}, contributes to the exponent in the rate expression in the same way as its internal counterpart. Assuming that CT is significantly slower than electronic polarization but much faster than the nuclear rearrangement of the environment, λ^{out} can be evaluated from the electric displacement fields created by the CT complex [66], provided that the Pekar factor is known. Alternatively, one can use polarizable force-fields [103] or quantum mechanics/molecular mechanics (QM/MM) methods [104]. It also turns out that the classical Marcus expression for the outersphere reorganization energy (inversely proportional to the molecular separation) can predict negative values of λ^{out} for small intramolecular separations, which are unphysical and hence should be used with care [45].

15.9 Electronic Coupling Elements

Electronic coupling elements, or transfer integrals , J_{ij}, entering the CT rate (Equations 15.23 and 15.24) are off-diagonal matrix elements, $J_{ij} = \left\langle \psi_i | \hat{H}^{el} | \psi_j \right\rangle$, of the electronic Hamiltonian, $\hat{H}^{el} = \hat{T}^{el} + \hat{V}^{el-el} + \hat{V}^{nuc-el}$, based on a diabatic (noninteracting) state, ψ_i [66]. A number of approaches can be used to evaluate electronic coupling elements. Their efficiency and accuracy depend on how the diabatic states and Hamiltonian are constructed, as well as how the matrix projection is performed.

Diabatic states are often approximated by the HOMO of monomers for hole transport, or the LUMO for electron transport ("frozen core" approximation) [105–107]. An approximate diabatic basis can also be constructed using constrained density functional theory [108].

The dimer Hamiltonian can be constructed using semiempirical methods, e.g., the Zerner's Intermediate Neglect of Differential Overlap (ZINDO) Hamiltonian [101,109–111]. This approach does not require a self-consistent evaluation of the dimer Hamiltonian and is therefore computationally very efficient [110]. One can also employ density functional theory (DFT) and use either the fully converged Hamiltonian or only the initial guess [105]. Another way to improve the efficiency is to reduce the number of orbitals for which electronic couplings are calculated. A detailed comparison of accuracy and efficiency of different approaches can be found in Refs. [105,112,113].

For (approximately) spherically shaped molecules the logarithm of the squared electronic coupling, $\log(J_{ij}^2)$, decays linearly as a function of the intermolecular separation *r* (at least for large distances), justifying the functional form of the Miller–Abrahams rate (see Equation 15.22). If molecular pairs are extracted from the respective dimers in the realistic atomistic morphology, $\log(J_{ij}^2)$ is often Gaussian distributed, with a mean and variance that depend on the molecular separation [55,114]. This observation can be used to parametrize coarse-grained (stochastic) charge transport models, as discussed in Section 15.11.1.

Since electronic couplings are related to the overlap of electronic orbitals participating in charge transport, they are very sensitive to the relative positions and orientations of molecules. Hence, an amorphous morphology of an organic material should be generated as precisely as possible. The corresponding methods are covered in the next section.

15.10 Morphology

In the previous sections, we explained how charge/exciton transport and recombination can be mapped on a series of events with specified rates. We have also emphasized that rates are very sensitive to both local and global molecular ordering. For computer-based predictions of material properties it is therefore important to simulate the material morphology as realistically as possible. In OLEDs, we usually deal with *amorphous* molecular arrangements—crystalline films are often less efficient in OLEDs because of the resulting large

exciton diffusion distances, leading to enhanced exciton loss due to quenching processes. Amorphous films have a well-defined local structure, which depends on molecular interactions and processing conditions. In this section, we describe how amorphous morphologies of 10^4–10^5 molecules can be simulated using atomistic force fields or coarse-grained models. These systems can then be used to study small-scale charge transport [45,115–119], or to parameterize mesoscale models, as discussed in Section 15.11.1.

15.10.1 Classical Force Fields

The role of classical force fields in simulations of organic semiconductors is two-fold. First, they are used to generate atomistically resolved morphologies of molecular assemblies. Second, they are employed to evaluate the solid-state electrostatic and induction contributions to site energies (see Section 15.7). In both cases, these classical molecular representations should be appropriately parametrized.

For site energy calculations, it is important to evaluate electrostatic and induction contributions as accurately as possible. The corresponding parametrization is rather straightforward: A perturbative expansion of the intermolecular interaction energies is based on distributed multipoles (electrostatic interaction) and distributed polarizabilities (induction interaction) [92,97,120–122]. Van der Waals interactions are normally ignored since we are interested in free energy differences between charged and neutral states of the system. To improve computational efficiency, one can also use machine-learning techniques to devise simple structure–property relations for, e.g., atomic multipoles of molecular conformers [123].

Polarizable force fields based on distributed multipole expansions are still computationally prohibitive for simulating molecular arrangements of large systems. Moreover, parametrizations of effective *pairwise* potentials (partial charges and Lennard–Jones parameters), which mimic many-body van der Waals and induction interaction energies, are a nontrivial task, which largely relies on experimental input [124,125]. Standard forces fields, e.g., those suitable for biosystems, are often not transferable to organic molecules with large π-conjugated subsystems. A representative example is the comparison of the Williams 99 and the optimized potential for liquid simulations (OPLS) force fields for the amorphous mesophase of Alq_3 [126], which predict fairly different densities, radial distribution functions, and glass transition temperatures.

That being said, the development of new force fields for organics is currently limited to refitting of partial charges using the Merz–Singh–Kollman [127] scheme or the CHarges from ELectrostatic potentials using a grid-based method (CHELPG) [128] scheme, and a parametrization of missing bonded interactions from first-principles scans (see Figure 15.3). The remaining parameters are often taken from standard force fields, such as OPLS [124]. Reparametrized force fields are then verified against experimentally available densities and glass transition temperatures.

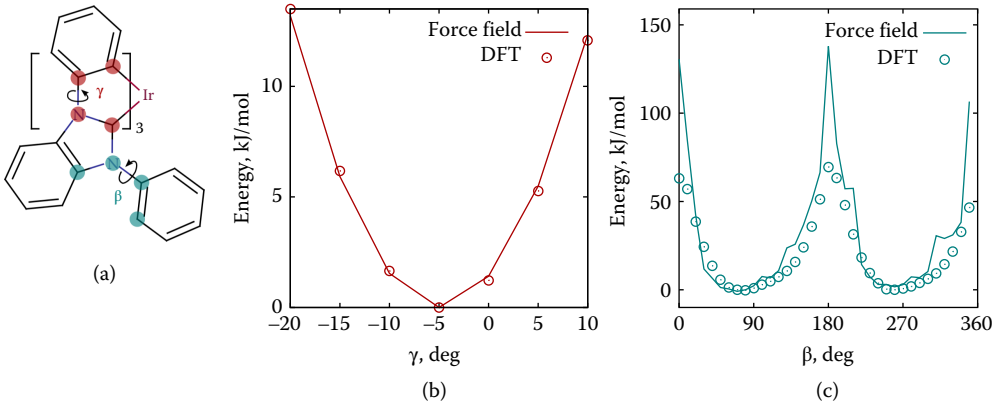

FIGURE 15.3 Potential energy scans of two dihedral angles of the DPBIC molecule, the chemical structure of which together with the definition of dihedral angles is shown in the inset (a). Symbols show DFT energies (B3LYP/6-311g(d,p)) and lines indicate the force-field energies after adjusting the force-field parameters. (Reproduced from P. Kordt et al., *Adv. Funct. Mater.*, 25, 1955–1971, 2015. With permission.)

15.10.2 Morphology Simulations

With the force field at hand, one can, in principle, simulate amorphous molecular assemblies. For systems with periodic boundaries in all directions this is done by first equilibrating the system above the glass transition temperature and then quenching it to room temperature in the *NPT* ensemble using molecular dynamics [126,129,130].

For thin slabs (2D-periodic systems) one can directly deposit molecules either using molecular dynamics [94] or Monte Carlo [131] techniques. In both cases the dynamics of deposition is not realistic. In other words, if surface diffusion plays an important role (e.g., in the case of guest aggregation in the host) these techniques cannot be applied in a straightforward manner, since the length scales and time scales required of host aggregation cannot be reached by atomistic molecular dynamics simulations. It is, however, possible to explore the fact that certain parts of molecules evolve on much slower time scales and larger-length scales and to combine several coherently moving atoms, connected via stiff degrees of freedom (e.g., bonds) into a single interaction site, as it is done, e.g., in the united atom force-fields with hydrogens incorporated into heavier atoms [132]. By doing this, we reduce the number of degrees of freedom to be propagated and, more importantly, obtain a much smoother potential energy landscape in terms of the *coarse-grained* degrees of freedom [133] (softer interaction potentials, less friction), allowing one to simulate 10–100 longer times and system sizes [132].

In order to perform correct statistical sampling of the coarse-grained degrees of freedom, the potential of mean force should be used as the interaction potential [134], which is inherently a many-body potential. To reduce computational cost, this potential is represented as the sum of a few functions, i.e., projected onto the basis functions of the force field. The accuracy of the coarse-grained model thus becomes sensitive to the way the projection is performed as well as the number of basis functions that are used to represent the coarse-grained force-field [135]. Existing projection schemes try to reproduce various pair distribution functions (structure-based coarse-graining [136–139]), to match the forces [134,140,141], to minimize the information loss in terms of relative entropy [142], or to use liquid state theory [143]. An extensive overview of such coarse-graining techniques is provided in Ref. [144]. Often, the accuracy of the coarse-grained model can be improved by explicitly incorporating information about macroscopic properties of the system, i.e., its equation of state or the symmetry of its mesophase [145–147]. Moreover, one can further reduce the number of degrees of freedom by introducing anisotropic interaction potentials. Eventually, atomistic details can be reintroduced into coarse-grained morphologies [148].

15.11 Scale Bridging

Now that we have seen how an OLED device can be modeled on different time and length scales, we eventually would like to transfer our knowledge about the system between the scales. For example, drift–diffusion equations (Section 15.3) can be used to calculate current–voltage curves of a device (micrometer scale). However, they require expressions for charge carrier mobility, the diffusion constant and, in the case recombination is taken into account, the recombination rate, all as a function of external field, carrier density, and temperature. Microscopic simulations (see Sections 15.4–15.10) can provide this information, at the same time retaining the link to the molecular structure. However, they are becoming computationally too demanding for high charge densities and large system sizes required to parametrize these dependencies. Our aim here is to provide several strategies, which can be used to link different time and length scales.

15.11.1 Stochastic Models

In an OLED, charges are inhomogeneously distributed [149] and charge density variations span several orders of magnitude. To cover the required density range in simulations, one needs to deal with relatively large systems—this quickly becomes computationally demanding if all rates are evaluated from

first principles, as described in Section 15.6. To remedy the situation, one can devise a phenomenological algorithm to parametrize the master equation. In this section we outline how such an algorithm can be constructed for charge transport simulations.

We first note that the master equation (Section 15.4) is completely determined by the event rates. In the case of charge transport, these rates depend on site energy differences, electronic coupling elements, and reorganization energies. In order to evaluate observables of interest (e.g., charge mobility), one additionally needs site positions. Hence, the algorithm should be able to reproduce (statistically) several distribution and correlation functions.

Let us start with molecular positions. In an amorphous solid, both positions and orientations are completely defined by a set of (many-body) spatial correlation functions. For (approximately) spherically shaped molecules, the pair correlation function, or radial distribution function, $g(r)$, which quantifies the density of molecules at a separation r, contains the most relevant structural information. To reproduce this function approximately, one can use "thinning of a Poisson process" [114,150]. More accurate coarse-graining techniques, such as iterative Boltzmann inversion [138,151] or inverse Monte Carlo [136,137,152], allow a (numerically) exact reproduction of the radial distribution function [55]. These methods optimize a pair interaction potential, $U(r)$, in a way that the corresponding $g(r)$ is reproduced. They rely on the Henderson theorem [153], which states that there is a unique correspondence between $U(r)$ and $g(r)$. An illustration of this algorithm for an amorphous layer of DPBIC is presented in Figure 15.4a. The approach can also be applied to nonspherical molecules, by using several interacting sites per molecule [135].

The second ingredient of the stochastic model is the connectivity: In the atomistic model only molecules within a certain cutoff distance are used for calculating CT rates. The rest of the rates are set to zero. This is justified by the fact that electronic coupling elements decrease roughly exponentially with molecular separation [55], see, e.g., Equation 15.22. The distance that determines whether or not two molecules are connected is given by their two closest atoms. Since this information is not present in the coarse-grained model, the resulting probability of two sites to be connected, as a function of their center-of-mass separation, is given by the corresponding probability extracted from the reference data. Figure 15.4b shows this probability for an amorphous DPBIC layer.

As explained in Section 15.7, the stochastic generation of site energies should reproduce both their distribution function (DOS) and their spatial correlation. This can be achieved by mixing in site energy contributions of neighboring sites [89]. Figure 15.4c shows the spatial correlation function for an amorphous DPBIC layer.

Electronic coupling elements can also be generated using appropriate distributions. These distributions are, however, separation dependent: The logarithm of squared transfer integrals, $\log J^2$ (which is often Gaussian distributed), depends on molecular separation. For DPBIC, the distance dependence of the mean and the standard deviation is shown in Figure 15.4d. In the stochastic model, transfer integrals are then drawn from such distant-dependent distributions.

With all necessary rate ingredients, one can now validate the model, e.g., by evaluating the distribution of rates, see Figure 15.4e, or by directly comparing charge carrier mobilities, as shown in Figure 15.4f. Note that charge transport in systems with large energetic disorder has pronounced finite size-effects [130]. Therefore, similar system sizes should be used to compare stochastic and reference simulations.

Since stochastic models are computationally significantly less demanding, they can serve as an intermediate step between atomistic and macroscopic (drift–diffusion) descriptions, as discussed in the next section.

15.11.2 Parametrization of Gaussian Disorder Models

As mentioned in Section 15.3, macroscopic OLED modeling requires the charge mobility as a function of external field, temperature, and carrier density, $\mu(T, F, \rho)$. Analytic expressions of these dependencies can be provided by the EGDM and ECDM, as discussed in Section 15.7.1. These generic expressions include

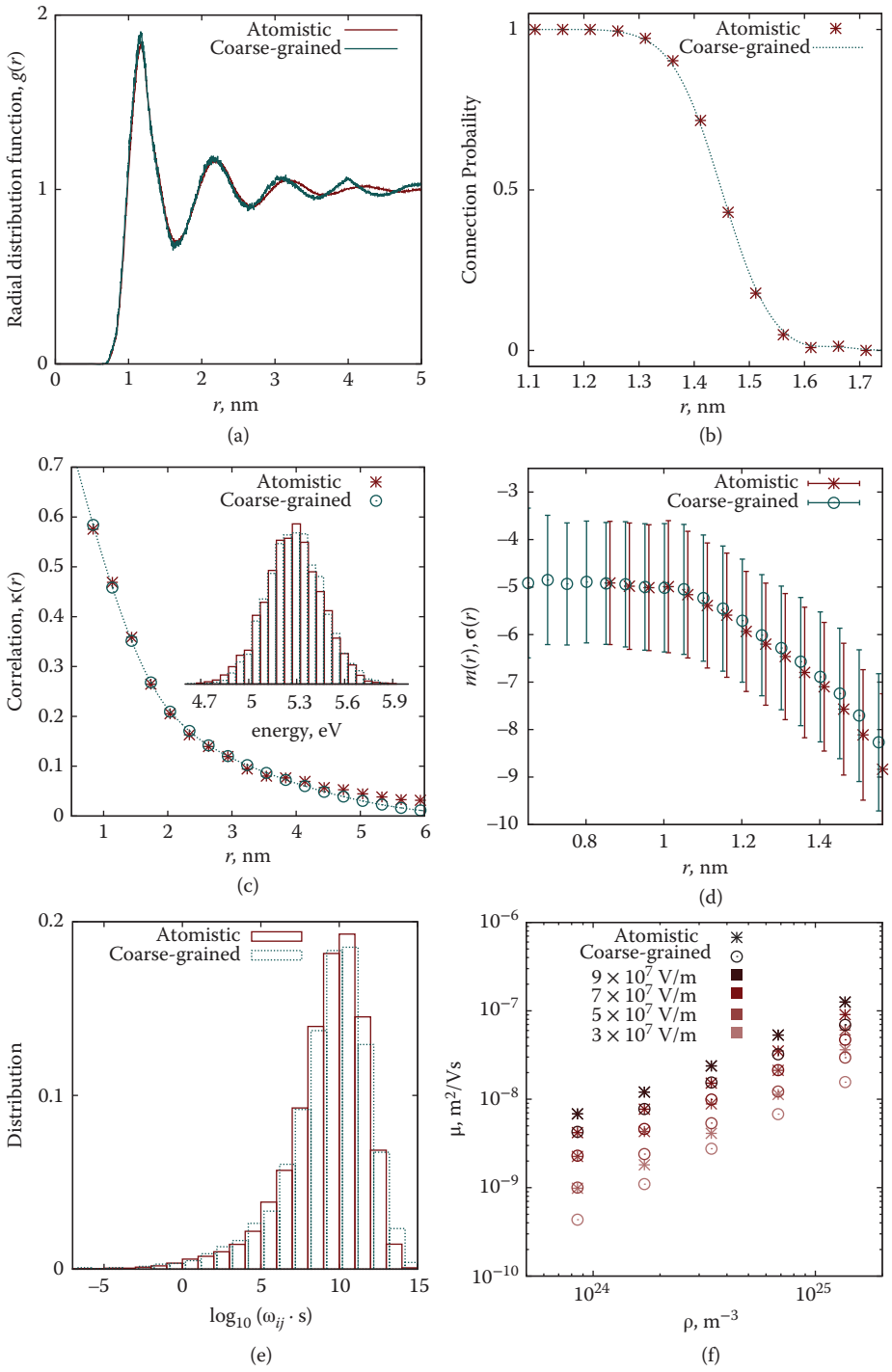

FIGURE 15.4 Comparison of hole states and hole mobilities in the atomistic reference system of DPBIC and in a stochastic model. (a) Radial distribution function, $g(r)$, where r is the center-of-mass distance. (b) Probability for sites to be connected. (c) Spatial correlation function, $\kappa(r)$, as defined in Equation 15.29; inset: site energy distribution with a mean (ionization potential) of 5.28 eV and an energetic disorder of $\sigma = 0.176$ eV. (d) Mean m and width σ of a distribution of the logarithm of squared electronic couplings, $\log_{10}(J^2/\text{eV}^2)$, for molecules at a fixed separation r. (e) Rate distributions. (f) Mobility as a function of hole density at different electric fields. (Reproduced from P. Kordt et al., *Adv. Funct. Mater.*, 25, 1955–1971, 2015, with permission.)

several material-specific parameters. In this section, we describe how to determine these parameters from simulations of small systems.

Both EGDM and ECDM depend parametrically on the lattice constant a, the energetic disorder σ (see Section 15.7), and a prefactor μ_0, which is related to the temperature-dependent mobility at zero field and charge density by $\mu_0(T) = 1.8 \times 10^{-9}\mu_0 \exp(-0.42\hat{\sigma}^2)$ for the EGDM (see Section 15.7.1) or $\mu_0(T) = 1.0 \times 10^{-9}\mu_0 \exp(-0.29\hat{\sigma}^2)$ in the case of ECDM.

In principle, both a and σ can be evaluated in a relatively small system: a as the mean distance between neighboring molecules and σ as the width of the DOS that results from perturbative energy calculations. μ_0 can be extracted from charge transport simulations performed at different temperatures. This approach, however, does not lead to reliable parameterizations [55,114]. Indeed, a multidimensional fit of simulated mobilities to the EGDM or ECDM expressions, for a wide range of temperatures, charge densities, and external fields, yields a very different set of parameters. A comparison of these two approaches for amorphous dicyanovinyl-substituted quaterthiophe (DCV4T) and DPBIC is given in Table 15.1. One can see, e.g., that the EGDM underestimates the energetic disorder, while the ECDM overestimates it. In both cases, spatial site energy correlations are responsible for this discrepancy: EGDM does not include correlations and compensates for higher mobility values by reducing the energetic disorder σ. On the other hand, ECDM overestimates spatial correlations and compensates this by reducing the lattice constant [89]. The discrepancy between microscopic values and fits to EGDM and ECDM teaches us that parameters of these models do not have a clear physical interpretation. Nevertheless, they still provide reasonable parametrizations and can eventually be used in conjunction with drift–diffusion equations; see Section 15.12.1. Figure 15.5 compares EGDM and ECDM fits to microscopic simulations for an amorphous mesophase of DCV4T.

On a technical side, the stochastic models described in Section 15.11.1 become very useful to perform the fits. They help cover the required range of charge carrier densities and reduce finite-size effects. Note, however, that finite-size effects in systems with small charge carrier densities and large disorder are so large that the actual value of mobility is overestimated by several orders of magnitude [130]. In this case, one needs to use the extrapolated mobility values [114].

15.11.3 Tabulated Mobilities

Fitting the results of KMC simulations to the parametrizations provided by the EGDM or ECDM imposes a constraint on the functional form of $\mu(\rho, \vec{F}, T)$. To avoid this, one can tabulate the mobility in a wide range of charge densities, temperatures, and electric fields. This tabulated function can then be used directly in the drift–diffusion equations solver [149]. The tabulation is computationally feasible only with the help of a stochastic model—otherwise it is not possible to reach the necessary system sizes and to span the wide density regime. Before using the tabulated function, it has to be interpolated and smoothed to ensure numeric stability [149]. Figure 15.6 shows the tabulated and smoothed mobility for amorphous DPBIC, which is eventually used to evaluate current–voltage characteristics of a DPBIC film; see Section 15.12.1.

TABLE 15.1 Lattice Spacing, Energetic Disorder, and Mobility at Zero Field and Density Extracted From a Microscopic System and From Fitting Simulated Hole Mobilities to Extend Gaussian Disorder Model (EGDM) and Extended Correlated Disorder Model (ECDM) for Amorphous Phases of (a) dicyanovinyl-substituted quaterthiophe (DCV4T) and (b) DPBIC

(a) DCV4T	a [nm]	σ [eV]	μ_0(300 K) [m^2/Vs]	(b) DPBIC	a [nm]	σ [eV]	μ_0(300 K) [m^2/Vs]
Microscopic	0.86	0.253	2.0×10^{-21}	Microscopic	1.06	0.176	3.4×10^{-12}
EGDM	1.79	0.232	2.1×10^{-21}	EGDM	1.67	0.134	2.1×10^{-11}
ECDM	0.34	0.302	3.3×10^{-22}	ECDM	0.44	0.211	1.8×10^{-13}

Source: DCV4T values: Reprinted from P. Kordt et al., *J. Chem. Theory Comput.*, 10, 2508–2513, 2014. With permission; DPBIC values: Reproduced from P. Kordt et al., *Adv. Funct. Mater.*, 25, 1955–1971, 2015. With permission.

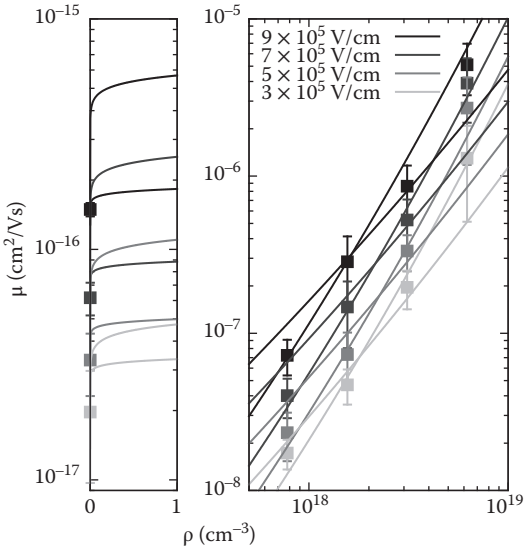

FIGURE 15.5 Parametrization of lattice models from simulated dicyanovinyl-substituted quaterthiophe (DCV4T) hole mobilities, μ, for different hole densities, ρ. Symbols are the simulated values for four different external fields, solid lines are the fit to the extended Gaussian disorder model (EGDM), and dashed lines are the fit to the extended correlated disorder model (ECDM). An extrapolation has been used to obtain non-dispersive values (i.e., without finite-size effects) in the limit of zero density. (Reprinted from P. Kordt et al., *J. Chem. Theory. Comput*., 10, 2508–2513, 2014. With permission.)

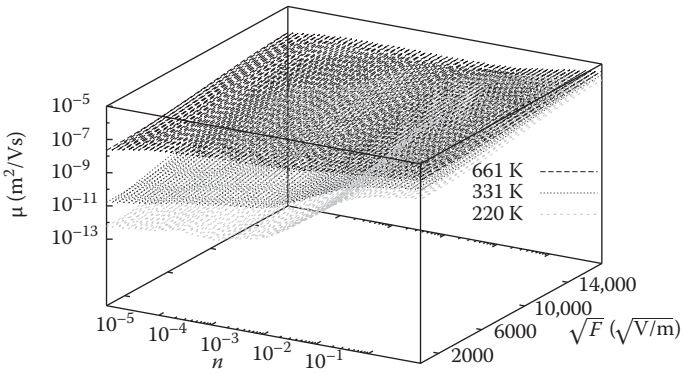

FIGURE 15.6 Hole mobilities for amorphous DPBIC as a function of the charge concentration n (number of charge carriers per site) after tabulation from simulations, smoothing, and interpolation. (Reprinted from P. Kordt et al., *Phys. Chem. Chem. Phys*., 17, 22778–22783, 2015. With permission.) Published by the PCCP Owner Societies.

15.12 Case Studies

So far we have described various methodological developments and simulation approaches, which can be used to simulate multilayered OLED structures. In the following sections, we show how these methods can be used to simulate steady-state current–voltage characteristics of an OLED, perform impedance spectroscopy simulations, estimate OLED efficiency, and study electroluminescence of a white OLED, as well as to gain insight into OLED stability.

15.12.1 Current–Voltage Characteristics

We start by showing how the steady-state current–voltage characteristics of a single layer device can be simulated starting from the chemical structure of an organic semiconductor. As an example, we use a thin layer of DPBIC, a hole-conducting material, which is sandwiched between an ITO and aluminum electrode.

After parametrizing the DPBIC force field [55], amorphous boxes of 4000 DPBIC molecules are simulated by molecular dynamics (MD) simulations, as described in Section 15.10.2. The DOS of an amorphous solid state is then evaluated for holes using the perturbative scheme (see Section 15.7.2), yielding a mean value of 5.28 eV and an energetic disorder of $\sigma = 0.176$ eV. Density functional theory calculations (B3LYP/6-311g(d,p), see also Section 15.8) yield a hole reorganization energy of $\lambda = 0.068$ eV.

The reference system of 4000 molecules is further used to parametrize a stochastic algorithm (see Section 15.11.1) and to generate larger systems of 40,000 sites (details can be found in Ref. [149]). KMC simulations in large systems allow us to tabulate charge mobility as a function of charge density, electric field, and temperature. The interpolated and smoothed tabulated function, depicted in Figure 15.6, is then used to solve the drift–diffusion equations, as described in Section 15.3. For the electrode IPs we use average values of experimental reports: 4.73 eV for ITO [154–157] and 4.16 eV for aluminum [158]. These values, together with the DPBIC IPs, provide the value of the injection barrier, which is required to solve the Poisson Equation 15.3.

Without further microscopic calculations, it is now possible to solve the drift–diffusion equations for film thicknesses of 203, 257, and 314 nm and temperatures of 233, 293, and 313 K, corresponding to different experiments. Figure 15.7 shows a comparison of simulated and experimental current–voltage curves in these situations. The agreement is remarkable for higher temperatures (293 and 313 K). A possible reason for the larger differences at low temperatures are nonequilibrium processes, such as charge relaxation, which are not accounted for in our approach that relies on mobility parametrizations under stationary conditions. Another possibility is that the Marcus rate expression, Equation 15.23, is no longer valid at these temperatures, since it is derived assuming a classical promoting mode and is valid only for high enough temperatures.

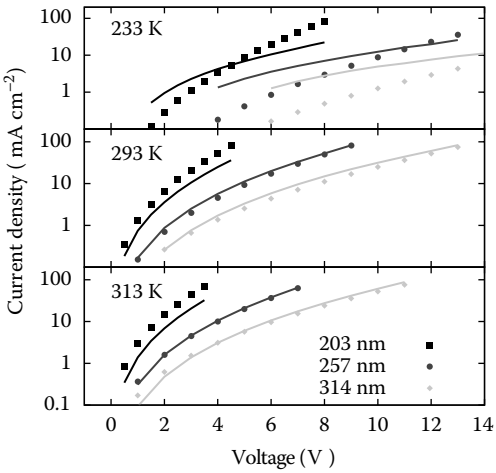

FIGURE 15.7 Current–voltage characteristics for DPBIC films of different thickness sandwiched between an indium tin oxide (ITO) anode and an aluminum cathode, measured at different temperatures. Theoretic predictions (lines) were obtained using parameter-free coupling of microscopic hole mobility data to drift–diffusion equations. Symbols are experimental results. (Reprinted from P. Kordt et al., *Phys. Chem. Chem. Phys.*, 17, 22778–22783, 2015. With permission.)

15.12.2 Impedance Spectroscopy

In addition to studies of steady-state current–voltage characteristics, valuable insight into the functioning of organic devices can be obtained from impedance spectroscopy studies. In particular, impedance spectroscopy can provide information about carrier relaxation in the DOS, can help distinguish between different trapping regimes in organic-semiconductor devices [36], and can be used to determine the width σ of the Gaussian DOS [159].

In impedance spectroscopy a dc bias V is applied over a device and, in addition to that, a small ac component $\Delta V(t) = \Delta V \exp(2\pi \mathrm{i} f t)$ is added, where f is the frequency. The impedance $Z = Z' + \mathrm{i}Z''$ is defined as the zero-amplitude limit of the ratio of $\Delta V(t)$ and the response $\Delta I \exp\left[2\pi \mathrm{i}(f + \phi)t\right]$ in the current, with ϕ a phase difference. Of particular interest is the capacitance–voltage, C–V, characteristic, with the capacitance given by $C = -Z''/2\pi f|Z|^2$.

Applying KMC simulations to extract the small response $\Delta I(t)$ is extremely cumbersome because of the noise present in such simulations. It has been shown, however, that for single-carrier organic devices the current–voltage characteristics obtained by solving the master equation are practically the same as those obtained from KMC simulations [160]. The influence of small perturbations can be rather easily evaluated using the master equation. Within the framework of the time-dependent master equation, the small ac component of the voltage in a single-carrier device leads to a time-dependent probability, $p_i(t)$, of the occupation of a site i by a charge, obeying Equation 15.14:

$$\frac{\mathrm{d}p_i}{\mathrm{d}t} = \sum_{j \neq i} \left[\omega_{ji} p_j (1 - p_i) - \omega_{ij} p_i (1 - p_j) \right] \equiv g_i(\vec{p}), \tag{15.34}$$

where $\vec{p}$ is the vector of occupational probabilities of all sites.

Using a perturbative approach, first the steady-state solution, $\vec{p}_0$, for $\mathrm{d}p_i/\mathrm{d}t = 0$ at the applied static voltage, V, has to be evaluated. The procedure for doing so has been described in Ref. [99]. Sheets of sites representing the electrodes are introduced at either side of a simulation box representing the device. An additional small-amplitude ac voltage with frequency f induces a small change, $\vec{\Delta p}$. Linearizing, we write $\vec{p}(t) \approx \vec{p}_0 + \exp\left(2\pi \mathrm{i} f t\right) \vec{\Delta p}$ and $\vec{g}(\vec{p}) \approx \vec{g}(\vec{p}_0) + \exp(2\pi \mathrm{i} f t) \left[\Delta V \partial \vec{g}/\partial V + \hat{J} \vec{\Delta p} \right]$, with the matrix elements of the Jacobian, $\hat{J}$, given by $J_{ij} = \partial g_i/\partial p_j \big|_{\vec{p}_0}$. Substituting these expressions into Equation 15.34 and linearizing leads to the equation

$$(2\pi \mathrm{i} f \hat{I} + \hat{J}) \vec{\Delta p} = -\Delta V \frac{\partial \vec{g}}{\partial V}, \tag{15.35}$$

with $\hat{I}$ denoting the identity matrix. Equation 15.35 can be solved for $\vec{\Delta p}$ with standard techniques, and from this the current, ΔI, and the capacitance, C, are readily obtained.

As an example, we consider two hole-only devices with the structure glass, ITO (100 nm), poly(3,4-ethylenedioxythiophene) polystyrene sulfonate (PEDOT:PSS) (100 nm), light-emitting polymer (LEP), Pd (100 nm). The LEP consists of polyfluorene with 7.5 mol% copolymerized triarylamine units for hole transport; see the Figure 15.8a inset. The LEP-layer thicknesses are $L = 97$ and 121 nm for the two devices and their areas are $A = 9 \times 10^{-6}\ \mathrm{m}^2$. According to EGDM modeling studies of the current density–voltage, J–V, characteristics of these devices [161,162] no injection barrier is present at the anode (PEDOT:PSS) and injection barriers of 1.65 and 1.90 eV are present at the cathode (Pd) for the $L = 97$ and $L = 121$ nm device, respectively. These modeling studies gave best fits for the J–V characteristics with $\sigma = 0.13$ eV.

We apply the above method to calculate the C–V characteristics of these two devices. As in the EGDM studies [161,162] we assume an uncorrelated Gaussian DOS with standard deviation σ. The use of Marcus rates would require knowledge of the reorganization energy λ_{ij}, which is not available here. We therefore assume Miller–Abrahams nearest-neighbor hopping with rates given by Equation 15.22. In the absence

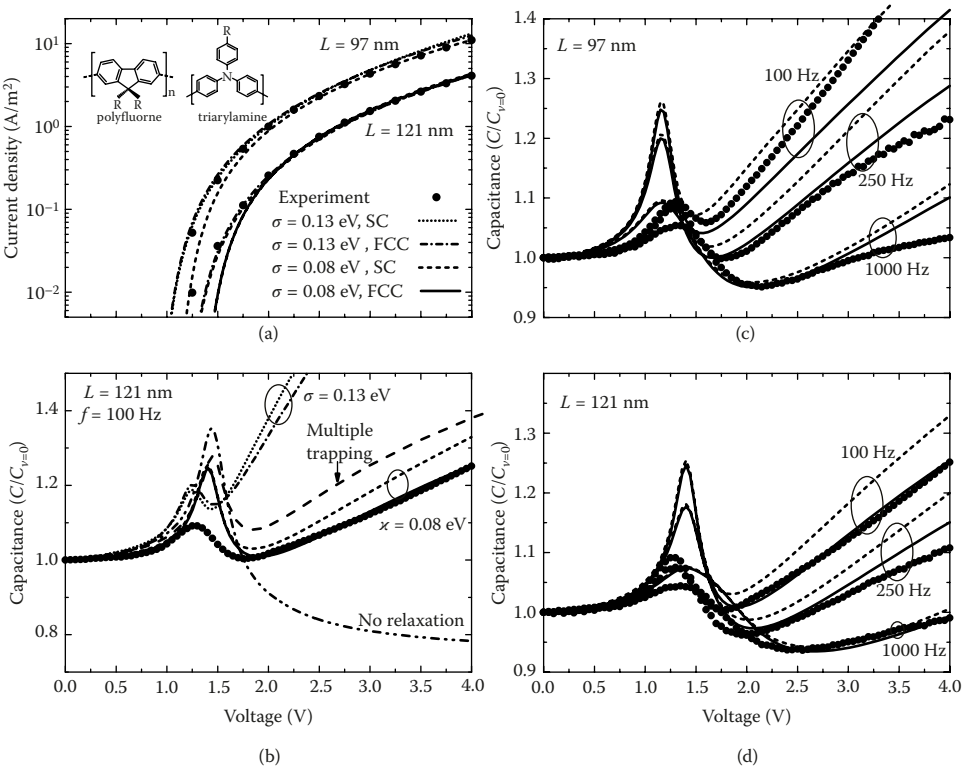

FIGURE 15.8 (a) Current density–voltage, J–V, characteristics of the devices with organic layer thicknesses $L = 97$ and $L = 121$ nm at $T = 295$ K. Dots: measurements [163]. Curves: solutions of the steady-state master equation with $\sigma = 0.13$ and 0.08 eV for simple cubic (SC) and face-centered cubic (FCC) lattices. Inset: the used hole-transporting copolymer. (b) Capacitance, C, normalized to its value at $V = 0$, as a function of V at a frequency $f = 100$ Hz and temperature $T = 295$ K for the $L = 121$ nm device. Dots: measurements [164]. Dash–dot–dotted curve: drift–diffusion calculation with an extended Gaussian disorder model (EGDM) mobility function for $\sigma = 0.13$ eV, neglecting relaxation. Long-dashed curve: multiple-trapping result, which includes relaxation [164]. Other curves: Solutions of the time-dependent master equation for $\sigma = 0.13$ and 0.08 eV and for SC and FCC lattices. (c) and (d) C–V characteristics at different frequencies for the $L = 97$ nm (c) and $L = 121$ (d) device. Dots: Measurements [164]. Curves: Solutions of the time-dependent master equation. (Reprinted with permission from M. Mesta, J. Cottaar, R. Coehoorn, and P. A. Bobbert, Study of charge-carrier relaxation in a disordered organic semiconductor by simulating impedance spectroscopy. *Appl. Phys. Lett.*, 104(21), 213301, 2014. Copyright [2014], American Institute of Physics.)

of any morphological information we assume a regular lattice of hole-transporting sites. To investigate a possible influence of morphology we investigate simple cubic (SC) as well as face-centered cubic (FCC) lattices with a lattice constant $a = 1.19$ nm for the SC lattice and $a = 1.88$ nm for the FCC lattice, in accordance with the known density 1.8×10^{26} m^{-3} of hole-transporting units. The simulation boxes have dimensions $L \times L_y \times L_z$, with $L_y = L_z = 50a$ and periodic boundary conditions in the y- and z-directions, yielding a sufficient lateral averaging. For further details we refer to Ref. [159].

As expected from the EGDM modeling [161,162] it can be observed in Figure 15.8a that the experimental J–V characteristics (dots) at room temperature ($T = 295$ K) of the two devices are very well described by the solution of the steady-state master equation for $\sigma = 0.13$ eV, both for the SC (dotted curve) and the FCC (dash-dotted curve) lattice (a prefactor in the hopping rates was adjusted in both cases to obtain an optimal fit). However, this is not at all true for the C–V characteristics. It is seen in Figure 15.8b that

for the $L = 121$-nm device solving the time-dependent master equation for $\sigma = 0.13$ eV at a frequency $f = 100$ Hz yields results that deviate strongly from the experimental C–V characteristic. The fact that the master-equation results for the SC and FCC lattice are quite comparable shows that this deviation is probably not due to a morphological issue.

In order to understand the problem better we first distinguish the different regimes in the C–V characteristics. (1) At low voltage all characteristics converge to the geometrical capacitance, because almost no carriers are present in the device. (2) With increasing voltage, a sheet of holes builds up by diffusion from the anode, but these cannot yet move to the cathode because the electric field is still directed from cathode to anode. As a result, the effective thickness of the device decreases and the capacitance rises. (3) When approaching the built-in voltage V_{bi} these holes start to move to the cathode, leading to a decrease of the capacitance. The result is a peak in the C–V curve before V_{bi} is reached [165]. (4) In the regime beyond V_{bi} the C–V curve rises again. Here relaxation effects play a dominant role and therefore this is the regime we want to focus on.

In order to identify the effects of relaxation we display in Figure 15.8b (dash–dot–dotted curve) the C–V characteristic obtained by solving the time-dependent drift–diffusion equation with the EGDM mobility function corresponding to $\sigma = 0.13$ eV. In this case, the local mobility $\mu(x; \rho, F, T)$ depends on the instantaneous local charge density ρ and electric field F, and therefore contains no relaxation effects. It is seen that without relaxation effects the capacitance decreases after V_{bi} to a value that is even smaller than the geometrical capacitance. The long-dashed curve is the result of a multiple-trapping model for relaxation [164]. With a fitted conduction-level energy $E_c = -0.75\sigma$, this model leads to a fair agreement with experiment.

Since solving the time-dependent master equation for $\sigma = 0.13$ eV apparently overestimates relaxation effects and since such effects decrease with decreasing σ we solved the time-dependent master equation for lower values of σ. With a value of $\sigma = 0.08$ eV we find a very satisfactory agreement with the experimental C–V characteristics, not only for the device and frequency considered in Figure 15.8b, but also for both devices and all considered frequencies; see Figure 15.8c and 15.8d. The only clear disagreement is in the peak, which is more pronounced in the calculations than in the experiment. This may be partially explained by lateral variations in V_{bi} of the devices [164]. The dashed (SC) and full (FCC) curves in Figure 15.8a are the corresponding J–V curves obtained by solving the steady-state master equation. It is observed that for high voltages the experimental J–V curves are very well described, but significant deviations occur at low voltages around V_{bi}.

The analysis brings up the question why there is such an apparent discrepancy between the description of steady-state and time-dependent charge transport. A possible explanation is that in steady-state transport the low-energy tail of the DOS is important, represented by a relatively large σ, while in time-dependent transport relaxing carriers probe a larger part of the DOS, represented by a smaller σ. This would mean that the shape of the DOS is more complicated than a single Gaussian. It would also explain the difference in the description of the J–V curves in Figure 15.8c. At low voltage, when carriers only occupy the low-energy tail of the DOS, $\sigma = 0.13$ eV gives a better description, while at higher voltage, when the DOS is filled up further, $\sigma = 0.08$ eV provides an excellent description, which is even slightly better than with $\sigma = 0.13$ eV. A value of $\sigma = 0.13$ eV at low voltage could also partially explain the lower peak in the C–V curve as compared to the calculations with $\sigma = 0.08$ eV in Figure 15.8b through d. We note that the position of the peak could be improved by adapting the used built-in voltages V_{bi}. These voltages were obtained from an EGDM fit of the J–V characteristics with $\sigma = 0.13$ eV [161,162], but should be optimized again in a fit with $\sigma = 0.08$ eV. In addition, the EGDM neglects spatial correlations of site energies. This can lead to discrepancies when analyzing an experimental system with correlations in terms of a model without correlations [89].

The present conclusion that $\sigma = 0.08$ eV should be used to describe carrier relaxation in the considered devices is fully in agreement with the conclusion that dark-injection experiments on the same devices, which also probe carrier relaxation, can be described by solving a time-dependent master equation with the same value of σ [163].

15.12.3 Efficiency

In OLEDs, electrical power is converted to a radiant flux (radiant energy emitted per unit time), Φ_e. The power efficiency, sometimes called the wall-plug efficiency, is given by

$$\eta_{power} = \frac{\Phi_e}{IV} = \frac{\int_0^\infty \Phi_{e,\lambda,OLED}(\lambda)d\lambda}{IV}, \tag{15.36}$$

where I is the current, V is the applied voltage, and $\Phi_{e,\lambda,OLED}$ is the total optical power that is emitted externally per unit wavelength λ. The power efficiency is generally limited by inevitable Ohmic losses in the electrodes and in the organic charge transport layers, and sometimes also by Ohmic losses due to the presence of internal organic–organic energy barriers outside the emissive layer. When judging the efficiency of the conversion process in the emissive layer, one therefore often focuses on a complementary quantity, the external quantum efficiency η_{EQE} (EQE), that is defined as the total number of externally emitted photons per charge carrier that has passed the device:

$$\eta_{EQE} = \frac{e}{I}\int_0^\infty \Phi_{e,\lambda,OLED}(\lambda)\frac{\lambda}{ch}d\lambda, \tag{15.37}$$

where e is the fundamental charge, c is the speed of light, and h is Planck's constant. Due to full or partial internal reflection of light, not all photons that are internally generated will escape from the microcavity that is formed by the OLED layer structure. It is therefore useful to introduce an additional quantity, the internal quantum efficiency (IQE) η_{IQE}, which is defined as the ratio of the total number of photons generated within the device and the number of electrons injected. The IQE is not directly measurable, but may be derived from the EQE using the expression

$$\eta_{IQE} = \frac{\eta_{EQE}}{\eta_{out}}, \tag{15.38}$$

where η_{out} is the light-outcoupling efficiency. For emission from a specific position in a planar OLED microcavity, under a specific angle and for a specific wavelength, the (*s* and *p*) polarization-dependent emitted light intensity may be obtained from optical simulations [166–169]. The light-outcoupling efficiency is thus an effective value, which is determined by averaging over the entire emission profile and the entire emission spectrum and which is sensitive to the precise angular dependence of the emission from the dye molecules. Application of advanced emission profile reconstruction techniques [170,171] and a measurement of the emitter orientation distribution [172] are required to determine η_{out} for a specific case with high precision. In the absence of such information, one often assumes that for well-designed phosphorescent OLEDs with a (glass | ITO | organic semiconductor | Al) layer structure and with a random emitter orientation η_{out} is approximately 0.2. Larger values, up to ∼0.25−0.30, are possible by optimizing all layer thicknesses [173]. Several methods, including the use of a roughened external glass surface or the use of internal high-refractive index scattering layers, have been developed to enhance the light-outcoupling efficiency to values above 0.5 [174,175].

Recently, much progress has been made in advanced molecular-scale KMC calculations of the IQE [176–178]. We focus in this section on applications of KMC simulations to phosphorescent OLEDs based on a small concentration of metal-organic emitter molecules in a matrix material. In general, the IQE may be expressed as [174]

$$\eta_{IQE} = \eta_{rec}\eta_{ST}q_{eff}, \tag{15.39}$$

with η_{rec} the recombination efficiency, defined as the fraction of injected charges which contributes to exciton formation, η_{ST} the singlet–triplet factor, defined as the fraction of generated excitons which is quantum-mechanically allowed to decay radiatively, and q_{eff} the effective radiative decay efficiency, defined

as the fraction of such excitons which actually decays radiatively. In phosphorescent OLEDs based on heavy metal-organic molecules, strong spin–orbit interaction gives rise to triplet states with some mixed-in singlet-character, so that also triplets are emissive and $\eta_{ST} = 1$. The recombination efficiency can be close to unity by making use of appropriate electron and hole blocking layers. A highly effective radiative decay efficiency may be obtained, first, by using emissive dye materials with a large radiative decay rate, Γ_{rad}, and a small nonradiative decay rate, Γ_{nr}. In the absence of other loss processes, the IQE is then equal to $\Gamma_{rad}/(\Gamma_{rad} + \Gamma_{nr}) \equiv \eta_{PL}$, the photoluminescence (PL) efficiency. Nonradiative decay is a result of the nuclear motion, so that the energy of the molecule in its excitonically excited state can be equienergetic with a highly vibrationally excited excitonic ground state [179]. Second, the matrix material and the adjacent blocking material should have a triplet energy level significantly larger than the dye triplet level so that the triplet excitons stay confined to the dye sites. These design rules are already relevant to the IQE at small current densities. Experimentally, the IQE is found to depend on the current density, J. At large J, η_{IQE} decreases with increasing J. For some devices, η_{IQE} is found to show a broad maximum before the decrease ("roll-off") sets in. A practical measure is the current density J_{90} at which the IQE has decreased to 90% of the maximum value. For efficient phosphorescent OLEDs, maximum reported values of J_{90} are approximately 300 A/m^2 [180]. In commercial white OLEDs for lighting conditions, operated at high luminance levels, the efficiency loss due to roll-off can be indeed of the order of 10%.

Understanding the roll-off is not only important as a first step toward enhancing the efficiency. Loss processes that limit the IQE at high J can also trigger local degradation processes with a certain probability, as will be discussed in Section 15.12.5. By building "virtual OLEDs" in which the interplay of all charge transport and excitonic processes is included mechanistically using KMC simulations, the functioning of OLEDs can be studied with subnanosecond time and molecular-scale spatial resolution. The first demonstration of the feasibility of such an approach was presented by van Eersel et al. [176]. We discuss their simulation results for OLEDs based on the green-emitting metal-organic molecule tris[2-phenylpyridine]iridium ($Ir(ppy)_3$) and the red phosphorescent dye platinum octaethylporphyrin (PtOEP).

The simulations were based on the three-dimensional (3D) KMC code Bumblebee (http:/simbeyond.com): For a detailed discussion of the model used, we refer to Refs. [176,177,181,182]. Ref. [176] also gives motivations for the parameter values used and provides analyses of the sensitivity of the simulation results to the parameter values. Briefly, the OLEDs were modeled as a collection of molecular sites on a simple cubic lattice. For each type of molecule, the site energies for electrons and holes were taken randomly from a Gaussian DOS with an average energy as given in Figure 15.9a. Charges were assumed to hop with a rate as described within the Miller–Abrahams formalism Equation 15.22. The hopping attempt frequency ω_0 and the wavefunction decay length $\lambda = 1/\gamma$ were taken equal for all pairs of sites. The simulations included the Coulomb interactions between all charge carrier pairs and with image charges in the metallic electrodes. In a natural way, the formation of space-charge layers near the injecting and organic–organic interfaces, and the resulting "band bending," was thus included. Cottaar et al. have demonstrated that in energetically disordered materials, as they are used in OLEDs, explicitly taking the individual 3D Coulomb interactions into account is important for properly treating charge accumulation near internal interfaces [183]. Instantaneous ISC was assumed, so that only triplet excitons were considered. Exciton generation and dissociation were treated in the same way as hops of electrons and holes, but including the triplet exciton binding energy. Radiative and nonradiative triplet exciton decay was included, as well as exciton transfer between the dye molecules, leading to exciton diffusion. The transfer rate was expressed as a sum of Förster-type and Dexter-type contributions, Equations 15.25 and 15.26, where $\tau = (\Gamma_{rad} + \Gamma_{nr})^{-1}$ is the effective decay time, R_F is the Förster radius for diffusion, and k_D is the Dexter prefactor. The two bimolecular loss processes that potentially contribute to the IQE roll-off, triplet–polaron quenching (TPQ) and triplet–triplet annihilation (TTA), were both included in a parameter-free manner, viz. by assuming an infinite (zero) rate when an exciton and a polaron or two excitons, respectively, are present on nearest-neighbor (more distant) sites. Table 15.2 gives an overview of the used parameter values.

Figure 15.10a and b shows a comparison of the calculated and experimental J–V and IQE roll-off curves, respectively, for a temperature of 300 K. The slope of the J–V curves is well described, but the absolute

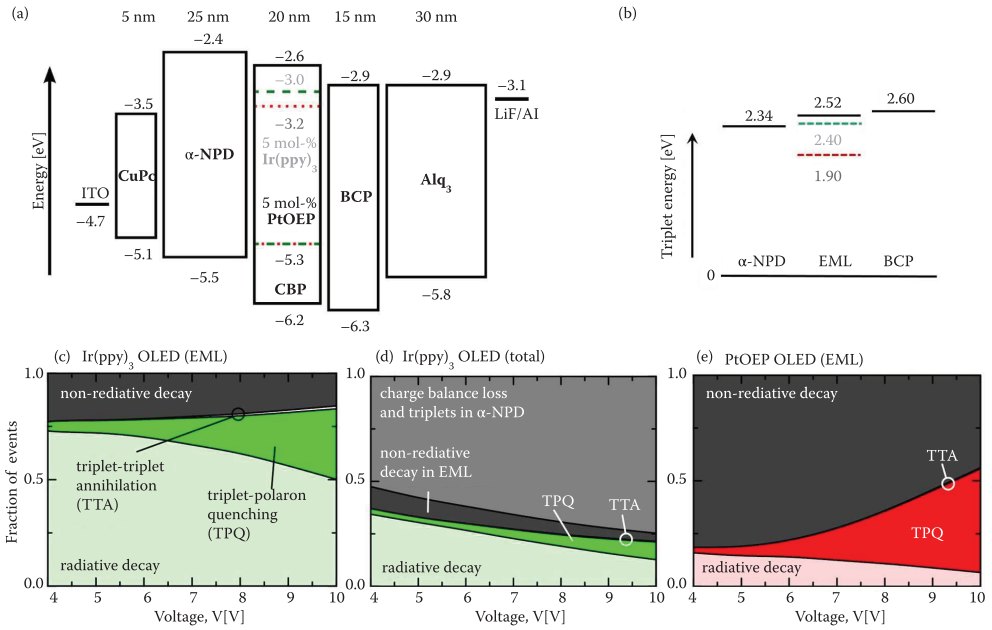

FIGURE 15.9 (a) Energy level and layer structure of the green and red organic light-emitting diodes (OLEDs) studied in Ref. [176]. The phosphorescent emissive layer (EML) is sandwiched in between materials facilitating hole and electron injection, transport and blocking: CuPc (copper phthalocyanine), α-NPD (4,4′-bis[N-(1-naphthyl)-N-phenyl-amino] biphenyl), BCP (2,9-dimethyl, 4,7-diphenyl, 1,10-phenanthroline), and Alq_3 (tris [8-hydroxyquinoline] aluminum). (b) Triplet energies for the materials used in the EML (solid line: CBP; dashed line: $Ir(ppy)_3$; dotted line: PtOEP) and the layers adjacent to the EML. (c)–(e) Contribution of the various exciton decay processes in the EML of the $Ir(ppy)_3$ device (b), in the entire $Ir(ppy)_3$ device (c), and in the EML of the PtOEP device (d). The figures show that even above 6 V only a small fraction of the efficiency loss (less than 2% (0.5%) for the $Ir(ppy)_3$ (PtOEP) devices) is due to triplet–triplet annihilation (TTA). (Reprinted with permission from H. van Eersel, P. A. Bobbert, R. A. J. Janssen, and R. Coehoorn. Monte Carlo study of efficiency roll-off of phosphorescent organic light-emitting diodes: Evidence for dominant role of triplet-polaron quenching. *Appl. Phys. Lett.*, 105(14):143303, 2014. Copyright [2014], American Institute of Physics.)

value of the current density is somewhat overestimated. As argued in Ref. [176], this might be related to an underestimation of the HOMO–LUMO gap, which was taken to be equal to the optical gap. This often-used approach neglects the exciton binding energy, which can be around 1 eV [184]. Such a correction would horizontally shift the J–V curves by 1 eV, giving rise to a significant reduction of the discrepancy. The roll-off curve would not be affected by such a correction. We note that (as mentioned above) the triplet exciton binding energy was included when calculating the rates of exciton generation and dissociation processes. For $Ir(ppy)_3$, the simulation results agree within the error margin with experiment, whereas for PtOEP the roll-off is slightly underestimated at high current densities. A sensitivity analysis was carried out to find out which uncertainties in the choice of parameter values have the largest impact. It was found, e.g., that the J–V characteristics are quite strongly determined by the energy-level differences at interfaces and between the host and guest states in the emissive layer (EML). The sensitivity to the hopping attempt frequency was found to be relatively small.

The simulations provided detailed views on the cause of the roll-off, as shown in Figure 15.9c through e. In both devices, most of the emission was found to occur near the anode-side of the EML. This may be understood from Figure 15.9a, from which the guest molecules are expected to give rise to stronger hole trapping than electron trapping. As a result, the effective electron mobility is larger than the hole mobility. In the emissive layer, most of the IQE loss was found to be due to TPQ, and only at high voltages

TABLE 15.2 Overview of the Simulation Parameters

Parameter	Description	Value	
Common			
ω_0	Hopping attempt frequency to the first neighbor	3.3×10^{10} s^{-1}	
σ	Width of the electron and hole Gaussian DOS	0.10 eV	
N_t	Site density	1.0×10^{27} m^{-3}	
$\lambda \equiv 1/\gamma$	Wavefunction decay length	0.3 nm	
ε_r	Relative dielectric permittivity	3.5	
σ_T	Width of the triplet exciton DOS	0.10 eV	
$E_{T,b}$	Triplet exciton binding energy	1.0 eV	
k_D	Prefactor for triplet exciton Dexter transfer	1.6×10^{10} s^{-1}	
Material-specific		$Ir(ppy)_3$	PtOEP
R_F	Förster radius for triplet exciton diffusion	1.5 nm	1.5 nm
Γ_{rad}	Radiative decay rate	0.816 μs^{-1}	0.1 μs^{-1}
Γ_{nr}	Nonradiative decay rate	0.249 μs^{-1}	0.525 μs^{-1}

Note: The highest occupied molecular orbital (HOMO) and lowest unoccupied molecular orbital (LUMO) energies are given in Figure 15.9a, and the triplet energies are given in Figure 15.9b.

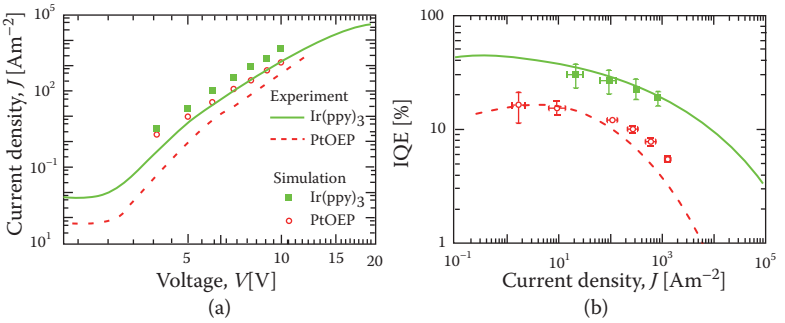

FIGURE 15.10 (a) Simulated and experimental J–V characteristics. (b) Simulated internal quantum energy (IQE) and experimental IQE (assuming 20% outcoupling efficiency) as a function of the current–density. Experimental data from Ref. [185]. (From N. C. Giebink and S. R. Forrest, *Phys. Rev.* B., 77, 235215, 2008; Reprinted with permission from H. van Eersel et al., Monte Carlo study of efficiency roll-off of phosphorescent organic light-emitting diodes: Evidence for dominant role of triplet–polaron quenching. *Appl. Phys. Lett.*, 105(14):143303, 2014. Copyright [2014], American Institute of Physics.)

was a small TTA contribution was found (see Figure 15.9c). For the $Ir(ppy)_3$ devices, the overall loss was found to be determined mostly by a nonideal recombination efficiency due to imperfect electron blocking and by triplet exciton diffusion to the α-NPD layer (see Figure 15.9d). The finding of electron loss to the α-NPD layer is consistent with the observation of some blue emission from that layer [185]. Due to subsequent TPQ, triplet transfer to the α-NPD layer also gives rise to a loss. From simulations, improving the devices by introducing perfect electron blocking was predicted to give rise to a 10% increase of the IQE at small voltages, from ∼ 34% to ∼ 44%. The IQE at small voltages was found to become equal to the PL efficiency assumed (77%) when the triplet transfer to the HTL also was eliminated. For the red devices, at low voltages no significant electron and triplet loss to the α-NPD layer was found, as may be understood from the lower LUMO energy and the smaller triplet energy of PtOEP. The IQE is then close to the PL efficiency assumed (16.4%).

15.12.4 Electroluminescence of a White OLED

The KMC simulations described in Section 15.5 can be employed to model all molecular-scale electronic processes that finally lead to electroluminescence of an OLED: injection, transport, and recombination of electrons and holes as well as diffusion and radiative decay of excitons. We consider here the white multilayer OLED stack of Figure 15.11a, which was studied experimentally and by KMC simulations in Ref. [181]. It concerns a so-called "hybrid" OLED, which combines red and green phosphorescent emission with blue fluorescent emission. Phosphorescent emission can be very efficient because of the harvesting of both singlet and triplet excitons (see Section 15.2.4). However, since stable blue phosphorescent emitters with long-term stability are to date unavailable, many commercial white OLEDs make use of blue fluorescent emission, despite the fact that then only singlet excitons are harvested.

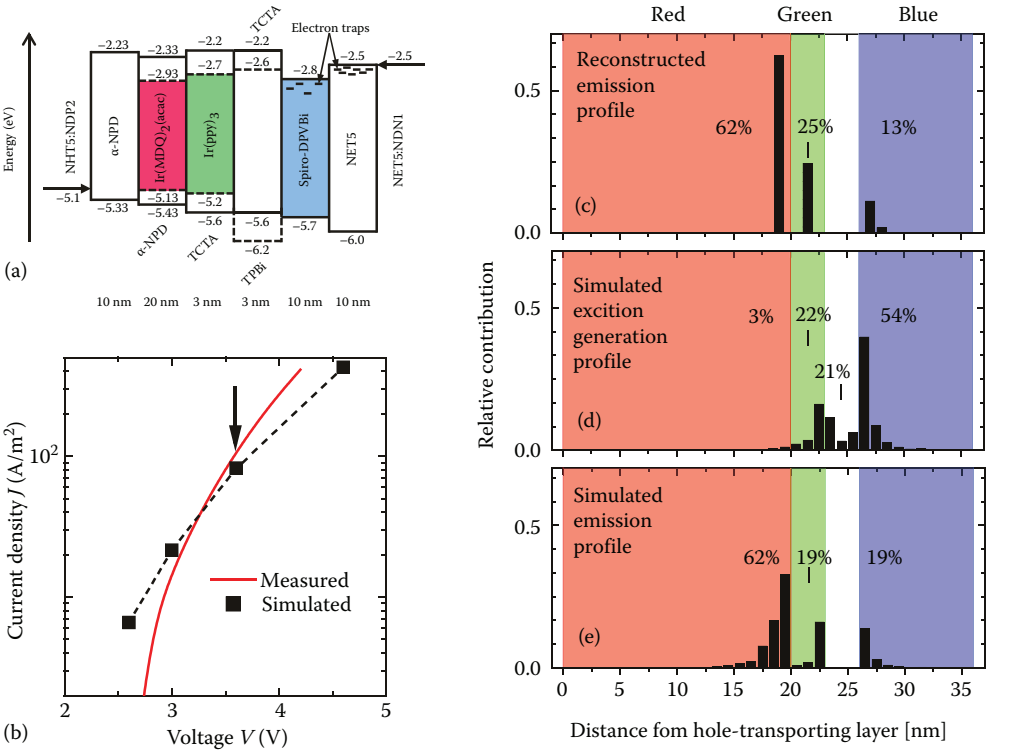

FIGURE 15.11 (a) Stack structure of the investigated white organic light-emitting diode (OLED) with the different layers and their thicknesses. The highest occupied molecular orbital (HOMO) and lowest unoccupied molecular orbital (LUMO) energy levels as used in the kinetic Monte Carlo (KMC) simulations are indicated by solid lines for host materials and dashed lines for guest materials. The horizontal arrows indicate the work functions used for the doped hole- and electron-injecting layers. Electron traps were assumed in the 2,2′,7,7′-tetrakis(2,2-diphenylvinyl)spiro-9,9′-bifluorene (spiro-DPVBi) and NET5 layers with concentrations of 10^{-3} and 5×10^{-3}, and characteristic energies $k_B T_0 = 0.2$ and 0.12 eV, respectively, of their exponential density of states (DOS). (b) Measured and simulated room-temperature current density-voltage, J–V characteristics. All other results presented are for the bias voltage $V = 3.6$ V, indicated by the arrow. The relative error in the simulated current density is estimated to be about 10%. (c) Emission profile reconstructed from the measured angle- and polarization-dependent emission spectrum [181]. The percentages of red, green, and blue emission are indicated. (d) Exciton generation profile as obtained from the KMC simulations. (e) Emission profile as obtained from the KMC simulations, after taking into account exciton diffusion and the radiative decay probabilities of the emitters. The error in the simulated profiles is about 1%. Reprinted by permission from Macmillan Publishers Ltd. *Nat. Mater.* [181] Copyright (2013).

The working principle of the OLED of Figure 15.11a is the following. Holes reach the light-emitting layers of the stack after being injected by an ITO layer into a 4 mol% *p*-doped injection layer of NHT5:NDP2 and transported through a hole-transporting and electron-blocking layer of α-NPD (N,N′-di(naphthalen-1-yl)-N,N′-diphenyl-benzidine). Electrons reach the emitting layers after being injected by an aluminum cathode into a 4 mol% *n*-doped electron-injection of NET5:NDN1 and transported through an electron-transporting and hole-blocking layer of NET5 (materials supplied by Novaled). Blue light is generated in a 10-nm-thick fluorescent layer of spiro-DPVBi (2,2′,7,7′-tetrakis(2,2-diphenylvinyl)spiro-9,9′-bi fluorene) adjacent to the NET5 layer. Green light is generated in a 3-nm-thick layer of TCTA (4,4′,4″-tris(N-carbazoyl)-triphenylamine) doped with 8 mol% of the green phosphorescent dye $Ir(ppy)_3$ (fac-tris(2-phenylpyridyl)iridium). Red light is generated in a 20-nm-thick α-NPD layer doped with 5 mol% of the red phosphorescent dye $Ir(MDQ)_2(acac)$ ((acelylacetonate)bis(2-methyldibenzo[f,h]quinoxinalate)iridium).

The green phosphorescent layer is separated from the blue fluorescent layer by a thin (3-nm) interlayer consisting of a mixture of the hole-transporter TCTA with 33 mol% of the electron-transporter TPBi (1,3,5-tris(N-phenylbenzimidazol-2-yl)benzene). This interlayer has several purposes [186]. It should block the transfer of singlet excitons from the blue to the green layer and of triplet excitons from the green to the blue layer (spiro-DPVBi has a triplet energy lower than that of $Ir(ppy)_3$). Also, this interlayer should allow the passage of electrons from the blue to the green layer (by the TPBi) and of holes from the green to the blue layer (by the TCTA). The red phosphorescent layer is purposely in direct contact with the green phosphorescent layer, allowing triplet excitons formed on the phosphorescent dye in the green layer to diffuse to the phosphorescent dye in the red layer. This diffusion is an important process in establishing the right color balance.

The measured room-temperature current density–voltage characteristic of the OLED is shown in Figure 15.11b. The arrow indicates the bias voltage of 3.6 V for most of the reported results. The CIE 1931 color point of the perpendicularly emitted light at this bias was measured to be $[x, y] = [0.47, 0.45]$, which corresponds to warm-white emission. The EQE, i.e., the fraction of emitted photons per injected electron–hole pair, is measured to be $5 \pm 1\%$ [181]. Figure 15.11c shows the color-resolved emission profile, which was reconstructed with a precision on the order of a nanometer from the measured angle- and polarization-dependent emission spectra [170,181]. It is observed that in the blue layer, the emission occurs close to the interface with the interlayer and in the red layer close to the interface with the green layer.

The KMC simulations of the charge dynamics were carried out in the following way. Point sites are arranged on a cubic lattice with a lattice constant $a = 1$ nm, the typical intermolecular distance of the used molecular semiconductors represent the molecules in the stack. A simulation box of $50 \times 50 \times 56$ sites with periodic boundary conditions in the lateral (x- and y-) directions was used, which turned out to yield sufficiently accurate results. Since charge transport in various small-molecule materials was found to be described well by the ECDM [161,191], correlated disorder caused by random dipoles was assumed for the electron and hole energies. The energetic disorder was taken to be $\sigma = 0.1$ eV, corresponding to the value found for hole transport in α-NPD [161]. With this value the charge transport in all materials in the stack is expected to be reasonably described. Red and green emitting guests were introduced according to the known concentration of the emitters, with appropriately adapted energy levels. Electron traps were introduced in the layers in which electron transport is important: the blue fluorescent layer and the ETL. They were modeled with an exponential DOS of trap energies [192–194] with a concentration c_{trap} and a characteristic energy $k_B T_0$.

Nearest-neighbor hopping of charges on the lattice using the Miller–Abrahams rate (see Equation 15.22) was assumed. The energy differences in the hopping rates contain, apart from the random site energies, an electrostatic contribution due to the bias applied to the OLED and the Coulomb energy due to all present charges. The doped injection layers were treated as metallic, injecting and collecting charges with an energy according to their work function, indicated by the arrows in Figure 15.11a. Exciton generation was assumed to occur by hopping of an electron to a site where a hole resides, or vice versa, and was assumed to be always an energetically downward process.

Table 15.3 gives the parameters of the stack materials used in the KMC simulations of the charge dynamics. The parameters were determined from charge-transport and spectroscopic studies of the various materials [181]. Exciton diffusion within the green and red layers and from the green to the red layer was included in the simulations. Since the red and green emitters trap electrons as well as holes (see the energy level scheme in Figure 15.11a), almost all excitons in the red and green layers are generated on the emitters. The diffusion of excitons among the emitters was described by Förster transfer (see Equation 15.25), made possible by the spin–singlet character that is mixed into the exciton wave function by the spin-orbit coupling of the heavy iridium atoms. Apart from being transferred, excitons can decay radiatively with a rate $\Gamma_{\mathrm{rad},i} = 1/\tau_{\mathrm{r},i}$ or nonradiatively with a rate $\Gamma_{\mathrm{nr},i} = 1/\tau_{\mathrm{nr},i}$. These rates are related to the radiative decay probabilities η_{r} by $\eta_{\mathrm{r},i} = \Gamma_{\mathrm{rad},i}/(\Gamma_{\mathrm{rad},i} + \Gamma_{\mathrm{nr},i})$. We took $\eta_{\mathrm{r}} = 0.84$ and 0.76 for the red and green phosphorescent emitters in their respective hosts [169]. Other parameters used in the exciton dynamics are given in Table 15.4. Only exciton transfer from green to red was taken into account. Transfer from red to green should be negligible.

TABLE 15.3 Highest Occupied Molecular Orbital (HOMO) and Lowest Unoccupied Molecular Orbital (LUMO) Energies, Room Temperature Hole- and Electron-Mobilities $\mu_{0,\mathrm{h}}$ and $\mu_{0,\mathrm{e}}$ at Low Field and Low Carrier Density, Electron-Trap Concentration c_{trap}, and Trap Temperature T_0 of the Exponential Trap Density of States (DOS) in the Different Layers of the Stack, as Used in the Kinetic Monte Carlo (KMC) Simulations

Material	E_{HOMO} (eV)	E_{LUMO} (eV)	$\mu_{0,\mathrm{h}}$ ($\mathrm{m^2/Vs}$)	$\mu_{0,\mathrm{e}}$ ($\mathrm{m^2/Vs}$)	c_{trap}	T_0 (K)
NHT5:NDP2	−5.10					
α-NPD	−5.43	−2.33	6×10^{-9}	6×10^{-10}		
$\mathrm{Ir(MDQ)_2(acac)}$	−5.13	−2.93	6×10^{-9}	6×10^{-10}		
TCTA	−5.60	−2.20	2×10^{-8}	2×10^{-9}		
$\mathrm{Ir(ppy)_3}$	−5.20	−2.70	2×10^{-8}	2×10^{-9}		
TCTA	−5.60	−2.20	2×10^{-8}	2×10^{-9}		
TPBi	−6.20	−2.60	2×10^{-8}	2×10^{-9}		
Spiro-DPVBi	−5.70	−2.80	6×10^{-9}	8×10^{-9}	0.001	2350
NET5	−6.00	−2.50	1.5×10^{-11}	1.5×10^{-10}	0.005	1400
NET5:NDN1		−2.50				

Source: M. Mesta et al., *Nat. Mater.*, 12, 652–658, 2013.
Note: Spiro-DPVBi, 2,2′,7,7′-tetrakis(2,2-diphenylvinyl)Spiro-9,9′-bifluorene; TCTA, 4,4′,4″-tris(N-carbazoyl)-triphenylamine; α-NPD, N,N′-di(naphthalen-1-yl)-N,N′-diphenyl-benzidine; MDQ, acelylacetonate)bis(2-methyldibenzo[f,h]quinoxinalate)iridium; TPBI, 1,3,5-tris(N-phenylbenzimidazol-2-yl)benzene. NHT5:NDP2 and NET5:NDN1 are doped hole and electron transporter materials supplied by Novaled.

TABLE 15.4 Radiative and Nonradiative Exciton Decay Rates Γ_{rad} and Γ_{nr}, Förster Radii R_F for Exciton Transfer between Phosphorescent Emitter Molecules and Triplet Energies E_{T}

Material	Γ_{rad} ($\mu\mathrm{s}^{-1}$)	Γ_{nr} ($\mu\mathrm{s}^{-1}$)	R_{F} (nm)	E_{T} (eV)
$\mathrm{Ir(MDQ)_2(acac)}$	0.588	0.112	1.5	2.0
$\mathrm{Ir(ppy)_3}$	0.816	0.249	1.5	2.4

Note: For the green emitter $\mathrm{Ir(ppy)_3}$ Γ_{rad} and Γ_{nr} were taken from Ref. [187]. For the red emitter $\mathrm{Ir(MDQ)_2(acac)}$, Γ_{rad} was obtained from Ref. [188], and Γ_{nr} from η_{r} is given in Ref. [169] and the relation $\eta_{\mathrm{r}} = \Gamma_{\mathrm{rad}}/(\Gamma_{\mathrm{rad}} + \Gamma_{\mathrm{nr}})$. The value $R_F = 1.5$ nm for transfer between equal emitter molecules is a typical value given in Ref. [189]. According to the estimate in Ref. [190] we took $R_{F,\mathrm{GR}} = 3.5$ nm for the transfer from a green to a red emitter. The triplet energies were taken from Ref. [186].

For each exciton generated in the red or green layer a separate simulation of its dye-to-dye diffusion and final radiative or nonradiative decay was performed. This diffusion was assumed to proceed independently from all other processes, which means that exciton quenching processes were neglected. Excitons generated on host sites in the red and green layers (a small fraction) were assumed to transfer instantaneously to an emitter in their neighborhood. Diffusion of excitons generated in the blue layer was not accounted for because the diffusion length of these excitons is short and because their transfer to the green layer is blocked by the interlayer. We assumed that in the blue fluorescent layer singlet and triplet excitons are generated in a quantum-statistical ratio of 1:3 and that the triplet excitons are lost. For the radiative decay probability of singlet excitons in the blue layer we took $\eta_r = 0.35$ [195]. Excitons generated in the interlayer were assumed to be lost by nonradiative decay or emission outside the visible spectrum.

The J–V characteristic following from the KMC simulations is given in Figure 15.11b. Considering the various simplifications that were made, the agreement with the measured characteristic is fair. The underestimation of J at high voltage could be due to a heating effect, while the overestimation of J at low voltage could result from a systematic underestimation of the LUMO energies due to neglect of the exciton binding energies. At the bias voltage of 3.6 V for most of the reported results, the experimental and simulated current density agree quite well.

Figure 15.11d and e present the simulated exciton generation profile and emission profile, respectively, at 3.6 V. The effect of exciton transfer from green to red is clearly observable and is very important for the color balance of this OLED. The simulated emission profile is in fair agreement with the reconstructed emission profile from Figure 15.11c. The broadening of the simulated emission profile over a few nanometers in the red layer found in the simulations is not seen in the reconstructed emission profile but this could be due to the limited resolution of the reconstruction procedure [170,181]. The total percentages of emission in the red, green, and blue agree quite well with the reconstructed emission profile.

The results presented in this section show that KMC simulations of all molecular-scale electronic processes leading to electroluminescence in quite complicated multilayer OLEDs with commercial relevance are feasible. The parameters in the present study were almost all obtained from experimental studies, but there is no obstacle for obtaining these from first-principles computational studies. Complete *in silico* studies of the functioning of commercial OLEDs therefore seem to be within reach.

15.12.5 Degradation

Developing an improved understanding of the mechanisms that limit the operational lifetime of OLEDs is of key importance toward the further adoption of OLED technology for display, lighting, and signage applications. Experimentally, given fixed current density and ambient temperature conditions, the luminance is often observed to decrease with time in an approximately exponential or stretched-exponential manner. For commercial white OLEDs for lighting applications, the time at which the luminance has dropped to 70% of the initial value (the so-called LT_{70} lifetime) can today be as large as 10,000 hours or more at a luminance of 8000 cd/m^2. However, long lifetimes are in practice often realized by making a trade-off with device efficiency (e.g., when using hybrid OLEDs, see Section 15.12.4) or production cost and ease of manufacturing (e.g., when using multiply vertically stacked OLEDs). Examples of other aspects of OLED reliability during prolonged operation are the voltage stability (at a fixed current density), the color point stability, the stability at high ambient temperatures, and the stability under prolonged exposure to the ambient atmosphere [196,197]. The latter issue, leading to so-called black-spot formation, has been mitigated by the development of improved encapsulation technologies [198]. Recently, an excellent overview of the degradation mechanisms and reactions in OLEDs has been given by Scholz et al. [199].

In this section, we focus on the use of molecular-scale OLED device modeling as a means to elucidate the role of various possible intrinsic degradation processes, i.e., processes that are caused by excitons (including those due to absorption of internally emitted photons), charges, and fields in the opto-electronically active organic semiconducting layer. When setting up an extension of the KMC simulations described in Sections 15.5 and 15.12.3, in order to include degradation, first an inventory should be made of the

processes which for a specific system are expected to give rise to degradation. Formally, a distinction should be made between (1) monomolecular degradation processes, which occur when the state of a single-specific molecule is modified, e.g., due to the presence of a polaron or an exciton or due to a local electric field, and (2) bimolecular degradation processes, which occur upon an interaction between charges and/or excitons on two different molecules, e.g., exciton–exciton annihilation or exciton–polaron quenching.

Including these two types of degradation processes requires a rate and a probability, respectively, as well as a description of the resulting changes of the KMC parameter values. Monomolecular degradation has been observed, e.g., in OLEDs containing the electron transport material Alq_3, in which the presence of holes gives rise to a reduction of the emission by the formation of fluorescence quenchers [200]. Degradation due to the presence of singlet excitons has been reported for 4,4′-bis(N-carbazolyl)biphenyl (CBP), used as a matrix material in the EML of green-emitting phosphorescent OLEDs [201]. The degradation products can act as nonradiative recombination centers and trap states. Degradation due to singlet excitons has also been reported for OLEDs based on the fluorescent emitter material spiro-DPVBi [202]. An example of a bimolecular process which can be accompanied by degradation is triplet–polaron quenching (TPQ). When TPQ is due to excitation of the polaron by the triplet exciton, followed by nonradiative decay of the excited polaron due to internal conversion, the locally dissipated energy may, with a certain probability, lead to a chemical change of the molecule. Such a mechanism was found for a blue phosphorescent OLED using 4,4′-bis(3-methylcarbazol-9-yl)-2,2′-biphenyl (mCBP) as the host material in the EML [203]. The mCBP defect sites were argued to act as deep charge traps and the dissociation products were argued to damage the guest (emitter) so that it becomes a nonradiative center and a luminescence quencher.

A KMC degradation study starts in general with running a simulation without degradation until dynamic equilibrium is achieved under the operational conditions (voltage, temperature) of interest. Subsequently, the simulations are continued while degradation is switched on. Monomolecular degradation due to a charge or an exciton on a sensitive site is included as a new possible process which competes with all other possible processes. Degradation that accompanies a bimolecular process such as TPQ is included by branching the end result of that process, so that a defect site is formed with a certain probability. We note that in KMC simulations only the primary event needs to be described. All subsequent effects (e.g., a shift of the recombination zone to a less favorable position due to a changed mobility balance, or an efficiency loss due to TPQ at charges residing on defect molecules which act as traps [204]) will follow "automatically" from the simulations. KMC lifetime simulations are necessarily strongly accelerated, as practically realistic simulated times are usually at most of the order of 1 ms, 6–11 orders of magnitude smaller than actual lifetimes in the range of 1–100,000 hours. This can be accomplished by assuming an enhanced value of the degradation rate (for monomolecular degradation) or the degradation probability (for degradation accompanying a bimolecular process). As in the case of experimental OLED lifetime tests, it is also possible to accelerate the simulations by carrying them out for high current densities and/or elevated temperatures, followed by extrapolation to application-relevant operational conditions.

A first demonstration of the feasibility of 3D KMC OLED lifetime simulations was presented in Ref. [177]. The simulations were carried out for a symmetric OLED with an energy level structure as shown in Figure 15.12a, for a temperature of 300 K. The mixed-matrix emissive layer contains equal concentrations of the HTL and ETL material as a host, and 4 mol% of emitter molecules (guest). The hole and electron blocking is excellent, so that the recombination efficiency is 100%. The simulation parameters are as given in Table 15.2, with the following exceptions: $\varepsilon_r = 3$, $\sigma_T = 0$ eV (no triplet energy disorder), and $\Gamma_{rad} = 0.544\ \mu s^{-1}$ and $\Gamma_{nr} = 0.181\ \mu s^{-1}$ (values typical for the orange-red emitter $Ir(MDQ)_2(acac)$ in an α-NPD matrix [205]). As in Section 15.12.3, TPQ and TTA were treated in a parameter-free manner as instantaneous nearest-neighbor processes.

Van Eersel et al. [176] showed that for these symmetric devices and with a dye trap depth $\Delta = 0.2$ – 0.3 eV the emission profile at small voltages is quite uniform across the emissive layer. This is illustrated by Figure 15.12b, which shows the emission profile at 3 V for the case studied ($\Delta = 0.2$ eV). On the one hand,

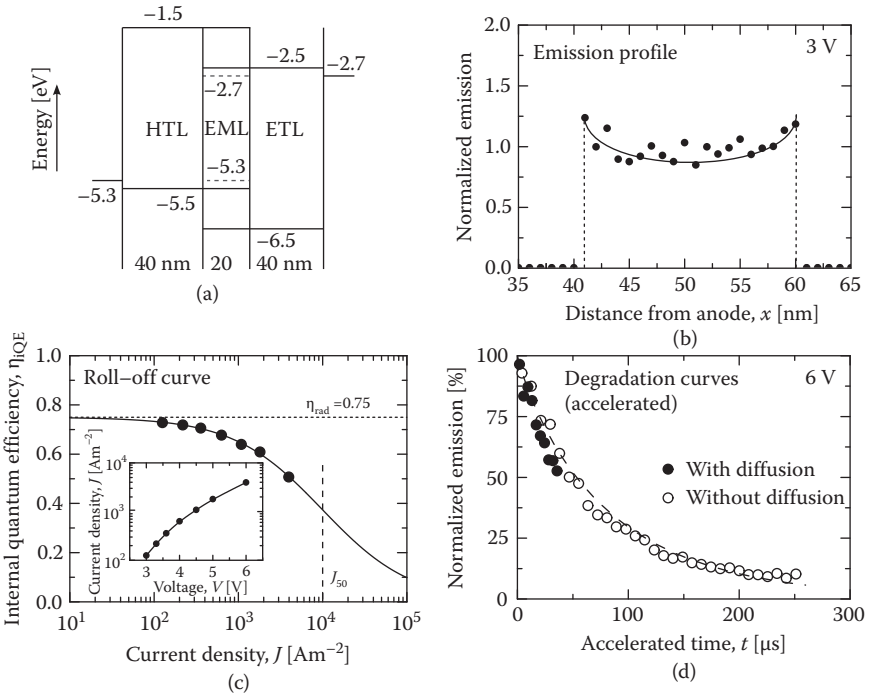

FIGURE 15.12 (a) Organic layer and energy level structure of the organic light-emitting diodes (OLEDs) considered in the kinetic Monte Carlo (KMC) degradation simulation studies. Dashed lines: Emitter energy levels (4 mol%). (b) Emission profile at 3 V. The full curve is a guide-to-the-eye. (c) Current density dependence of the internal quantum efficiency. The full curve is a fit using Equation 15.41 with $m = 0.80$ and $J_{50} = 10$ kA/m^2. The vertical dashed line indicates the current density at which the IQE has dropped to 50% (J_{50}) of the value at small current densities. (d) Time-dependence of the normalized emission at 6 V as obtained from a KMC degradation study which assumes that (1) the degradation occurs on triplet-excited dye sites upon triplet-polaron quenching processes (TPQ-t, see the main text) resulting from the displacement of a polaron on a neighbor site to the dye site, and which assumes that (2) each quenching is followed by a degradation process upon which the involved dye molecule becomes nonemissive. The simulations were carried out with and without exciton diffusion (closed and open symbols, respectively). The dashed curve is a stretched-exponential fit to the simulation results without diffusion (see the main text). (Reproduced from R. Coehoorn et al., *Adv. Funct. Mater.*, 25, 2024–2037, 2015. With permission.)

this choice avoids a large Ohmic loss due to deep trapping and a large overvoltage due to the enhanced built-in voltage (for large Δ). On the other hand, it also avoids a large loss due to strong roll-off caused by emission from thin zones near the blocking layer interfaces (for small Δ). The optimum value of Δ will depend on the application goal (see Ref. [176]) and is sensitive to the detailed mechanisms of the TPQ and TTA processes. Figure 15.12c shows the calculated IQE roll-off curve before degradation. The J–V curve is given in the inset. At small voltages, the IQE is equal to the assumed radiative decay (PL) efficiency ($\eta_{rad} = \Gamma_{rad}/(\Gamma_{rad} + \Gamma_{nr}) = 0.75$). The J_{90} current density, defined in Section 15.12.3, is approximately 640 A/m^2. This is larger than the largest experimental value obtained so far ($\sim$ 300 A/m^2, see Section 15.12.3). The full curve gives a fit through the data points, discussed below (see Equation 15.41). For the nearest-neighbor TPQ and TTA mechanisms considered, the roll-off is almost completely due to TPQ.

Figure 15.12d shows the dependence of the normalized emission on the simulated time, at constant voltage conditions (6 V), for a degradation scenario in which, upon a TPQ process, the polaron involved is displaced to the site at which the triplet exciton resides (a "TPQ-t process," see Figure 2 in Ref. [177]), which then becomes nonemissive with a degradation probability $p_{degr} = 1$. The simulations thus employ the

largest possible acceleration factor. All other parameters are kept identical. The current density was found to remain essentially unchanged during the degradation process. Furthermore, it was found that choosing a smaller degradation probability does not significantly change the results, apart from changing the time scale. That indicates that the lifetime is still much larger than all other relevant time scales. Simulations including exciton diffusion are computationally more expensive, and were stopped when a 50% emission reduction had been obtained. Simulations without diffusion were continued until a reduction to only about 10% of the initial emission was reached. An approximate description of the decay is given by the stretched exponential curve, shown in the figure, with the form $I(t) = I(0)\exp[-(t/\tau_{\mathrm{sim,acc}})^{\beta}]$, with a simulated accelerated (1/e) lifetime $\tau_{\mathrm{sim,acc}} \cong 80$ μs and a stretching exponent $\beta = 0.91$. The actual (1/e) lifetime as predicted from the simulations is given by

$$\tau_{\mathrm{sim}} = \frac{\tau_{\mathrm{sim,acc}}}{p_{\mathrm{degr}}}. \tag{15.40}$$

Conversely, using Equation 15.40 the value of p_{degr} could be deduced from a degradation simulation and a measurement of the lifetime. Such an analysis led in Ref. [177] to an estimated order-of-magnitude value of $p_{\mathrm{degr}} \sim 10^{-8}$ when assuming that the simulations discussed above are relevant to state-of-the-art white OLEDs.

Within a refined approach, the probability that a TPQ process gives rise to degradation could be treated stochastically, e.g., by treating it as a thermally activated process with an activation energy with a Gaussian distribution. Such an approach is expected to give rise to a smaller value of the stretching exponent β. Experimentally, values of β around 0.5 have been observed. The effect is so far generally explained in a rather phenomenological manner [206]. It should be noted that the decay can also become more stretched-exponential like for OLEDs with imbalanced electron and hole mobilities, resulting in a highly nonuniform emission profile. If this picture is correct, lifetime studies could also, albeit indirectly, provide information about the shape of the emission profile.

We envisage that, using Equation 15.40, lifetime predictions can be obtained from KMC lifetime simulations if the parameter p_{degr} (or its distribution) can be determined from a few well-chosen combined experimental and KMC calibration studies. Subsequently, KMC-based lifetime predictions can be obtained for other measurement conditions (e.g., current density and temperature), device architectures, and other dye concentrations in the same host. So far, studies which could validate this view have not been carried out. As a first step, KMC simulations were carried out of the iridium dye concentration dependence of the lifetime, for otherwise identical simulation parameters [177]. Figure 15.13a shows the accelerated LT_{90} lifetime obtained from the simulations (large open circles), and results from simulations in which exciton diffusion was switched off (small open circles). For iridium dye concentrations above about 7 mol%, exciton diffusion is found to yield a significant decrease of the lifetime, up to a factor ∼4. The exciton diffusion length is then larger than the average distance to a degraded site, so that a large fraction of the excitons which have been generated on nondegraded sites is lost due to diffusion and subsequent nonradiative decay on degraded sites. Interestingly, the IQE obtained when exciton diffusion was switched off was found to be slightly reduced for all systems studied, namely by 1%–2%. The reduction cannot be due to switching off the transfer of triplets from matrix to guest sites, as due to the large energy gap of the matrix materials all excitons are generated directly on the dye sites. We surmise that the reduction is due to switching off the possibility that excitons diffuse to molecular sites in regions with a slightly smaller average polaron density, in which the loss due to TPQ is reduced. This effect may for the highest iridium dye concentrations considered be compensated in part by an increase of the IQE due to a smaller exciton diffusion contribution to the TPQ loss. For the devices and the parameter values employed, the latter effect was found to be almost negligible for iridium dye concentrations around and below 8 mol% [177].

The simulation data shown in Figure 15.13a may be analyzed more quantitatively with the help of a useful relationship which has been established between the lifetime and the IQE roll-off [177]. It is based on a model which assumes uniform electron, hole, and exciton densities in the emissive layer. If

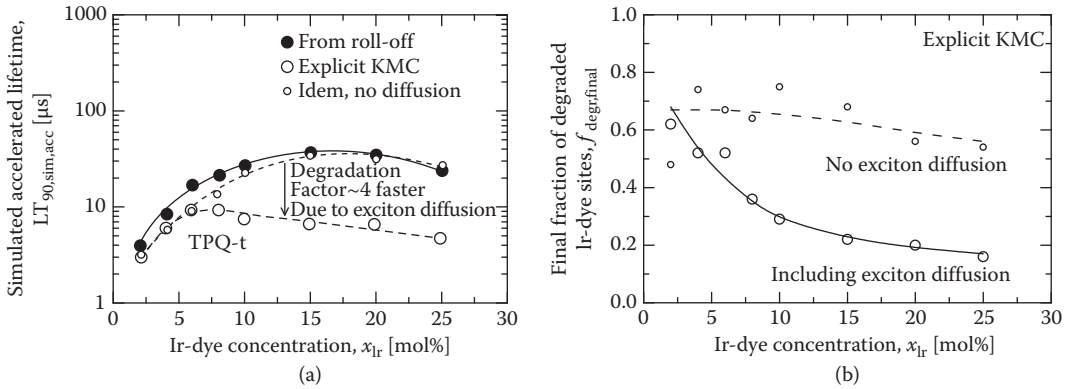

FIGURE 15.13 (a) Ir-dye concentration dependence of the LT_{90} lifetime as obtained from accelerated KMC lifetime simulations for the OLEDs shown in Figure 15.12a, at 6 V, for the TPQ-t degradation scenario discussed in the main text (large open circles) and when switching off exciton diffusion (small open circles). The closed circles show the accelerated lifetime which is expected from the simulated roll-off using Equation 15.42. The curves are guides-to-the-eye. The predicted actual LT_{90} lifetime is equal to $LT_{90,sim,acc}/p_{degr}$, with p_{degr} the probability that upon a TPQ event degradation takes place. Note that due to a variation of the current density with the Ir-dye concentration (see the main text) the lifetime as obtained using Equation 15.42 is not proportional to that concentration. (b) Ir-dye concentration dependence of the fraction of degraded Ir-dye molecules in the $t \to \infty$ limit, as obtained from the same KMC simulations including and without exciton diffusion. (Reproduced from R. Coehoorn et al., *Adv. Funct. Mater.*, 25, 2024–2037, 2015. With permission.)

the charge carrier mobility is charge carrier concentration (c) dependent and proportional to c^b, with b a positive exponent which increases with an increasing width of the polaron DOS, the IQE roll-off curve is given by

$$\eta_{IQE} = \frac{\eta_{rad}}{1 + \left(\frac{J}{J_{50}}\right)^m}, \tag{15.41}$$

with η_{rad} the PL efficiency (0.75 for the case studied in this section), J_{50} the current density at which the IQE has dropped to 50% of its low-voltage value, and $m = (1 + b)/(2 + b)$. In the absence of exciton diffusion, the simulated accelerated (1/e) lifetime is then given by

$$\tau_{sim,acc} = edn_{dye}J_{50}^m/J^{m+1}, \tag{15.42}$$

where d is the EML layer thickness and n_{dye} is the dye molecule volume density. From this formalism, a current density acceleration exponent $m + 1 = (3 + 2b)/(2 + b)$ in the range 1.5–2 is expected, as is indeed often observed for phosphorescent OLEDs. The linear n_{dye} dependence is due to the linear increase with increasing dye concentration of the probability that a dye is still emissive after a certain period of operation. The closed circles in Figure 15.13 show the simulated accelerated LT_{90} lifetime, obtained using Equation 15.42 under the assumption of exponential decay ($\beta = 1$), so that $LT_{90,sim,acc} = -\ln(0.9) \times \tau_{sim,acc} \cong 0.105 \times \tau_{sim,acc}$. These predictions from the roll-off curves agree quite well with the explicit KMC simulation results, obtained when exciton diffusion is switched off (small open circles). We note that at the constant voltage (6-V) condition employed, the current density shows a weak but nonnegligible nonmonotonic dependence on the Ir-dye concentration. It shows a broad minimum around a concentration of approximately 10 mol%, and is ~15% larger for the 2 and 25 mol% systems. Within the concentration range studied, the transport shows a cross-over from a low-concentration guest–host–guest hopping regime, in which the guest molecule states act as traps, to a high-concentration regime, in which

the transport is predominantly due to direct guest–guest hopping (see, e.g., Figure 15.5a in Ref. [100]). The lifetime at a fixed voltage, as obtained from Equation 15.42, is therefore not proportional to the Ir-dye concentration.

Various optical and chemical analytical techniques have been used to investigate the degradation mechanisms of OLEDs and to quantify the concentrations of degraded molecules (see Ref. [199] and references therein). In future studies, it would be useful to compare such experimental results with the results of KMC simulations. Interestingly, the simulations for the model systems discussed in this section revealed that, to a good approximation, the fraction of degraded Ir-dye molecules increases with time as

$$f_{\mathrm{degr}}(t) = f_{\mathrm{degr,final}}\left[1 - \exp\left(-\left(t/\tau_{\mathrm{sim,acc}}\right)^{\beta}\right)\right], \tag{15.43}$$

with $f_{\mathrm{degr,final}}$ the final ($t \to \infty$) fraction of degraded Ir-dye molecules.

The values of $\tau_{\mathrm{sim,acc}}$ and β are, within the numerical uncertainty, equal to the values describing the luminance decay. Figure 15.13b shows the Ir-dye concentration dependence of $f_{\mathrm{degr,final}}$ as obtained from KMC simulations with and without exciton diffusion. The figure shows that in the absence of diffusion, only 60%–70% of the sites have degraded in the $t = \infty$ limit. This result indicates that, on a significant fraction (30%–40%) of the sites, excitons are either never formed, or that excitons on those sites are well protected against TPQ due to a position of those sites well outside the somewhat filamentary electron and hole current density pathways. The first explanation is consistent with the finding that in systems with a monomodal Gaussian DOS exciton generation preferentially takes place on sites with a low-lying electron or hole state [73]. When exciton diffusion is included, $f_{\mathrm{degr,final}}$ is found to decrease significantly with increasing Ir-dye concentration, to only approximately 0.16 for 25 mol% systems. This is consistent with the view that, due to the energetic disorder, the average polaron density and the local polaron diffusivity will be quite nonuniform [99,207], so that in the case of strong exciton diffusion, degradation will occur predominantly on the relatively small fraction of sites that are located in a region with a large average polaron density and diffusivity.

We emphasize that the degradation scenario assumed in the case study discussed in this section was only chosen for the purpose of giving a demonstration of the feasibility of KMC lifetime simulations. Including monomolecular or other bimolecular scenarios or including refinements (e.g., a degradation probability distribution, conversion of the degraded molecules to polaron trap sites, or an extension of the TPQ interaction range so that the role of the degraded molecule as an exciton quencher is enhanced) is straightforward. It also will be useful to develop analytical models as discussed above for other degradation scenarios and to extend these to a more realistic nonuniform emission from the EML, in which the lifetime becomes position-dependent, so that the emission decay becomes more stretched-exponential like.

15.13 Outlook

In this chapter, we have reviewed multiscale techniques used to simulate organic LEDs and demonstrated the feasibility of full 3D OLED modeling. As an outlook, we would like to mention areas where substantial method development is still required in order to achieve a parameter-free modeling of realistic devices. Refined studies that aim at developing a final view on the detailed performance of specific devices should consider the following: (1) first-principles evaluations of charge injection rates, (2) explicit treatment of the induction interaction when solving the master equation, (3) quantitative treatment of excited states embedded in a heterogeneous polarizable molecular environment, (4) more quantitative descriptions of charge–exciton and exciton–exciton interactions, and (5) descriptions of TTA and TPQ as longer-range Förster and Dexter-type interactions. Advancements in all these directions are absolutely vital for devising accurate structure–property relationships for organic semiconductors used in OLEDs.

Acknowledgments

This work was supported in part by the Federal Ministry of Education and Research BMBF grants MEDOS (FKZ 03EK3503B), MESOMERIE (FKZ 13N10723), and InterPhase (FKZ 13N13661). The project received funding from the NMP-20-2014—"Widening materials models" program under Grant Agreement No. 646259 (MOSTOPHOS). The project was also supported by NanoNextNL, a nanotechnology program of the Dutch Ministry of Economic Affairs. Part of the work of one of the authors (RC) was carried out at the Philips Research Laboratories (Eindhoven, The Netherlands).

References

1. I. Akasaki. Nobel lecture: Fascinated journeys into blue light. *Rev. Mod. Phys.*, 87(4):1119–1131, 2015.
2. H. Amano. Nobel lecture: Growth of GaN on sapphire via low-temperature de-posited buffer layer and realization of p-type GaN by Mg doping followed by low-energy electron beam irradiation. *Rev. Mod. Phys.*, 87(4):1133–1138, 2015.
3. S. Nakamura. Nobel lecture: Background story of the invention of efficient blue InGaN light emitting diodes. *Rev. Mod. Phys.*, 87(4):1139–1151, 2015.
4. S. R. Forrest. The path to ubiquitous and low-cost organic electronic appliances on plastic. *Nature*, 428(6986):911–918, 2004.
5. C. W. Tang and S. A. VanSlyke. Organic electroluminescent diodes. *Appl. Phys. Lett.*, 51(12):913–915, 1987.
6. R. Friend, J. Burroughes, and D. Bradley. Electroluminescent devices, Patent WO9013148 (A1). Cambridge Research and Innovation Limited; Cambridge Capital Management Limited; Lynxvale Limited, 1990.
7. R. Coehoorn, V. van Elsbergen, and C. Verschuren. High efficiency OLEDs for lighting applications. In E. Cantatore, editor, *Applications of Organic and Printed Electronics, Integrated Circuits and Systems*, pages 83–100. New York, NY: Springer, 2013.
8. T. Otani. Samsung shows new 4 flexible AMOLED that is so thin (0.05 mm) it "aps" in the wind textbar OLED-Info. 2008. https://www.oled-info.com/samsung-shows-new-4-flexible-amoled-so-thin-005mm-it-flaps-wind
9. R. Mertens. Continental shows a prototype dual-screen flexible AMOLED display for the automotive market—OLED-Info, 2014. https://www.oled-info.com/continental-shows-prototype-dual-screen-flexible-amoled-display-automotive-market
10. R. Olivares-Amaya, C. Amador-Bedolla, J. Hachmann, S. Atahan-Evrenk, R. S. Sanchez-Carrera, L. Vogt, and A. Aspuru-Guzik. Accelerated computational discovery of high-performance materials for organic photovoltaics by means of cheminformatics. *Energy Environ. Sci.*, 4(12):4849–4861, 2011.
11. P. Deglmann, A. Schaefer, and C. Lennartz. Application of quantum calculations in the chemical industry—An overview. *Int. J. Quantum Chem.*, 115(3):107–136, 2015. WOS:000346654700001.
12. H. Yersin. Triplet emitters for OLED applications. Mechanisms of exciton trapping and control of emission properties. In *Transition Metal and Rare Earth Compounds*, Yersin, Hartmut (Ed.) No. 241 in Topics in Current Chemistry, Hartmut Yersin (Ed.), pp. 1–26. Berlin: Springer, 2004.
13. L. Duan, K. Xie, and Y. Qiu. Review paper: Progress on efficient cathodes for organic light-emitting diodes. *J. Soc. Inf. Disp.*, 19(6):453–461, 2011.
14. A. Fukase, K. Luan Thanh Dao, and J. Kido. High-efficiency organic electroluminescent devices using iridium complex emitter and arylamine-containing polymer buffer layer. *Polymer. Adv. Tech.*, 13(8):601–604, 2002.
15. X. Zhou, D. S. Qin, M. Pfeiffer, J. Blochwitz-Nimoth, A. Werner, J. Drechsel, B. Maen-nig, K. Leo, M. Bold, P. Erk, and H. Hartmann. High-efficiency electrophosphorescent organic light-emitting diodes with double light-emitting layers. *Appl. Phys. Lett.*, 81(21):4070–4072, 2002.

16. C. Adachi, M. A. Baldo, S. R. Forrest, S. Lamansky, M. E. Thompson, and R. C. Kwong. High-efficiency red electrophosphorescence devices. *Appl. Phys. Lett.*, 78(11):1622–1624, 2001.
17. Y. Kawamura, S. Yanagida, and S. R. Forrest. Energy transfer in polymer electrophosphorescent light emitting devices with single and multiple doped luminescent layers. *J. Appl. Phys.*, 92(1):87–93, 2002.
18. B. A. Gregg, S.-G. Chen, and R. A. Cormier. Coulomb forces and doping in organic semiconductors. *Chem. Mater.*, 16(23):4586–4599, 2004.
19. K. Walzer, B. Maennig, M. Pfeiffer, and K. Leo. Highly efficient organic devices based on electrically doped transport layers. *Chem. Rev.*, 107(4):1233–1271, 2007.
20. M. A. Baldo, S. Lamansky, P. E. Burrows, M. E. Thompson, and S. R. Forrest. Very high-efficiency green organic light-emitting devices based on electrophosphorescence. *Appl. Phys. Lett.*, 75(1):4–6, 1999.
21. Md. K. Nazeeruddin, R. Humphry-Baker, D. Berner, S. Rivier, L. Zuppiroli, and M. Graetzel. Highly phosphorescence iridium complexes and their application in organic light-emitting devices. *J. Am. Chem. Soc.*, 125(29):8790–8797, 2003.
22. Q. Peng, W. Li, S. Zhang, P. Chen, F. Li, and Y. Ma. Evidence of the reverse intersystem crossing in intra-molecular charge transfer fluorescence based organic light-emitting devices through magneto-electroluminescence measurements. *Adv. Opt. Mater.*, 1(5):362–366, 2013.
23. S. Hirata, Y. Sakai, K. Masui, H. Tanaka, S. Youn Lee, H. Nomura, N. Nakamura, M. Yasumatsu, H. Nakanotani, Q. Zhang, K. Shizu, H. Miyazaki, and C. Adachi. Highly efficient blue electroluminescence based on thermally activated delayed fluorescence. *Nat. Mater.*, 14(3):330–336, 2015.
24. S. Reineke, K. Walzer, and K. Leo. Triplet–exciton quenching in organic phosphorescent light-emitting diodes with Ir-based emitters. *Phys. Rev. B*, 75(12):125328, 2007.
25. E. Knapp, R. Husermann, H. U. Schwarzenbach, and B. Ruhstaller. Numerical simulation of charge transport in disordered organic semiconductor devices. *J. Appl. Phys.*, 108(5):054504, 2010.
26. B. Ruhstaller, E. Knapp, B. Perucco, N. Reinke, D. Rezzonico, and F. Mueller. Advanced numerical simulation of organic light-emitting devices. In O. Sergiyenko, editor, *Optoelectronic Devices and Properties*. InTech, 2011. Rijeka, Croatia: DOI: 10.5772/14626.
27. G. Kaniadakis and P. Quarati. Kinetic equation for classical particles obeying an exclusion principle. *Phys. Rev. E*, 48(6):4263–4270, 1993.
28. Y. Roichman and N. Tessler. Generalized Einstein relation for disordered semiconductors-implications for device performance. *Appl. Phys. Lett.*, 80(11):1948–1950, 2002.
29. G. Lakhwani, A. Rao, and R. H. Friend. Bimolecular recombination in organic photovoltaics. *Annu. Rev. Phys. Chem.*, 65(1):557–581, 2014.
30. B. Perucco, N. A. Reinke, D. Rezzonico, E. Knapp, S. Harkema, and B. Ruhstaller. On the exciton profile in OLEDs-seamless optical and electrical modeling. *Org. Electron.*, 13(10):1827–1835, 2012.
31. G.-J. A. H. Wetzelaer and P. W. M. Blom. Diffusion-driven currents in organic-semiconductor diodes. *NPG Asia Mater.*, 6(7):e110, 2014.
32. J.-H. Lee, S. Lee, S.-J. Yoo, K.-H. Kim, and J.-J. Kim. Langevin and trap-assisted recombination in phosphorescent organic light emitting diodes. *Adv. Funct. Mater.*, 24(29):4681–4688, 2014.
33. R. R. Chance, A. Prock, and R. Silbey. Molecular fluorescence and energy transfer near interfaces. In I. Prigogine and S. A. Rice, editors, *Advances in Chemical Physics*, pp. 1–65. Hoboken, NJ: John Wiley & Sons, 1978.
34. H. K. Gummel. A self-consistent iterative scheme for one-dimensional steady state transistor calculations. *IEEE Trans. Electron Dev.*, 11(10):455–465, 1964.
35. D. L. Scharfetter and H. K. Gummel. Large-signal analysis of a silicon Read diode oscillator. *IEEE Trans. Electron Dev.*, 16(1):64–77, 1969.
36. E. Knapp and B. Ruhstaller. Numerical impedance analysis for organic semiconductors with exponential distribution of localized states. *Appl. Phys. Lett.*, 99(9), 093304, 2011.

37. E. Tuti, I. Batisti, and D. Berner. Injection and strong current channeling in organic disordered media. *Phys. Rev. B*, 70(16):161202, 2004.
38. K. D. Meisel, W. F. Pasveer, J. Cottaar, C. Tanase, R. Coehoorn, P. A. Bobbert, P. W. M. Blom, D. M. de Leeuw, and M. A. J. Michels. Charge-carrier mobilities in disordered semiconducting polymers: Effects of carrier density and electric field. *Phys. Status Solidi C*, 3(2):267–270, 2006.
39. N. Rappaport, Y. Preezant, and N. Tessler. Spatially dispersive transport: A mesoscopic phenomenon in disordered organic semiconductors. *Phys. Rev. B*, 76(23):235323, 2007.
40. J. Honerkamp. *Stochastische Dynamische Systeme : Konzepte, Numerische Methoden, Datenanalysen*. Weinheim: Wiley-VCH, 1990.
41. B. Derrida. Velocity and diffusion constant of a periodic one-dimensional hopping model. *J. Stat. Phys.*, 31(3):433–450, 1983.
42. K. Seki and M. Tachiya. Electric field dependence of charge mobility in energetically disordered materials: Polaron aspects. *Phys. Rev. B*, 65:014305, 2001.
43. S. D. Baranovskii. Theoretical description of charge transport in disordered organic semiconductors. *Phys. Status Solidi B*, 251(3):487–525, 2014.
44. J. Cottaar and P. A. Bobbert. Calculating charge-carrier mobilities in disordered semiconducting polymers: Mean field and beyond. *Phys. Rev. B*, 74(11):115204, 2006.
45. V. Rühle, A. Lukyanov, F. May, M. Schrader, T. Vehoff, J. Kirkpatrick, B. Baumeier, and D. Andrienko. Microscopic simulations of charge transport in disordered organic semiconductors. *J. Chem. Theory Comput.*, 7(10):3335–3345, 2011.
46. M. Heidernaetsch, M. Bauer, and G. Radons. Characterizing N-dimensional anisotropic Brownian motion by the distribution of diffusivities. *J. Chem. Phys.*, 139(18):184105, 2013.
47. J. L. Doob. Topics in the theory of Markoff chains. *Trans. Am. Math. Soc.*, 52(1):37–64, 1942.
48. J. L. Doob. Markoff chains—Denumerable case. *Trans. Am. Math. Soc.*, 58(3):455, 1945.
49. W M. Young and E. W. Elcock. Monte Carlo studies of vacancy migration in binary ordered alloys: I. *Proc. Phys. Soc.*, 89(3):735–746, 1966.
50. A. B. Bortz, M. H. Kalos, and J. L. Lebowitz. A new algorithm for Monte Carlo simulation of Ising spin systems. *J. Comput. Phys.*, 17(1):10–18, 1975.
51. D. T. Gillespie. A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *J. Comput. Phys.*, 22(4):403–434, 1976.
52. D. T. Gillespie. Exact stochastic simulation of coupled chemical reactions. *J. Phys. Chem.*, 81(25):2340–2361, 1977.
53. K. A. Fichthorn and W. H. Weinberg. Theoretical foundations of dynamical Monte Carlo simulations. *J. Chem. Phys.*, 95(2):1090, 1991.
54. A. P. J. Jansen. Monte Carlo simulations of chemical reactions on a surface with time-dependent reaction-rate constants. *Comput. Phys. Commun.*, 86(1–2):1–12, 1995.
55. P. Kordt, J. J. M. van der Holst, M. Al Helwi, W. Kowalsky, F. May, A. Badinski, C. Lennartz, and D. Andrienko. Modeling of organic light emitting diodes: From molecular to device properties. *Adv. Funct. Mater.*, 25(13):1955–1971, 2015.
56. M. A. Gibson and J. Bruck. Efficient exact stochastic simulation of chemical systems with many species and many channels. *J. Phys. Chem. A.*, 104(9):1876–1889, 2000.
57. B. D. Lubachevsky. Efficient parallel simulations of dynamic Ising spin systems. *J. Comput. Phys.*, 75(1):103–122, 1988.
58. Y. Shim and J. G. Amar. Semirigorous synchronous sublattice algorithm for parallel kinetic Monte Carlo simulations of thin film growth. *Phys. Rev. B*, 71(12), 125432, 2005.
59. M. Merrick and K. A. Fichthorn. Synchronous relaxation algorithm for parallel kinetic Monte Carlo simulations of thin film growth. *Phys. Rev. E*, 75(1), 011606, 2007.
60. E. Martnez, J. Marian, M. H. Kalos, and J. M. Perlado. Synchronous parallel kinetic Monte Carlo for continuum diffusion-reaction systems. *J. Comput. Phys.*, 227(8):3804–3823, 2008.

61. N. J. van der Kaap and L. J. A. Koster. Massively parallel kinetic Monte Carlo simulations of charge carrier transport in organic semiconductors. *J. Comput. Phys.*, 307:321–332, 2016.
62. A. Miller and E. Abrahams. Impurity conduction at low concentrations. *Phys. Rev.*, 120(3):745–755, 1960.
63. H. Baessler. Charge transport in disordered organic photoconductors: A Monte Carlo simulation study. *Phys. Status Solidi B*, 175(1):15–56, 1993.
64. R. A. Marcus. Electron transfer reactions in chemistry. Theory and experiment. *Rev. Mod. Phys.*, 65(3):599–610, 1993.
65. G. R. Hutchison, M. A. Ratner, and T. J. Marks. Hopping transport in conductive heterocyclic oligomers: Reorganization energies and substituent effects. *J. Am. Chem. Soc.*, 127(7):2339–2350, 2005.
66. V. May and O. Kuehn. *Charge and Energy Transfer Dynamics in Molecular Systems*, 3rd ed. Weinheim: Wiley-VCH, 2011.
67. M. Bixon and J. Jortner. Electron transfer from isolated molecules to biomolecules. In I. Prigogine and S. A. Rice, editors, *Advances in Chemical Physics*, pp. 35–202. Hoboken, NJ: John Wiley & Sons, 2007.
68. R. Egger, C. H. Mak, and U. Weiss. Quantum rates for nonadiabatic electron transfer. *J. Chem. Phys.*, 100(4):2651–2660, 1994.
69. M. P. A. Fisher and A. T. Dorsey. Dissipative quantum tunneling in a biased double-well system at finite temperatures. *Phys. Rev. Lett.*, 54(15):1609–1612, 1985.
70. H. Grabert and U. Weiss. Quantum tunneling rates for asymmetric double-well systems with Ohmic dissipation. *Phys. Rev. Lett.*, 54(15):1605–1608, 1985.
71. A. J. Leggett, S. Chakravarty, A. T. Dorsey, M. P. A. Fisher, A. Garg, and W. Zwerger. Dynamics of the dissipative two-state system. *Rev. Mod. Phys.*, 59(1):1–85, 1987.
72. K. Asadi, A. J. Kronemeijer, T. Cramer, L. J. A. Koster, P. W. M. Blom, and D. M. de Leeuw. Polaron hopping mediated by nuclear tunnelling in semiconducting polymers at high carrier density. *Nat. Commun.*, 4:1710, 2013.
73. J. J. M. van der Holst, F. W. A. van Oost, R. Coehoorn, and P. A. Bobbert. Electron–hole recombination in disordered organic semiconductors: Validity of the Langevin formula. *Phys. Rev. B*, 80(23):235202-1–235202-8, 2009.
74. T. H. Foerster. Zwischenmolekulare Energiewanderung und Fluoreszenz. *Ann. Phys.*, 437(1–2):55–75, 1948.
75. D. L. Dexter. A theory of sensitized luminescence in solids. *J. Chem. Phys.*, 21(5):836–850, 1953.
76. G. D. Scholes and K. P. Ghiggino. Rate expressions for excitation transfer I. Radiationless transition theory perspective. *J. Chem. Phys.*, 101(2):1251–1261, 1994.
77. G. D. Scholes, R. D. Harcourt, and K. P. Ghiggino. Rate expressions for excitation transfer. III. An ab initio study of electronic factors in excitation transfer and exciton resonance interactions. *J. Chem. Phys.*, 102(24):9574–9581, 1995.
78. R. D. Harcourt, G. D. Scholes, and K. P. Ghiggino. Rate expressions for excitation transfer. II. Electronic considerations of direct and through configuration exciton resonance interactions. *J. Chem. Phys.*, 101(12):10521–10525, 1994.
79. S. Speiser. Photophysics and mechanisms of intramolecular electronic energy transfer in bichromophoric molecular systems: Solution and supersonic jet studies. *Chem. Rev.*, 96(6):1953–1976, 1996.
80. R. F. Fink, J. Pfister, H. Mei Zhao, and B. Engels. Assessment of quantum chemical methods and basis sets for excitation energy transfer. *Chem. Phys.*, 346(1–3):275–285, 2008. WOS:000256142000032.
81. C. Tanase, P. W. M. Blom, D. M. de Leeuw, and E. J. Meijer. Charge carrier density dependence of the hole mobility in poly(p-phenylene vinylene). *Phys. Status Solidi A*, 201(6):1236–1245, 2004.

82. F. Laquai and D. Hertel. Influence of hole transport units on the efficiency of polymer light emitting diodes. *Appl. Phys. Lett.*, 90(14):142109, 2007.
83. S. V. Novikov, D. H. Dunlap, V. M. Kenkre, P. E. Parris, and A. V. Vannikov. Essential role of correlations in governing charge transport in disordered organic materials. *Phys. Rev. Lett.*, 81(20):4472–4475, 1998.
84. D. H. Dunlap, P. E. Parris, and V. M. Kenkre. Charge-dipole model for the universal field dependence of mobilities in molecularly doped polymers. *Phys. Rev. Lett.*, 77(3):542–545, 1996.
85. S. W. de Leeuw, J. W. Perram, and E. R. Smith. Simulation of electrostatic systems in periodic boundary conditions. I. Lattice sums and dielectric constants. *Proc. Royal Soc. A: Math. Phys. Eng. Sci.*, 373(1752):27–56, 1980.
86. P. P. Ewald. Die Berechnung optischer und elektrostatischer Gitterpotentiale. *Ann. Phys.*, 369(3):253–287, 1921.
87. R. H. Young. Dipolar lattice model of disorder in random media analytical evaluation of the Gaussian disorder model. *Philos. Mag. B*, 72(4):435–457, 1995.
88. S. V. Novikov and A. V. Vannikov. Cluster structure in the distribution of the electrostatic potential in a lattice of randomly oriented dipoles. *J. Phys. Chem.*, 99(40):14573–14576, 1995.
89. P. Kordt and D. Andrienko. Modeling of spatially correlated energetic disorder in organic semiconductors. *J. Chem. Theory Comput.*, 12(1):36–40, 2016.
90. W. F. Pasveer, J. Cottaar, C. Tanase, R. Coehoorn, P. A. Bobbert, P. W. M. Blom, D. M. de Leeuw, and M. A. J. Michels. Unified description of charge-carrier mobilities in disordered semiconducting polymers. *Phys. Rev. Lett.*, 94(20), 2005.
91. M. Bouhassoune, S. L. M. van Mensfoort, P. A. Bobbert, and R. Coehoorn. Carrier-density and field-dependent charge-carrier mobility in organic semiconductors with correlated Gaussian disorder. *Org. Electron.*, 10(3):437–445, 2009.
92. A. J. Stone. *The Theory of Intermolecular Forces*. Oxford: Clarendon Press, 1997.
93. C. Poelking and D. Andrienko. Long-range embedding of molecular ions and excitations in a polarizable molecular environment. *J. Chem. Theory Comput.*, 12(9):4516–4523, 2016.
94. C. Poelking, M. Tietze, C. Elschner, S. Olthof, D. Hertel, B. Baumeier, F. Wrthner, K. Meerholz, K. Leo, and D. Andrienko. Impact of mesoscale order on open-circuit voltage in organic solar cells. *Nat. Mater.*, 14(4):434–439, 2014.
95. C. Poelking and D. Andrienko. Design rules for organic donor-acceptor heterojunctions: Pathway for charge splitting and detrapping. *J. Am. Chem. Soc.*, 137(19):6320–6326, 2015.
96. D. Andrienko. Simulations of Morphology and Charge Transport in Supramolecular Organic Materials. In *Supramolecular Materials for Opto-Electronics*, Norbert Koch (Ed.). Cambridge, UK: The Royal Society of Chemistry, 2014.
97. B. T. Thole. Molecular polarizabilities calculated with a modified dipole interaction. *Chem. Phys.*, 59(3):341–350, 1981.
98. S. L. M. van Mensfoort and R. Coehoorn. Effect of Gaussian disorder on the voltage dependence of the current density in sandwich-type devices based on organic semiconductors. *Phys. Rev. B*, 78(8):085207, 2008.
99. J. J. M. van der Holst, M. A. Uijttewaal, B. Ramachandhran, R. Coehoorn, P. A. Bobbert, G. A. de Wijs, and R. A. de Groot. Modeling and analysis of the three-dimensional current density in sandwich-type single-carrier devices of disordered organic semiconductors. *Phys. Rev. B*, 79(8):085203, 2009.
100. R. Coehoorn and P. A. Bobbert. Effects of Gaussian disorder on charge carrier transport and recombination in organic semiconductors. *Phys. Status Solidi A*, 209(12):2354–2377, 2012.
101. J.-L. Bredas, D. Beljonne, V. Coropceanu, and J. Cornil. Charge-transfer and energy-transfer processes in pi-conjugated oligomers and polymers: A molecular picture. *Chem. Rev.*, 104(11):4971–5004, 2004.
102. J. R. Reimers. A practical method for the use of curvilinear coordinates in calculations of normal-mode-projected displacements and Duschinsky rotation matrices for large molecules. *J. Chem. Phys.*, 115(20):9103–9109, 2001.

103. D. P. McMahon and A. Troisi. Evaluation of the external reorganization energy of polyacenes. *J. Phys. Chem. Lett.*, 1(6):941–946, 2010.
104. J. E. Norton and J.-L. Bredas. Polarization energies in oligoacene semiconductor crystals. *J. Am. Chem. Soc.*, 130(37):12377–12384, 2008.
105. B. Baumeier, J. Kirkpatrick, and D. Andrienko. Density-functional based determination of intermolecular charge transfer properties for large-scale morphologies. *Phys. Chem. Chem. Phys.*, 12(36):11103, 2010.
106. J. Huang and M. Kertesz. Validation of intermolecular transfer integral and bandwidth calculations for organic molecular materials. *J. Chem. Phys.*, 122(23):234707, 2005.
107. E. F. Valeev, V. Coropceanu, D. A. da Silva Filho, S. Salman, and J.-L. Bredas. Effect of electronic polarization on charge-transport parameters in molecular organic semiconductors. *J. Am. Chem. Soc.*, 128(30):9882–9886, 2006.
108. T. Van Voorhis, T. Kowalczyk, B. Kaduk, L.-P. Wang, C.-L. Cheng, and Q. Wu. The diabatic picture of electron transfer, reaction barriers, and molecular dynamics. *Annu. Rev. Phys. Chem.*, 61(1):149–170, 2010.
109. J. Ridley and M. Zerner. An intermediate neglect of differential overlap technique for spectroscopy: Pyrrole and the azines. *Theor. Chim. Acta*, 32(2):111–134, 1973.
110. J. Kirkpatrick. An approximate method for calculating transfer integrals based on the ZINDO Hamiltonian. *Int. J. Quantum Chem.*, 108(1):51–56, 2008.
111. V. Coropceanu, J. Cornil, D. A. da Silva Filho, Y. Olivier, R. Silbey, and J.-L. Bredas. Charge transport in organic semiconductors. *Chem. Rev.*, 107(4):926–952, 2007.
112. A. Kubas, F. Hoffmann, A. Heck, H. Oberhofer, M. Elstner, and J. Blumberger. Electronic couplings for molecular charge transfer: Bench-marking CDFT, FODFT, and FODFTB against high-level ab initio calculations. *J. Chem. Phys.*, 140(10):104105, 2014.
113. F. Gajdos, S. Valner, F. Hoffmann, J. Spencer, M. Breuer, A. Kubas, M. Dupuis, and J. Blumberger. Ultrafast estimation of electronic couplings for electron transfer between pi-conjugated organic molecules. *J. Chem. Theory Comput.*, 10(10):4653–4660, 2014.
114. P. Kordt, O. Stenzel, B. Baumeier, V. Schmidt, and D. Andrienko. Parametrization of extended Gaussian disorder models from microscopic charge transport simulations. *J. Chem. Theory Comput.*, 10(6):2508–2513, 2014.
115. S. Athanasopoulos, J. Kirkpatrick, D. Martnez, J. M. Frost, C. M. Foden, A. B. Walker, and J. Nelson. Predictive study of charge transport in disordered semiconducting polymers. *Nano Lett.*, 7(6):1785–1788, 2007.
116. J. Kirkpatrick, V. Marcon, J. Nelson, K. Kremer, and D. Andrienko. Charge mobility of discotic mesophases: A multiscale quantum and classical study. *Phys. Rev. Lett.*, 98(22):227402, 2007.
117. J. Nelson, J. J. Kwiatkowski, J. Kirkpatrick, and J. M. Frost. Modeling charge transport in organic photovoltaic materials. *Acc. Chem. Res.*, 42(11):1768–1778, 2009.
118. Y. Nagata and C. Lennartz. Atomistic simulation on charge mobility of amorphous tris(8-hydroxyquinoline) aluminum (Alq[sub 3]): Origin of Poole–Frenkel-type behavior. *J. Chem. Phys.*, 129(3):034709, 2008.
119. A. Fuchs, T. Steinbrecher, M. S. Mommer, Y. Nagata, M. Elstner, and C. Lennartz. Molecular origin of differences in hole and electron mobility in amorphous Alq3—A multiscale simulation study. *Phys. Chem. Chem. Phys.*, 14(12):4259–4270, 2012.
120. A. J. Stone. Distributed polarizabilities. *Mol. Phys.*, 56(5):1065–1082, 1985.
121. A. J. Stone. Distributed multipole analysis—Stability for large basis sets. *J. Chem. Theory Comput.*, 1(6):1128–1132, 2005.
122. A. J. Misquitta and A. J. Stone. Distributed polarizabilities obtained using a constrained density-fitting algorithm. *J. Chem. Phys.*, 124(2):024111, 2006.
123. T. Bereau, D. Andrienko, and O. Anatole von Lilienfeld. Transferable atomic multipole machine learning models for small organic molecules. *J. Chem. Theory Comput.*, 11(7):3225–3233, 2015.

124. W. L. Jorgensen and J. Tirado-Rives. The OPLS [optimized potentials for liquid simulations] potential functions for proteins, energy minimizations for crystals of cyclic peptides and crambin. *J. Am. Chem. Soc.*, 110(6):1657–1666, 1988.
125. W. L. Jorgensen, D. S. Maxwell, and J. Tirado-Rives. Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *J. Am. Chem. Soc.*, 118(45):11225–11236, 1996.
126. A. Lukyanov, C. Lennartz, and D. Andrienko. Amorphous films of tris(8-hydroxyquinolinato)aluminium: Force-field, morphology, and charge transport. *Phys. Status Solidi (A)*, 206(12):2737–2742, 2009.
127. B. H. Besler, K. M. Merz, and P. A. Kollman. Atomic charges derived from semiempirical methods. *J. Comput. Chem.*, 11(4):431–439, 1990.
128. C. M. Breneman and K. B. Wiberg. Determining atom-centered monopoles from molecular electrostatic potentials. The need for high sampling density in formamide conformational analysis. *J. Comput. Chem.*, 11(3):361–373, 1990.
129. F. May, M. Al-Helwi, B. Baumeier, W. Kowalsky, E. Fuchs, C. Lennartz, and D. Andrienko. Design rules for charge-transport efficient host materials for phosphorescent organic light-emitting diodes. *J. Am. Chem. Soc.*, 134(33):13818–13822, 2012. WOS:000307699000042.
130. A. Lukyanov and D. Andrienko. Extracting nondispersive charge carrier mobilities of organic semiconductors from simulations of small systems. *Phys. Rev. B*, 82(19):193202, 2010.
131. T. Neumann, D. Danilov, C. Lennartz, and W. Wenzel. Modeling disordered morphologies in organic semiconductors. *J. Comput. Chem.*, 34(31):2716–2725, 2013.
132. M. Moral, W.-J. Son, J. C. Sancho-Garca, Y. Olivier, and L. Muccioli. Cost-effective force field tailored for solid-phase simulations of OLED materials. *J. Chem. Theory Comput.*, 11(7):3383–3392, 2015.
133. O. Anatole von Lilienfeld, I. Tavernelli, U. Rothlisberger, and D. Sebastiani. Optimization of effective atom centered potentials for London dispersion forces in density functional theory. *Phys. Rev. Lett.*, 93(15):153004, 2004.
134. W. G. Noid, J.-W. Chu, G. S. Ayton, V. Krishna, S. Izvekov, G. A. Voth, A. Das, and H. C. Andersen. The multiscale coarse graining method. 1. A rigorous bridge between atomistic and coarse-grained models. *J. Chem. Phys.*, 128:244114, 2008.
135. V. Rühle, C. Junghans, A. Lukyanov, K. Kremer, and D. Andrienko. Versatile object-oriented toolkit for coarse-graining applications. *J. Chem. Theory Comput.*, 5(12):3211–3223, 2009.
136. A. K. Soper. Empirical potential Monte Carlo simulation of fluid structure. *Chem. Phys.*, 202(2–3):295–306, 1996.
137. A. P. Lyubartsev and A. Laaksonen. Calculation of effective interaction potentials from radial distribution functions: A reverse Monte Carlo approach. *Phys. Rev. E*, 52(4):3730–3737, 1995.
138. D. Reith, M. Pütz, and F. Müller-Plathe. Deriving effective mesoscale potentials from atomistic simulations. *J. Comput. Chem.*, 24(13):1624–1636, 2003.
139. M. Jochum, D. Andrienko, K. Kremer, and C. Peter. Structure-based coarse-graining in liquid slabs. *J. Chem. Phys.*, 137(6):064102, 2012.
140. F. Ercolessi and J. B. Adams. Interatomic potentials from first-principles calculations: The force-matching method. *EPL*, 26(8):583, 1994.
141. S. Izvekov, M. Parrinello, C. J. Burnham, and G. A. Voth. Effective force fields for condensed phase systems from ab initio molecular dynamics simulation: A new method for force-matching. *J. Chem. Phys.*, 120(23):10896–10913, 2004.
142. M. Scott Shell. The relative entropy is fundamental to multiscale and inverse thermo-dynamic problems. *J. Chem. Phys.*, 129(14):144108, 2008.
143. W. G. Noid, J.-W. Chu, G. S. Ayton, and G. A. Voth. Multiscale coarse-graining and structural correlations: Connections to liquid-state theory. *J. Phys. Chem. B*, 111(16):4116–4127, 2007.

144. W. G. Noid. Perspective: Coarse-grained models for biomolecular systems. *J. Chem. Phys.*, 139(9):090901, 2013.
145. K. Ch. Daoulas and M. Müller. Comparison of simulations of lipid membranes with membranes of block copolymers. In W. Peter Meier and W. Knoll, editors, *Polymer Membranes/Biomembranes*, No. 224 in Advances in Polymer Science, pp. 43–85. Berlin: Springer, 2010.
146. P. Gemnden, C. Poelking, K. Kremer, D. Andrienko, and K. Ch. Daoulas. Nematic ordering, conjugation, and density of states of soluble polymeric semiconductors. *Macromolecules*, 46(14):5762–5774, 2013.
147. C. Poelking, K. Daoulas, A. Troisi, and D. Andrienko. Morphology and charge transport in P3ht: A theorist's perspective. In S. Ludwigs, editor, *P3HT Revisited—From Molecular Scale to Solar Cell Devices*, No. 265 in Advances in Polymer Science, pp. 139–180. Berlin: Springer, 2014. doi:10.1007/12-2014-277.
148. P. Gemnden, C. Poelking, K. Kremer, K. Daoulas, and D. Andrienko. Effect of mesoscale ordering on the density of states of polymeric semiconductors. *Macromol. Rapid Commun.*, 36(11):1047–1053, 2015.
149. P, Kordt, S. Stodtmann, A. Badinski, M. Al Helwi, C. Lennartz, and D. Andrienko. Parameter-free continuous drift–diffusion models of amorphous organic semiconductors. *Phys. Chem. Chem. Phys.*, 17(35):22778–22783, 2015.
150. B. Baumeier, O. Stenzel, C. Poelking, D. Andrienko, and V. Schmidt. Stochastic modeling of molecular charge transport networks. *Phys. Rev. B*, 86(18):184202, 2012.
151. W. Tschoep, K. Kremer, J. Batoulis, T. Buerger, and O. Hahn. Simulation of polymer melts. I. Coarse-graining procedure for polycarbonates. *Acta Polym.*, 49(2–3):61–74, 1998.
152. T. Murtola, A. Bunker, I. Vattulainen, M. Deserno, and M. Karttunen. Multiscale modeling of emergent materials: Biological and soft matter. *Phys. Chem. Chem. Phys.*, 11(12):1869–1892, 2009.
153. R. L. Henderson. A uniqueness theorem for fluid pair correlation functions. *Phys. Lett. A*, 49(3):197–198, 1974.
154. W.-K. Oh, S. Qamar Hussain, Y.-J. Lee, Y. Lee, S. Ahn, and J. Yi. Study on the ITO work function and hole injection barrier at the interface of ITO/a-Si:H(p) in amorphous/crystalline silicon heterojunction solar cells. *Mater. Res. Bull.*, 47(10):3032–3035, 2012.
155. R. Schlaf, H. Murata, and Z. H. Kafafi. Work function measurements on indium tin oxide films. *J. Electron Spectrosc. Relat. Phenom.*, 120(1–3):149–154, 2001.
156. G. M. Wu, H. H. Lin, and H. C. Lu. Work function and valence band structure of tin-doped indium oxide thin films for OLEDs. *Vacuum*, 82(12):1371–1374, 2008.
157. C. Hyung Kim, C. Deuck Bae, K. Hee Ryu, B. Ki Lee, and H. Jung Shin. Local work function measurements on various inorganic materials using Kelvin probe force spectroscopy. *Solid State Phenom.*, 124–126:607–610, 2007.
158. D. R. Lide. *CRC Handbook of Chemistry and Physics: A Ready-Reference Book of Chemical and Physical Data*, 79th ed. Boca Raton, FL: CRC Press, 1998.
159. M. Mesta, J. Cottaar, R. Coehoorn, and P. A. Bobbert. Study of charge-carrier relaxation in a disordered organic semiconductor by simulating impedance spectroscopy. *Appl. Phys. Lett.*, 104(21):213301-1–213301-4, 2014.
160. J. J. M. van der Holst, F. W. A. van Oost, R. Coehoorn, and P. A. Bobbert. Monte Carlo study of charge transport in organic sandwich-type single-carrier devices: Effects of Coulomb interactions. *Phys. Rev. B*, 83(8):085206, 2011.
161. S. L. M. van Mensfoort, S. I. E. Vulto, R. A. J. Janssen, and R. Coehoorn. Hole transport in polyfluorene-based sandwich-type devices: Quantitative analysis of the role of energetic disorder. *Phys. Rev. B*, 78(8):085208, 2008.
162. R. J. de Vries, S. L. M. van Mensfoort, R. A. J. Janssen, and R. Coehoorn. Relation between the built-in voltage in organic light-emitting diodes and the zero-field voltage as measured by electroabsorption. *Phys. Rev. B*, 81(12):125203, 2010.

163. M. Mesta, C. Schaefer, J. de Groot, J. Cottaar, R. Coehoorn, and P. A. Bobbert. Charge-carrier relaxation in disordered organic semiconductors studied by dark injection: Experiment and modeling. *Phys. Rev. B*, 88(17):174204, 2013.
164. W. C. Germs, J. J. M. van der Holst, S. L. M. van Mensfoort, P. A. Bobbert, and R. Coehoorn. Modeling of the transient mobility in disordered organic semiconductors with a Gaussian density of states. *Phys. Rev. B*, 84(16):165210, 2011.
165. S. L. M. van Mensfoort and R. Coehoorn. Determination of injection barriers in organic semiconductor devices from capacitance measurements. *Phys. Rev. Lett.*, 100(8):086802, 2008.
166. K. A. Neyts. Simulation of light emission from thin-film microcavities. *J. Opt. Soc. Am. A*, 15(4):962–971, 1998.
167. V. Bulovi, V. B. Khalfin, G. Gu, P. E. Burrows, D. Z. Garbuzov, and S. R. Forrest. Weak microcavity effects in organic light-emitting devices. *Phys. Rev. B*, 58(7):3730–3740, 1998.
168. Z. B. Wang, M. G. Helander, X. F. Xu, D. P. Puzzo, J. Qiu, M. T. Greiner, and Z. H. Lu. Optical design of organic light emitting diodes. *J. Appl. Phys.*, 109(5):053107, 2011.
169. M. Furno, R. Meerheim, S. Hofmann, B. Luessem, and K. Leo. Efficiency and rate of spontaneous emission in organic electroluminescent devices. *Phys. Rev. B*, 85(11):115205, 2012.
170. S. L. M. van Mensfoort, M. Carvelli, M. Megens, D. Wehenkel, M. Bartyzel, H. Greiner, R. A. J. Janssen, and R. Coehoorn. Measuring the light emission profile in organic light-emitting diodes with nanometre spatial resolution. *Nat. Photon.*, 4(0):329–335, 2010.
171. M. Flaemmich, D. Michaelis, and N. Danz. Accessing OLED emitter properties by radiation pattern analyses. *Org. Electron.*, 12(1):83–91, 2011.
172. J. Frischeisen, D. Yokoyama, C. Adachi, and W. Bruetting. Determination of molecular dipole orientation in doped fluorescent organic thin films by photoluminescence measurements. *Appl. Phys. Lett.*, 96(7):073302, 2010.
173. S.-Y. Kim and J.-J. Kim. Outcoupling efficiency of organic light emitting diodes and the effect of ITO thickness. *Org. Electron.*, 11(6):1010–1015, 2010.
174. W. Bruetting, J. Frischeisen, T. D. Schmidt, B. J. Scholz, and C. Mayr. Device efficiency of organic light-emitting diodes: Progress by improved light outcoupling. *Phys. Status Solidi A*, 210(1):44–65, 2013.
175. M. C. Gather and S. Reineke. Recent advances in light outcoupling from white organic light-emitting diodes. *J. Photon. Energy*, 5(1):057607–057607, 2015.
176. H. van Eersel, P. A. Bobbert, R. A. J. Janssen, and R. Coehoorn. Monte Carlo study of efficiency roll-offof phosphorescent organic light-emitting diodes: Evidence for dominant role of triplet–polaron quenching. *Appl. Phys. Lett.*, 105(14):143303, 2014.
177. R. Coehoorn, H. van Eersel, P. Bobbert, and R. Janssen. Kinetic Monte Carlo study of the sensitivity of OLED efficiency and lifetime to materials parameters. *Adv. Funct. Mater.*, 25(13):2024–2037, 2015.
178. Y. Shen and N. C. Giebink. Monte Carlo simulations of nanoscale electrical in-homogeneity in organic light-emitting diodes and its impact on their efficiency and lifetime. *Phys. Rev. Appl.*, 4(5):054017, 2015.
179. G. Lanzani. *The Photophysics behind Photovoltaics and Photonics*. Weinheim: Wiley-VCH, 2012. OCLC: ocn768072884.
180. C. Murawski, K. Leo, and M. C. Gather. Efficiency roll-off in organic light-emitting diodes. *Adv. Mater.*, 25(47):6801–6827, 2013.
181. M. Mesta, M. Carvelli, R. J. de Vries, H. van Eersel, J. J. M. van der Holst, M. Schober, M. Furno, B. Luessem, K. Leo, P. Loebl, R. Coehoorn, and P. A. Bobbert. Molecular-scale simulation of electroluminescence in a multilayer white organic light-emitting diode. *Nat. Mater.*, 12(7):652–658, 2013.
182. H. van Eersel, P. A. Bobbert, and R. Coehoorn. Kinetic Monte Carlo study of triplet–triplet annihilation in organic phosphorescent emitters. *J. Appl. Phys.*, 117(11):115502, 2015.

183. J. Cottaar, R. Coehoorn, and P. A. Bobbert. Modeling of charge transport across disordered organic heterojunctions. *Org. Electron.*, 13(4):667–672, 2012.
184. H. Yoshida and K. Yoshizaki. Electron affinities of organic materials used for organic light-emitting diodes: A low-energy inverse photoemission study. *Org. Electron.*, 20:24–30, 2015.
185. N. C. Giebink and S. R. Forrest. Quantum efficiency roll-off at high brightness in fluorescent and phosphorescent organic light emitting diodes. *Phys. Rev. B*, 77(23):235215, 2008.
186. G. Schwartz, S. Reineke, T. Conrad Rosenow, K. Walzer, and K. Leo. Triplet harvesting in hybrid white organic light-emitting diodes. *Adv. Funct. Mater.*, 19(9):1319–1333, 2009.
187. S. Mladenovski, S. Reineke, and K. Neyts. Measurement and simulation of exciton decay times in organic light-emitting devices with different layer structures. *Opt. Lett.*, 34(9):1375–1377, 2009.
188. K.-C. Tang, K. Lin Liu, and I.-C. Chen. Rapid intersystem crossing in highly phosphorescent iridium complexes. *Chem. Phys. Lett.*, 386(4–6):437–441, 2004.
189. Y. Kawamura, J. Brooks, J. J. Brown, H. Sasabe, and C. Adachi. Intermolecular interaction and a concentration-quenching mechanism of phosphorescent Ir(III) complexes in a solid film. *Phys. Rev. Lett.*, 96(1):017404, 2006.
190. F. S. Steinbacher, R. Krause, A. Hunze, and A. Winnacker. Triplet exciton transfer mechanism between phosphorescent organic dye molecules. *Phys. Stat. Sol. (A)*, 209(2):340–346, 2012.
191. S. L. M. van Mensfoort, R. J. de Vries, V. Shabro, H. P. Loebl, R. A. J. Janssen, and R. Coehoorn. Electron transport in the organic small-molecule material BAlq—The role of correlated disorder and traps. *Org. Electron.*, 11(8):1408–1413, 2010.
192. M. M. Mandoc, B. de Boer, G. Paasch, and P. W. M. Blom. Trap-limited electron transport in disordered semiconducting polymers. *Phys. Rev. B*, 75(19):193202, 2007.
193. F. May, B. Baumeier, C. Lennartz, and D. Andrienko. Can lattice models predict the density of states of amorphous organic semiconductors? *Phys. Rev. Lett.*, 109(13):136401, 2012.
194. S. Olthof, S. Mehraeen, S. K. Mohapatra, S. Barlow, V. Coropceanu, J.-L. Bredas, S. R. Marder, and A. Kahn. Ultralow doping in organic semiconductors: Evidence of trap filling. *Phys. Rev. Lett.*, 109(17):176601, 2012.
195. G. Li, C. H. Kim, Z. Zhou, J. Shinar, K. Okumoto, and Y. Shirota. Combinatorial study of exciplex formation at the interface between two wide band gap organic semiconductors. *Appl. Phys. Lett.*, 88(25):253505, 2006.
196. H. Aziz and Z. D. Popovic. Degradation phenomena in small-molecule organic light-emitting devices. *Chem. Mater.*, 16(23):4522–4532, 2004.
197. F. So and D. Kondakov. Degradation mechanisms in small-molecule and polymer organic light-emitting diodes. *Adv. Mater.*, 22(34):3762–3777, 2010.
198. J.-S. Park, H. Chae, H. Kyoon Chung, and S. I. Lee. Thin film encapsulation for flexible AM-OLED: A review. *Semicond. Sci. Technol.*, 26(3):034001, 2011.
199. S. Scholz, D. Kondakov, B. Luessem, and K. Leo. Degradation mechanisms and reactions in organic light-emitting devices. *Chem. Rev.*, 115(16):8449–8503, 2015.
200. H. Aziz, Z. D. Popovic, N.-X. Hu, A.-M. Hor, and G. Xu. Degradation mechanism of small molecule-based organic light-emitting devices. *Science*, 283(5409):1900–1902, 1999.
201. D. Y. Kondakov, W. C. Lenhart, and W. F. Nichols. Operational degradation of organic light-emitting diodes: Mechanism and identification of chemical products. *J. Appl. Phys.*, 101(2):024512, 2007.
202. R. Seifert, S. Scholz, B. Luessem, and K. Leo. Comparison of ultraviolet- and charge-induced degradation phenomena in blue fluorescent organic light emitting diodes. *Appl. Phys. Lett.*, 97(1):013308, 2010.
203. N. C. Giebink, B. W. D'Andrade, M. S. Weaver, P. B. Mackenzie, J. J. Brown, M. E. Thompson, and S. R. Forrest. Intrinsic luminance loss in phosphorescent small-molecule organic light emitting devices due to bimolecular annihilation reactions. *J. Appl. Phys.*, 103(4):044509, 2008.

204. T. D. Schmidt, L. Jaeger, Y. Noguchi, H. Ishii, and W. Bruetting. Analyzing degradation effects of organic light-emitting diodes via transient optical and electrical measurements. *J. Appl. Phys.*, 117(21):215502, 2015.
205. T. D. Schmidt, D. S. Setz, M. Flaemmich, J. Frischeisen, D. Michaelis, B. C. Krummacher, N. Danz, and W. Bruetting. Evidence for non-isotropic emitter orientation in a red phosphorescent organic light-emitting diode and its implications for determining the emitter's radiative quantum efficiency. *Appl. Phys. Lett.*, 99(16):163302, 2011.
206. C. Fry, B. Racine, D. Vaufrey, H. Doyeux, and S. Cin. Physical mechanism responsible for the stretched exponential decay behavior of aging organic light-emitting diodes. *Appl. Phys. Lett.*, 87(21):213502, 2005.
207. A. Mass, R. Coehoorn, and P. A. Bobbert. Universal size-dependent conductance fluctuations in disordered organic semiconductors. *Phys. Rev. Lett.*, 113(11), 2014.

16

Tunnel-Junction Light-Emitting Diodes

Yen-Kuang Kuo
Jih-Yuan Chang
Ya-Hsuan Shih
Fang-Ming Chen
and
Miao-Chan Tsai

16.1 Introduction

The use of p-n tunnel junctions for applications in GaN-based electronic and optoelectronic devices has become increasingly attractive. The use of tunnel junctions has conceptually been desired for the reuse of carriers for coupled active regions, enabling high quantum efficiencies and improved vertical transport (Ozden et al., 2001). Concepts of the tunnel field-effect transistors, multiple active region light-emitting diodes (LEDs) and laser diodes, high-conductivity hole injection layers, and multijunction solar cells may be realized with the III-nitride material system providing that a sufficiently low-resistivity tunnel junction can be obtained. However, the feasibility of forming low-resistivity tunnel junctions in the wide-bandgap III-nitride devices has been a challenge due to the high hole and electron concentrations required for band alignment. Substantial effort is still required to resolve this critical issue.

16.1.1 Polarization-Assisted Tunneling Effect

The probability of interband tunneling across a potential barrier is mainly governed by the tunneling barrier height determined by the bandgap and the tunneling barrier thickness (Simon et al., 2009). For wide-bandgap III-nitrides, tunneling is difficult due to the high barrier height and the difficulty in achieving degenerate p-type impurity doping (Simon et al., 2009; Krishnamoorthy et al., 2010). However, the strong polarization-induced electric field available in the III-nitrides and other highly polar semiconductor materials provides a new design approach for tunneling structures. In hetero-structures of highly polar materials, the polarization-induced sheet charges can create significantly high electric fields resulting in large band bending over a small distance. The insufficient band-bending for interband tunneling from doping-induced built-in electric field (E_{bi}) can thus be supplemented by the polarization-induced

electric field (E_p) assuming that the two electric fields are oriented in the same direction. The probability of interband tunneling can therefore be increased.

Recently, GaN-based tunnel junctions, which are supplemented by the polarization-induced electric field, have been demonstrated (Simon et al., 2009; Krishnamoorthy et al., 2010, 2011, 2013; Schubert, 2010a; Grundmann and Mishra, 2007). The use of large spontaneous and piezoelectric polarization fields present in the III-nitrides enables the formation of thin depletion region and relaxes the strict requirement for degenerate doping. Structures including the GaN/AlN/GaN, AlN/GaN/AlN, and GaN/InGaN/GaN have been investigated theoretically and experimentally (Simon et al., 2009; Krishnamoorthy et al., 2010; Schubert, 2010a). In particular, tunnel-junction structures using an intermediate InGaN layer have been developed due to the advantages of low bandgap and large piezoelectric polarization. It was shown by estimation of tunneling probability using the Wentzel–Kramers–Brillouin (WKB) approximation that the optimal performance, with n-doping $N_D = 5\times10^{19}$ cm^{-3} and p-doping $N_A = 1\times10^{19}$ cm^{-3}, would occur when the indium composition in InGaN is greater than 30% and the thickness is larger than 3 nm, which was achieved on the N-polar GaN with molecular beam epitaxy (MBE) growth method (Grundmann and Mishra, 2007; Krishnamoorthy et al., 2011, 2013). In many of the envisioned applications in optoelectronics, an indium composition of greater than approximately 20% in InGaN is not favorable due to the high optical absorption caused by the InGaN and the problem in crystal growth for the layers subsequent to the tunnel junction. Moreover, the doping levels are limited by the Si-induced roughening and strain, ionization efficiency, and solubility of Mg dopants. In this chapter, design and analysis for the low-resistivity tunnel junctions, which are with adequate indium composition and doping density for a typical Ga-polar InGaN multiple-quantum well (MQW) LED structure, are introduced.

16.1.2 Efficiency Droop in III-Nitride LEDs

For the application in solid-state lighting, the development of high-efficiency and high-power III-nitride LEDs is required. In this demand, the major issue that needs to be resolved might be the problem of efficiency droop, i.e., the reduction in illumination efficiency at high current density. It is believed that heating is not the major cause for efficiency droop due to the fact that the efficiency droop occurs in both pulsed and continuous wave (CW) conditions and the droop becomes severe with the decrease of ambient temperature (Laubsch et al., 2009; Kim et al., 2007). To date, the physical origin of efficiency droop remains debatable even though numerous possible physical mechanisms, such as the carrier delocalization (Monemar and Sernelius, 2007; Chichibu et al., 2006), Auger recombination (Shen et al., 2007; Delaney et al., 2009), insufficient hole injection (Rozhansky and Zakheim, 2006), carrier leakage (Shim et al., 2012; Lin et al., 2012), and polarization effect (Kim et al., 2007; Kuo et al., 2009), have been proposed and demonstrated to be related to the efficiency droop in GaN-based LEDs. Even though continuous efforts have been made, an overall solution for the efficiency droop of III-nitride LEDs is still lacking. In this chapter, instead of trying to promote the quantum efficiency at high current density, another approach in circumventing the issue of efficiency droop is proposed, which is based on the use of GaN-based tunnel junctions. Specifically, in this approach it is suggested to operate the III-nitride LEDs within the range of high efficiency, i.e., to operate the LEDs at low current density, with the insertion of a low-loss tunnel junction between two active regions, which is beneficial in reducing the carrier density in each active region and mitigating the problem of efficiency droop.

16.2 Physical Models and Simulation Parameters

16.2.1 Bulk Band Structure

The characteristics of III-nitride tunnel junction and tunnel-junction LEDs constructed along the *c*-axis are numerically studied using a self-consistent simulation program Advanced Physical Models of Semiconductor Devices (APSYS). APSYS employs the 6×6 $\mathbf{k}\cdot\mathbf{p}$ model, which was developed for the strained

wurtzite semiconductor by Chuang and Chang (1996), to calculate the energy-band structures of the non-active bulk regions. The unstrained bandgap energies of the $In_xGa_{1-x}N$ and $Al_yGa_{1-y}N$ ternary alloys can be expressed by the following formula (Vurgaftman et al., 2001)

$$E_g(In_xGa_{1-x}N) = x \cdot E_g(GaN) + (1-x) \cdot E_g(InN) - x \cdot (1-x) \cdot B(InGaN), \tag{16.1}$$

$$E_g(Al_yGa_{1-y}N) = y \cdot E_g(GaN) + (1-y) \cdot E_g(AlN) - y \cdot (1-y) \cdot B(AlGaN), \tag{16.2}$$

where $E_g(GaN)$, $E_g(InN)$, and $E_g(AlN)$ are the bandgap energies of GaN, InN, and AlN, which have values of 3.42, 0.64, and 6.0 eV at 300 K, respectively (Nepal et al., 2005; Wu et al., 2003). $B(InGaN)$ and $B(AlGaN)$ are the bandgap bowing parameters of InGaN and AlGaN, respectively. The bandgap bowing parameters of InGaN and AlGaN are assumed to be 2.1 and 1.0 eV, respectively (Gorczyca et al., 2011a, b). The band-offset ratio is assumed to be 0.7/0.3 for the III-nitride material systems.

16.2.2 MQW Model: Approximation of Effective Mass

A model of effective mass for the computation of density of states (DOS) and for a simplified MQW model is needed because, in general, the dispersion of bulk valence bands is nonparabolic and anisotropic with strong mixing or anticrossing behavior in the direction perpendicular to the *c*-axis. Based on the method of Chuang and Chang (1996), an analytical model of effective mass is implemented for the valence band.

Within a range of small k, a situation when the valence band is lightly populated by holes, the following effective masses hold:

$$m_{hh}^{z} = -m_0(A_1 + A_3)^{-1}, \tag{16.3}$$

$$m_{lh}^{z} = -m_0 \left[A_1 + \left(\frac{E_2^0 - \lambda_e}{E_2^0 - E_3^0} \right) A_3 \right]^{-1}, \tag{16.4}$$

$$m_{ch}^{z} = -m_0 \left[A_1 + \left(\frac{E_3^0 - \lambda_e}{E_3^0 - E_2^0} \right) A_3 \right]^{-1}, \tag{16.5}$$

$$m_{hh}^{t} = -m_0(A_2 + A_4)^{-1}, \tag{16.6}$$

$$m_{lh}^{t} = -m_0 \left[A_2 + \left(\frac{E_2^0 - \lambda_e}{E_2^0 - E_3^0} \right) A_4 \right]^{-1}, \tag{16.7}$$

$$m_{ch}^{t} = -m_0 \left[A_2 + \left(\frac{E_3^0 - \lambda_e}{E_3^0 - E_2^0} \right) A_4 \right]^{-1}, \tag{16.8}$$

where E_i^0 $(i = 1, 2, 3)$ are the valence-band edge (values) at $k = 0$.

For a large range of k, i.e., a situation when the valence band is heavily populated by holes, the following effective masses are valid:

$$m_{hh}^{z} = -m_0(A_1 + A_3)^{-1}, \tag{16.9}$$

$$m_{lh}^{z} = -m_0(A_1 + A_3)^{-1}, \tag{16.10}$$

$$m_{ch}^{z} = -m_0 A_1^{-1}, \tag{16.11}$$

$$m_{hh}^{t} = -m_0(A_2 + A_4 - A_5)^{-1}, \tag{16.12}$$

$$m_{lh}^{t} = -m_0(A_2 + A_4 - A_5)^{-1}, \tag{16.13}$$

$$m_{ch}^{t} = -m_0 A_2^{-1}. \tag{16.14}$$

In the present simulation, a compromise of the above two models is used, i.e., to average them. Note that the choice of model is based on how heavily the valence bands are populated by holes. Other parameters of band structure for the binary nitride wurtzite semiconductors employed in the simulation are listed in Table 16.1 (Wu, 2009; Vurgaftman and Meyer, 2007).

16.2.3 Incomplete Ionization of Impurities

The degree of ionization is described by the occupancies f_D and f_A. Shallow impurities are assumed to be in equilibrium with the local carriers and therefore the occupancy of the shallow impurities can be described by

$$f_D = \frac{1}{1 + g_d^{-1} \exp[(E_D - E_{fn})/kT]}, \tag{16.15}$$

$$f_A = \frac{1}{1 + g_a \exp[(E_A - E_{fp})/kT]}, \tag{16.16}$$

where the subscripts D and A are used to denote shallow donors and acceptors, respectively. The degeneracy levels are $g_d = 2$ and $g_a = 4$ (Tiwari, 1992).

TABLE 16.1 Material Parameters of the Binary Semiconductors GaN, AlN, and InN (Wu, 2009; Vurgaftman and Meyer, 2007)

Parameter	Symbol (Unit)	GaN	AIN	InN
Electron effective mass (*c*-axis)	$m_c^z\,(m_0)$	0.21	0.33	0.068
Electron effective mass (transverse)	$m_c^t\,(m_0)$	0.19	0.32	0.065
Hole effective mass parameter	A_1	−7.21	−3.86	−8.21
	A_2	−0.44	−0.25	−0.68
	A_3	6.68	3.58	7.57
	A_4	−3.46	−1.32	−5.23
	A_5	−3.40	−1.47	−5.11
	A_6	−4.90	−1.64	−5.96
Spin–orbit split energy	Δ_{so}(eV)	0.017	0.019	0.005
Crystal-field split energy	Δ_{cr}(eV)	0.010	–0.169	0.040
Lattice constant *a*	a_0(Å)	3.189	3.112	3.533
Lattice constant *c*	c_0(Å)	5.185	4.982	5.693
Elastic stiffness constant	C_{11} (GPa)	390	396	223
Elastic stiffness constant	C_{12} (GPa)	145	137	115
Elastic stiffness constant	C_{13} (GPa)	106	108	92
Elastic stiffness constant	C_{33} (GPa)	398	373	224
Elastic stiffness constant	C_{44} (GPa)	105	116	48
Hydrost. deform. potential (*c*-axis)	a_z (eV)	−7.1	−3.4	−4.2
Hydrost. deform. potential (transverse)	a_t (eV)	−9.9	−11.8	−4.2
Shear deform. potential	D1 (eV)	−3.6	−2.9	−3.6
	D2 (eV)	1.7	4.9	1.7
	D3 (eV)	5.2	9.4	5.2
	D4 (eV)	−2.7	−4.0	−2.7
	D5 (eV)	−2.8	−3.3	−2.8
	D6 (eV)	−4.3	−2.7	−4.3
Dielectric constant	ε	8.9	8.5	10.5

Note: $\Delta_{cr} = \Delta_1$, $\Delta_{so} = 3\Delta_2 = 3\Delta_3$.

In the simulation, the acceptor activation energy (E_A) of Mg doped in InGaN is scaled linearly from 60 meV (InN) (Khan et al., 2007) to 166 meV (GaN) (Kumakura et al., 2003). The acceptor activation energy of Mg doped in AlGaN is scaled linearly from 166 meV (GaN) (Kumakura et al., 2003) to 510 meV (AlN) (Nam et al., 2003).

16.2.4 Electronic Polarization and Interface Charges

In typical III-nitride semiconductor devices, the device structure is usually constructed of ternary AlGaN and InGaN alloys. In order to consider the influences of internal polarization in tunnel junction and tunnel-junction LEDs, the method developed by Fiorentini et al. (2002) is employed to estimate the internal polarization in Ga-face configuration, which is represented by fixed surface charges at hetero-interfaces (Fiorentini et al., 2002). The composition of the compound determines the net polarization charges that remain at each hetero-interface. The nonlinear model described in the papers by Fiorentini et al. (2002) and Vurgaftman and Meyer (2003) is used here. The spontaneous polarization P_{sp} (Cm^{-2}) is calculated as

$$P_{sp}(Al_xGa_{1-x}N) = -0.090 \cdot x - 0.034 \cdot (1-x) + 0.019 \cdot x \cdot (1-x), \tag{16.17}$$

$$P_{sp}(In_xGa_{1-x}N) = -0.042 \cdot x - 0.034 \cdot (1-x) + 0.038 \cdot x \cdot (1-x). \tag{16.18}$$

For binary compounds, the piezoelectric polarization P_{pz} (Cm^{-2}) is given as the nonlinear function of the transverse strain ε_{xx} by

$$P_{pz}(GaN) = -0.918 \cdot \varepsilon_{xx} + 9.541 \cdot \varepsilon_{xx}^2, \tag{16.19}$$

$$P_{pz}(InN) = -1.373 \cdot \varepsilon_{xx} + 7.559 \cdot \varepsilon_{xx}^2, \tag{16.20}$$

$$P_{pz}(AlN) = -1.808 \cdot \varepsilon_{xx} - 7.888 \cdot \varepsilon_{xx}^2 \ (\varepsilon_{xx} > 0), \tag{16.21}$$

$$P_{pz}(AlN) = -1.808 \cdot \varepsilon_{xx} + 5.624 \cdot \varepsilon_{xx}^2 \ (\varepsilon_{xx} < 0). \tag{16.22}$$

The piezoelectric polarization of a specific ternary compound is linearly interpolated by the piezoelectric polarizations of the corresponding binary compounds.

The built-in electric fields of III-nitride devices obtained from experimental measurements are usually smaller than the values obtained from theoretical calculation, ranging from 20% to 80% (Piprek, 2007), due mainly to partial compensation of the polarization field by fixed defect and interface charges. In this study, the polarization-induced surface charge densities are assumed to be 50% of the values determined from theoretical calculation.

16.2.5 Interband Tunneling Model

Direct interband tunneling in a p-n junction between the valence and conduction bands in an external applied electric field is commonly referred to as a tunnel junction, or an Esaki junction, or a Zener tunneling diode under reverse bias. The probability of tunneling within the WKB approximation, following the derivation of band-to-band tunneling in the papers by Kane (1960), Duke (1969), and Moll (1964), is calculated as follows:

$$D = \exp(-2J) = P_0 \exp\left(\frac{-E_\perp}{E}\right), \tag{16.23}$$

where

$$J\left(E_{\parallel}\right)=\int_{x_2}^{x_1}\left|\left(\frac{2m^*}{\hbar^2}\right)\frac{\left[\left(E_g/2\right)^2-\left(\varepsilon_c\right)^2\right]}{E_g}+E_{\perp}\right|^{1/2}\mathrm{d}x, \tag{16.24}$$

$$P_0=\exp\left[\frac{\pi(m^*)^{1/2}(E_g)^{3/2}}{2(2)^{1/2}qF\hbar}\right]=\exp\left(-\frac{E_g}{4E}\right)$$
$$\underline{E}=\frac{(2)^{1/2}qF\hbar}{2\pi(m^*)^{1/2}(E_g)^{1/2}}, \tag{16.25}$$

$$m^*=\frac{2m_cm_v}{(m_c+m_v)}. \tag{16.26}$$

$E_{\parallel}$ and $E_{\perp}$ are the electron kinetic energies in the directions along and perpendicular to the tunneling direction, respectively. P_0 is the tunneling probability with a momentum of zero perpendicular to the x-direction. The electrons that can tunnel through the bandgap barrier are those electrons with perpendicular momentum near zero if $\underline{E}$ is small. $\underline{E}$ is a measure of significance of perpendicular momentum range and m^* is the effective tunneling mass.

16.3 III-Nitride Tunnel Junction

For the applications in Ga-polarity InGaN-based LEDs, with the use of InGaN as the barrier material between the heavily doped n- and p-type GaN layers, the vectors of the polarization and normal built-in fields have the same direction for an n-on-p tunnel diode, which helps the generation of the electric field required for tunneling. Furthermore, the reduced barrier height of InGaN also promotes the tunneling probability of carriers since the bandgap energy of InGaN is relatively low when compared to GaN. Figure 16.1 shows the schematic diagram of n-GaN/p-GaN and n-GaN/InGaN/p-GaN tunnel-junction structures. The illustrations for the directions of built-in field and polarization field are also indicated in Figure 16.1 for better reference. Here, let us begin with the n-GaN/InGaN/p-GaN tunnel-junction structure, which is composed of a heavily doped n-GaN layer, a 7-nm-thick InGaN barrier layer, and a heavily doped p-GaN layer (Krishnamoorthy et al., 2011), and calculate the tunneling current as a function of reverse bias voltage, as shown in Figure 16.2. In the simulation, the dopant concentrations of n- and p-type GaN layers are assumed to be identical, while the ionization energies for the computation of incomplete ionization of donors and acceptors are quite distinct, making the electron and hole concentrations also different. It is observed in Figure 16.2a that the tunneling current increases dramatically with the increase of indium composition. It is required that the Fermi levels are in the valence band and the conduction band

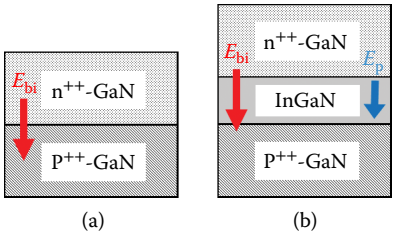

FIGURE 16.1 Schematic diagrams of (a) n-GaN/p-GaN and (b) n-GaN/InGaN/p-GaN tunnel-junction structures.

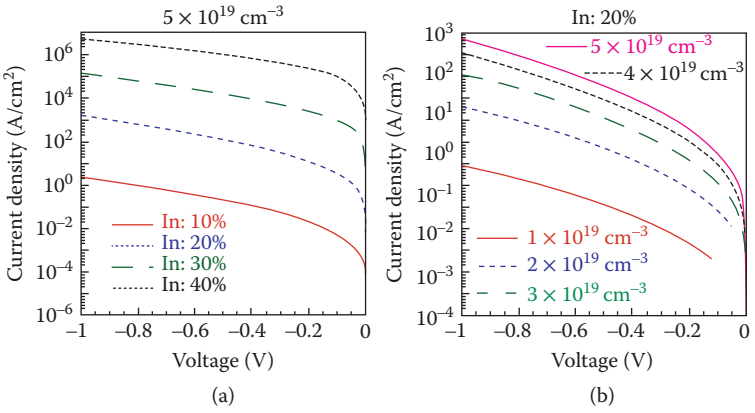

FIGURE 16.2 Electrical characteristics of the n-i-p tunnel diode in the reversed bias regime under various conditions of (a) indium composition and (b) acceptor and donor densities.

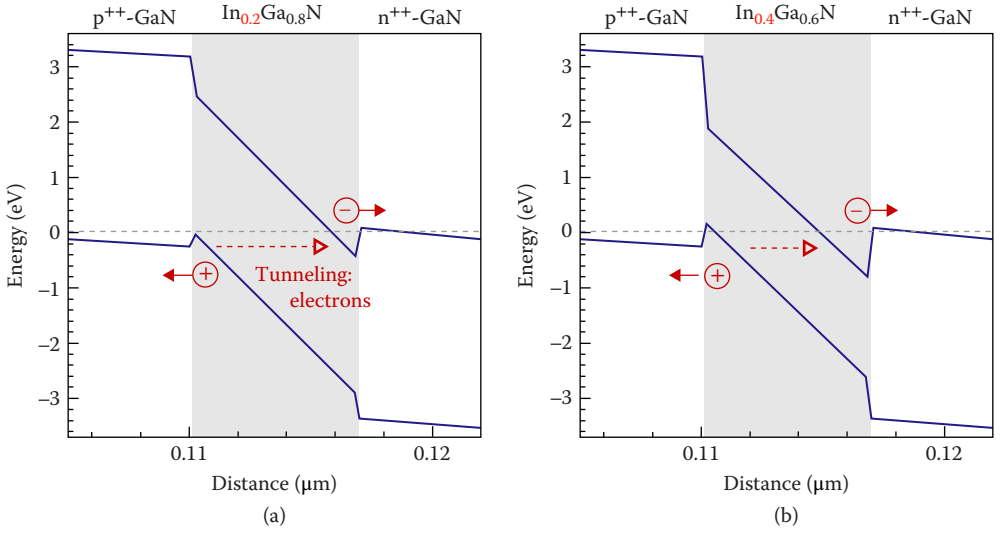

FIGURE 16.3 Energy band diagrams of (a) n-GaN/$In_{0.2}Ga_{0.8}N$/p-GaN and (b) n-GaN/$In_{0.4}Ga_{0.6}N$/p-GaN tunnel diodes at equilibrium.

on the p- and n-side, respectively, so that the carriers can tunnel between the bands as in a Zener diode. The improved tunneling characteristics with the increase of indium composition are thus ascribed to the promotion of the polarization field and the reduction of potential barrier height, as shown in Figure 16.3. However, it is noteworthy that when the composition of indium is high, the crystalline layer tends to relax and the film quality may deteriorate due to the large lattice mismatch and thermal mismatch between the InN and GaN. Under these circumstances, the strain-related polarization field reduces and the defect density increases, which is harmful for the applications in optoelectronics since the high defect density may act as the nonradiative recombination centers (Cherns et al., 2001). Moreover, for many applications in optoelectronics, it is usually required that the tunnel junction is with a reduced percentage of indium due to the problem of optical absorption. In the present study, the indium composition is fixed at 20% for subsequent simulations.

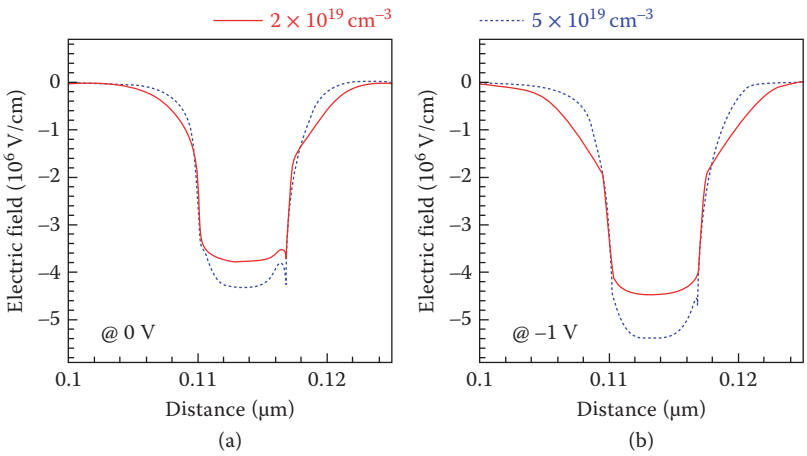

FIGURE 16.4 Electric fields of n-GaN/$In_{0.2}Ga_{0.8}N$/p-GaN tunnel diode with various acceptor and donor densities (a) at equilibrium and (b) at −1 V.

Another approach to improve the performance of tunnel junction is to increase the built-in field and reduce the depletion width by supporting more space charges through the increase of acceptor and donor densities of heavily doped p-type and n-type GaN layers. In Figure 16.4, it is found that the total electric field of the $In_{0.2}Ga_{0.8}N$ layer can be effectively enhanced when the acceptor and donor densities are increased from 2×10^{19} cm^{-3} to 5×10^{19} cm^{-3}. Consequently, the tunneling current increases with the increase of doping level, as shown in Figure 16.2b. Note that there is a limit for the above-mentioned improvement because the doping level of III-nitrides cannot be unlimitedly increased, especially for p-type doping. It is also noteworthy that the tunneling characteristics can be significantly promoted by increasing the Si doping level, while the tunneling current cannot be appreciably increased by increasing the Mg doping (Krishnamoorthy et al., 2013; Tsai et al., 2013). The findings provide another design approach that might be applicable for the realization of III-nitride tunnel junctions since heavy doping can be more easily achieved for the n-type materials than the p-type materials.

16.4 Tunnel-Junction LED

16.4.1 Single LED

Characterization of the blue single LED structure under study, which is used as a reference for the following investigations in stacking LEDs with tunnel junctions, is introduced first in this section. The LED structure is composed of a 3-μm-thick n-GaN layer (1×10^{19} cm^{-3}), four pairs of QWs with 3-nm-thick $In_{0.2}Ga_{0.8}N$ wells and 6-nm-thick GaN barriers, a 20-nm-thick p-$Al_{0.15}Ga_{0.85}N$ electron-blocking layer (EBL) (1×10^{19} cm^{-3}), and 100-nm-thick p-GaN layer (1×10^{19} cm^{-3}). The layer structure of the reference blue single LED is listed in Table 16.2. In the simulation, to focus on the topic explored and to eliminate the issue of current crowing, ideal ohmic contacts are assumed to cover the full top and bottom surfaces of the simulated vertical LED structure. The Shockley–Read–Hall (SRH) recombination lifetime and Auger recombination coefficient are set to be 50 ns and 2×10^{-30} cm^6/s, respectively (Yoshida et al., 2010; Piprek and Li, 2013). It is also assumed that the light extraction efficiency is 70%.

Figure 16.5 shows the simulated electrical and optical characteristics of the blue single LED. In the energy band diagram, as shown in Figure 16.5a, obvious band deformation caused by the polarization field is observed. In particular, the energy bands of QWs and barriers are severely sloped, forming triangular-shaped band profiles, which impact the transport, injection, and confinement of carriers in the MQW active region (Kim et al., 2007; Kuo et al., 2009). The polarization-induced sloped well results in spatial

TABLE 16.2 Layer Structure of the Reference Blue Single Light-Emitting Diode (LED) under Study

Parameter (unit)	d(nm)	$N_{dop}(1/cm^3)$
p-GaN	100	1×10^{19}
p-$Al_{0.15}Ga_{0.85}N$ (EBL)	20	1×10^{19}
i-GaN (barrier)	6	–
i-$In_{0.2}Ga_{0.8}N$ (well)	3	–
i-GaN (barrier)	6	–
i-$In_{0.2}Ga_{0.8}N$ (well)	3	–
i-GaN (barrier)	6	–
i-$In_{0.2}Ga_{0.8}N$ (well)	3	–
i-GaN (barrier)	6	–
i-$In_{0.2}Ga_{0.8}N$ (well)	3	–
i-GaN (barrier)	6	–
n-GaN	3000	1×1019

Note: d, layer thickness; N_{dop}, dopant density.

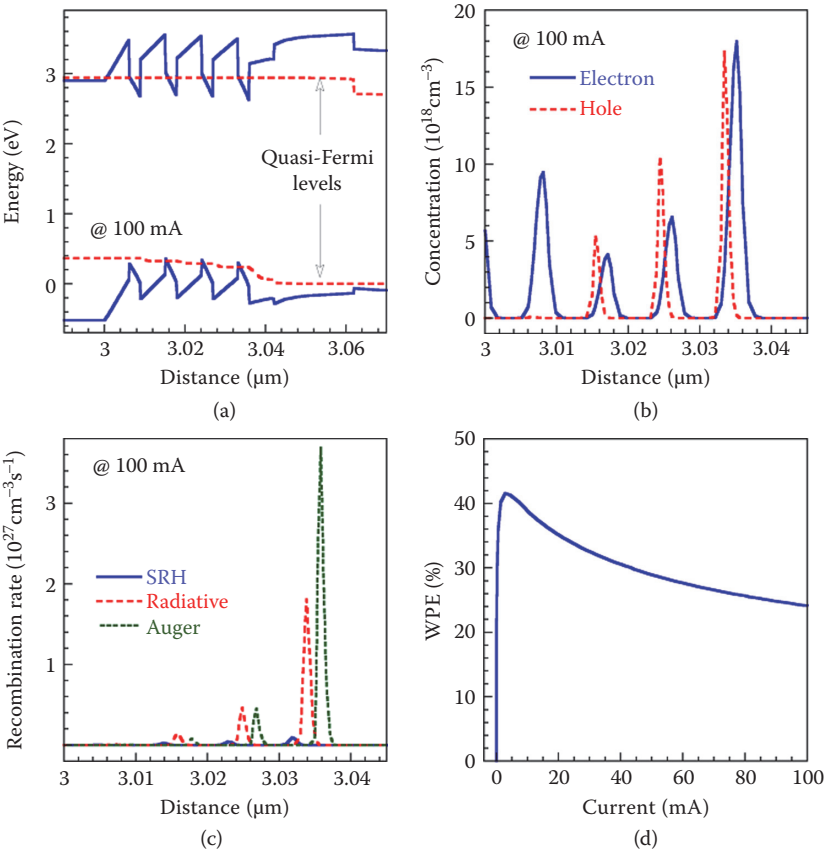

FIGURE 16.5 (a) Energy band diagram, (b) carrier concentrations, (c) recombination rates at 100 mA, and (d) wall-plug efficiency (WPE) of the reference blue single light-emitting diode (LED).

separation of electron and hole wave functions in the well, which is believed to be one of the key issues that deteriorate the radiative recombination. Moreover, the electron leakage current is negligibly small, which indicates that the capability of carrier confinement of the MQW active region is superior. As shown in Figure 16.5b, the distribution of carriers in the QWs is quite nonuniform. Most of the carriers accumulate in the QWs close to the p-side. Similar nonuniform carrier distribution in the InGaN LED has also been reported, which is attributed to the poor injection and transportation of holes (David et al., 2008). Under these circumstances, the illumination efficiency of the blue single LED is severely limited by the huge transition loss of Auger recombination. The droop of wall-plug efficiency (WPE), defined as $(\mathrm{WPE}_{\max} - \mathrm{WPE}_{100\mathrm{mA}})/\mathrm{WPE}_{\max}$ where $\mathrm{WPE}_{\max}$ is the maximum WPE and $\mathrm{WPE}_{100\mathrm{mA}}$ is the WPE at 100 mA, is as large as 42%, as shown in Figure 16.5c and d. Note that the WPE is referred to the energy conversion efficiency with which the LED converts the electrical power into optical power, as defined below:

$$\mathrm{WPE} = \frac{\text{Light output power}}{\text{Current} \times \text{Voltage}}. \tag{16.27}$$

The influence of Auger recombination on the efficiency of LEDs is typically analyzed based on the reduced rate equation. Assuming that the injection efficiency is 100%, the internal quantum efficiency (IQE), which is directly related to the light output power, can be expressed as

$$\mathrm{IQE} = \frac{B \cdot n^2}{A \cdot n + B \cdot n^2 + C \cdot n^3}, \tag{16.28}$$

where A, B, C, and n are the SRH coefficient, radiative coefficient, Auger coefficient, and carrier density, respectively. According to Equation 16.28, the Auger recombination dominates the total recombination rate at high carrier density because it scales with the cubic power of the free carrier density, which is believed to be one of the key factors contributing to the efficiency droop (Shen et al., 2007). If the carrier density in the active region can be reduced while the total carriers confined within the active region remains unchanged, the efficiency at high current may be improved due to the suppression of Auger recombination. It might be helpful to use more QWs in the active region because the carrier density per QW may be reduced under this situation. Figure 16.6 shows the simulated carrier concentrations and recombination rates of the single LED structure with 8 pairs and 12 pairs of QWs. It is observed that the carrier distribution of the LED structures with more QWs is still quite nonuniform and most of the carriers accumulate in the QWs close to the p-side. Therefore, the goal of reducing the carrier density in the QWs cannot be satisfactorily achieved. Under this circumstance, the structure with more QWs does not benefit from the suppressed Auger recombination and hence, similar radiative-Auger relations are observed for the 4-, 8-, and 12-QW LED structures as shown in Figure 16.6c and d. As a result, the efficiency characteristics do not exhibit much difference when the number of QWs changes.

16.4.2 Blue Tunnel-Junction LED

The monolithic structure with multiple LEDs stacked by tunnel junctions was introduced recently (Akyol et al., 2013; Piprek, 2014; Tsai et al., 2014). The stacking of multiple LED structures with tunnel junctions allows for the repeated use of electrons and holes for photon generation in each individual single LED. Ideally, if every individual LED in the stacking structure possesses identical carrier transport and recombination mechanism, the total photon number of the tunnel-junction LED should be approximately N times to that obtained from the single LED if the number of the unit LED in the stacking structure is N, provided that the tunnel junctions possess negligible loss. Therefore, an IQE of more than 100% is possible for the tunnel-junction LEDs. Note that, accompanying with the multiplication of output power, bias and input power multiply with the number of unit LED in the stacking structure simultaneously. Hence, the WPE of tunnel-junction LEDs remains below 100% (Piprek, 2014).

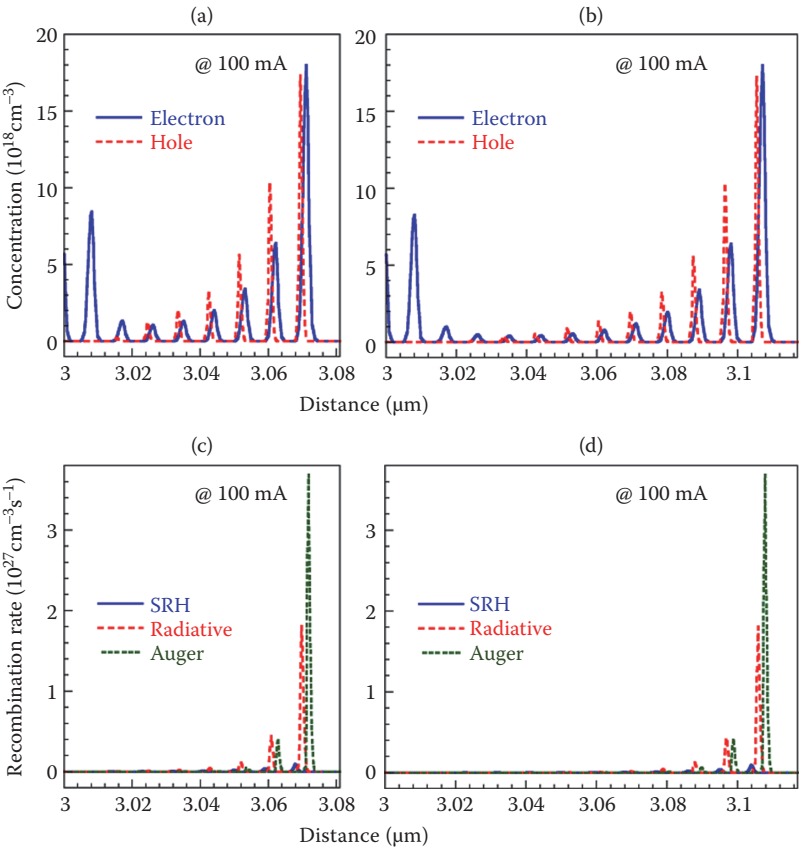

FIGURE 16.6 (a) and (b) Carrier concentrations and (c) and (d) recombination rates of the blue single light-emitting diode (LED) with 8 pairs and 12 pairs of QWs at 100 mA.

The simulated energy band diagram and radiative recombination rate at 100 mA for the blue tunnel-junction LEDs are depicted in Figure 16.7. Note that, in this figure, the n-type layer is located in the left-hand side of each unit LED. In the simulation, the n-GaN/$In_{0.2}Ga_{0.8}N$/p-GaN tunnel junction with a dopant concentration of 5×10^{19} cm^{-3} is utilized to ensure good tunneling characteristics, as shown in Figure 16.2c. When the tunnel-junction LED is turned on, electrons are injected and then transported in the conduction band of n-type layer of the first stacked LED. In the meantime, in the valance band of the p-type layer, electrons are transported and then transferred into the conduction band of the second stacked LED, which is regarded as holes moving in the opposite direction. The electrons and holes then recombine in the MQW active region of the first stacked LED to generate photons. The carrier transport and photon generation repeat in subsequent stacked LEDs. The emission profiles of all unit LEDs are almost identical, as shown in Figure 16.7.

Figure 16.8 shows the light output power as a function of current (L–I) and current as a function of voltage (I–V) characteristics of the blue single LED and tunnel-junction LEDs. It is observed that the light output power and operation voltage of the stacked LEDs are almost equivalent to the light output power and operation voltage of the blue single LED multiplied by the number of unit LED in the stacked LED structure. Under these circumstances, the WPE of the tunnel-junction LEDs would not vary too much comparing to that of the single LED. As mentioned previously, the approach of stacking LEDs with tunnel junctions is not intended to promote the quantum efficiency at high current density; instead, it finds a way to operate the LED devices at low current density that possesses relatively high IQE. Figure 16.9 shows the

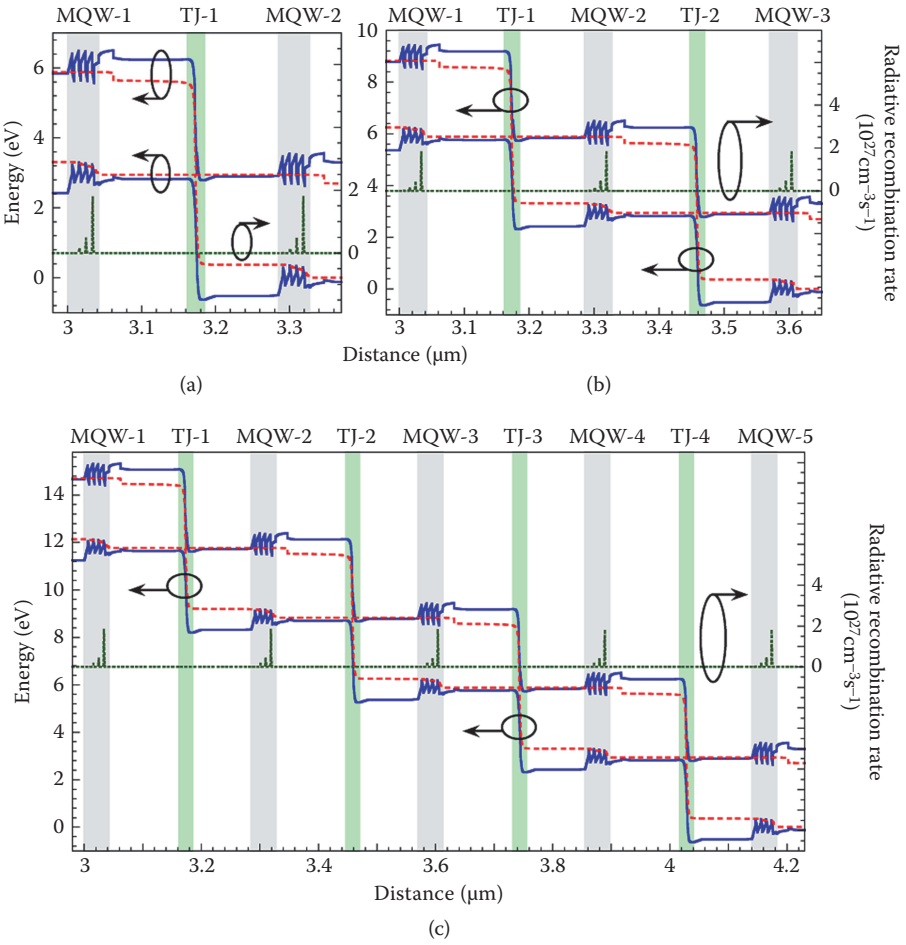

FIGURE 16.7 Energy band diagram and radiative recombination rate at 100 mA for the blue tunnel-junction light-emitting diodes (LEDs) with (a) two unit LEDs, (b) three unit LEDs, and (c) five unit LEDs.

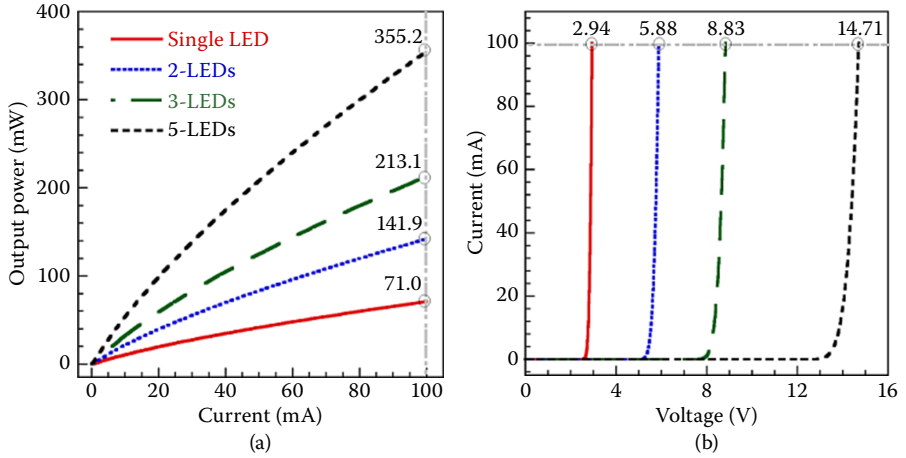

FIGURE 16.8 L–I and I–V characteristics of the blue single light-emitting diode (LED) and tunnel-junction LEDs.

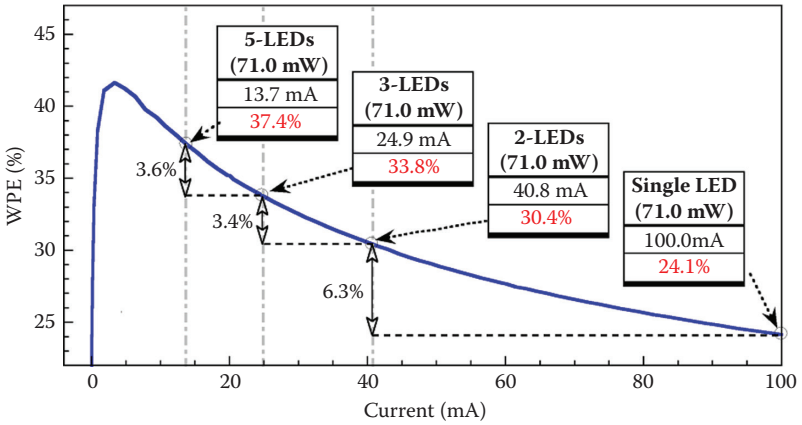

FIGURE 16.9 WPE of the blue light-emitting diode (LED) structures under study.

WPE of the blue LED structures under study. Only one curve is presented in this figure since all the WPE profiles are quite similar. For the single LED structure, the output power is 71 mW at 100 mA. For the tunnel-junction LEDs with two and three stacked unit LEDs, in order to obtain the same output power, the operation currents are 40.8 and 20.9 mA, and the WPE are 30.4% and 33.8%, respectively. As for the stacked structure of five unit LEDs, it can be operated at a relatively low current of 13.7 mA, which has a WPE of 37.4%. Consequently, if the tunnel-junction LED is stacked by more unit LEDs, the input power required to obtain a specific output power can be markedly reduced because of the relatively high WPE, as shown in Figure 16.10.

16.4.3 Green Tunnel-Junction LED

The concept of stacking identical LED structures with tunnel junctions is also investigated for green InGaN LEDs in this chapter to probe its effectiveness. The green single LED structure under study is identical to the blue LED except that the indium composition of InGaN wells is increased to 29%. Moreover, considering the enlarged lattice mismatch between the wells and barriers and the limited solubility of indium in GaN, the SRH recombination lifetime is reduced to 20 ns in the simulation. Figure 16.11 shows the energy band diagram, carrier concentrations, and recombination rates at 100 mA and WPE of the green single LED. The electrical and optical characteristics of the green single LED are quite similar to those of the blue single LED shown in Figure 16.5. However, since the indium composition of the InGaN wells in the green single LED is higher than that in the blue single LED, deeper QWs are observed in the band profile. Due to the deeper wells, the feature of nonuniform carrier distribution becomes much more severe, in which only the QW closest to the p-side contributes to interband transitions, as shown in Figure 16.11b and c. Under these circumstances, the droop of WPE in the green LED is more severe than that in the blue LED (48% versus 42%).

Since the holes almost accumulate in the last QW closest to the p-side, similar WPE is obtained no matter how many QWs are utilized in the green single LED under study, as shown in Figure 16.12a. Moreover, it was reported that the defect-related SRH recombination is an important interband transition loss that slashes the low current efficiency (Schubert et al., 2007). The WPEs of the green single LED structures with various SRH lifetimes are plotted in Figure 16.12b. The peak efficiency increases with the increase of SRH lifetime, which is similar to the results reported in the published literature (Schubert et al., 2007). The WPE droop deteriorates when the SRH lifetime increases, due mainly to the marked variation of peak efficiency. Since the high current efficiency remains almost unchanged, the high-power applications cannot benefit much from the improved film quality in the single LED. However, it is not the case for the tunnel-junction LED because the design concept for tunnel-junction LEDs is to operate the LEDs at low current

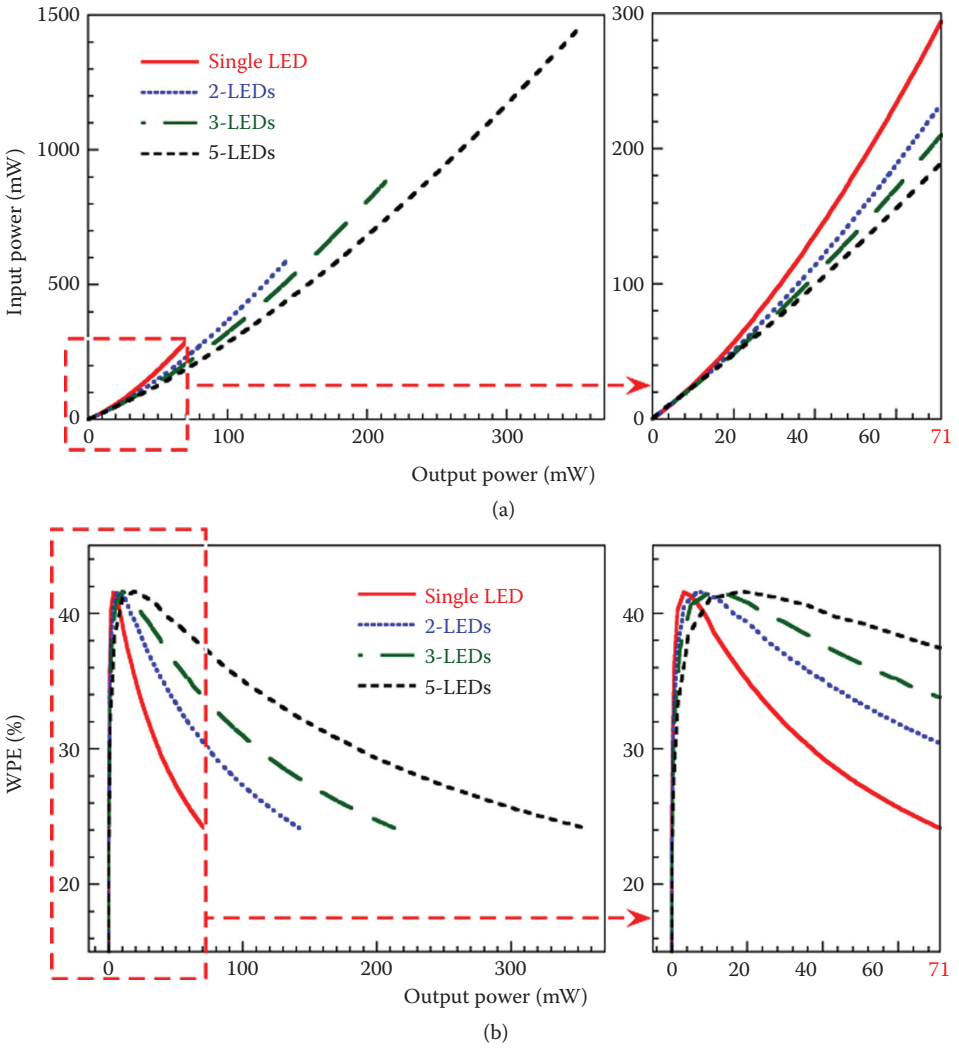

FIGURE 16.10 Required input power and corresponding wall-plug efficiency (WPE) of the blue single light-emitting diode (LED) and tunnel-junction LEDs as a function of the light output power.

density. Figure 16.13 shows the WPE of the green single LED and tunnel-junction LEDs when the SRH lifetimes are 20 and 50 ns. It is observed that the WPE of the tunnel-junction LED with 5 unit LEDs in the stacked structure is almost twice that of the single LED, especially when the SRH lifetime is high. The design strategy of tunnel-junction LED is thus applicable to the green LEDs. Since the efficiency of green LEDs is difficult to upgrade due to the nature of deeper InGaN wells, the idea of tunnel-junction LED may be of great benefit in improving the physical and optical performance.

16.5 Conclusion

In this chapter, several design approaches of the GaN-based polarization-assisted tunnel junctions are explored. With appropriate design via polarization engineering, low-resistivity and excellent tunneling characteristics can be achieved in the structures with practical indium composition and doping density. Note that the tunneling currents of the tunnel-junction structures obtained from experiments typically

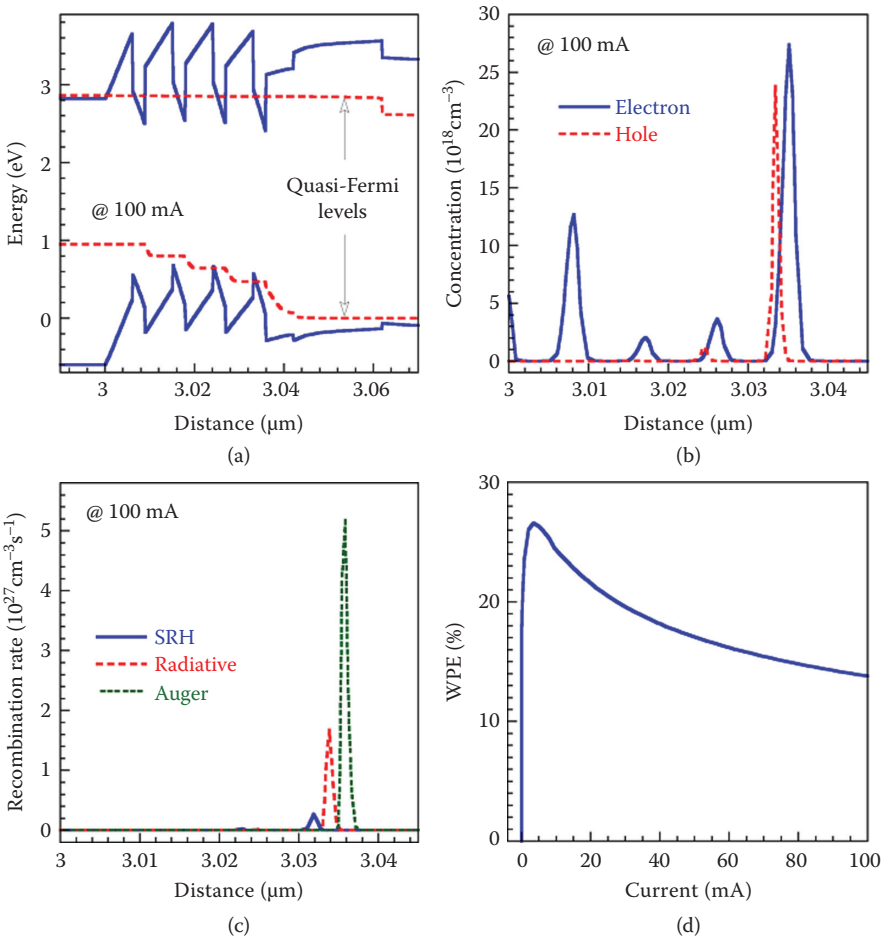

FIGURE 16.11 (a) Energy band diagram, (b) carrier concentrations, (c) recombination rates at 100 mA, and (d) wall-plug efficiency (WPE) of the green single light-emitting diode (LED).

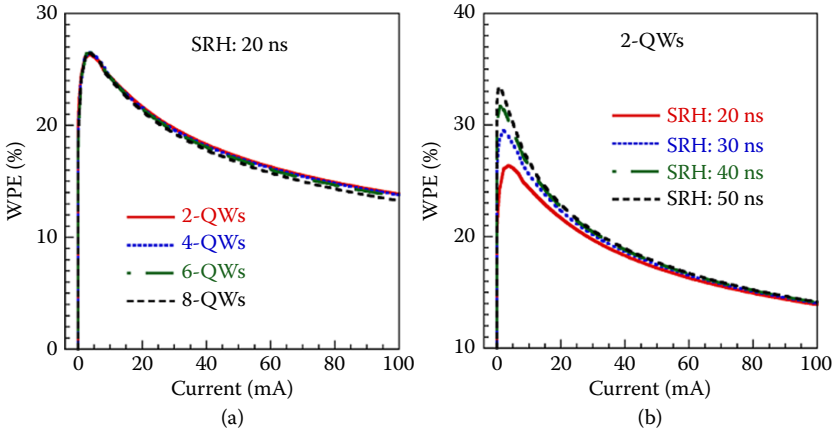

FIGURE 16.12 Wall-plug efficiency (WPE) of the green single light-emitting diode (LED) structures with (a) various pairs of QWs and (b) various Shockley–Read–Hall (SRH) lifetimes.

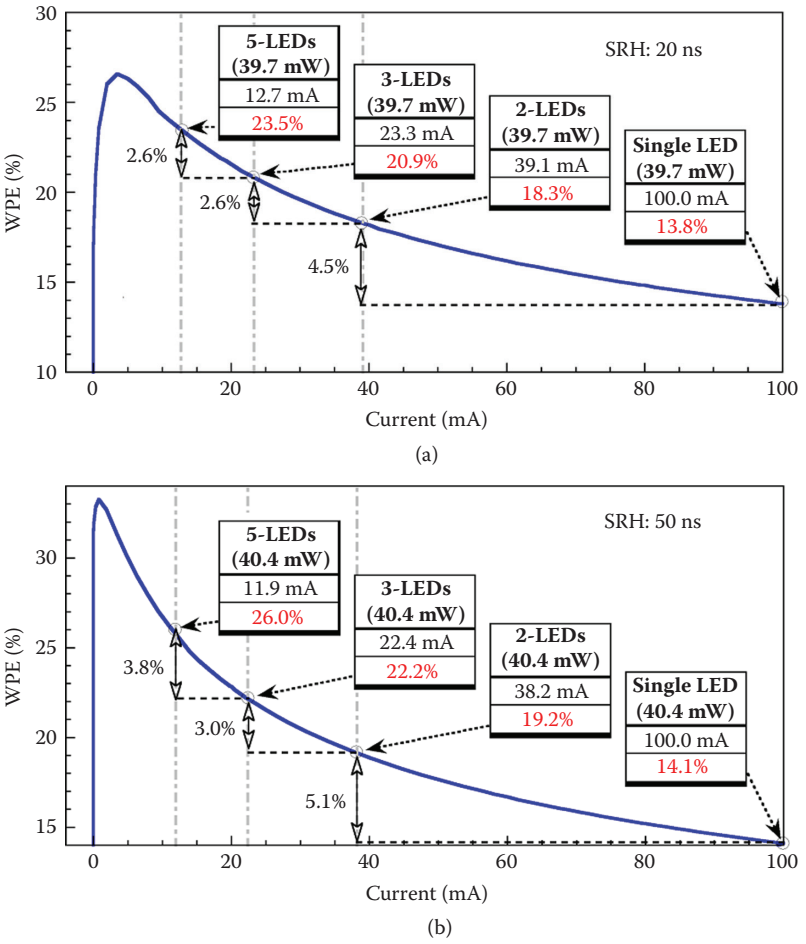

FIGURE 16.13 Wall-plug efficiency (WPE) of the green single light-emitting diode (LED) and tunnel-junction LEDs when the Shockley–Read–Hall (SRH) lifetimes are (a) 20 ns and (b) 50 ns.

exceed the values obtained from theoretical calculation. The additional tunneling current may be owing to the defect and impurity-related mid-gap electronic states (Schubert, 2010b), which are not considered in the present study. Hence, the utilization of tunnel junction in nitride-based applications is quite feasible, which enables the development of high-efficiency GaN-based devices, such as the multijunction solar cells, multijunction optoelectronic devices, and multijunction electronic devices. Specifically, in the present chapter, the application of low-resistivity tunnel junction in both the blue and green tunnel-junction LEDs is explored. Through the stacking of LEDs, it is shown that marked improvement in WPE can be achieved by operating the stacked LED devices at low current density. Furthermore, the tunnel-junction LED structure allows further design modification for better electrical and optical characteristics, such as the removal of the EBL and optimization in number of QWs. It is our belief that the method introduced in this chapter may be of great help in obtaining cost-effective and high-efficiency LED devices for solid-state lighting.

Acknowledgments

This work is supported by the Ministry of Science and Technology of Taiwan under grants NSC-99-2119-M-018-002-MY3 and NSC-102-2112-M-018-004-MY3.

References

Akyol F, Krishnamoorthy S, and Rajan S (2013) Tunneling-based carrier regeneration in cascaded GaN light emitting diodes to overcome efficiency droop. *Applied Physics Letters*. 103:081107-1–081107-4.

Cherns D, Henley SJ, and Ponce FA (2001) Edge and screw dislocations as nonradiative centers in InGaN/GaN quantum well luminescence. *Applied Physics Letters*. 78:2691–2693.

Chichibu SF et al. (2006) Origin of defect-insensitive emission probability in In-containing (Al,In,Ga)N alloy semiconductors. *Nature Materials*. 5:810–816.

Chuang SL and Chang CS (1996) k·p method for strained wurtzite semiconductors. *Physical Review B*. 54:2491–2504.

David A et al. (2008) Carrier distribution in (0001) InGaN/GaN multiple quantum well light-emitting diodes. *Applied Physics Letters*. 92: 053502-1–053502-3.

Delaney KT, Rinke P, and Van de Walle CG (2009) Auger recombination rates in nitrides from first principles. *Applied Physics Letters*. 94:191109-1–191109-3.

Duke CB (1969) *Tunneling in Solids*. New York, NY: Academic Press.

Fiorentini V, Bernardini F, and Ambacher O (2002) Evidence for nonlinear macroscopic polarization in III–V nitride alloy heterostructures. *Applied Physics Letters*. 80:1204–1206.

Gorczyca I, Suski T, Christensen NE, and Svane A (2011a) Band gap bowing in quaternary nitride semiconducting alloys. *Applied Physics Letters*. 98:241905-1–241905-3.

Gorczyca I, Suski T, Christensen NE, and Svane A (2011b) Size effects in band gap bowing in nitride semiconducting alloys. *Physical Review B*. 83:153301-1–153301-4.

Grundmann MJ and Mishra UK (2007) Multi-color light emitting diode using polarization-induced tunnel junctions. *Physica Status Solidi C*. 4:2830–2833.

Kane EO (1960) Zener tunneling in semiconductors. *Journal of Physics and Chemistry of Solids*. 12:181–188.

Khan N, Nepal N, Sedhain A, Lin JY, and Jiang HX (2007) Mg acceptor level in InN epilayers probed by photoluminescence. *Applied Physics Letters*. 91:012101-1–012101-3.

Kim MH et al. (2007) Origin of efficiency droop in GaN-based light-emitting diodes. *Applied Physics Letters*. 91:183507-1–183507-3.

Krishnamoorthy S, Akyol F, Park PS, and Rajan S (2013) Low resistance GaN/InGaN/GaN tunnel junctions. *Applied Physics Letters*. 102:113503-1–113503-5.

Krishnamoorthy S, Nath DN, Akyol F, Park PS, Esposto M, and Rajan S (2010) Polarization-engineered GaN/InGaN/GaN tunnel diodes. *Applied Physics Letters*. 97:203502-1–203502-3.

Krishnamoorthy S, Park PS, and Rajan S (2011) Demonstration of forward inter-band tunneling in GaN by polarization engineering. *Applied Physics Letters*. 99:233504-1–233504-3.

Kumakura K, Makimoto T, and Kobayashi N (2003) Mg-acceptor activation mechanism and transport characteristics in p-type InGaN grown by metal organic vapor phase epitaxy. *Journal of Applied Physics*. 93:3370–3375.

Kuo YK, Chang JY, Tsai MC, and Yen SH (2009) Advantages of blue InGaN multiple-quantum well light-emitting diodes with InGaN barriers. *Applied Physics Letters*. 95:011116-1–011116-3.

Laubsch A et al. (2009) On the origin of IQE-"droop" in InGaN LEDs. *Physica Status Solidi C*. 6:S913–S916.

Lin GB, Meyaard D, Cho J, Schubert EF, Shim H, and Sone C (2012) Analytic model for the efficiency droop in semiconductors with asymmetric carrier-transport properties based on drift-induced reduction of injection efficiency. *Applied Physics Letters*. 100:161106-1–161106-4.

Moll JL (1964) *Physics of Semiconductors*. New York, NY: McGraw-Hill.

Monemar B and Sernelius BE (2007) Defect related issues in the "current roll-off" in InGaN based light emitting diodes. *Applied Physics Letters*. 91:181103-1–181103-3.

Nam KB, Nakarmi ML, Li J, Lin JY, and Jiang HX (2003) Mg acceptor level in AlN probed by deep ultraviolet photoluminescence. *Applied Physics Letters*. 83:878–880.

Nepal N, Li J, Nakarmi ML, Lin JY, and Jiang HX (2005) Temperature and compositional dependence of the energy band gap of AlGaN alloys. *Applied Physics Letters*. 87:242104-1–242104-3.

Ozden I, Makarona E, Nurmikko AV, Takeuchi T, and Krames M (2001) A dual-wavelength indium gallium nitride quantum well light emitting diode. *Applied Physics Letters*. 79:2532–2534.

Piprek J, Li ZM, Farrell R, DenBaars SP, and Nakamura S (2007) Electronic properties of InGaN/GaN vertically-cavity lasers, In: *Nitride Semiconductor Devices: Principles and Simulation*, Piprek J (ed.), pp. 423–445. Weinheim: John Wiley & Sons.

Piprek J (2014) Origin of InGaN/GaN light-emitting diode efficiency improvements using tunnel-junction-cascaded active regions. *Applied Physics Letters*. 104:051118-1–051118-4.

Piprek J and Li SZM (2013) Origin of InGaN light-emitting diode efficiency improvements using chirped AlGaN multiquantum barriers. *Applied Physics Letters*. 102:023510-1–023510-4.

Rozhansky IV and Zakheim DA (2006) Analysis of the causes of the decrease in the electroluminescence efficiency of AlGaInN light-emitting-diode heterostructures at high pumping density. *Semiconductors*. 40:839–845.

Schubert MF (2010a) Interband tunnel junctions for wurtzite III-nitride semiconductors based on heterointerface polarization charges. *Physical Review B*. 81:035303-1–035303-8.

Schubert MF (2010b) Polarization-charge tunnel junctions for ultraviolet light emitters without p-type contact. *Applied Physics Letters*. 96:031102-1–031102-3.

Schubert MF et al. (2007) Effect of dislocation density on efficiency droop in GaInN/GaN light-emitting diodes. *Applied Physics Letters*. 91:231114-1–231114-3.

Shen YC et al (2007) Auger recombination in InGaN measured by photoluminescence. *Applied Physics Letters*. 91:141101-1–141101-3.

Shim JI et al. (2012) Efficiency droop in AlGaInP and GaInN light-emitting diodes. *Applied Physics Letters*. 100:111106-1–111106-4.

Simon J et al. (2009) Polarization-induced Zener tunnel junctions in wide-band-gap heterostructures. *Physical Review Letters*. 103:026801-1–026801-4.

Tiwari S (1992) *Compound Semiconductor Device Physics*. London: Academic Press.

Tsai MC, Leung B, Hsu TC, and Kuo YK (2013) Low resistivity GaN-based polarization-induced tunnel junctions. *Journal of Lightwave Technology*. 31:3575–3581.

Tsai MC, Leung B, Hsu TC, and Kuo YK (2014) Tandem structure for efficiency improvement in GaN based light-emitting diodes. *Journal of Lightwave Technology*. 32:1801–1806.

Vurgaftman I and Meyer JR (2003) Band parameters for nitrogen-containing semiconductors. *Journal of Applied Physics*. 94:3675–3696.

Vurgaftman I and Meyer JR (2007) Electron band structure parameters. In: *Nitride Semiconductor Devices: Principles and Simulation*, Piprek J (ed.), pp. 13–48. Weinheim: John Wiley & Sons.

Vurgaftman I, Meyer JR, and Ram-Mohan LR (2001) Band parameters for III–V compound semiconductors and their alloys. *Journal of Applied Physics*. 89:5815–5875.

Wu J (2009) When group-III nitrides go infrared: New properties and perspectives, *Journal of Applied Physics*. 106:011101-1–011101-27.

Wu J et al. (2003) Temperature dependence of the fundamental band gap of InN. *Journal of Applied Physics*. 94:4457–4460.

Yoshida H, Kuwabara M, Yamashita Y, Uchiyama K, and Kan H (2010) Radiative and nonradiative recombination in an ultraviolet GaN/AlGaN multiple-quantum-well laser diode. *Applied Physics Letters*. 96:211122-1–211122-3.

17

Quantum Disk Nanowire Light-Emitting Diodes

Fabio Sacconi

17.1 Introduction

Due to their properties, nanowire (NW) based light-emitting diode (LED) structures offer very interesting advantages, such as defect-free material, due to small footprint on the substrate, which allows higher LED efficiency, and three-dimensional (3D) geometry features, such as lateral strain relaxation, which can also be exploited to increase efficiency (Ristic et al., 2005). Group III nitride semiconductors have attracted much attention for quite a long time especially for their light-emitting device applications. A new approach for reaching exceptionally high efficiencies of LEDs for the whole visible spectrum is therefore based on nanostructured InGaN emitters, where nanorods, which have been shown to be defect-free, serve as active light-emitting structures. Light-emitting nanorod arrays based on InGaN as a material for the active region have been investigated (Kim et al., 2004); in principle they could cover the whole visible spectrum from blue to red and be incorporated into a single device for phosphor-free white light emission. By controlling indium content in the InGaN quantum disk (QD) embedded in the NW, nanorod emitters are expected to exhibit efficient green, yellow, and red emission, allowing coverage of the full visible spectrum with GaN-based LEDs. This could overcome a major drawback of today's red–green–blue (RGB) light sources, that is, the lack of an efficient green emitter due to the efficiency gap existing between the blue InGaN and the red InGaAlP materials system (Piprek, 2010).

In this chapter, we discuss several simulation approaches for the numerical study of the transport and optoelectronic properties of GaN NW diode structures with an embedded InGaN QD. Nanostructure-based emerging electronic and optoelectronic devices, such as LEDs or photodetectors, need a fully

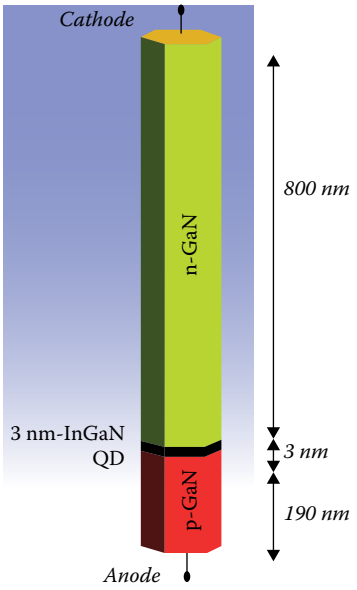

FIGURE 17.1 Nanowire (NW) quantum dot (QD) light-emitting diode (LED) model structure.

quantum mechanical or atomistic approaches to obtain a correct description of structural, electrical, and optical properties. Usually, the active region of a device, which needs such an advanced investigation, is small compared to the overall simulation domain. The computational cost of more accurate quantum mechanical models, however, makes their application to the whole device unfeasible. It is, therefore, necessary to implement a multiscale simulation approach, which couples the semiclassical models describing the bulk regions of the device to the quantum mechanical or atomistic models, acting only on the nanostructured regions of the device.

We see in the rest of the chapter how an integrated multiscale and multiphysics simulation environment (see Auf der Maur et al., 2011, 2013; Auf der Maur, 2015) may be capable of coupling different models on different scales, ranging from macroscopical to atomistic representations.

In the following, we first describe briefly the different physical models that need to be applied for an adequate description of these devices. Then we show the main results obtained from these numerical methods in the simulation of QD-based LEDs. If not stated otherwise, in the rest of the text we refer to the model NW LED structure with a single QD described in Figure 17.1. This model structure is a GaN NW usually around 1 μm high with a hexagonal base and a radius between 25 nm and several hundreds of nanometers. The embedded QD, which constitutes the LED active region, is a 3-nm-thick disk made of InGaN alloy, with In composition varying between 10% and 40%.

17.2 Overview of Numerical Methods

17.2.1 Strain

Strain has a critical influence on the behavior of heterostructures due to its effect on the band energies and the strain-induced piezoelectric polarization. The latter is particularly important in nitride-based devices. A straightforward approach for the calculation of strain in lattice mismatched heterostructures is the one based on the linear elasticity theory of solids. Usually, in this approach it is assumed that pseudomorphic interfaces are present between different materials (Povolotskyi and Di Carlo, 2006). One also assumes small deformations, such that the strain is a linear function of deformation and that Hooke's law, which linearly relates stress to strain, can be used. The strain and deformation field are found by minimizing the elastic

energy of the system. As a result, one obtains the strain tensor in any point of the structure, the shape deformation, and the piezoelectric polarization, which to the first order depends linearly on strain. Self-consistent electromechanical simulations can be carried out by including the converse piezoelectric effect.

This continuum elastostatic model described here has several limitations.

In particular, it does not contain any information on internal strain, which can be very important, particularly in wurtzite crystals. A more accurate description of strain can be obtained by means of the valence force field (VFF) method, which offers the best compromise between accuracy and computational effort in the study of semiconductor nanostructures. One of the implementations is based on the Keating model (Keating, 1966), for diamond structures, as generalized by Camacho and Niquet (2010) to treat nonideal wurtzite structures. As VFF is computationally more intensive than the continuous media model, a possibly convenient implementation is the one that allows coupling of the two models in a multiscale simulation (see Auf der Maur et al., 2011, 2013).

17.2.2 Transport

Usually transport of electrons and holes in LED structures is treated in a semiclassical picture based on the drift–diffusion model (Sze, 1981). The particle fluxes are written in terms of the electrochemical potentials. The carrier statistics are given by Boltzmann or Fermi–Dirac statistics, assuming as usual local equilibrium. The conduction and valence band edges and effective masses may be obtained from bulk **k**·**p** calculations; in this way, the local corrections due to strain can be easily included.

To account for quantum effects, a correction to this purely classical approach is needed. Several methods can be applied, such as the density gradient quantum correction model (Ancona, 1990) up to the more computationally demanding nonequilibrium Green's function (NEGF) (Datta, 2000).

A self-consistent approach in many cases may be a good compromise. Some applications of self-consistent techniques will be described in the following.

17.2.3 Quantum Models

Quantum mechanical models are used for the calculation of eigenstates of confined particles in nanostructures, either based on the envelope function approximation (EFA) or on atomistic approaches. In the former, the Hamiltonian of the system is constructed in the framework of single-band and multiband **k**·**p** theory (Chuang, 1995; Chuang and Chang, 1996). The single particle wave functions are expanded in terms of bulk Bloch functions (usually taken at the zone center $\boldsymbol{k} = 0$ for direct band gap materials) of the constituent materials, leading to a system of equations for the envelope functions. The atomistic approaches are typically based on the empirical tight binding (ETB) method (Di Carlo, 2003). In this method, the electronic states are written as a linear combination of atomic orbitals (LCAO). In this case, an atomistic structure describing the heterostructure studied has to be generated according to the macroscopic device description and crystallographic orientation. The solution of the eigenvalue problems resulting from the EFA and ETB models provides the energy spectrum, the particle densities, and the probabilities of optical transitions. The particle densities are calculated by populating the electron and hole states according to the expectation value of the corresponding electrochemical potential. The particle densities may then be fed back to the Poisson/drift–diffusion model for a self-consistent Schrödinger–Poisson/drift–diffusion calculation.

17.3 Strain Maps

It is interesting to study the behavior of the strain tensor in an GaN NW embedding an InGaN QD. As previously discussed, strain induced by the lattice mismatch in a heterostructure can be calculated, for example, by means of an elasticity model. In the case of the QD NW LED, strain obtained from this numerical

model is due to the lattice mismatch between the material of the NW or nanocolumn and the material that constitutes the QD. It has been shown (Hugues et al., 2013) that a GaN NW grown bottom-up becomes strain free after around 100 nm. Thus, we can expect that a typical GaN NW, with a height of 1 μm or more, is completely relaxed beyond the region surrounding the QD. This justifies the choice of the NW material (GaN in this case) as the substrate reference material for the strain calculations. This means that the components of the strain tensor as resulting from the minimization of the elastic energy in the system composed by the whole heterostructure will tend to vanish in the substrate material.

If we look at the calculated in-plane strain tensor component ε_{xx} in an NW, we clearly see a strong compressive strain inside the InGaN QD (Sacconi et al., 2010, 2012). In fact, the InN lattice constant is higher than that of the reference GaN lattice material; thus, a negative in-plane strain is determined in the pseudomorphic InGaN/GaN heterostructure. Besides, due to 3D effects, a tensile strain is present in the GaN barrier, close to the interface with the quantum well, while it vanishes in the bulk of the GaN NW. A peculiar feature of the NW structure is the strong reduction of the in-plane strain close to the column surface. The strain relaxation due to the surface boundary is visible in Figure 17.2 where ε_{xx} is shown, along the z-axis (growth direction), for two different radial positions: in the center of the QD ($r = 0$) and close to the surface boundary ($r = 25$). Strain is clearly higher in the center of the column, while it tends to vanish at the surface boundary. Piezoelectric polarization is obtained from the strain field and exhibits a similar behavior: polarization is higher at the center of the QD and tends to decrease toward the lateral surface of the NW. Piezoelectric polarization (together with spontaneous or pyroelectric polarization) has a direct effect on the band profiles and therefore on the quantum and optical properties.

It is interesting to analyze the dependence of the ε_{xx} strain tensor component in the QD on the column size. It is found (Sacconi et al., 2012) that, by increasing the NW size, the (tensile) strain in the barrier regions outside the QD tends to vanish, while the compressive strain slightly increases inside the QD. Moreover, it can clearly be seen that the effect of strain relaxation is more evident for columns with a radius lower than 50 nm and it tends to vanish for very large columns.

The dependence on QD thickness of the ε_{xx} strain tensor component in the NW is shown in Figure 17.3, for an $In_{0.1}GaN$ QD in an NW with radius 25 nm. The value of ε_{xx} is shown along the z-axis, close to the surface boundary. It can be noted that by increasing QD thickness from 2 to 5 nm, the compressive strain slightly decreases inside the QD, while the tensile strain in the GaN surrounding the QD slightly increases. On the other hand, by comparing strain behavior in the center and close to the lateral surface (Figure 17.3)

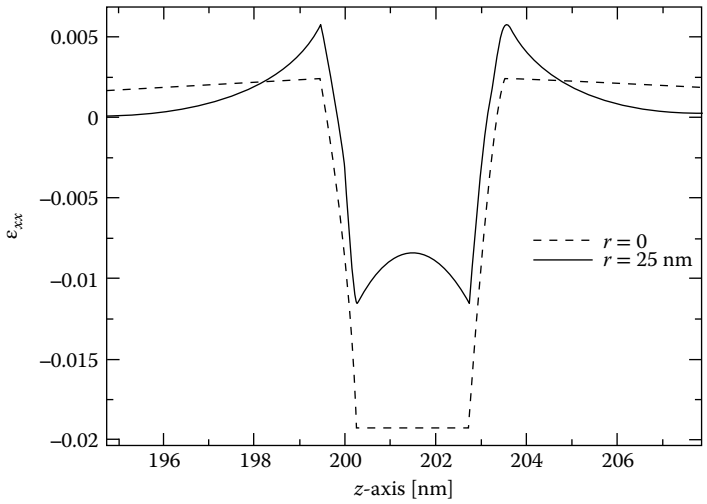

FIGURE 17.2 ε_{xx} strain tensor component along the z-axis, respectively, at the center of the column (dashed line) and close to the column surface (solid line), for an $In_{0.2}GaN$ quantum dot (QD) with $r = 25$ nm.

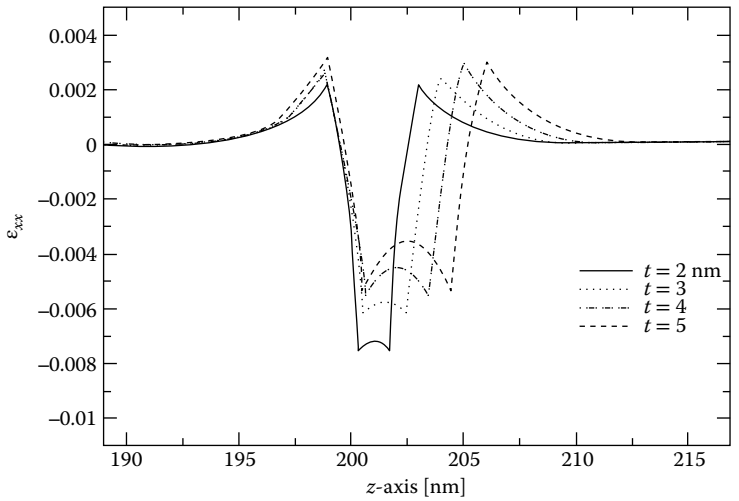

FIGURE 17.3 ε_{xx} strain tensor component along the z-axis, close to the lateral surface, for an $In_{0.1}GaN$ quantum dot (QD) thickness from 2 to 5 nm.

of the NW, the strain relaxation is clearly noticed, which is more evident for thicker QDs (50% reduction for 5-nm-thick QD). We see in the following how these peculiar strain features of the NW structure can be put in relation with the emission properties of the QD LED.

17.4 Transport Properties

Transport in NW QD LED may be studied in a first approach with a classical model. In this way, the IV characteristic of the forward-biased p-i-n LED diode is calculated by imposing ohmic boundary conditions to the contact regions at the two sides of the NW (anode and cathode) and by solving the Poisson/drift–diffusion model, taking care to include the effects of polarization (both spontaneous and piezoelectric), by means of an appropriate polarization vector P in the current equation. Figure 17.4 shows the conduction and valence band profiles and the electron and hole densities along the z-axis in the center of the NW, as obtained from these classical calculations. The dependence of the calculated IV characteristics on the In molar fraction in the InGaN QD is shown in Figure 17.5 for a column with radius $r = 25$ nm (Sacconi et al., 2010). As expected, the threshold voltage decreases from 3.2 to around 2.2 V with an increase of In concentration from $x = 0.1$ to $x = 0.4$, since the energy gap of the alloy material in the QD is decreasing. In addition, from the IV characteristics for a 20% In composition in the QD we obtain an increase of the output current by about two orders of magnitude for increasing values of the column size, from 25 to 200 nm. In the following, we see how a self-consistent calculation can be used to couple these classical results with a quantum model for the QD.

17.5 Quantum Calculations: EFA Models

For the calculations of the eigenvalues of a confined system, an approach based on EFA is usually a good trade-off between computational load and precision of the results. A solution based on the **k·p** model allows us to take into account the electronic structure details close to the valley minima, which is considered a reasonable approximation, provided that particle energies are not too high, as could be the case for hot electrons. The most complete model, the so-called 8×8 **k·p** model, takes into account the three valence bands: heavy-hole (HH), light-hole (LH), and crystal-field split-off (CH), together with the first conduction band (CB).

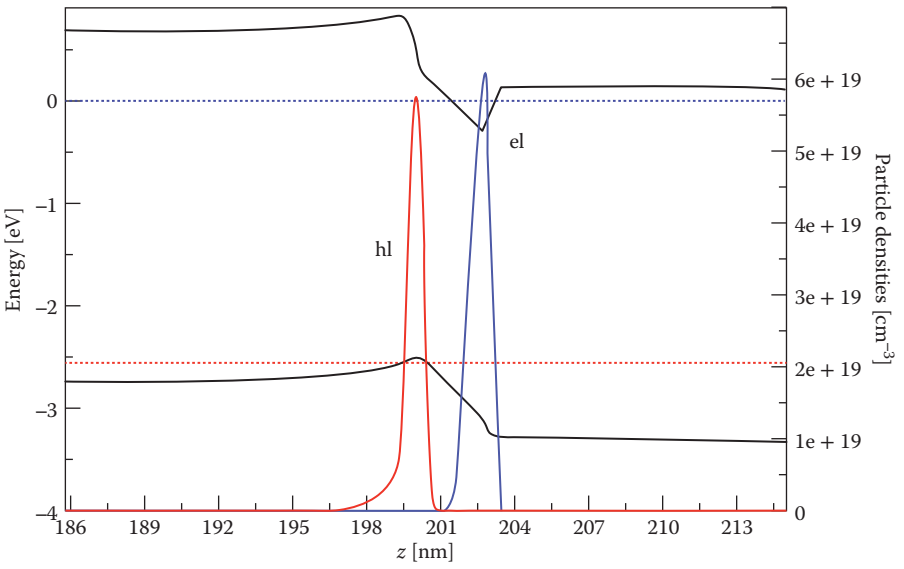

FIGURE 17.4 Band profiles and electron and hole densities from classical calculations.

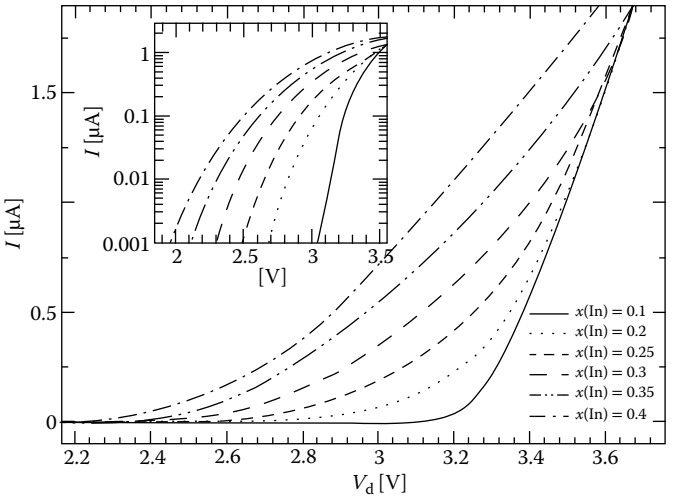

FIGURE 17.5 IV characteristics of the p-i-n diode with radius $r = 25$ nm for several values of In concentration in the InGaN quantum dot (QD).

However, when the energy gap of the material is wide enough, as is the case of GaN and its alloys with AlN and InN (for low InN contents), the interaction between conduction and valence bands is so weak that it can be neglected. Thus, one can apply separately a single band (effective mass) model for CB and a 6×6 $\mathbf{k}\cdot\mathbf{p}$ model for the three VBs, thus reducing the computational time needed without significant loss of accuracy in calculations.

To perform quantum calculations with the EFA model for an LED device, usually the drift–diffusion model is first solved, in order to apply a bias ramp to the diode until the desired operation conditions are obtained. Then a Schrödinger solver is applied for the solution of the eigenvalue problem, restricted

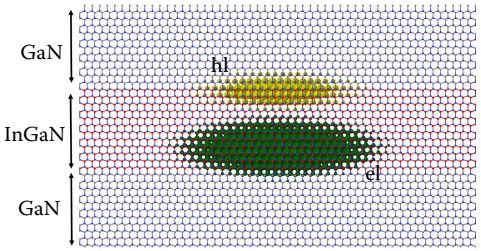

FIGURE 17.6 Conduction and valence band tight-binding states in a 14-nm-wide $In_{0.1}GaN$ quantum dot (QD).

to the active region of the device, comprising the QD and part of the GaN barrier regions. From the calculated conduction and valence ground states, the transition energies between the first electron and hole states can be obtained, for several values of In content (Sacconi et al., 2012). It is found that the transition energy decreases with higher In content mainly due to a lower energy gap in the QD. The transition energy dependence on geometrical and material parameters will be analyzed in more detail later in this chapter.

If we express the conduction and valence band ground states by means of the isosurface of probability density for each quantum state, it is evident that in both cases the quantum states are confined in the QD region (see also Figure 17.6). However, a spatial separation is clearly present between the electron and hole states due to the polarization fields that cause the so-called quantum confined stark effect (QCSE), contrary to what usually happens with GaAs-based structures.

In the following, we see how these results can be corrected by a self-consistent solution of Schrödinger and Poisson/drift–diffusion equations.

17.6 Quantum Calculations: Empirical Tight-Binding Approach

As we have seen before, a more accurate approach for the study of quantum properties of nanostructures is based on atomistic methods. It has been shown recently that this approach is able to provide a description of a realistic InGaN alloy features, contributing to an explanation of the decreasing emission efficiency on nitride-based quantum well (QW) LEDs (green droop) (Auf der Maur et al., 2016). We therefore present the results obtained when ETB calculations are performed to find confined electronic states in an InGaN/GaN QD NW.

In a first example, the simulation model is applied to an GaN NW with an $In_{0.1}GaN$ QD about 14 nm wide and with a thickness of 3 nm. First, an atomistic structure comprising around 77,000 atoms is generated, based on the finite element (FE) grid used for the discretization of continuous partial differential equation (PDE) models (Auf der Maur et al., 2011, 2013). This association with FE allows projection to the atomic positions the displacements obtained from strain calculation in the heterostructure with the PDE-based elasticity model.

Then eigenstates and eigenfunctions of the system are calculated on the strained atomistic system with an ETB model based on an $sp^3d^5s^*$ parameterization (Jancu et al., 2002).

The calculation is restricted to the atomistic structure, which describes the materials included in the LED active region consisting of the InGaN QD (see Figure 17.6) and an appropriate portion of the GaN NW surrounding the QD. As in the EFA $\mathbf{k}\cdot\mathbf{p}$ case, here a Poisson/drift–diffusion calculation also is performed earlier on the NW model (that is its FE grid). Then the potential profile obtained at a given bias, which includes polarization fields, is projected on the atomistic representation to be used for ETB calculations. This is generally accomplished with a shifting of the on-site energies of the ETB Hamiltonian matrix (Auf der Maur et al., 2011, 2013).

In Figure 17.6, we show the first hole and electron quantum states, calculated at an applied bias of $V_d = 3.5$ V. As in the previous EFA results, it is clearly seen that the polarization-induced electric field leads to a spatial separation of electron and hole states.

We see in the following how a multiscale simulation can be performed where both the atomistic (ETB) and the continuous (EFA) models are applied together for a better description of the QD electronic structure.

As a second example, we see an ETB calculation based on a random alloy approach.

Given a ternary alloy in the form $A_xB_{1-x}C$, such as $In_xGa_{1-x}N$, there are two ways to model it. The first is virtual crystal approximation (VCA), where we consider the alloy as a fictitious material whose properties are a weighted average of the properties of the binary components AC and BC (e.g. InN and GaN), in function of the molar fraction x, according to Vegard's law. The second is the random alloy approach, which consists of studying an ensemble of stochastic realizations of structures where we substitute the anion according to a probability given by the concentration of anion A in the alloy.

Let us consider then a QD NW with a diameter of 10.7 nm with a 2-nm-thick QD. The total number of atoms in the atomistic structure is around 51,000.

A set of atomistic structures of the GaN/InGaN/GaN QD region, including the InGaN QD and two 2-nm-thick GaN barrier regions, are created on the GaN substrate lattice. Random alloy is generated in the InGaN region by including In atoms with a probability given by the chosen In concentration. Then, as an initial guess, macroscopic strain is calculated and the displacements obtained are projected on the atom positions. VFF relaxation is applied on this structure until the atoms' equilibrium positions are found. Finally, ETB calculation is performed on the relaxed structure, taking into account the appropriate scaling of ETB parameters according to the bond length obtained by the relaxation procedure.

Electron and hole ground states have been calculated. Figure 17.7 shows the transition energy E_t for increasing In concentration in the InGaN QD. The average value of E_t for the set of calculations is plotted, together with the range between the minimum and maximum values. Random alloy results are compared with those obtained with ETB for the same model structure, but with a VCA approach and just a macroscopic strain. It can be seen that there is a significant discrepancy between the two models. In fact, they differ by about 50 meV for $x(\text{In}) = 0.1$ and the difference gets larger with increasing In content. These results show that VCA may not be adequate to describe nanostructures based on InGaN alloys due to the

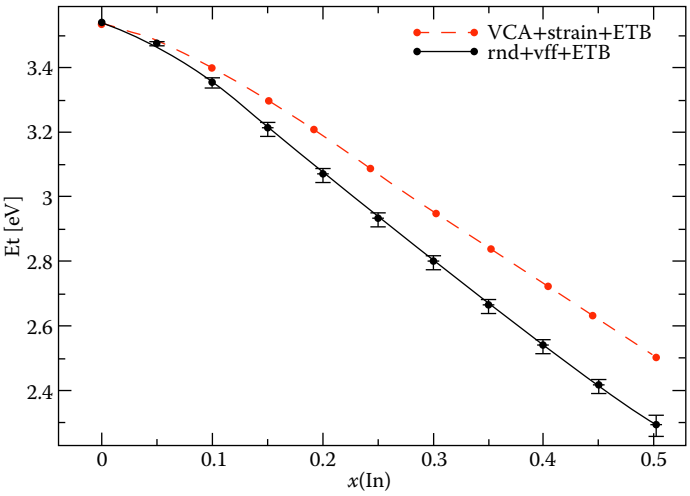

FIGURE 17.7 Dependence of transition energy on In concentration in the quantum dot (QD): Virtual crystal approximation and random alloy model.

great difference in the lattice properties of its component. The random alloy approach, with an atomistic treatment of strain with VFF, may be required to study these devices.

17.7 Self-Consistent Calculations

As discussed previously, a correction to a purely classical description of transport in the NW QD LEDs is needed to account for quantum effects. In this section, we describe a convenient procedure to take into account quantum charge densities of electrons and holes, obtained from an EFA model as described earlier, in the transport calculations of an NW QD LED. The solution to this issue described here consists of a self-consistent coupling of the Schrödinger equation in the EFA framework and the Poisson/drift–diffusion equations. This approach presents the advantage of a lower computational load with respect to more accurate methods for the calculation of quantum transport (e.g., NEGF; see Datta, 2000) while providing all the information about scattering features included in the drift–diffusion model. Thus, even if several approximations must be adopted, such as the assumption of quasiequilibrium of Fermi levels in the quantum regions, the results obtained may be considered reasonably accurate and have been validated by several experimental benchmarks. For more details on the method applied, see, e.g., Auf der Maur et al. (2011).

Let us see now an example of the application of a self-consistent model to an NW QD LED device. It is convenient to perform first purely classical drift–diffusion calculations to ramp the diode voltage up to a desired operating point (e.g., just above its threshold voltage V_{th}). In this way, the LED is biased at a point where the first quantized levels in the QD are populated; this guarantees that the calculated particle quantum density resulting from quantum EFA calculations is not vanishing and thus can be considered comparable with the classical density of the particle resulting from the Fermi level position in the QD. Then, the drift–diffusion model is coupled with a Schrödinger solver for the self-consistent solution of the eigenvalue problem restricted to the active region of the device. The quantum mechanical electron and hole densities are fed back into the Poisson/drift–diffusion equations for a self-consistent Schrödinger–Poisson/drift–diffusion calculation. A self-consistent loop may be implemented by using a simple predictor–corrector scheme that assumes that the quantum density n_q varies with the potentials as the classical density n_{cl} (for more details see Sacconi et al., 2012). As the electron and hole states in the system are calculated applying closed boundary conditions, the quantum densities near the interface between quantum mechanical and classical simulation domains suffer from artificial behavior. To obtain a continuous transition from the purely classical densities far away from the QD to the quantum mechanical densities, one may define an embracing region with an extension of a few nanometers where a linear mixing of classical and quantum density can be applied (Sacconi et al., 2012). Figure 17.8 shows the conduction and valence band profiles along the *z*-axis in the center of the NW, together with the particle densities obtained after the self-consistent cycle.

By comparing the self-consistent electron and hole charge densities of Figure 17.8 with those of Figure 17.4, obtained with purely classical calculations, the effect of quantum correction is evident. The charge density distribution is no longer peaked at the heterointerface between the QD and the GaN barrier material, as determined by the band bending in classical calculations. Instead, it is moved toward the center of the QD, accordingly with the behavior of the wave functions of the particle confined states.

The distribution of the electron quantum densities on a *xy*-plane orthogonal to the growth direction is shown in Figure 17.9. The planar quantum confinement toward the center of the QD is clearly visible for the electrons and it is found even more pronounced for holes, due to their larger effective mass. This effect is even more evident when we compare the classical and self-consistent electron density in the QD region along the *y*-axis on a slice on a *xy*-plane orthogonal to the growth direction (Sacconi et al., 2012). As for the classical results, for both conduction and valence band a band bending forms close to the lateral surface due to the strain relaxation. A lower strain determines a lower band-gap at the surface than in the center of the NW. This in turn determines a higher value of classical electron and hole densities close to the surface.

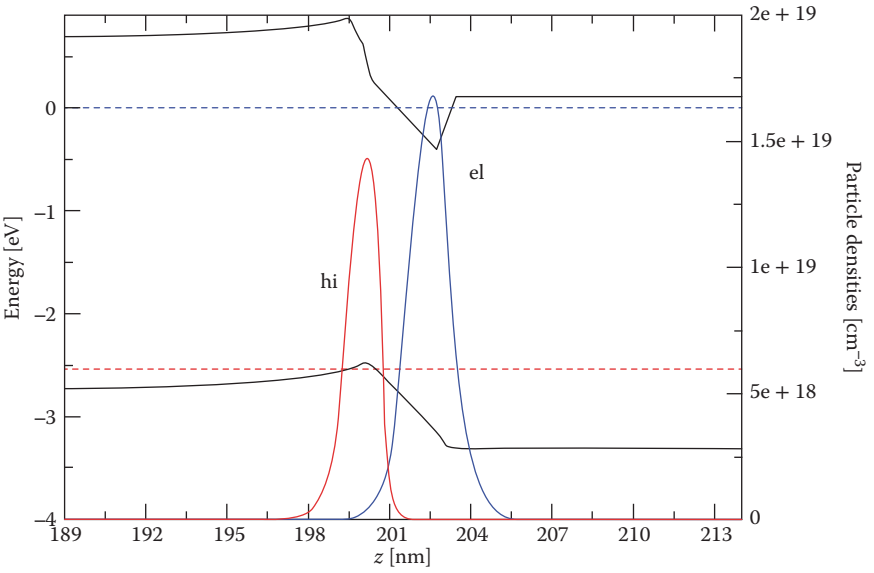

FIGURE 17.8 Self-consistent band profiles and densities along growth direction z in the quantum dot (QD).

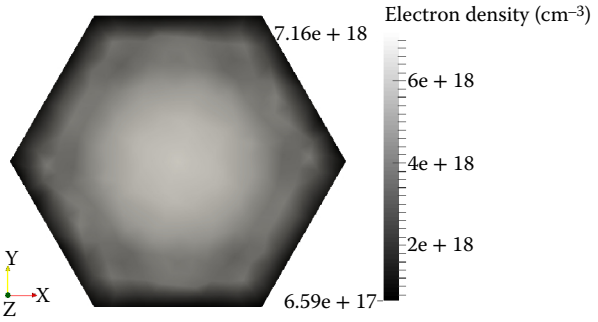

FIGURE 17.9 Electron quantum densities in the xy-plane in the middle of the quantum dot (QD).

On the other hand, the self-consistent densities show the effect of a lateral quantum confinement, so that the carriers are mainly confined to the center region of the QD, even if for electrons this behavior is slightly more complicated.

Due to the behavior of both hole and electron density, we can expect a very different qualitative result when comparing the radiative recombination rate $R = Bnp$ in the active region, that is, the QD, calculated with classical and quantum model. In fact, it is found that from the classical results the emission appears to originate from a region very close to the surface of the column; on the other hand, when quantum mechanical particle densities are used the radiative recombination is mainly concentrated at the center of the QD due to the spatial confinement of the carriers. In Figures 17.10 and 17.11, we show a comparison of the calculated current density through the NW QD, in both cases. Current flow lines in the self-consistent case (Figure 17.11) indicate that the confinement of the quantum states leads to a current crowding at the center of the NW. In the purely classical calculation (Figure 17.10), on the contrary, flow lines focus toward the lateral surfaces, where the classical densities are higher. This result can explain the difference in the IV characteristics obtained with the two approaches (Sacconi et al., 2012). In fact, it turns out that the self-consistent results show a lower threshold voltage and an integrated current value at the contacts of around

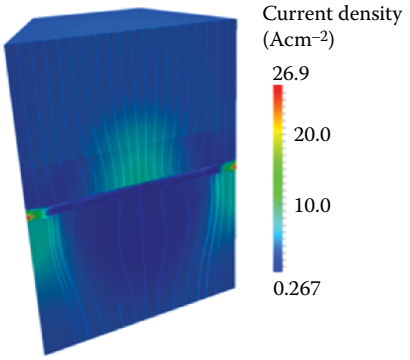

FIGURE 17.10 Classical calculations: Current flow is focused to quantum dot (QD) edges.

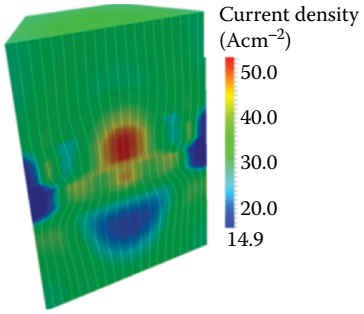

FIGURE 17.11 Self-consistent calculations: Current flow is focused to quantum dot (QD) center.

10 times higher with respect to the classical outcome. This discrepancy can be attributed to the increased recombination rate in the QD active region. In fact, on the one hand, the classical particle densities show higher peak values, but on the other hand, the average value of densities and thus of recombination is higher after quantum correction, which leads to a higher level of output current.

17.7.1 Multiscale Simulation

We now consider an example of a multiscale simulation applied to a NW QD LED (Auf der Maur et al., 2011). Here, a self-consistent classical/quantum calculation is again performed, but this time the information on quantum density of electrons and holes is obtained by an atomistic model. In this case, the device model is a p-i-n LED made of an AlGaN NW with an embedded GaN QD. The simulation is performed as follows.

First, the strain in the NW is calculated based on continuum elasticity theory. This result is then used as a guess for an atomistic VFF method applied for the relaxation of the atomistic structure generated according to the FEM model of the GaN device active region.

Then the electronic transport in the NW LED is simulated based on the self-consistent solution of the Schrödinger/Poisson/drift–diffusion equations using both the continuous media EFA model and the atomistic ETB approach.

To perform the transport calculation, first the drift–diffusion model is applied to get a bias where the diode is nearly in flat band condition. At this bias point the actual multiscale simulation is performed by concurrently solving the Poisson/drift–diffusion and the quantum mechanical models. For the latter both a $\mathbf{k}\cdot\mathbf{p}$ 6×6 model and an ETB model (using an $sp^3d^5s^*$parameterization) are used. Since the hole states

result is very dense for this device structure, a large number of states are needed to obtain a reasonably converged quantum hole density. Moreover, the ETB calculation is computationally much more intensive than an EFA calculation. Therefore, one possible choice is to calculate only the first three electron states using ETB and the higher ones with EFA, while EFA alone is used for the hole states.

This approach is a good compromise between the accuracy of full-band ETB calculations of the particle quantum states and an affordable computational load. In fact, it turns out that the hole state energies calculated by EFA and ETB are very similar and thus we can rely on the less computationally demanding EFA model for them. On the other hand, the higher electron states are less populated and therefore they have less impact on the properties of the system. We limit then our ETB approach to the first more populated electron states. This can be considered a decomposition of the problem in energy space since EFA and ETB are used to calculate energetically different particle states.

The coupling between the different models constituting the multiscale system is obtained by a concurrent solution implemented by applying a self-consistent cycle, similar to that shown earlier.

This coupling can be viewed as a quantum correction of the local density of states (LDOS) in the classical expression for the particle density, where the local densities can be easily calculated based on the envelope functions. In addition to the purely continuous case, here we have also the particle density obtained by ETB, which needs to be projected onto the FE mesh that has been used for the discretization of the continuous media models. This projection may be obtained by using an exponentially decaying function centered on each atomic site.

The results of these self-consistent calculations show also in this case that the confinement of the quantum states leads to a current crowding toward the center of the NW (Auf der Maur et al., 2011). We can conclude that a multiscale approach is in principle able to describe with higher accuracy the confinement effects that are critical for the NW LED devices while keeping the simulation domain for the quantum mechanical models reasonably small.

17.8 Modeling of Surface States

Surface states are expected to be present on the surface of GaN NWs due to defects and dangling bonds (Van de Walle and Segev, 2007).

The effect of surface states can be taken into account in a numerical simulation through the implementation in the drift–diffusion model of an acceptor trap model so that the trap is negatively charged if not occupied by a hole. Following the results obtained in Calarco et al. (2005) we can place the trap at an energy value of 0.6 eV below the conduction band, assuming a trap charge density of 1×10^{12} cm^{-2}. If we apply this model to a GaN NW we obtain a conduction band profile that shows a pinning of the Fermi level at the surface at around 0.6 eV below the conduction band. As is shown in Calarco et al. (2005), when a doping is present the pinning of the Fermi level determines a depletion region that, in this case, covers the whole column size. In this way, we have fitted the surface trap model to the experimental value of the Fermi level pinning in a GaN NW. In the following, we assume that, as a first approximation, the same model holds for the InGaN/GaN QD NW structure.

To investigate the influence of surface states on the electronic properties of the QD NW LED, a Shockley–Read–Hall (SRH) surface recombination model has to be associated to the traps. In this way, an SRH model for nonradiative recombination takes into account the surface traps, with a given density and recombination time, and their contribution to the recombination part of the current (Sacconi et al., 2012). Now, if one performs classical drift–diffusion calculations, taking into account the SRH surface recombination component, results (Sacconi et al., 2012) show that the presence of the surface states induces a large recombination current in the diode, which provides an output current around 60 times larger than in the case where surface states are neglected.

As is reported in Sacconi et al. (2012), the flow lines focus close to the surface, where there is the peak of the classical particle densities, which, thus, largely increases the surface recombination contribution.

On the other hand, when self-consistent quantum densities are calculated, this effect almost vanishes, and the self-consistent current at the contacts is just 20% higher than in the case without surface recombination. In fact, it is found that the self-consistent current flow lines tend again to crowd inside the QD since contribution from surface states is now limited by the quantum confinement effect.

17.9 Sensitivity to Geometrical and Material Parameters

We discuss in the following the dependence of the transition energies and thus of the optical emission from the QD on the NW geometrical and material parameters.

The optical emission spectra from spontaneous recombination may be calculated in the following way (Chuang, 2009):

$$P(\hbar\omega) = \sum_{i,j}^{\infty} n \frac{1}{2\pi^2} \frac{\omega_{ij}^2 e^2}{m^2 c^3} \left| M_{i,j} e \right|^2 f_i(E_i) \left[1 - f_j(E_j)\right] \frac{\Gamma/2}{\left(\hbar\omega_{ij} - \hbar\omega\right)^2 + (\Gamma/2)^2} \mathrm{d}\Omega,$$

where f_i and f_j are the Fermi distributions and $M_{i,j}$ is the optical matrix element between the states i and j. The matrix elements may be obtained with different quantum models, such as ETB and EFA $\mathbf{k}\cdot\mathbf{p}$, and provide the probability of optical transitions between each couple of states. The sum is extended to all the quantum states calculated for conduction and valence band. In this way, the optical transition probability is weighted with the occupation functions for each couple of particles, calculated according to the Fermi distribution at each state energy. The result is the total emitted power for transitions in a given energy range, given the particle population of the considered quantum states.

Usually, we are interested in the emission resulting from the spontaneous recombination of electrons in conduction band and holes in valence band. From the simulation point of view, this means that the device conduction band has to be populated with an optical or electrical pumping. In the following, the latter method has been performed, that is, a ramp bias has been applied until the LED is in conduction regime and the conduction and valence band are populated. At this point, a quantum model, such as $\mathbf{k}\cdot\mathbf{p}$, is applied for the calculation of quantum states. Following that, the optical emission spectra are calculated based on the previous steps.

Figure 17.12 shows the optical spectra obtained for a GaN column with a radius of 25 nm and a 3-nm-thick InGaN QD (Sacconi et al., 2010). The results show a decrease of the peak emission energy with the increase of In concentration in the InGaN QD from 10% to 40%, indicating that an emission wavelength range between 400 and 680 nm can be covered in principle by these devices.

Generally speaking, in NW LEDs, similarly to QWs LED structures, emission energy decreases with higher In content due to a lower energy gap in the QD and a higher QCSE. The variation of the QCSE is caused by the increase of polarization fields with lattice mismatch between InGaN and GaN (piezoelectric polarization) and with In content (pyroelectric polarization). Moreover, it can be seen (Sacconi et al., 2012) that the emission power tends to decrease with higher In concentration due to the increase of spatial separation of the carriers caused by higher strain-induced QCSE. On the other hand, the calculated optical spectra for a fixed In composition and variable column size for the LED considered earlier show an increase in the value of peak emission energy of around 60 meV when the NW width is reduced from 400 to 50 nm. This result is a combination of several effects related to strain relaxation close to the NW boundaries and quantum confinement in the QD; we discuss this point in detail in the following paragraph.

As for the effect of InGaN QD thickness on the optical emission spectra, the behavior is again similar to what occurs in a QW LED. For example, the emission peak in a 50-nm-wide QD increases by 150 meV when the thickness of the QD decreases from 5 to 2 nm. This is mainly due to the higher quantum confinement in the thinner QD; moreover, for the same polarization field, the QCSE is lower for a smaller

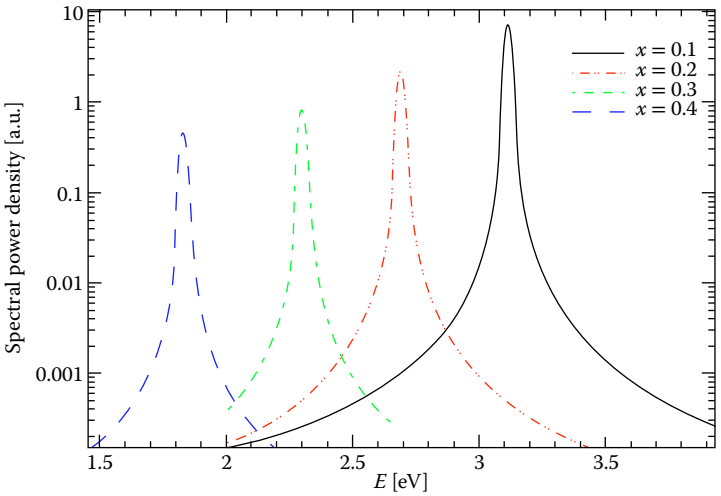

FIGURE 17.12 Optical emission spectra for a column with a radius of 25 nm and several In concentrations.

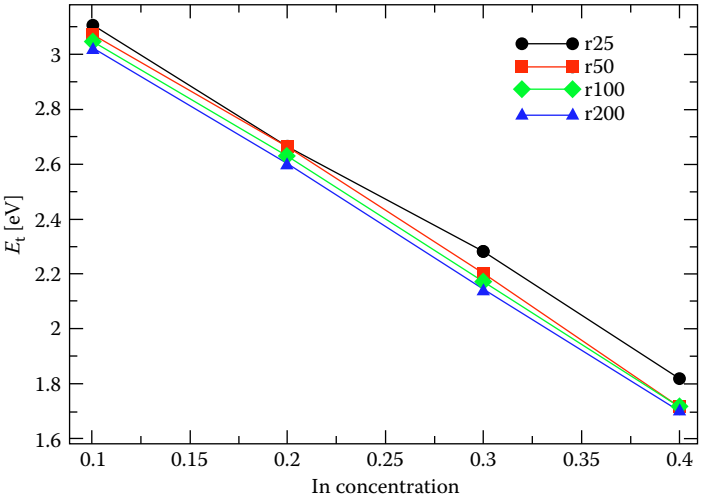

FIGURE 17.13 Dependence of emission energy on In concentration for several column widths.

thickness, and thus QCSE-induced red-shift is reduced. This also determines an increased emission power due to the reduced spatial separation of electron and hole quantum states.

Let us see now in more detail how the transition energies depend on the geometry parameters of the QD NW, in particular, how the lateral scaling of the NW affects the emission properties.

Figure 17.13 shows the dependence of emission energy on In concentration in an InGaN QD NW for several values of the column radius (Sacconi et al., 2012). It can be seen that, for all the column widths, the emission energy has a linear dependence on the In molar fraction. On the other hand, Figure 17.13 shows that the transition energy increases slightly with decreasing width, with a larger slope for radii lower than 50 nm. The effect gets larger with increased In concentration.

Calculations performed in a study by Sacconi et al. (2012) by applying a full self-consistent quantum/drift–diffusion coupling shows an emission energy increase of around 64 meV, from 2.440 to 2.504 eV, when the column radius decreases from 100 to 25 nm. These results show a trend in qualitative

agreement with experiments (Ramesh et al., 2010), even if it must be noted that a comparison of calculations with experimental results is quite difficult because the results depend critically on the geometrical and structure details, as well as on the particle density distribution in the NW.

To point out the role of confinement and of strain distribution on this lateral size effect, the dependence of emission energy on NW geometry has been studied in Sacconi et al. (2012) for several structures, beginning with simulations for undoped structures at equilibrium. In Figure 17.14, the result of a 1D calculation for a QW structure, which can be assumed as a model of a NW with an infinite radius, is compared to the result for a 50-nm-wide NW. It is clear that the geometry of the NW increases the transition energy for a given value of In molar fraction. Moreover, calculations for a very wide column (10 μm) yield results very close to the thin film case. This result can be put in relation with the polarization-induced electric field obtained in the QD shown in Figure 17.15. It can be seen that the electric field gets lower in the NW QD

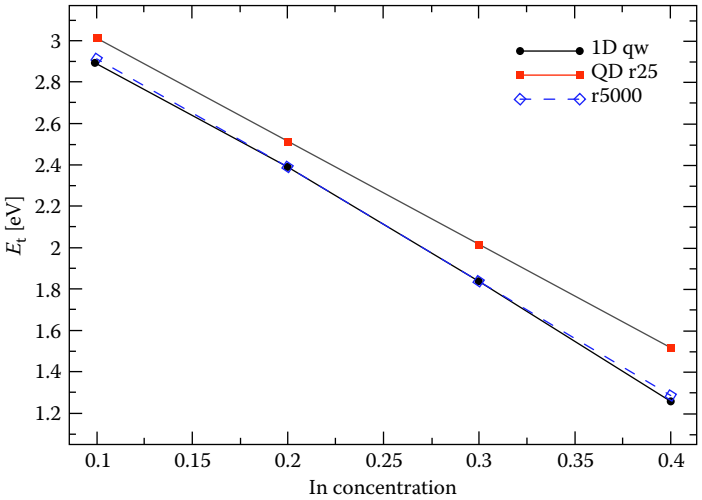

FIGURE 17.14 Dependence of emission energy on In content for undoped structures.

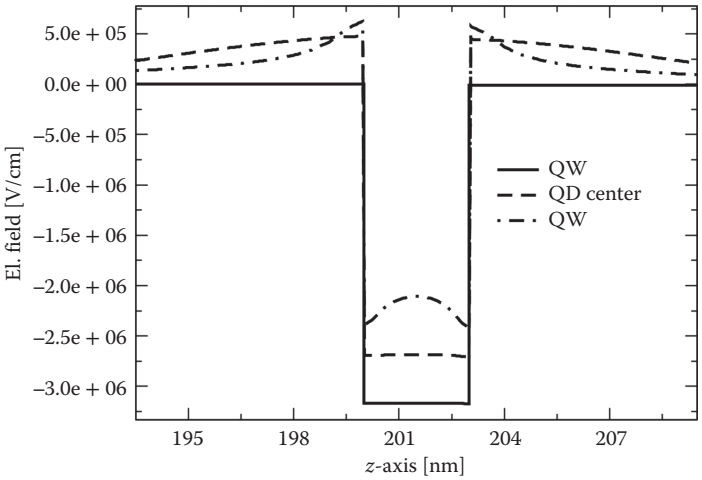

FIGURE 17.15 Electric field along the *z*-axis in an undoped nanocolumn.

with respect to the QW value due to the 3D strain effects, causing lower QCSE and therefore higher transition energy. Furthermore, the electric field in the QD is even more reduced close to the lateral surface due to the strain relaxation effect. The situation is anyway less clear in the doped structures, where screening effects become appreciable. Moreover, it is found that if polarization and strain are not taken into account, the lateral confinement has a significant effect only when the NW radius is reduced under 10 nm (Sacconi et al., 2012).

From these considerations, it seems reasonable to conclude that the main effect of the lateral scaling of the NW is the reduction of the strain inside the QD due to its relaxation at the NW lateral surfaces. In fact, the smaller the NW size, the more pronounced is this reduction of the strain. A lower strain implies less piezoelectric polarization and thus a lower value of QCSE, which causes a blue shift in emission energies. In addition, a smaller size induces a stronger quantum confinement, whose effect on transition energy would be, however, significant only at an extremely reduced size. Anyway, as we have seen earlier, an important role of quantum confinement is to keep particles far from the lateral surfaces, thus limiting the direct influence of NW surface effects, such as strain relaxation and surface states, on the behavior of current in the QD and thus on the LED emission properties.

We have seen the effects of scaling down the NW geometry; on the other hand, one can wonder what happens when the NW size is scaled up. It has been found experimentally (Kawakami et al., 2010) that for a 500-nm-wide NW, emission seems to occur at the strain relaxed region close to the lateral boundary. From simulations performed on an NW model with this size (Sacconi et al., 2012) it turns out that higher recombination rates are obtained for the regions close to the lateral surface, suggesting that for very wide NWs the confinement of quantum states on the NW lateral plane is not sufficient to focus the recombination at the center of the QD. Thus, emission takes place mainly in the lateral regions, where the strain relaxation induces a lower energy gap. However, it has been found from simulations that already for 200-nm-wide NWs the quantum confinement begins to affect the charge distribution, focusing the emission at the center of the QD.

17.10 Conclusions

After having discussed the several aspects of NW QD behavior, we can now underline the main beneficial features provided by NW geometry compared with planar LED structure.

First of all, the NW geometry can be grown almost without defects, yielding in principle high-emission efficiency. NWs also offer the possibility of covering a wide emission spectrum by means of band-gap engineering through variations of alloy concentration in the QD.

Another peculiar NW feature is the relaxation of strain induced in the QD, which occurs at the lateral surface of the NW. This is part of a 3D effect, which determines a lower strain in the QD active region, with respect to an analogous QW in a planar LED structure. This in turns causes less QCSE and higher overlap of quantum wave functions, increasing recombination rate, and internal quantum efficiency, with respect to the planar case.

A lower strain at lateral boundaries determines a lower energy gap with respect to the rest of the QD. This may affect the LED efficiency by reducing the recombination rate. However, as we have shown earlier, the quantum confinement on the lateral plane is sufficient, at least for NW size lower than 200 nm, to focus particle densities in a limited region at the center of the QD. The resulting current crowding in the QD and the increase of radiative recombination are then further advantages of NW geometry versus planar LED. It must be noted also that very often the embedded QD obtained by the NW growth is surrounded by the NW material (yielding a so-called InGaN inclusion in a GaN NW). In this case, the active region is even less affected by surface effects, such as surface states, while the confinement is increased and thereby also the current focusing and the radiative recombination.

For all of these features, one can expect NW QD structures to be very promising for the design of future high-efficiency LED devices.

References

M. G. Ancona, Density-gradient theory analysis of electron distributions in heterostructures, *Superlattices Microstruct.*, 7(2), 119–130 (1990).

M. Auf der Maur, Multiscale approaches for the simulation of InGaN/GaN LEDs, *J. Comput. Electron.*, 14(2), 398–408 (June 2015).

M. Auf der Maur et al., The multiscale paradigm in electronic device simulation, *IEEE Trans. Electron. Devices*, 58(5), 1425–1432 (May 2011).

M. Auf der Maur, A. Pecchia, G. Penazzi, F. Sacconi, and A. Di Carlo, Coupling atomistic and continuous media models for electronic device simulation, *J. Comput. Electron.*, 12(4), 553–562 (2013).

M. Auf der Maur, A. Pecchia, G. Penazzi, W. Rodrigues, and A. Di Carlo, Efficiency drop in green InGaN/GaN light emitting diodes: The role of random alloy fluctuations. *Phys. Rev. Lett.*, 116, 027401 (2016).

R. Calarco et al., Size-dependent photoconductivity in MBE-grown GaN-nanowires. *Nano Lett.*, 5, 981–984 (2005).

D. Camacho and Y. M. Niquet, Application of Keating's valence force field model to non-ideal wurtzite materials. *Phys. E,* 42, 1361–1364 (2010).

S. L. Chuang, *Physics of Optoelectronic Devices*, 1st ed. Wiley Series in Pure and Applied Optics. New York, NY: Wiley (1995).

S. L. Chuang, *Physics of Photonic Devices*, 2nd ed. New York, NY: Wiley (2009).

S. L. Chuang and C. Chang, k•p method for strained wurtzite semiconductors, *Phys. Rev. B,* 54, 2491–2504 (1996).

S. Datta, Nanoscale device simulation: The Green's Function Method, *Superlattices Microstruct.*, 28, 253–278 (2000).

A. Di Carlo, Microscopic theory of nanostructured semiconductor devices: Beyond the envelope-function approximation, *Semicond. Sci. Technol.,* 18, 1 (2003).

M. Hugues et al., Strain evolution in GaN nanowires: From free-surface objects to coalesced templates, *J. Appl. Phys.*, 114, 084307 (2013).

J.-M. Jancu et al., Transferable tight-binding parametrization for the group-III nitrides. *Appl. Phys. Lett.*, 81, 4838, (2002).

Y. Kawakami et al., Optical properties of InGaN/GaN nanopillars fabricated by postgrowth chemically assisted ion beam etching, *J. Appl. Phys.*, 107(2), 023522-1–023522-7 (January 2010).

P. N. Keating, Effect of invariance requirements on the elastic strain energy of crystals with application to the diamond structure. *Phys. Rev.,* 145(2), 26 (1966).

H. Kim et al., High-brightness light emitting diodes using dislocation-free indium gallium nitride/gallium nitride multiquantum-well nanorod arrays, *Nano Lett.*, 4(6), 1059 (2004).

J. Piprek, Efficiency drop in nitride-based light-emitting diodes, *Phys. Stat. Sol. A*, 207(10), 2217–2225 (October 2010).

M. Povolotskyi and A. Di Carlo, Elasticity theory of pseudomorphic heterostructures grown on substrates of arbitrary thickness, *J. Appl. Phys.*, 100, 063514 (2006).

V. Ramesh, A. Kikuchi, K. Kishino, M. Funato, and Y. Kawakami, Strain relaxation effect by nanotexturing InGaN/GaN multiple quantum well, *J. Appl. Phys.*, 107(11), 114303-1–114303-6 (June 2010).

J. Ristic et al., Carrier-confinement effects in nanocolumnar GaN/AlGaN quantum disks grown by molecular-beam epitaxy, *Phys. Rev. B*, 72, 085330 (2005).

F. Sacconi, M. Auf der Maur, and A. Di Carlo, Optoelectronic properties of nanocolumn InGaN/GaN LEDs, *IEEE Trans. Electron. Dev.*, 59(11), (November 2012).

F. Sacconi, G. Penazzi, A. Pecchia, M. Auf der Maur and A. Di Carlo, Optoelectronic and transport properties of nanocolumnar InGaN/GaN quantum disk LEDs, *Proc. SPIE 7597, Phys. Sim. Optoelectron. Dev.*, XVIII, 75970D (February 25, 2010); doi:10.1117/12.840533.

F. Sacconi, M. Auf der Maur, A. Pecchia, M. Lopez, and A. Di Carlo, Optoelectronic properties of nanocolumnar InGaN/GaN quantum disk LEDs, *Phys. Status Solidi C*, 9(5), 1315–1319 (2012); doi:10.1002/pssc.201100205.

S. M. Sze, *Physics of Semiconductor Devices*, 2nd ed. Hoboken, NJ: Wiley (1981).

C. Van de Walle and D. Segev, Microscopic origins of surface states on nitride surfaces, *J. Appl. Phys.*, 101, 081704 (2007).

18

Influence of Random InGaN Alloy Fluctuations on GaN-Based Light-Emitting Diodes

Chen-Kuo Wu
Tsung-Jui Yang
and
Yuh-Renn Wu

18.1 Introduction to Random Alloy Distribution

Unlike the GaAs-based light-emitting diodes (LEDs), the carrier transport and recombination mechanism in nitride-based LEDs are relatively more complicated. As we know, nitride-based wurtzite structures have a strong spontaneous polarization field inside the material. In addition, when the InGaN quantum well (QW) is grown on the GaN buffer layer, it also suffers extra strain, which induces the piezoelectric polarization. Therefore, the polarization field difference at the interface will induce the quantum confined stark

effect (QCSE) and increase the radiative lifetime. The lack of native substrate also makes the crystal quality of the GaN and InGaN layer relatively poor, where a typical 10^8 cm^{-2} dislocation density can be observed in the InGaN QW, which is expected to affect device performance significantly. The efficiency droop, where the light output power increases sublinearly with the drive current, has also been an issue in limiting LED efficiency. There are many studies focused on determining the origin of droop, such as electron overflow [1–3], Auger recombination [4–9], or defects [10–14]. The defect or phonon-assisted Auger recombination is also proposed to explain the larger Auger coefficient. Recently, the direct measurement of Auger electrons [7,15] and a clear correlation between the Auger current and droop seem to indicate that the Auger process might be the main cause of droop. In addition, as mentioned, the QCSE will increase the radiative lifetime and the contribution of nonradiative Auger recombination or carrier delocalization toward droop will become more significant.

Despite these issues, the nitride-based blue LED still has very high internal quantum efficiency (IQE) even under this high dislocation density. The 80% peak IQE of blue LEDs still can be easily achieved in today's commercial LEDs. The reason for this high quantum efficiency has been attributed to the carrier localization effects, where carriers are localized at a local potential minimum so that it will not diffuse into the dislocation center for nonradiative recombination. The spectrum broadening effect is also much stronger compared to GaAs-based LEDs. This is attributed to the random distributed localized state. The origin of carrier localization has been an issue whether it is from indium clustering effect, random alloy fluctuation, or even charged dislocation line-induced potential barrier, etc. More and more evidence shows that random alloy fluctuation should be the main reason for this effect. First, this localization and spectrum broadening is still observed in GaN-negative substrate, which could exclude the effect of a charged dislocation-induced potential barrier. In addition, the three-dimensional (3D) atom probe tomography (APT) data [16–24] show that the indium composition distribution is naturally disordered. Figure 18.1a shows the randomly fluctuated indium distribution in the lateral direction from the APT data [17]. The in-plane map reveals that there are some high indium composition locations corresponding to the red regions and relatively low indium composition sites scattered in the 2D map. Figure 18.1b also indicates that the average indium composition along the growth direction is not uniform and decreases from the middle of the QW to the interface of the InGaN/GaN. Therefore, the QW structure possesses an indium distribution closer to a Gaussian shape rather than the ideal "top-hat" function. A similar idea of the influence of Gaussian shape QW on the IQE has been discussed by Hader et al. [6]. However, their analyses were one dimensional (1D) and lateral fluctuations in indium composition were not considered.

Many simulation tasks [22,24] suggested that carrier localization, induced by these fluctuations, has a strong influence on the broadening of the light emission spectrum. As discussed in Refs. [6,22,23], due to

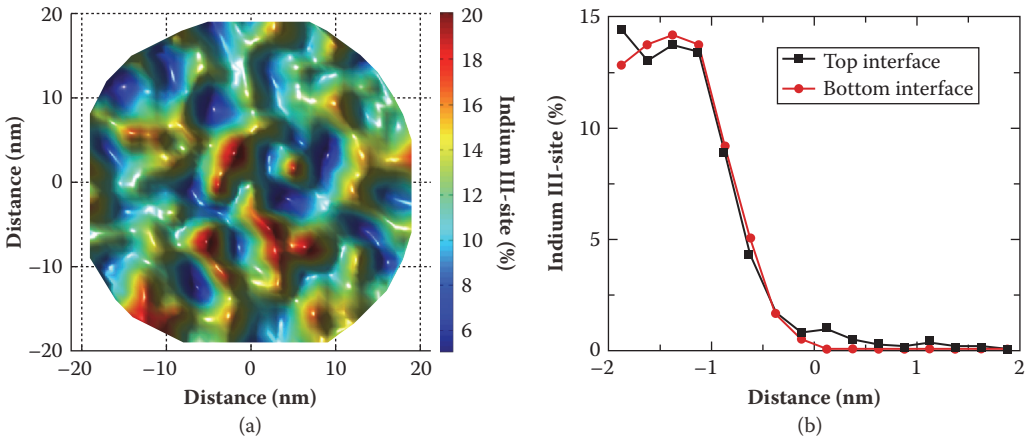

FIGURE 18.1 (a) The in-plane indium distribution of the quantum well (QW). (b) The average indium composition along the growth direction of one 3-nm QW. (From D. A. Browne et al., *Journal of Applied Physics*, 117, 185703, 2015.)

the indium fluctuation, the electrons and holes injected into InGaN QWs will localize in the high indium regions because of the deep localized potential. Additionally, the localized energy states originating from quantum confinement will vary due to fluctuations in indium composition and the QW width, thus broadening the emission spectrum [22,24]. In addition, the carrier localization will increase the local carrier density and the strong screening of the polarization field might be achieved at a smaller current density. The high local carrier density might increase the influence on Auger recombination or defect-assisted recombination [11] at a relatively smaller current density. In addition, the local polarization value induced by composition fluctuation will also influence polarization potential in the QW [25].

Besides these tasks that study the influence on radiative and nonradiative recombination due to the localization effect, there are a few other efforts that analyze the impact on carrier transport by considering random alloy potential fluctuation. In the past, many simulation studies on multiple QW LEDs have been performed under the assumption of a uniform composition and thus potential in the QWs. However, the simulations reveal a large deviation from the experimental data, especially in the InGaN MQWs cases. Due to the strong GaN barrier induced by the piezoelectric charges, a much higher applied bias is usually obtained in the simulation than in the experimental data. The deviation of simulation to experimental results becomes very significant, especially in MQW cases where many triangular shapes of GaN barriers exist in the system. Some studies reduced the theoretical polarization value to fit the experimental data. However, these assumptions might lead to more problems since they might improve electron–hole overlap and result in no droop or droop at a much larger bias, or need a much larger Auger coefficient to reproduce the experimentally observed droop behavior. Some used the tunneling effect to explain this phenomenon. However, the tunneling effect is significant only when a smaller polarization field was employed since the potential barrier induced by the 100% theoretical polarization filed could be very high. If the simulation only tries to fit the current versus IQE and neglects the fitting of I–V curve, usually a higher voltage is obtained in simulation and a conclusion of overflow or overshot effects are often obtained. This is due to that the extra-voltage lowered down polarization barrier induced by polarization field and potential barrier in electron blocking layer and make carrier to overflow much easier. However, this extra voltage is not observed in experimental result in a good commercial LED. Therefore, the conclusion would be wrong since it is based on a non-existence factor.

As mentioned earlier, the influence of random alloy fluctuation on the transport is seldom considered due to the need of large amounts of computational power. Atomic simulation of carrier transport with the 3D random indium fluctuation, especially in MQWs at current computation power, is also impossible. Even modeling a small area of InGaN QW (10 nm × 10 nm × 3 nm) will require a few months calculation with a super computer. Therefore, some quasi-classical simulation work is needed to approach these issues. Our past work [23,24,26] indicated that the results obtained by including the indium fluctuation into the 2D and 3D Poisson, drift–diffusion, and Schrödinger equation simulation model could be closer to experimental data, without the necessity for the assumptions mentioned earlier. The randomly fluctuated alloy composition of ternary epilayers will result in irregular energy bandgap and piezoelectric potential distribution, which will make carrier percolation and confinement more complicated compared to the traditional 1D assumption.

Figure 18.2 shows a sketch of a carrier percolation in the random alloy system. The red regions refer to the high potential region and the blue regions refer to the low potential area. As the sketch shows, the potential distribution is fluctuated either in the QW region or in the electron-blocking layer (EBL) region, where carriers are much easier to percolate through smaller barrier sites and localize in the local low-energy areas. While the large number of localized carriers accumulate in QWs, the local polarization charges will be screened and the effect of Auger recombination will be enhanced at the same time. In addition, the inherent complex potential distribution might provide carriers a path to avoid flow into defect-related regions formed by threading dislocations (TDs) in GaN LEDs. Our past research [23,24] has shown that the calculated turn-on voltages and IQE performances can be better predicted without any parameter reductions by considering random alloy fluctuations. A more reasonable Auger coefficient is used in the model [15], and a broad emission spectrum is also observed due to different localized states produced by alloy fluctuations. Moreover, the emission spectrum shift can be modeled well in the paper [24]. In this chapter, we will introduce how we use the traditional 3D Poisson and drift–diffusion solver to address these issues.

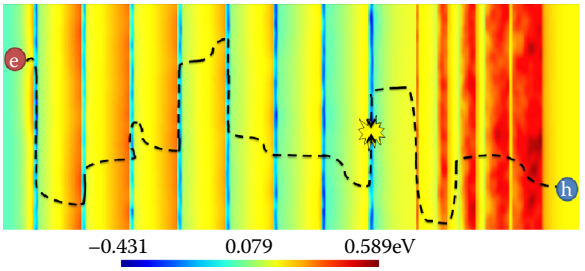

FIGURE 18.2 The sketch illustrates the carrier percolation in the random alloy system. The red regions refer to the high potential region, and the blue regions are low potential area.

18.2 Methodology in Random Alloy Modeling

To model the 3D structure properly, we need to use a full 3D model to examine the device performances of the different structures. The 3D finite element method (FEM) Poisson, drift–diffusion, and Schrödinger solver developed by our lab (named the 3D-DDCC) is mainly used for solving the 3D carrier transport issue of the semiconductor. This program can correlate the electrical characteristic with optical problems and give a self-consistent solution. Our lab has tested and verified the model [24,26–28] and the model has been further developed to apply in the 3D transport simulation. In addition, 3D drift-diffusion charge control (3D-DDCC) program is a very versatile software, in that we can easily use inserted functions in special cases to model the structures or the physics more precisely. In this chapter, the externally inserted functions have the decisive position in the simulation. The details of the functions will be introduced later.

Figure 18.3 shows the full diagram of the simulation flow chart in this modeling process. First, the Gmsh program is used to construct the 3D mesh structure [29]. Afterwards, we need to use the random number

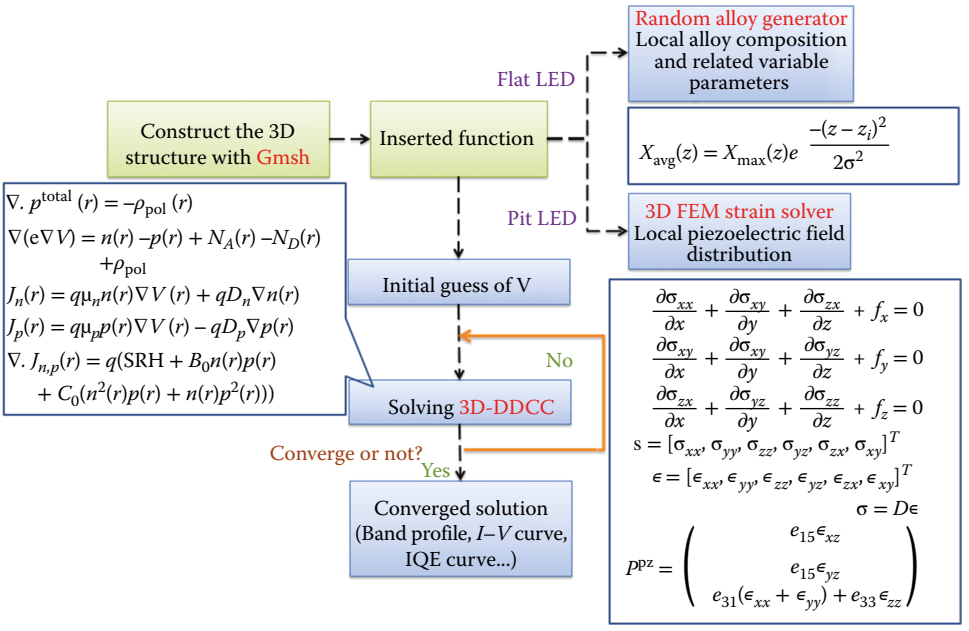

FIGURE 18.3 The full diagram of the simulation flow chart in this modeling process.

generator to generate the random alloy composition which is utilized in the examination of carrier transport in the random alloy system. After obtaining the indium map, we assign the indium composition in each node of the mesh element and all physical parameters change with the indium composition. To obtain a proper consideration of the piezoelectric polarization under the random alloy conditions, the 3D FEM-based strain solver can be used to account for the piezoelectric polarization and calculate the polarization charge. The 3D-DDCC developed by our lab is then used to solve the transport equations with the appropriate input parameters generated by the external modules until we get the converged solutions. Finally, we can obtain the physical profiles such as the potential distributions, current–voltage (I–V) curve, and IQE curve.

18.2.1 Method for Generating the Random Alloy Distribution

As shown in Figure 18.4a, first the In(Al) and Ga atoms are randomly assigned by the random number generator according to the average indium composition and are aligned to the cation lattice site. The lattice site size is decided by the atom density. In addition, according to the APT data, the average alloy composition of each lateral plane along the z-direction, $\chi_{avg}(z)$, is like a Gaussian shape distribution as shown in Figure 18.5a and b, especially when the QW is very thin [6,23,24,30]. Note that the adenosine triphosphate (ATP) resolution in the depth is about one monolayer and lateral resolution is around 2 nm. Therefore, we can obtain the average indium composition of the ith QW by the following equation:

$$\chi_{avg}(z) = \chi_{max} e^{\frac{-(z-z_i)^2}{2\sigma_s^2}}, \tag{18.1}$$

where χ_{max} is the peak average composition of the epilayer, z_i is at the middle of the ith QW, and σ_s is the half width of the Gaussian broadening coefficient. The value of σ_s is around the half width of the epilayer in the z-direction. When atoms are assigned according to Equation 18.1, we need to choose a volume size to get the average local alloy composition. The volume size we use to calculate the local composition is around 2 nm × 2 nm × 0.667 nm, which is close to the APT resolution. Then the alloy composition is weighted at this volume region with Gaussian shape weighting criteria.

As shown in Figure 18.6, the local alloy composition extracted on volumes as mentioned before is following a binomial distribution, which is similar to Refs. [16,18–22].

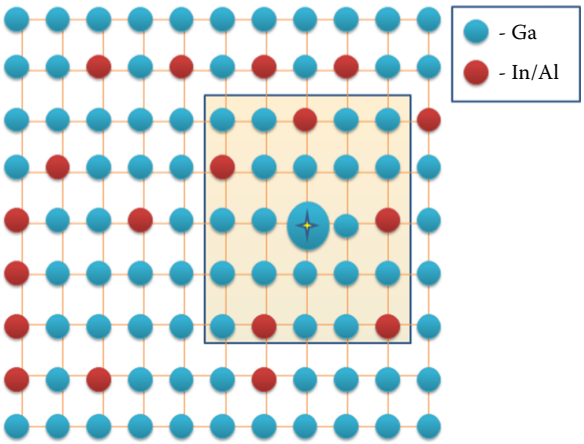

FIGURE 18.4 The concept of the alloy generator. The In, Al, or Ga atom at each lattice site (all with N atoms) is assigned by the random number. Then we choose a proper volume to average the local alloy composition. (From C.-K. Wu et al., *Journal of Computational Electronics*, 14, 416–424, 2015.)

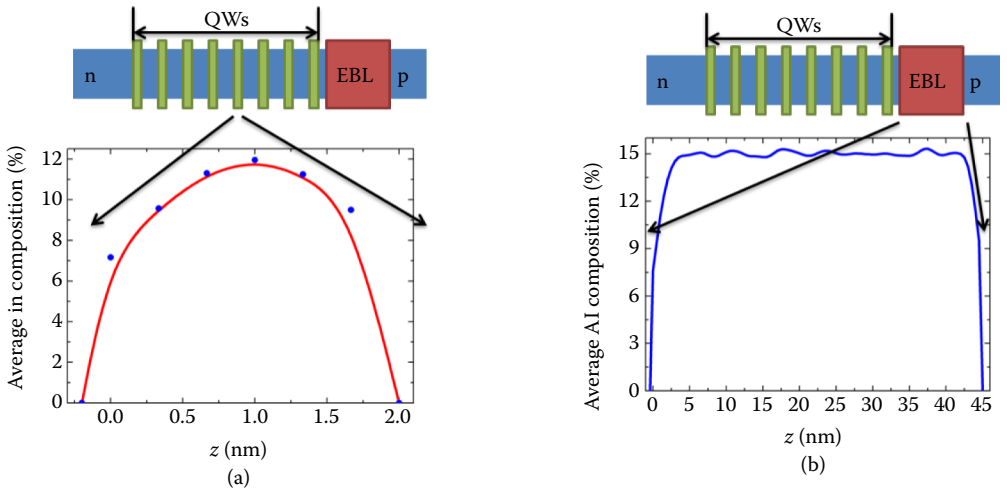

FIGURE 18.5 (a) The average alloy composition distributes along the 2-nm quantum well. (b) The average alloy composition distributes along the 45-nm electron blocking layer (EBL). (From C.-K. Wu et al., *Journal of Computational Electronics*, 14, 416–424, 2015.)

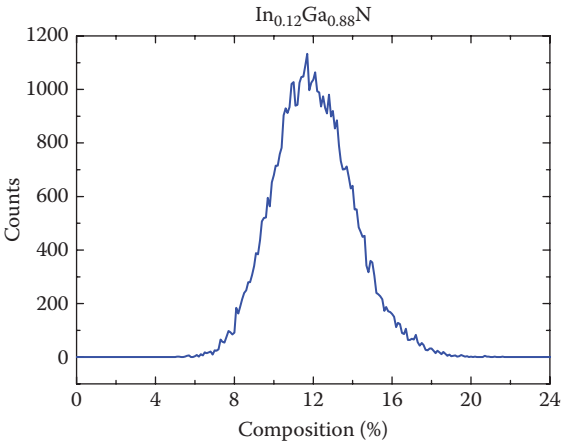

FIGURE 18.6 The counting distribution of certain alloy compositions of the $In_{0.12}Ga_{0.88}N$ quantum well (QW) [26].

After assigning the alloy landscape as shown in Figure 18.7, we combine the mesh file of the structures with the alloy maps and produce the local input parameters simultaneously. The material parameters of each mesh node are assigned according to the local alloy composition map $x(r)$. All parameters (e.g., bandgap [31], effective mass [32,33]) in each node are calculated locally with the indium composition map $x(r)$ by our in-house 3D simulation solver.

18.2.2 Strain Modeling

18.2.3 Calculation of Strain with 3D FEM Method

To understand how the polarization is induced in the random alloy fluctuation condition, we need to do a 3D finite element strain analysis. We need to solve the displacement field by the FEM after considering the incorporation of incoherent lattice constants.

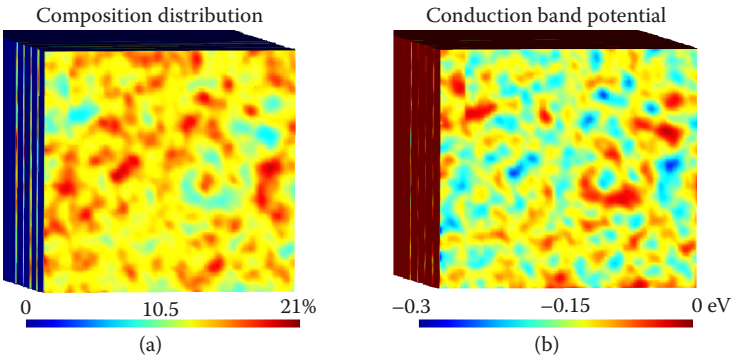

FIGURE 18.7 (a) One random alloy distribution of the n-i-n InGaN quantum wells (QW). (b) The section view of the fluctuated conduction band potential, which is corresponds to the random alloy distribution, in the n-i-n InGaN quantum well case at 0 V bias. (From C.-K. Wu et al., *Journal of Computational Electronics*, 14, 416–424, 2015.)

18.2.3.1 Equilibrium Equations

When the element is located in the *x,y,z* coordinates and some regions are constrained, the element deforms under the balanced force. We can describe the deformation of a point x ($= [x, y, z]^T$) in three components of its displacement:

$$u = [u, v, w]^T. \tag{18.2}$$

The force per unit volume in the vector form can be expressed as:

$$f = [f_x, f_y, f_z]^T. \tag{18.3}$$

In Figure 18.8, we can find the stresses and the body force distributed on the elemental volume dV. Here, we describe the stress by six components as

$$\boldsymbol{\sigma} = [\sigma_{xx}, \sigma_{yy}, \sigma_{zz}, \sigma_{yz}, \sigma_{xz}, \sigma_{xy}]^T, \tag{18.4}$$

where σ_x, σ_y, σ_z are the normal stresses and σ_{yz}, σ_{xz}, σ_{xy} are the shear stresses. Under the equilibrium conditions, the total forces along each direction are zero in each elemental volume. In Figure 18.8, we can pay attention to the x-axis and multiply the stresses to the corresponding areas and list the following equation:

$$\frac{\partial \sigma_{xx}}{\partial x}\mathrm{d}x(dydz) + \frac{\partial \sigma_{xy}}{\partial y}\mathrm{d}y(dxdz) + \frac{\partial \sigma_{zx}}{\partial z}\mathrm{d}z(\mathrm{d}x\mathrm{d}y) + f_x \mathrm{d}V = 0. \tag{18.5}$$

Knowing the dV = dxdydz, Equation 18.5 can be simplified as the following:

$$\frac{\partial \sigma_{xx}}{\partial x} + \frac{\partial \sigma_{xy}}{\partial y} + \frac{\partial \sigma_{zx}}{\partial z} + f_x = 0. \tag{18.6}$$

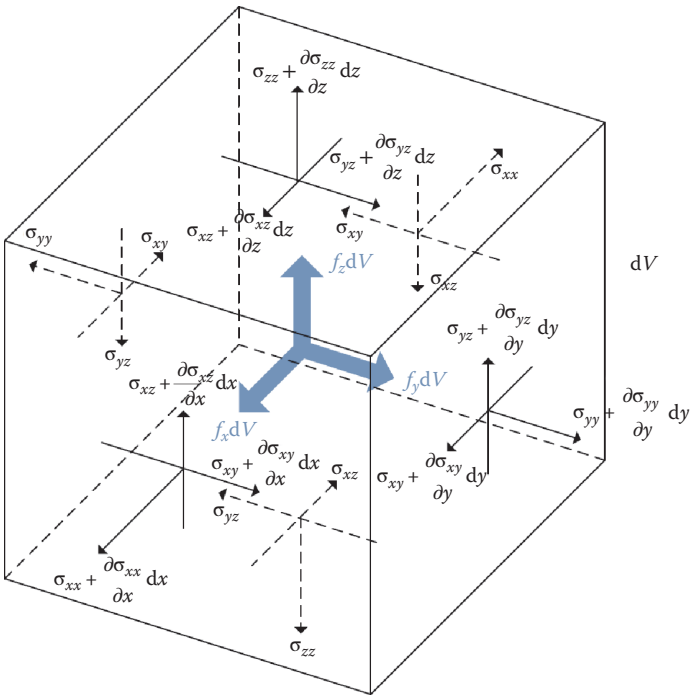

FIGURE 18.8 Equilibrium condition of an elemental volume. (From T. R. Chandrupatla and A. D. Belegundu. *Introduction to Finite Elements in Engineering*. Pearson Education International, 2002.)

After considering the total force along x-, y-, and z-directions, we can list the total equilibrium equations as follows:

$$\frac{\partial \sigma_{xx}}{\partial x} + \frac{\partial \sigma_{xy}}{\partial y} + \frac{\partial \sigma_{zx}}{\partial z} + f_x = 0 \tag{18.7}$$

$$\frac{\partial \sigma_{xy}}{\partial x} + \frac{\partial \sigma_{yy}}{\partial y} + \frac{\partial \sigma_{yz}}{\partial z} + f_y = 0 \tag{18.8}$$

$$\frac{\partial \sigma_{zx}}{\partial x} + \frac{\partial \sigma_{yz}}{\partial y} + \frac{\partial \sigma_{zz}}{\partial z} + f_z = 0. \tag{18.9}$$

18.2.3.2 Stress–Strain–Displacement Relations

In Equations 18.2 and 18.4, we know the form of the stresses and displacement. The strains have the corresponding form as following:

$$\boldsymbol{\epsilon} = [\epsilon_{xx}, \epsilon_{yy}, \epsilon_{zz}, \epsilon_{yz}, \epsilon_{xz}, \epsilon_{xy}]^T, \tag{18.10}$$

where ϵ_{xx}, ϵ_{yy}, and ϵ_{zz} are normal strains and ϵ_{yz}, ϵ_{xy}, and ϵ_{xz} are shear strains. We can also write the strain–displacement relation as

$$\boldsymbol{\epsilon} = \left[\frac{\partial u}{\partial x}, \frac{\partial v}{\partial y}, \frac{\partial w}{\partial z}, \frac{\partial v}{\partial z} + \frac{\partial w}{\partial y}, \frac{\partial u}{\partial z} + \frac{\partial w}{\partial x}, \frac{\partial u}{\partial y} + \frac{\partial v}{\partial x}\right]^T. \tag{18.11}$$

These strain–displacement relations hold under small deformations.

In our research, we assume that the linearly electric stress–strain relations come from the generalized Hooke's law.

$$\boldsymbol{\sigma} = \boldsymbol{D}\boldsymbol{\epsilon}. \tag{18.12}$$

$\boldsymbol{D}$ is the material matrix. In our research, we focused on the wurtzite structure, AlN, GaN, and InN, and $\boldsymbol{D}$ is given by

$$\boldsymbol{D} = \begin{pmatrix} c_{11} & c_{12} & c_{13} & 0 & 0 & 0 \\ c_{12} & c_{11} & c_{13} & 0 & 0 & 0 \\ c_{13} & c_{13} & c_{33} & 0 & 0 & 0 \\ 0 & 0 & 0 & c_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & c_{44} & 0 \\ 0 & 0 & 0 & 0 & 0 & c_{66} \end{pmatrix}. \tag{18.13}$$

18.2.4 Calculation of Piezoelectric Polarization and the Implementation Method

Due to the random alloy fluctuation, each element has its own spontaneous polarization $P^{\mathrm{sp}}(r)$ and piezoelectric polarization $P^{\mathrm{pz}}(r)$ depending on the local alloy composition. After calculating the strain energy, we can further analyze the piezoelectric polarization field distribution using the following equation:

$$P^{\mathrm{pz}} = \begin{pmatrix} 0 & 0 & 0 & 0 & e_{15} & 0 \\ 0 & 0 & 0 & e_{15} & 0 & 0 \\ e_{31} & e_{31} & e_{33} & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \epsilon_{xx} \\ \epsilon_{yy} \\ \epsilon_{zz} \\ \epsilon_{yz} \\ \epsilon_{xz} \\ \epsilon_{xy} \end{pmatrix} = \begin{pmatrix} e_{15}\epsilon_{xz} \\ e_{15}\epsilon_{yz} \\ e_{31}(\epsilon_{xx} + \epsilon_{yy}) + e_{33}\epsilon_{zz} \end{pmatrix}. \tag{18.14}$$

The e_{15}, e_{31}, and e_{33} are the piezoelectric coefficients, which are listed in Table 18.1.

After obtaining the piezoelectric polarization, the total polarization of each element can be calculated by

$$P^{\mathrm{total}}(r) = P^{\mathrm{sp}}(r) + P^{\mathrm{pz}}(r), \tag{18.15}$$

where P^{sp} is the spontaneous polarization of InGaN alloy where the value can be found in Ref. [35]. After calculating the polarization $P^{\mathrm{total}}(r)$, the polarization charge, $\rho^{\mathrm{pol}}(r)$, induced at each element can be calculated by

$$\nabla \cdot P^{\mathrm{total}}(r) = -\rho_{\mathrm{pol}}(r). \tag{18.16}$$

The induced fixed polarization charges at different positions will be put into the 3D Poisson and drift–diffusion solver to obtain the potential inside the devices.

TABLE 18.1 Piezoelectric Coefficients

	$e_{33}(\mathrm{cm}^{-2})$	$e_{31}(\mathrm{cm}^{-2})$	$e_{15}(\mathrm{cm}^{-2})$
GaN	0.73	−0.49	−0.4
InN	0.73	−0.49	−0.4

18.2.5 Modeling of Carrier Transport under the Random Alloy Fluctuation with Drift–Diffusion Solver

To study the 3D carrier transport simulation under the random alloy fluctuation, we need to apply our 3D FEM-based Poisson and drift–diffusion solver. Although the 3D Poisson and drift–diffusion solver cannot describe the tunneling effect, which can estimate the device performance more accurately, it has been proved that it is a suitable model to describe the carrier transport behavior with proper physical parameters, and a solver that yields an acceptable calculation time and required computation memory [24,26–28]. The solver is based on solving the following equations:

$$\nabla \cdot P^{\text{total}}(r) = -\rho_{\text{pol}}(r) \tag{18.17}$$

$$\nabla(\epsilon \nabla V(r)) = n(r) - p(r) + N_A(r) - N_D(r) + \rho_{\text{pol}}(r), \tag{18.18}$$

$$J_n(r) = q\mu_n n(r)\nabla V(r) + qD_n \nabla n(r), \tag{18.19}$$

$$J_p(r) = q\mu_p p(r)\nabla V(r) - qD_p \nabla p(r), \tag{18.20}$$

$$\nabla \cdot J_{n,p}(r) = q(\text{SRH} + B_0 n(r)p(r) + C_0(n^2(r)p(r) + n(r)p^2(r))), \tag{18.21}$$

$$\text{SRH} = \frac{n(r)p(r) - n_{\text{i}}^2}{\tau_n\left(p(r) + n_{\text{i}} e^{\frac{E_{\text{i}} - E_t}{k_{\text{B}}T}}\right) + \tau_p\left(n(r) + n_{\text{i}} e^{\frac{E_t - E_{\text{i}}}{k_{\text{B}}T}}\right)}, \tag{18.22}$$

where V is the potential, and ϵ is the static dielectric constant. N_A^- and N_D^+ are the doping density. n and p are the free carrier concentration of the electron and hole. $\rho_{\text{pol}}(r)$ is the local analytic polarization charge which varies with local indium composition or the calculated results by the 3D FEM elastic strain solver. q is 1.6×10^{-19}C. $\mu_{n,p}$ are electron and hole mobility. $D_{n,p}$ are the coefficients of diffusion, and $J_{n,p}(r)$ are the electron and hole current, respectively. The Shockley–Read–Hall (SRH) is the defect-assisted nonradiative recombination, where τ_n and τ_p are the nonradiative carrier lifetime depending on the crystal quality. E_{t} is the trap energy level located at the midgap. E_{i} and n_{i} are the intrinsic energy level and intrinsic carrier density, respectively. k_{B} is the Boltzmann constant and T is the temperature (300 K here). B_0 is the radiative recombination coefficient. C_0 is the Auger recombination coefficient. We solved those equations until we obtained a converged solution.

18.2.6 Schrödinger Equation and Emission Rate

To analyze the emission properties of the LED, we need to solve the time-independent Schrödinger equation:

$$H\Psi\left(x, y, z\right) = E\Psi\left(x, y, z\right), \tag{18.23}$$

where H is the Hamiltonian operator, which is a second-order differential operator. The time-independent Schrödinger equation can be also expressed as

$$\left[\frac{-\hbar}{2m^*}\nabla^2 + E_{\text{c,v}}\left(x, y, z\right)\right]\Psi\left(x, y, z\right) = E\Psi\left(x, y, z\right). \tag{18.24}$$

$E_{\text{c,v}}$ is the conduction and valence band potential calculated by the 3D-DDCC. Ψ is the carrier wavefunction. m^* is the effective mass and $\hbar$ is the Planck constant divided by 2π. Since the Schrödinger equation is an eigenvalue problem, we obtain the eigenenergy E corresponding to the eigenwave function.

The formula we used to calculate the spontaneous emission rate [36] can be expressed as

$$R_{\mathrm{sp}}(\hbar\omega) = \frac{e^2 n_{\mathrm{r}} \hbar\omega}{m_0^2 \varepsilon_0 c^3 \hbar^2} \frac{1}{V} \sum_{i,j} |\mathbf{a} \cdot \mathbf{p_{i,j}}|^2 \times \frac{1}{\sigma_{\mathrm{b}}\sqrt{2\pi}} \exp\left[\frac{-\left(E_{i,j} - \hbar\omega\right)^2}{2\sigma_{\mathrm{b}}^2}\right] \times f_{\mathrm{e}}\left(E_i^e\right) f_{\mathrm{h}}\left(E_j^h\right) \left(\mathrm{cm}^{-3}\mathrm{s}^{-1}\mathrm{eV}^{-1}\right), \tag{18.25}$$

where n_{r} is the refractive index and c is the light velocity. The $|\mathbf{a} \cdot \mathbf{p_{i,j}}|$ is the momentum matrix element, which includes the overlapping of the electron and hole wave function. E_i^e and E_j^h are electron and hole eigenenergy states, respectively. $E_{i,j}$ is the effective bandgap and is equal to $E_i^e - E_j^h$. f_{e} and f_{h} are the Fermi–Dirac distribution of electrons and holes, respectively. The equations are shown below:

$$f_{\mathrm{e}}\left(E_i^e\right) = \frac{1}{1 + \exp\left(\left(E_i - E_{fn}\right)/k_{\mathrm{B}}T\right)} \tag{18.26}$$

$$f_{\mathrm{h}}\left(E_j^h\right) = \frac{1}{1 + \exp\left(\left(E_{fp} - E_j\right)/k_{\mathrm{B}}T\right)}. \tag{18.27}$$

The $|\mathbf{a} \cdot \mathbf{p_{i,j}}|$ is the momentum matrix element term where electron and hole overlapping is taken into account in both the localization by fluctuations. The effect of QCSE is already accounted for in solving the Poisson equation. A Gaussian broadening σ_{b} was used in the modeling. The value of σ_{b} used in modeling was 10 meV, which is slightly smaller than $k_{\mathrm{B}}T$ to limit the spectrum broadening due to σ_{b}. Therefore, the calculated emission spectrum broadening will be mainly determined by the different energy levels of localized states.

18.3 Unipolar Transport for Random Alloy System

In this section, the percolation transport study of the random alloy system will be presented. To understand how the piezoelectric field affects the transport, we investigate the pure electron transport in the n-GaN/i-InGaN/n-GaN QW structures and the simulation results will also be compared to the experimental work from Browne et al. [17,37]. The 3D numerical model considering random alloy fluctuations will be applied.

18.3.1 Electron Transport in n-GaN/i-InGaN/n-GaN Structures

The existence of piezoelectric polarization and spontaneous polarization has been confirmed by many studies [31,38,39]. As shown in Figure 18.9, if the InGaN QW is grown in between the GaN, it will induce a huge electric field (>1 MV/cm) in the QW and cause the potential bending as shown in Figure 18.9b. If the quantum barrier (QB) thickness is 10 nm, the induced potential barrier peak could be more than 1 eV without electron screening. Theoretically, under this large electric field, we need to apply a strong bias to overcome this barrier for carrier to go through the multiple quantum wells (MQWs). For the single QW case, the influence of the polarization-induced barrier is not significant, especially as it will be screened by carriers from the n-GaN cap layer. However, under the MQW condition, it could cause a huge resistance for carrier transport. For the n-i-n structure, there is no influence of radiative or nonraditive recombination. Therefore, it is a good platform to test how alloy fluctuation and the polarization-induced potential barrier affect the carrier transport. In this section, in order to understand the carrier transport mechanism across the InGaN QWs, we show that we must study how carriers overcome the piezoelectric barriers before we start to investigate the LED structures.

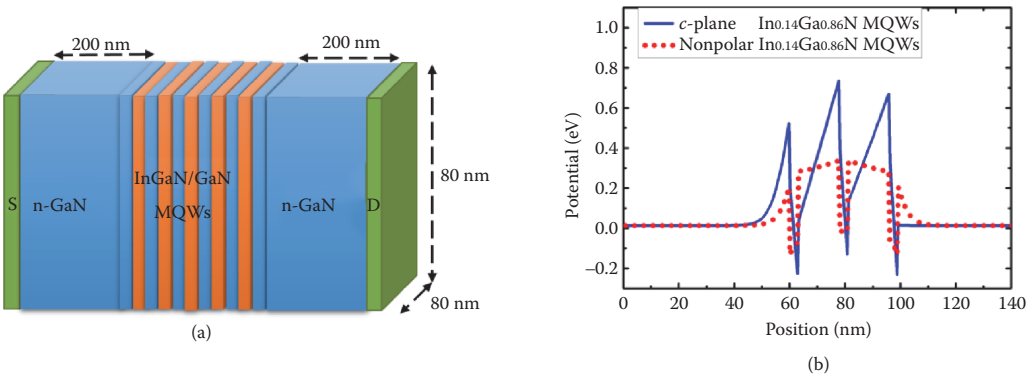

FIGURE 18.9 (a) The modeled device paradigm of n-GaN/i-InGaN/n-GaN quantum well structures for modeling the pure electron transport. (b) The potential distribution of $In_{0.14}Ga_{0.86}N$ *c*-plane and nonpolar piezoelectric barriers.

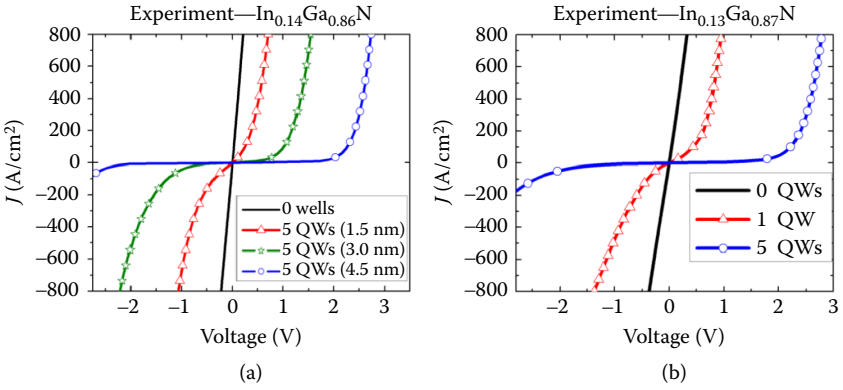

FIGURE 18.10 (a) The experimental *I*–*V* curves of *c*-plane $In_{0.14}Ga_{0.86}N$ with different thicknesses of quantum wells (QWs). (b) The experimental *I*–*V* curves of *c*-plane $In_{0.13}Ga_{0.87}N$ with different numbers of QWs. (From D. A. Browne et al., *Journal of Applied Physics*, 117, 185703, 2015; D. A. Browne et al., Investigation of electron transport through InGaN quantum well structures. In *14th Electronic Materials Conference*, Santa Barbara, CA, June 25–27, 2014.)

The traditional *c*-plane MQWs have a larger degree of piezoelectric barriers than nonpolar (*m*-plane) MQWs, as shown in Figure 18.9b. That means the performance of different traditional *c*-plane MQW structures has a higher correlation with their level of piezoelectric barriers. Browne et. al. [17,37] have found a clear rectifying *I*–*V* curve in *c*-plane MQWs which is related to the QW thicknesses and numbers, where the equivalent piezoelectric barriers will increase with increasing the QW thicknesses and numbers. Figure 18.10 shows that the higher piezoelectric barriers in the devices will increase the driving voltages at the same current density. However, if the QW model is used to study the vertical transport by considering either a full theoretical polarization value or a zero polarization value (nonpolar MQWs), there is a large deviation of predicted driving voltages between experimental work at the same current density, as shown in Figure 18.11.

Besides, it should not be Auger-assisted hot carriers overflow since no Auger recombination could occur in an n-i-n structure and an investigation of the n-i-n structure can be regarded as one that focuses on the carrier transport without any recombination mechanism. The percolation transport through the random alloy system might be a possible reason. A similar effect is also observed experimentally in n-GaN/

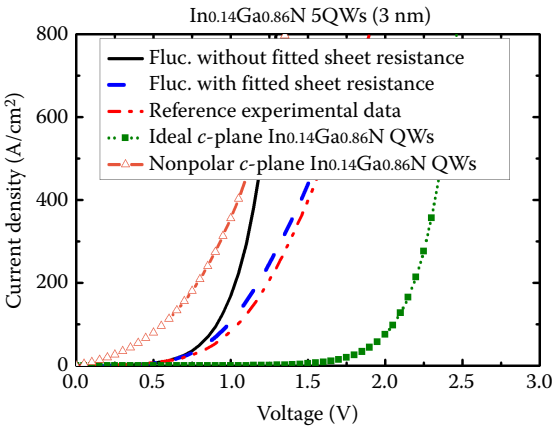

FIGURE 18.11 The *I–V* curves of five experimental and simulated 3-nm quantum wells (QWs) in the n-i-n system. (From D. A. Browne et al., *Journal of Applied Physics*, 117, 185703, 2015; D. A. Browne et al., Investigation of electron transport through InGaN quantum well structures. In *14th Electronic Materials Conference*, Santa Barbara, CA, June 25–27, 2014.)

TABLE 18.2 Detailed Parameters of Each Epilayer for n-GaN/i-InGaN/n-GaN Structures

Epilayer	n-GaN	i-InGaN/GaN
μ_e (cm^2/Vs)	200	600
Doping ($1/cm^3$)	5.0×10^{18}	1.0×10^{17}
E_a (meV)	25	–

i-AlGaN/n-GaN EBL cases [40]. To verify this, we apply the 3D program by considering the random alloy fluctuation in the InGaN QW.

The structures modeled consiste of two n-type doping sides (5×10^{18} cm^{-3}) that sandwich an intrinsic MQW region. The thicknesses of the QWs are 1.5, 3, and 4.5 nm. The barriers in all the cases are 10 nm. The numbers of QWs are 1 and 5. The configuration of the n-i-n structures is shown in Figure 18.9a. In order to present the nanoscale random alloy fluctuations, a small mesh element size is needed. Therefore, we need to limit our chip to 80 nm × 80 nm in the lateral direction due to computational limitations. The random alloy distribution is based on the rule mentioned in the previous section. The average indium composition is 14% according to Refs. [17,37]. The fluctuation range is around 8%–21%, which follows the binomial distribution. We have simulation results of the ideal $In_{0.14}Ga_{0.86}N$ QW cases for comparison. The physical parameters are listed in Table 18.2.

One of the alloy distributions is shown in Figure 18.7a, and the corresponding fluctuated conduction band potential at 0 V is shown in Figure 18.7b, where a local composition site will induce a reversed trend of potential distribution. Moreover, the lower indium regions will induce relatively low polarization barriers and the electrons could percolate through the relatively lower barrier, as plotted in Figure 18.12.

Hence, the fluctuated case performs smaller turn-on voltages than the uniform cases, as shown in Figure 18.11. Compared with the experimental data [17,37], the turned-on voltage of the *I–V* curve with indium fluctuations is close to the experimental results. Since the experimental result has additional sheet resistance not modeled by our vertical transport study due to the size limitation, with the fitted sheet resistance, the calculated results can further match the experimental results. Figure 18.13 shows a comparison of

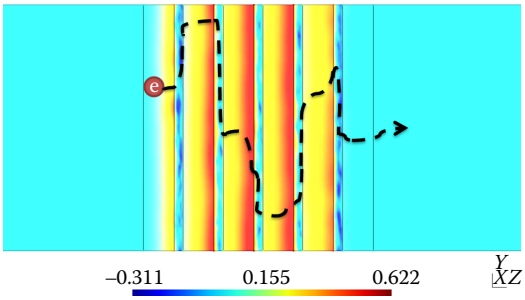

FIGURE 18.12 The side view of conduction band potential in the n-i-n InGaN quantum wells and the scheme of the carrier transport along the fluctuated barriers.

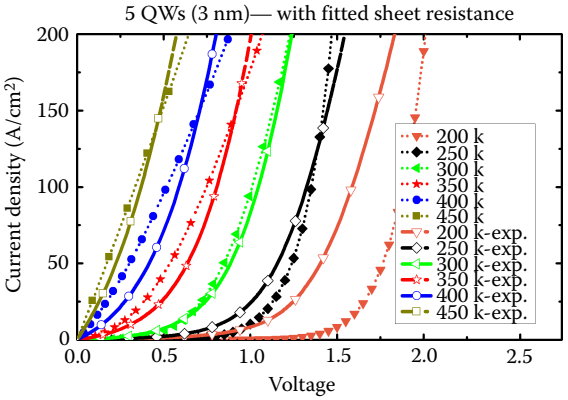

FIGURE 18.13 The *I–V* curves of the experimental and calculated work at various temperatures. (From D. A. Browne et al., Investigation of electron transport through InGaN quantum well structures. In *14th Electronic Materials Conference*, Santa Barbara, CA, June 25–27, 2014.)

the experimental work with the calculated outcomes at different temperatures. At low temperature, the calculations do not fit well. The result reveals that a small portion of tunneling current might dominate at the low temperature rather than the thermionic emission current in the random alloy system because the tunneling current should have a weaker temperature dependence. Consequently, the deviation between predicted current behavior and experimental results might be due to the lack of tunneling process in our transport model.

On the other hand, Figure 18.14 shows the increasing turn-on voltage with increasing the thickness or number of QWs. As we know, for the same polarization electric field, a thicker QW will cause a much larger potential band bending, making it harder for carriers to go across the junction. The turn-on voltage of a single QW is very low because it has only one barrier to be overcome by applying the bias voltage. Even though the thickness of the QW increases to 4.5 nm, the turn-on voltage is still less than 0.5 V. When the number of QWs increases to five, the positive turn-on voltages for 1.5, 3, and 4.5 nm at 20 A/cm^2 current density are 0.10, 0.68, and 1.73 V, respectively.

In addition, there is an asymmetric *I–V* behavior induced by the polarization field, which needs a larger negative voltage to reduce the barriers at reverse bias range. However, since the thermionic current is lower at the reverse bias region, the tunneling current might play a more important role, which requires a program that can handle tunneling transport to model this problem in the future.

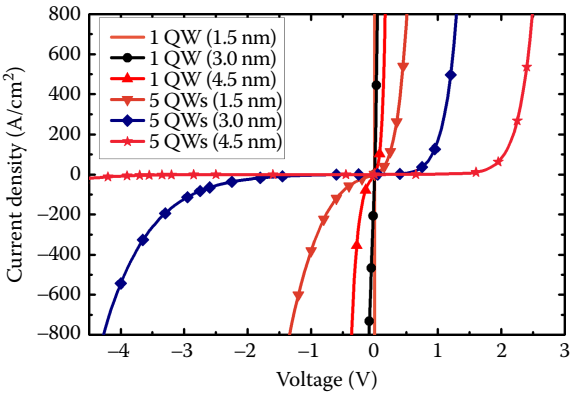

FIGURE 18.14 The *I–V* curves of various thicknesses and numbers of the fluctuated quantum wells (QWs).

18.4 Modeling Result of MQW InGaN LED and the Comparison to the Traditional Model

From the unipolar transport, we find that the polarization-induced barrier does play an important role in limiting the carrier transport behavior. However, without using the random alloy fluctuation, it is hard to approach the experimental *I–V* behavior. In this section, we focus on how indium fluctuation affects the carrier transport in InGaN LEDs.

As we know, InGaN QW LED has become a popular technology in solid-state lighting. Due to the strong lattice mismatch between InN and GaN layers, the self-formed random indium fluctuation has played an important role in influencing LED electrical and optical properties. In this section, we first study the influence of nanoscale indium fluctuation on the emission spectrum and then the relation of carrier transport and radiative efficiency in LEDs. The randomly generated indium fluctuation in the QW in this simulation will be used again in the study with 3D modeling.

18.4.1 Simulation Structure

As we mentioned previously, we divide our study into two parts: the influence of indium fluctuation on (1) emission spectrum and (2) carrier transport. In these two parts, we use different simulation structures. In part (1), since the 3D eigenvalue solver requires huge computer memory and calculating time, the simulation structure is relatively small. The simulation domain for a double QW LED was 30 nm × 30 nm, as shown in Figure 18.15a. The detail parameters setting can be found in Table 18.3. On the other hand, the simulation structure is an 80 nm × 80 nm 6-pair QW LED and included a 40-nm AlGaN EBL in part (2), as shown in Figure 18.15b. The detailed setting can also be found in Table 18.4. We considered a MQW with a 100-nm p-doped GaN layer and a 200-nm n-doped GaN layer, and the QW and barrier width are 3 and 10 nm, respectively. We focus pm a 450-nm MQW blue LED ($In_{0.17}Ga_{0.83}N$). In the uniform QW case, the indium composition is uniformly chosen to be 17%. In our case, as we mentioned previously, we assigned the indium fluctuation in each QW randomly. The maximum local indium composition, which is determined by the random number generator, was around 18%–19%.

A constant nonradiative lifetime was assumed to be 5.0×10^{-8} s, and the radiative recombination coefficient B_0 was assumed to be 2.0×10^{-11} cm^3/s [5,15,41]. Note that recombination rate is decided by B_0 $n(r) \cdot p(r)$ where the electron–hole overlap term is in the term of $n(r) \cdot p(r)$. In all cases, a 100% theoretical polarization value was applied [38]. The EBL thickness was 40 nm with 15% Al content AlGaN. For studying the carrier transport, we solved the 3D Poisson and drift–diffusion solver with the classical particle model. A larger area can be used in the modeling because we do not need to solve for the Schrödinger equation for the eigenstates for transport properties, which needs a much longer computing

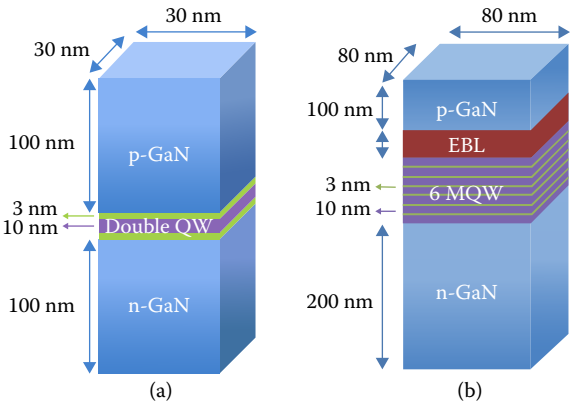

FIGURE 18.15 (a) The simulation structure in analyzing the emission spectrum. (b) The simulation structure in analyzing the carrier transport.

TABLE 18.3 The Material Parameter Settings of the Structure in Analyzing the Emission Spectrum

Unit	Thickness (nm)	Doping ($1/cm^3$)	E_a (meV)	Impurity ($1/cm^3$)	e Mobility ($cm^2/V \cdot s$)	h Mobility ($cm^2/V \cdot s$)
p-GaN	100	2.0×10^{19}	170	0.0	200	5
InGaN	3	0.0	0.0	1.0×10^{17}	600	10
GaN	10	0.0	0.0	1.0×10^{17}	200	10
InGaN	3	0.0	0.0	1.0×10^{17}	600	10
n-GaN	100	5.0×10^{18}	25	0.0	200	10

Note: The unit of thickness is *nm*. The unit of doping and impurity is $1/cm^3$. The unit of activation energy (E_a) is meV. The unit of electron and hole mobility is $cm^2/V{\cdot}s$.

time. We modeled 10 different fluctuation maps and took the averaged I–V and IQEs. Due to memory limitation, we cannot model the whole area of LED. Therefore, we only modeled the limited area in the p-i-n region, as shown in Figure 18.15b. Therefore, the sheet resistance in our simulation was not included due to the memory limitation. In general, the sheet resistance will cause voltage drops in the p-GaN and n-GaN layers so the calculated voltage across the active QW region is smaller than the externally applied bias in experiments. In addition, the real device structures also suffer current crowding effects, which will lead to a more serious droop effect than the result we obtain here.

18.4.2 The Influence of Indium Fluctuation on the Emission Spectrum

In this section, we discuss the influence of indium fluctuation on the emission spectrum. First, as we mentioned previously, the simulation structure we used in this section is a double QW LED and the simulation domain was 30 nm × 30 nm due to computer memory limitations required for 3D simulation, especially for the eigenvalue solver. To avoid the result being limited by the restricted area, we ran a set of different random cases and took the average results. First, the 3D Poisson and drift–diffusion solver developed by our lab was used to obtain a converged band potential at a fixed current density of 20 A/cm^2. Second, we solved the 3D Schrödinger equation with the calculated potential and obtained the confined eigenstate $E_{i,j}$.

Figure 18.16 shows the effects of randomly generated fluctuations in the QW for five different random maps. A total of 22 different random cases were run and averaged (dashed line in Figure 18.16). As we can see, the emission spectra broadened when the indium fluctuations were included in the simulation. The main reason is that the indium fluctuations form indium-rich regions in the QW (Figure 18.7a and b) with

TABLE 18.4 The Material Parameter Settings of the Structure in Analyzing the Carrier Transport

Unit	Thickness (nm)	Doping ($1/cm^3$)	E_a (meV)	Impurity ($1/cm^3$)	e Mobility ($cm^2/V \cdot s$)	h Mobility ($cm^2/V \cdot s$)
p-GaN	100	2.0×10^{19}	170	0.0	200	5
EBL	40	2.0×10^{19}	200	0.0	200	5
p-GaN	10	1.0×10^{18}	170	0.0	200	5
InGaN	3	0.0	0.0	1.0×10^{17}	600	10
GaN	10	0.0	0.0	1.0×10^{17}	200	10
InGaN	3	0.0	0.0	1.0×10^{17}	600	10
GaN	10	0.0	0.0	1.0×10^{17}	200	10
InGaN	3	0.0	0.0	1.0×10^{17}	600	10
GaN	10	0.0	0.0	1.0×10^{17}	200	10
InGaN	3	0.0	0.0	1.0×10^{17}	600	10
GaN	10	0.0	0.0	1.0×10^{17}	200	10
InGaN	3	0.0	0.0	1.0×10^{17}	600	10
GaN	10	0.0	0.0	1.0×10^{17}	200	10
InGaN	3	0.0	0.0	1.0×10^{17}	600	10
n-GaN	10	5.0×10^{18}	25	0.0	200	10
n-GaN	200	5.0×10^{18}	25	0.0	200	10

Note: The unit of thickness is *nm*. The unit of doping and impurity is $1/cm^3$. The unit of activation energy (E_a) is meV. The unit of electron and hole mobility is ($cm^2/V \cdot s$).

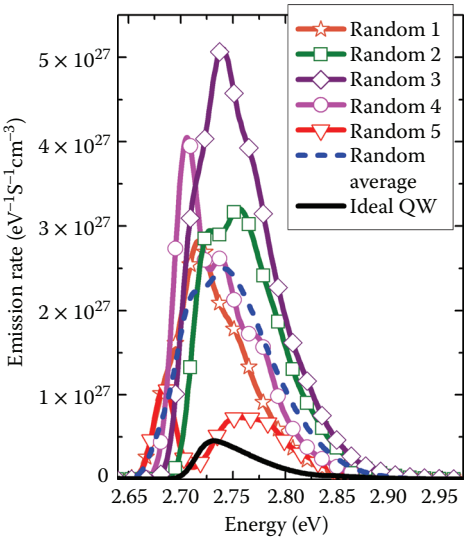

FIGURE 18.16 The calculated emission spectra when including the indium fluctuation. The dashed line is the average result of 22 random cases.

different bandgaps and confined energy levels. However, local peaks are observed in some cases (the random 2 and random 3 cases in Figure 18.16). This might be the result of the small sampling volume used in the simulation. In real devices with much larger area, these different local peaks merge into the overall spectral broadening. Figure 18.17b shows the calculated full width half maximum (FWHM) of the emission spectrum is around 80–120 meV at room temperature, which is close to the experimental results.

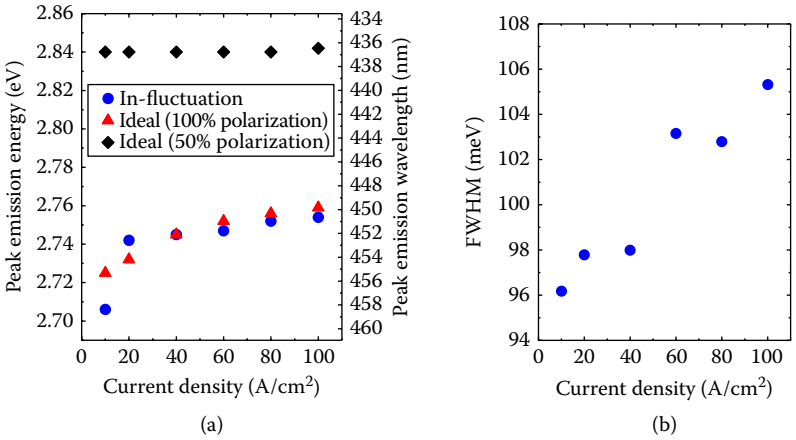

FIGURE 18.17 (a) The position of emission peak shift at different current density. (b) The full width half maximum (FWHM) of the emission spectrum. (From T.-J. Yang et al., *Journal of Applied Physics*, 116, 113104, 2014.)

Homogeneous broadening mechanisms usually will not lead to such large linewidths. If we compare the calculated intensity to the uniform QW case, the emission is stronger due to three factors occurring in the high In content regions where carriers localize in high indium region (deep potential): (1) there is less electron–hole separation; (2) there is more screening of electric field at a given current; and (3) there is an increased QW occupancy factor due to carrier localization. Figure 18.17a shows the blue shift of the spectrum when increasing the current density. A 40 meV (~ 7 nm) blue shift is observed when the current density increases from 10 to 100 A/cm^2, which is close to most experimental observations. Comparing with simulations of ideal QWs, assuming 50% polarization only leads to a very small blue shift due to a very reduced QCSE and the emission peak shifts to much shorter wavelength, while assuming a 100% polarization gives a reasonable shift, while, however, leading to very large forward bias (see the discussion in Section 18.4.3.1).

18.4.3 The Influence of Indium Fluctuation on the Carrier Transport and the Efficiency Droop

To model the vertical carrier transport in the 3D indium fluctuation cases, we used LEDs with six InGaN QWs with GaN barriers and included an AlGaN EBL, as shown in Figure 18.15b. We considered an MQW with a 100-nm p-doped GaN layer and a 200-nm n-doped GaN layer. The QW and barrier widths are 3 and 10 nm, respectively. We focus on 450-nm MQW blue LEDs ($In_{0.17}Ga_{0.83}N$). In the ideal QW case, the indium composition was uniformly 17%. In our case, we assigned the indium fluctuation in each QW randomly as described above. The maximum indium composition, which is determined by the random number generator, was around 18%–19%.

18.4.3.1 The Simulation Result of *c*-Plane LED

The simulation result is shown in Figure 18.18a. This is the calculated conduction band potential landscape. The indium-rich regions correspond to regions with a lower potential. The carriers will localize in the relatively lower potential regions because carriers inherently tend to stay in low-energy regions, as shown in Figure 18.18b. Consequently, the radiative recombination rate increases in the indium-rich region because of the carrier screening of the polarization fields, as shown in Figure 18.18c.

18.4.3.2 The *I*–*V* Curve

Next, we focused on the comparison of transport between the QW including indium fluctuation and uniform QW. Figure 18.19a shows the *I*–*V* curves. The forward voltage V_f used here is for a current density of

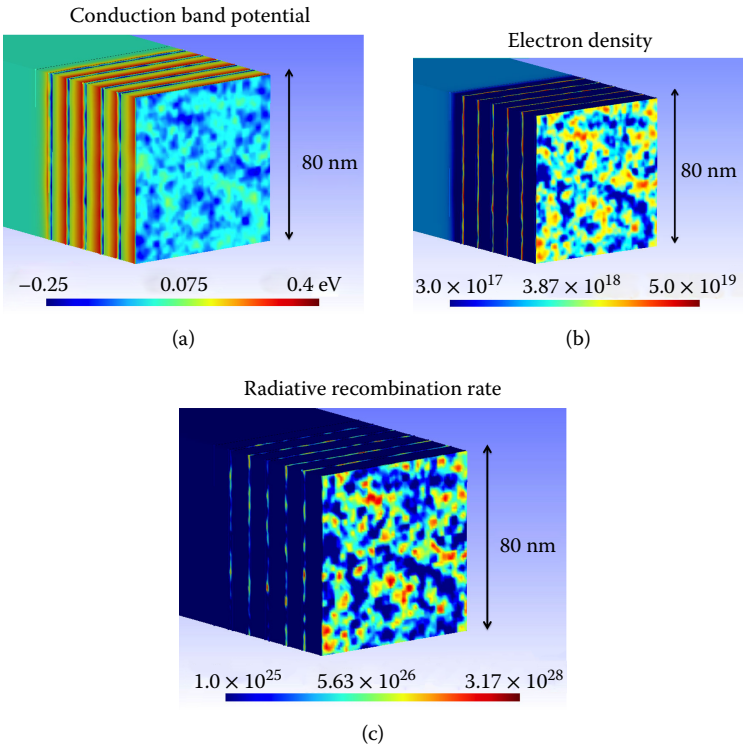

FIGURE 18.18 (a) The conduction band potential at 3.1 V. (b) The electron density at 3.1 V. (c) The radiative recombination rate at 3.1 V. (From T.-J. Yang et al., *Journal of Applied Physics*, 116, 113104, 2014.)

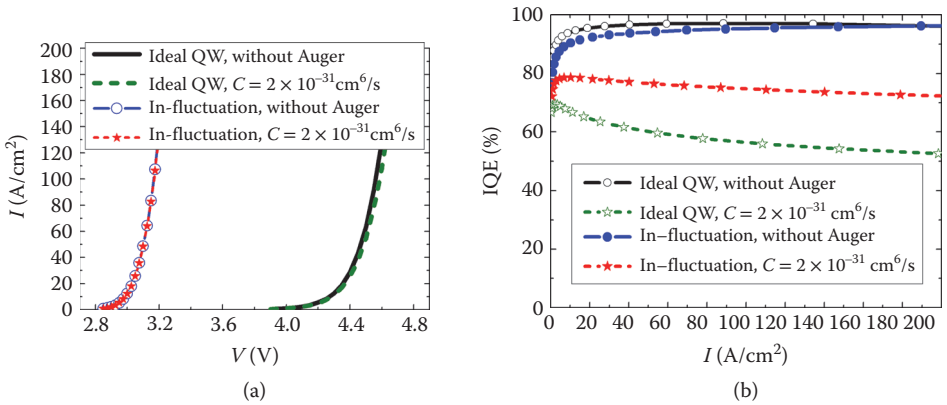

FIGURE 18.19 (a) The influence of indium fluctuation on the *I*–*V* curves. (b) The internal quantum efficiency (IQE) curves. (From T.-J. Yang et al., *Journal of Applied Physics*, 116, 113104, 2014.)

20 A/cm^2. For commercial blue light LEDs, V_f is around 2.8–3.0 V. For the ideal QW simulation (with uniform indium composition and 100% theoretical polarization value), V_f is almost 4.4 V, which is far larger than experimental results. This is also observed in most commercial simulation software with the same parameters set, which was also discussed in our twodimensional (2D) modeling results [23]. Some suggest including tunneling in the model to get a smaller V_f. But our calculation shows that using Wentzel–Kramers–Brillouin (WKB) tunneling between ideal QWs and 100% polarization cannot push V_f

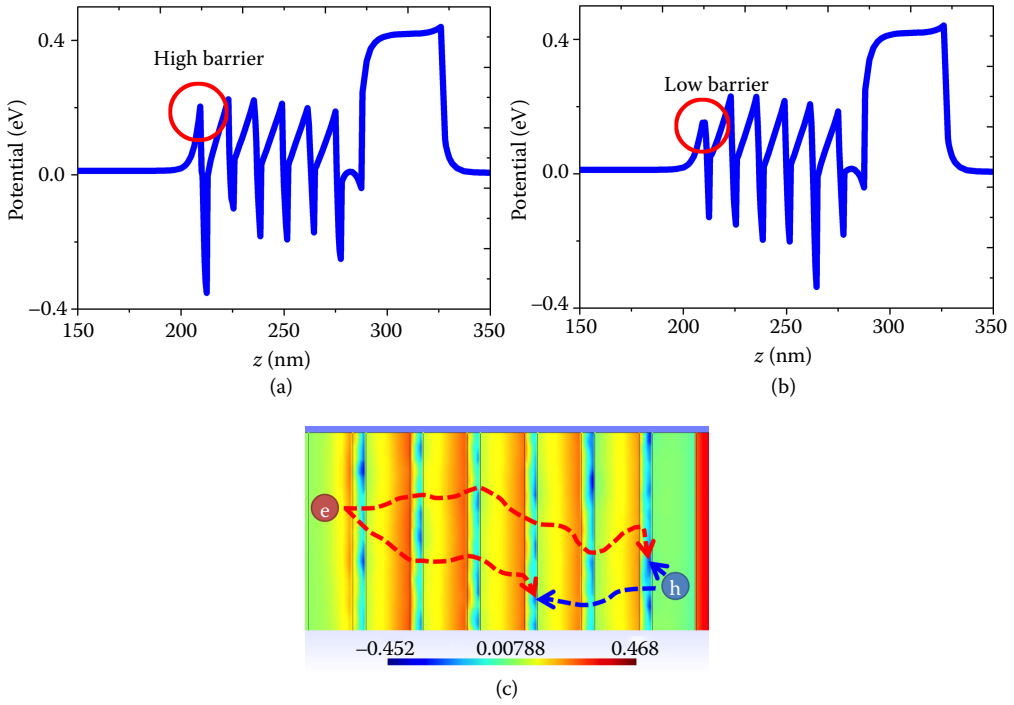

FIGURE 18.20 (a) The high barrier in higher indium region in the first quantum well (QW). (b) The low barrier in low indium region in the first QW. (c) The illustration of carrier transport in the device.

lower than 3.7 V. On the other hand, when including indium fluctuations, we calculated a V_f of 3.05 V even without considering tunneling.

The reason why the V_f shifts to a reasonable value is that the indium composition will directly affect the bandgap potential in the QW due to different strength of the polarization field that in turn affects the bandgap potential in the QW. As we can see for the first QW in Figure 18.20a and b, the potential at the high indium region is lower than the potential at the lower indium region. However, due to the strong lattice mismatch, the piezo polarization charge will induce a triangular shape barrier between the GaN barrier and InGaN QW. The piezo polarization-inducing barrier height also depends on the indium composition. As shown in picture, the barrier at the high indium region is high due to the large polarization difference. On the other hand, the barrier at the low indium side will be reduced. For carriers, they are much easier to flow through the low barrier (low indium side) and localize in the low potential region (high indium side). This is the main reason why the I–V curve shifts to a value that matches experimental data more closely. Because the carriers will inherently find a percolation path and transport in the device, the indium fluctuation will strongly influence the carrier transport. Figure 18.20c illustrates the carrier transport in the device. Nevertheless, the V_f calculated is still slightly higher than the experimental result. Recent studies show that QW LEDs without V-pit or GaN substrate LED with very low dislocation density have a V_f around 3.2–3.4 eV, which is close to our prediction because the influence of V-pit has not yet been included in our simulation. In addition, the indium fluctuations of the piezoelectric field–induced GaN barrier might enhance the tunneling since the induced barrier height will be smaller at the lower indium composition site. This might further lower the V_f.

18.4.3.3 The Droop Behavior of IQE

Turning to the dependence of IQE on current, there are several reasons proposed for origins of droop, such as overflow [1–3], Auger recombination [4–9], and defects [10–14]. For the idealized uniform QW without considering the Auger effect and with EBL, the efficiency peak occurs at 80 A/cm^2 and at a very large

bias (~ 4.5V). (Figure 18.19b) (considering only vertical transport and neglecting the influence of sheet resistance). Although the drift process becomes dominant at this high bias, the electrons are blocked by the EBL, and the overflow is weak. Even including the Auger recombination, the droop effect only occurs at 4.0 V, which is very large compared to commercial data sheets. Note that this large bias condition across the junction should not occur in the real device application. If we observe the most recent experimental results, we can find that the droop occurs at very low current density (~10 A/cm^2) and the applied bias is typically lower than 3.1 V. In the past, most researchers focused on the current density and disregarded the voltage except when discussing the wall plug efficiency. However, here the voltage plays a key role in understanding physical processes in LEDs.

When Auger recombination is excluded from the simulation, there is no droop even when the current is ~200 A/cm^2 (Figure 18.19b). This is due to the blocking action of EBL. With In fluctuations included, a small droop occurs when the current density is above 400 A/cm^2 and the applied voltage above 3.5 V is larger than V_{BI} and ΔE_c, which is unphysical as the resistive voltage drop should only occur in n and p-layer surrounding the depletion region. The biased voltage larger than built-in voltage V_{BI} including the ΔE_c should not be exist in the depletion region. Such voltages are very different from observed ones since the experimental droop typically occurs at 2.8–3.0 V from commercial data sheets (Figure 18.24).

At 3.0 V, the bias is smaller than the built-in voltage, V_{BI} (3.3 V in our case), and the carrier transport is mainly dominated by the diffusion process, as we can see in Figure 18.22a. The potential in the p-region

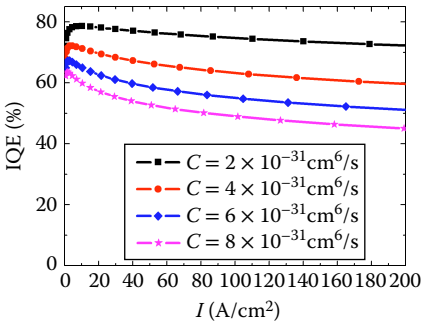

FIGURE 18.21 The internal quantum efficiency (IQE) curves with different Auger coefficients.

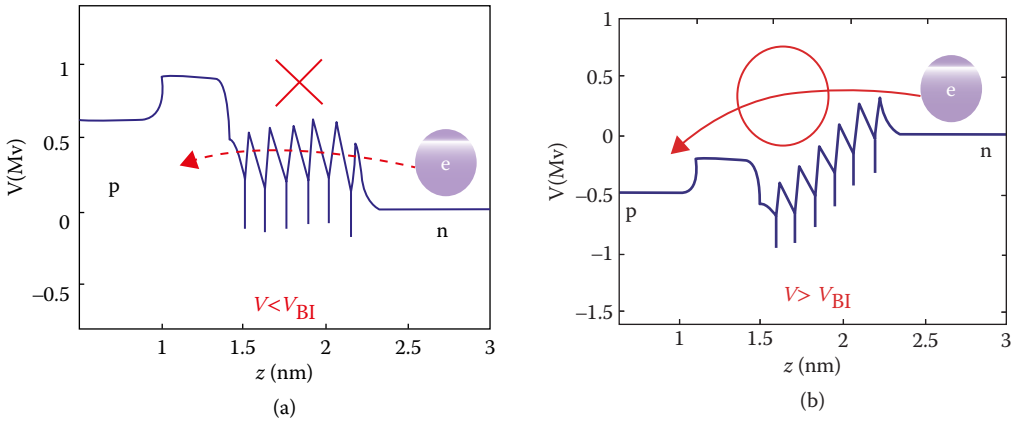

FIGURE 18.22 (a) When $V < V_{BI}$, it is mainly dominated by the diffusion process. It is harder for the carrier to overflow or overshoot the active region. (b) When $V > V_{BI}$, the carrier start to overflow or overshoot the device. However, the droop of a commercial grade light-emitting diode (LED) occurs at $V << V_{BI}$.

including the extra barrier height from the AlGaN EBL is much higher than the n-region and forms a large barrier to the carriers. Therefore, it is harder for carriers to spill over the barrier especially with the existence of the AlGaN barrier. Even with the polar optical phonon absorption process to gain the electron energy (~90 meV), it is still hard to overcome the 0.3–0.4 eV barrier. On the other hand, when the voltage exceeds the V_{BI}, as we can see in Figure 18.22b, the band structure reverses. In this situation, the carrier transport turns from the diffusion process to the drift process. The carriers start to overflow the device. Therefore, the simulation only sees the droop caused by overflow at a voltage higher than the built-in voltage (>3.3 V), if the Auger recombination is not considered. Therefore, the droop caused by the overflow only happens when the bias is larger than the built-in voltage or even larger bias when the AlGaN EBL is added.

Since the actual droop occurs at a very low bias in the experiments, we need to consider other factors that might cause droop effect at low bias. When the Auger recombination is included, the result shows that the droop occurs at 3.0 V, which matches closely to the experimental data. Consequently, we conclude that the droop effect is mainly dominated by the Auger recombination rate or other reasons such as carrier density dependent nonradiative recombination by defects.

As we can see in Figure 18.19a and b, when Auger recombination is included, droop occurs in both cases of ideal QW or QW incorporating In fluctuations. However, in the former case the bias voltage is again an unphysical 4.0 V, while in the latter case it is 3.0 V, which matches experimental data. Consequently, we conclude that the droop effect is mainly dominated by Auger recombination, enhanced by the effect of indium fluctuations. In addition, it should be pointed that due to the memory limitation, we calculated the vertical carrier transport within a limited area. In a real device, current crowding [42] issues can make local current density much higher than expected, making the droop effect worse, as discussed in earlier studies [43,44]. Figure 18.21 shows the IQE curves with the different Auger recombination coefficients, C. With larger values of C, the IQE curves show earlier droop onset, lower IQE peak, and more severe droop effects.

18.4.3.4 The Comparison of IQE between the Ideal QW and In-Fluctuation QW

In Figure 18.19b, the indium fluctuation case shows a good IQE performance with the higher IQE peak value and lower droop effect. Due to the carrier localization, the QCSE will be screened by the localized carrier charge. As a result, the better electron–hole overlapping will lead to a higher efficiency peak. On the other hand, in the ideal QW case, the IQE curve reaches the peak value when the applied bias is over 4.0 V, which is much larger than V_{BI} (3.3V). Therefore, the carrier leakage starts to affect the droop effect so the IQE curve shows a severe droop effect in the ideal QW case.

In conclusion, we prove that by considering the indium fluctuation in MQW LEDs, the electrical and optical properties are much closer to the commercial blue light LEDs. The droop behavior might be dominated by the Auger recombination at the lower current density since the voltage is much smaller than V_{BI}, and the ratio of the overflow mechanism might increase only when the bias is close to or larger than V_{BI} and the current density gets larger.

18.5 Carrier Transport in Green LEDs: Influence of Possible Imperfection in QWs

18.5.1 Introduction

Simulations with a proper model can provide some suitable evidence or clues to improve the optimization of devices. Before the 3D random alloy fluctuation model was proposed, scientists had focused on the current behavior of blue LEDs, because the I–V behavior could not be modeled well with 1D transport computer-aided design (CAD) software without considering the random alloy fluctuation. To explore the carrier transport and recombination mechanism in green LEDs, we conduct a preliminary examination of the electric property of green p-n LEDs.

18.5.2 Carrier Transport in Green LEDs Considering Indium Fluctuations Only

Initially, we constructed a green emitting p-n LEDs with five pairs of uniform 3-nm/10-nm $In_{0.27}Ga_{0.73}N$/GaN MQWs (517-nm green emission) and a fluctuated case for comparison. Both the cases are composed of 295-nm n-doped layer with 5×10^{18} cm^{-3} doping density and a 100-nm p-type layer with 2×10^{19} cm^{-3} doping density. Figure 18.23a shows the epilayers of all the simulated structure here. The chip size for the 3D case is 250 nm × 40 nm. The average composition for the 3D fluctuated case is 21%, as shown in Figure 18.23b, where the maximum composition of the fluctuated case is around 27%. Table 18.5 lists the input parameters of all the cases in this section. Figure 18.24 shows the calculated *I*–*V* curves of green LEDs. As expected, the 1D vertical transport model with a 100% theoretical polarization value failed to describe the experimental observation because of strong polarization fields. Even when the ideal polarization is reduced to 50%, the calculated result still shows a large difference with the extracted experimental values. However, the simulated green emitting LED with random alloy fluctuations also performs a larger driving voltage compared to the experimental result. Obviously, carriers still could not percolate through the fluctuated piezoelectric barriers in the simulation because the average compositions of QWs are so high that the piezoelectric fields will result in high potential barriers. The assumption of average 21% fluctuated QWs with 27% maximum composition might be far removed from the real devices. Our past study also shows that a much larger period of composition fluctuation will further reduce the turn-on voltage [23]. Therefore, it is worth taking further investigation.

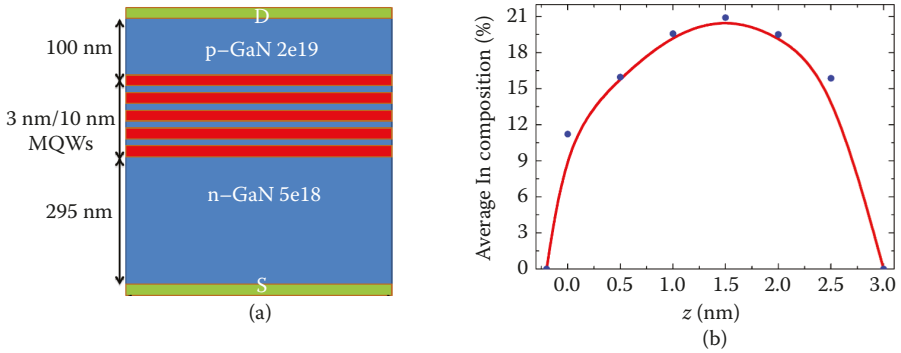

FIGURE 18.23 (a) The structure of the simulated green light-emitting diode (LED). (b) The average composition of the quantum well along the growth direction in the 3D fluctuated case.

TABLE 18.5 Detailed Parameters of Each Epilayer for Green Emission LEDs

Epilayer	n-GaN	i-InGaN/GaN	p-GaN
μ_h (cm^2/Vs)	10	10	5.0
μ_e (cm^2/Vs)	200	600	200
Doping ($1/cm^3$)	5.0×10^{18}	1.0×10^{17}	2.0×10^{19}
E_a (meV)	25	–	170
τ_n (s^{-1})	5.0×10^{-8}	5.0×10^{-8}	5.0×10^{-8}
τ_p (s^{-1})	5.0×10^{-8}	5.0×10^{-8}	5.0×10^{-8}
B_0(cm^3/s)	2.0×10^{-11}	2.0×10^{-11}	2.0×10^{-11}
C_0(cm^6/s)	2.0×10^{-31}	2.0×10^{-31}	2.0×10^{-31}

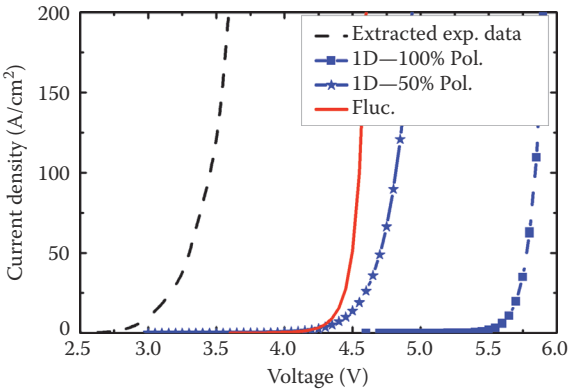

FIGURE 18.24 The *I–V* curves of the several models and the extracted experimental data. (From C.-H. Lu et al., *Journal of Alloys and Compounds*, 555, 250–254, 2013.)

18.5.3 Imperfection of QWs in Green LEDs

Some studies [47,48] have shown that the QWs are not perfect. The studies in Refs. [47,48] further indicated that large-scale well-width fluctuation exists in the commercial green LED samples. Inconsistent growth temperature between GaN layer and $In_xGa_{1-x}N$ might be the main reason for the interrupted void regions among the QWs. Especially when the indium composition increases, the difference of the required growth temperature with GaN will be larger. To further understand how the imperfect QWs affect the electrical property of green p-n LEDs, we apply the 3D model in examining the imperfection of QW in the green LED device.

18.5.4 3D Examination of Green LEDs with Imperfect QW Model and Random Indium Fluctuations Model

To examine the carrier transport in the void structures, we construct a 3D structure with random alloy fluctuations that is much closer to the real physics. The structure is with different void densities. The distance of void is 30 and 40 nm. The input parameters are listed in Table 18.5. Due to the huge computation time in the 3D calculation, the calculation area is limited to 250 nm × 40 nm in the lateral direction. All the epilayers are the same as mentioned previously. The side view and section view composition distributions are shown in Figure 18.25. Compared to the normal fluctuated QWs, the map with imperfect QWs will have some void regions in the QW region. The void regions are also not aligned with one another according to experimental observations [47,48]. Figure 18.26a shows that the 3D model with different void lengths can better model the V_f of the experimental data. The conduction band potential at 3.05 V of 3D fluctuation with 40-nm l_{void} shows that the piezoelectric barriers are smaller in the void region, which is shown in Figure 18.26b. In addition the piezoelectric fields near the normal fluctuated QW region have the insufficient impact on the void region. Hence, the carriers would prefer to percolate through the void regions into the active area, which might be closer to the real devices.

Figure 18.27 shows the calculated vector profile of both electron and hole currents. The results indicate that the void regions dominate the current flow path for both electron and hole currents. While carriers percolate into the void regions, other carriers might prefer to flow into the low potential areas locatied at the stripe QWs. As a result, the smaller V_fs in imperfect QW cases are attributed to an alternative percolation path provided by the void regions.

To conclude this section, we applied the 3D model to show that the void regions in the green emission LEDs indeed affect the carrier injection. The imperfect regions provide an additional injection path for the

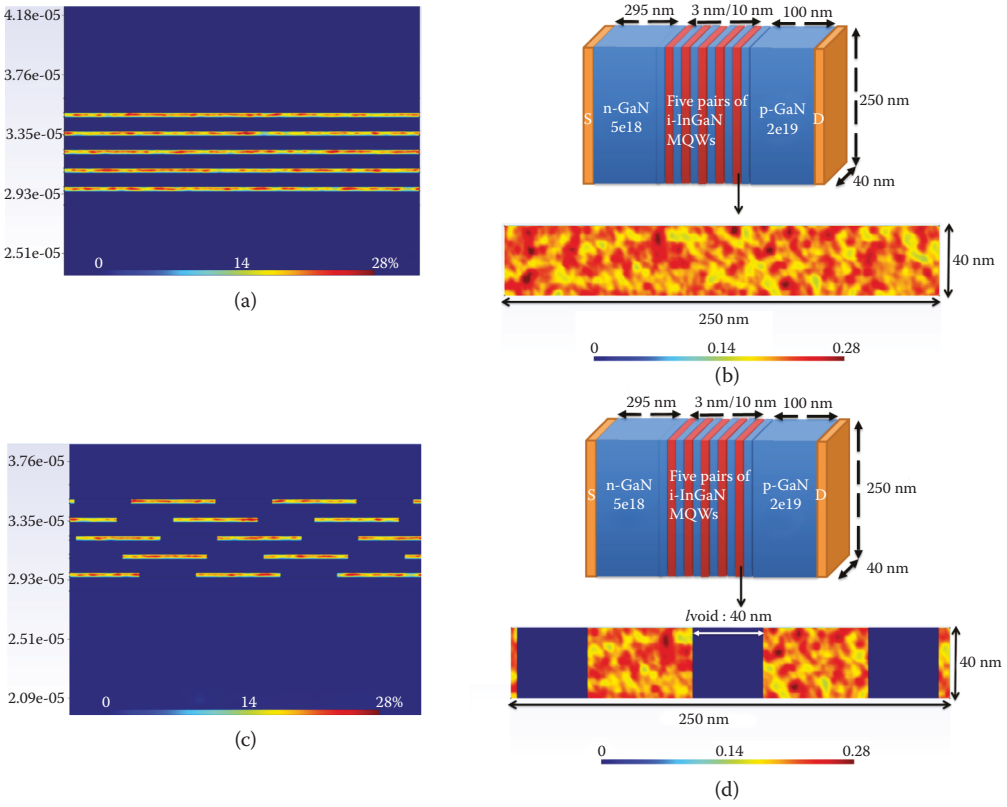

FIGURE 18.25 (a) The side view of the composition map with normal fluctuated QWs. (b) The section view of the composition map with normal fluctuated QWs. (c) The side view of the composition map with imperfect fluctuated QWs. (d) The section view of the composition map with imperfect fluctuated QWs.

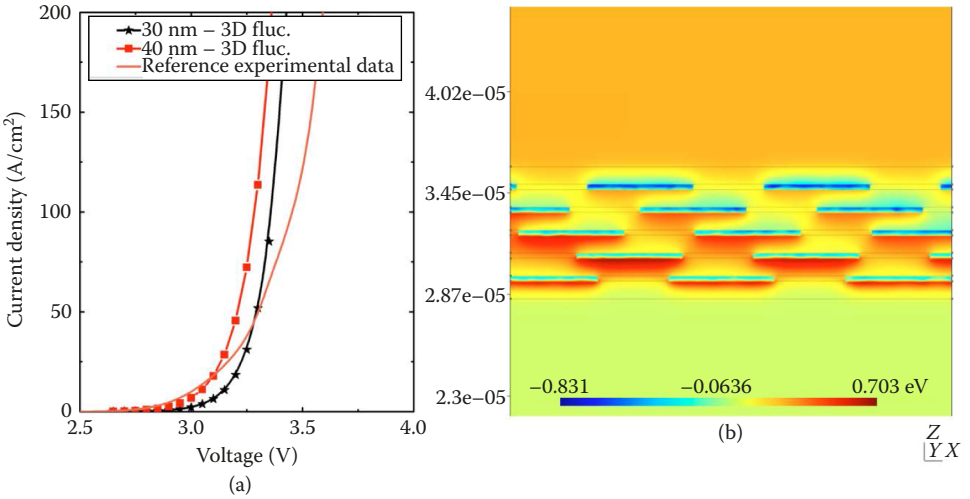

FIGURE 18.26 (a) The *I–V* curves of the different widths of the void region taking into consideration alloy fluctuations [45]. (b) The conduction band potential of the 40 nm case, where the location of the void regions is consistent, is at 3.05 V, which is near the turn-on point at 20 A/cm^2.

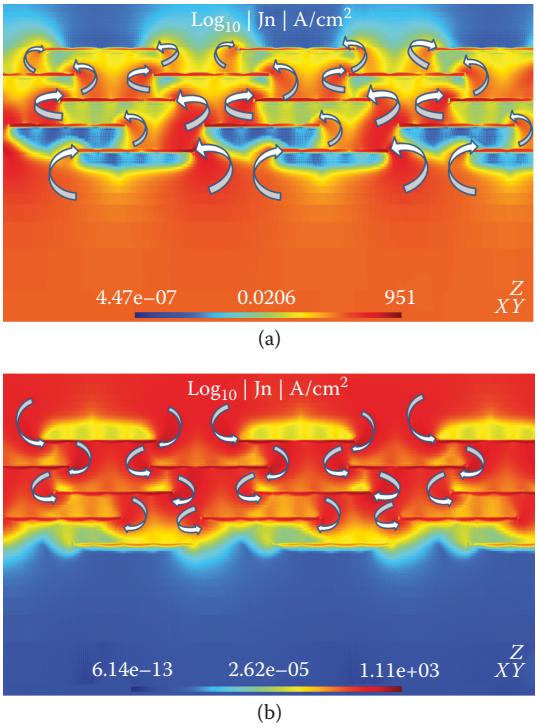

FIGURE 18.27 (a) The calculated vector profile of electron current. (b) The calculated vector profile of hole current.

carrier percolation. Although more reports of microscope observations in green emitting QWs are needed to prove that the structure of green LEDs are commonly imperfect, this preliminary simulated survey can still provide some hints for future modeling of the physical properties of green LEDs.

18.6 Summary

For the sake of simplicity, the simulation of the InGaN/GaN multiple QW LED usually assumes the uniform QW. However typical simulation results, such as the extremely high forward voltage (V_f), were far from the experimental data. These extra voltages will make the interpretation of carrier transport inaccurate. Because the extra voltage applied in the junction does not appear in the real device and will change the current transport mechanism from diffusion process into drift process. Therefore, it is important to make voltage fitting at accurate as possible. The simulation results show that by including the random indium fluctuation distribution in the QW, we can get a more reasonable fitting in the I–V curve.

For the transport issue with the random alloy fluctuation in the unipolar system, the simulation results of n-i-n GaN/InGaN/GaN MQW structures show that the fluctuated piezoelectric barrier induced by random alloy fluctuations will provide some current flow paths for the percolation transport, and the fluctuation model can give a reliable explanation for how the carriers perform in real devices. In modeling the emission spectrum, we found that the indium fluctuation will lead to a broader emission spectrum. The main reason is that the indium fluctuation forms several In-rich clusters, which look like quantum dots. The carriers localize in the QD-like regions and lead to different levels of quantum confinement. The effective bandgaps at each region vary from the different indium composition and the confined size of the indium clusters, which are the two important roles to affecting the discrete energy of quantum confinement. In addition, the emission strength will also be enhanced due to the better electron–hole localization effects.

We also analyzed the influence of indium fluctuations on the carrier transport. The simulation result of the *I–V* curve shifts to a value that is close to the experimental data. It indicates that the indium fluctuation does strongly affect the transport in the device. The main reason is that the low indium composition region will reduce the large potential barrier induced by the QCSE. For the uniform QW structure, the large V_f mainly comes from the strong band bending, which makes carriers hard to travel across the quantum barrier. For the indium fluctuation case, the reduction of the band bending at low indium regions provides a leaky way for carriers to flow through. The tunneling effect is not included in the indium fluctuation model but this will be a future project and may push the *I–V* curve to an even lower value. Furthermore, the IQE peaks occur at about 3.0 V, which also matches up to the experimental observation. Based on these simulation results, we tried to explain the cause of the droop effect. At a low bias condition, the droop effect is mainly dominated by the Auger recombination since the voltage is much lower than V_{BI}. When the bias gets higher and reaches the built-in voltage, the ratio of the droop influenced by carrier leakage will increase since the carrier transport turns from diffusion process into drift process at the reverse band condition.

Finally, we analyzed the green LED. We found that green emitting LEDs are difficult to use the traditional model to describe their physical property. Even for simulations that consider nanoscale alloy disorders, the calculated result differs from the experimental result. According to some reports, the green LEDs might exist in large-scale fluctuations, such as imperfect QWs in devices. Hence, we applied 3D models to examine the performance of the LEDs with imperfect QWs. The results show that the void region in the QWs provide an alternative path for carrier percolation. Thus, the IV of our simulation cases may approach the experimental observation more closely. In the future, we can model the green emission LEDs more elaborately based on the imperfect QW model.

References

1. H.-Y. Ryu, D.-S. Shin, and J.-I. Shim. Analysis of efficiency droop in nitride light-emitting diodes by the reduced effective volume of InGaN active material. *Applied Physics Letters*, 100(13):131109, 2012.
2. X. Ni, X. Li, J. Lee, S. Liu, V. Avrutin, A. Matulionis, U. Ozgur, and H. Morkoc. Pivotal role of ballistic and quasi-ballistic electrons on led efficiency. *Superlattices and Microstructures*, 48:133–153, 2010.
3. F. Akyol, D. N. Nath, S. Krishnamoorthy, P. S. Park, and S. Rajan. Suppression of electron overflow and efficiency droop in n-polar GaN green light emitting diodes. *Applied Physics Letters*, 100(11):111118, 2012.
4. F. Bertazzi, M. Goano, and E. Bellotti. A numerical study of Auger recombination in bulk InGaN. *Applied Physics Letters*, 97(23):231118, 2010.
5. K. T. Delaney, P. Rinke, and C. G. V. de Walle. Auger recombination rates in nitrides from first principles. *Applied Physics Letters*, 94(19):191109, 2009.
6. J. Hader, J. V. Moloney, B. Pasenow, S. W. Koch, M. Sabathil, N. Linder, and S. Lutgen. On the importance of radiative and Auger losses in GaN-based quantum wells. *Applied Physics Letters*, 92(26):261103, 2008.
7. J. Iveland, L. Martinelli, J. Peretti, J. S. Speck, and C. Weisbuch. Direct measurement of Auger electrons emitted from a semiconductor light-emitting diode under electrical injection: Identification of the dominant mechanism for efficiency droop. *Physical Review Letters*, 110:177406, April 2013.
8. H.-Y. Ryu, H.-S. Kim, and J.-I. Shim. Rate equation analysis of efficiency droop in InGaN light-emitting diodes. *Applied Physics Letters*, 95(8):081114, 2009.
9. Y. C. Shen, G. O. Mueller, S. Watanabe, N. F. Gardner, A. Munkholm, and M. R. Krames. Auger recombination in InGaN measured by photoluminescence. *Applied Physics Letters*, 91(14):141101, 2007.

10. J. Hader, J. V. Moloney, and S. W. Koch. Density-activated defect recombination as a possible explanation for the efficiency droop in GaN-based diodes. *Applied Physics Letters*, 96(22):221106, 2010.
11. J. Hader, J. V. Moloney, and S. W. Koch. Temperature-dependence of the internal efficiency droop in GaN-based diodes. *Applied Physics Letters*, 99(18):181127, 2011.
12. N. Okada, H. Kashihara, K. Sugimoto, Y. Yamada, and K. Tadatomo. Controlling potential barrier height by changing V-shaped pit size and the effect on optical and electrical properties for InGaN/GaN based light-emitting diodes. *Journal of Applied Physics*, 117(2):025708, 2015.
13. J. Smalc-Koziorowska, E. Grzanka, R. Czernecki, D. Schiavon, and M. Leszczyski. Elimination of trench defects and V-pits from InGaN/GaN structures. *Applied Physics Letters*, 106(10):101905, 2015.
14. H. Wang, X. Wang, Q. Tan, and X. Zeng. V-defects formation and optical properties of InGaN/-GaN multiple quantum well LED grown on patterned sapphire substrate. *Materials Science in Semiconductor Processing*, 29:112–116, 2015.
15. E. Kioupakis, P. Rinke, K. T. Delaney, and C. G. Van de Walle. Indirect Auger re-combination as a cause of efficiency droop in nitride light-emitting diodes. *Applied Physics Letters*, 98(16):161107, 2011.
16. S. E. Bennett, D. W. Saxey, M. J. Kappers, J. S. Barnard, C. J. Humphreys, G. D. Smith, and R. A. Oliver. Atom probe tomography assessment of the impact of electron beam exposure on In*x*Ga1−*x*N/GaN quantum wells. *Applied Physics Letters*, 99(2):021906, 2011.
17. D. A. Browne, B. Mazumder, Y.-R. Wu, and J. S. Speck. Electron transport in unipolar InGaN/GaN multiple quantum well structures grown by NH_3 molecular beam epitaxy. *Journal of Applied Physics*, 117(13):185703, 2015.
18. M. J. Galtrey, R. A. Oliver, M. J. Kappers, C. J. Humphreys, P. H. Clifton, D. Larson, D. W. Saxey, and A. Cerezo. Three-dimensional atom probe analysis of green and blue-emitting In*x*Ga1−*x*N/GaN multiple quantum well structures. *Journal of Applied Physics*, 104(1):013524, 2008.
19. M. J. Galtrey, R. A. Oliver, M. J. Kappers, C. J. Humphreys, D. J. Stokes, P. H. Clifton, and A. Cerezo. Three-dimensional atom probe studies of an In*x*Ga1−*x*N/GaN multiple quantum well structure: Assessment of possible indium clustering. *Applied Physics Letters*, 90(6):061903, 2007.
20. J. R. Riley, T. Detchprohm, C. Wetzel, and L. J. Lauhon. On the reliable analysis of indium mole fraction within In*x*Ga1−*x*N quantum wells using atom probe tomography. *Applied Physics Letters*, 104(15):152102, 2014.
21. R. Shivaraman, Y. Kawaguchi, S. Tanaka, S. DenBaars, S. Nakamura, and J. Speck. Comparative analysis of 2021 and 2021 semipolar GaN light emitting diodes using atom probe tomography. *Applied Physics Letters*, 102(25):251104, 2013.
22. D. Watson-Parris, M. J. Godfrey, P. Dawson, R. A. Oliver, M. J. Galtrey, M. J. Kappers, and C. J. Humphreys. Carrier localization mechanisms in In*x*Ga1−*x*N/GaN quantum wells. *Physical Review B*, 83:115321, March 2011.
23. Y.-R. Wu, R. Shivaraman, K.-C. Wang, and J. S. Speck. Analyzing the physical properties of InGaN multiple quantum well light emitting diodes from nano scale structure. *Applied Physics Letters*, 101(8):083505, 2012.
24. T.-J. Yang, R. Shivaraman, J. S. Speck, and Y.-R. Wu. The influence of random indium alloy fluctuations in indium gallium nitride quantum wells on the device behavior. *Journal of Applied Physics*, 116(11):113104, 2014.
25. M. A. Caro, S. Schulz, and E. P. O'Reilly. Theory of local electric polarization and its relation to internal strain: Impact on polarization potential and electronic properties of group-III nitrides. *Physical Review B*, 88:214103, December 2013.
26. C.-K. Wu, C.-K. Li, and Y.-R. Wu. Percolation transport study in nitride based LED by considering the random alloy fluctuation. *Journal of Computational Electronics*, 14(2):416–424, 2015.

27. C.-K. Li, H.-C. Yang, T.-C. Hsu, Y.-J. Shen, A.-S. Liu, and Y.-R. Wu. Three dimensional numerical study on the efficiency of a core-shell InGaN/GaN multiple quantum well nanowire light-emitting diodes. *Journal of Applied Physics*, 113(18):183104, 2013.
28. C.-K. Li, P.-C. Yeh, J.-W. Yu, L.-H. Peng, and Y.-R. Wu. Scaling performance of Ga_2O_3/GaN nanowire field effect transistor. *Journal of Applied Physics*, 114(16):163706, 2013.
29. C. Geuzaine and J.-F. Remacle. Gmsh: A 3-D finite element mesh generator with built-in pre- and post-processing facilities. *International Journal for Numerical Methods in Engineering*, 79:1309–1331, 2009.
30. M. Sabathil, A. Laubsch, and N. Linder. Self-consistent modeling of resonant PL in InGaN SQW LED-structure. *SPIE Proceedings*, 6486:64860V–648609, 2007.
31. I. Vurgaftman, J. R. Meyer, and R. L. R. Mohan. Band parameters for III–V compound semiconductors and their alloys. *Journal of Applied Physics*, 89(11):5815–5875, 2001.
32. S. L. Chuang and C. S. Chang. k p method for strained wurtzite semiconductors. *Physical Review B*, 54:2491–2504, July 1996.
33. J. Wu. When group-III nitrides go infrared: New properties and perspectives. *Journal of Applied Physics*, 106(1):011101, 2009.
34. T. R. Chandrupatla., et al. *Introduction to Finite Elements in Engineering.* Vol. 2. Upper Saddle River, NJ: Prentice Hall, 2002.
35. A. Romanov, T. Baker, S. Nakamura, and J. Speck. Strain-induced polarization in wurtzite III-nitride semipolar layers. *Journal of Applied Physics*, 100(2):023522, 2006.
36. Y.-R. Wu, Y.-Y. Lin, H.-H. Huang, and J. Singh. Electronic and optical properties of InGaN quantum dot based light emitters for solid state lighting. *Journal of Applied Physics*, 105(1):013117, 2009.
37. D. A. Browne, B. Mazumder, Y.-R. Wu, and J. S. Speck. Investigation of electron transport through InGaN quantum well structures. In *14th Electronic Materials Conference,* Santa Barbara, CA, June 25–27, 2014.
38. O. Ambacher, J. Majewski, C. Miskys, A. Link, M. Hermann, M. Eickhoff, M. Stuzmann, F. Bernardini, V. Fiorentini, V. Tilak, B. Schaff, and L. F. Eastman. Pyroelectric properties of Al(In)GaN/GaN hetero- and quantum well structures. *Journal of Physics: Condensed Matter*, 14:3399–3434, 2002.
39. O. Ambacher, J. Smart, J. R. Shealy, N. G. Weimann, K. Chu, M. Murohy, W. J. Schaff, L. F. Eastman, R. Dimitrov, L. Wittmer, M. Stutzmann, W. Rieger, and J. Hilsenbeck. Two-dimensional electron gases induced by spontaneous and piezo-electric polarization charges in N- and Ga-face AlGaN/GaN heterostructures. *Journal of Applied Physics*, 85(6):3222–3233, March 1999.
40. D. N. Nath, Z. C. Yang, C.-Y. Lee, P. S. Park, Y.-R. Wu, and S. Rajan. Unipolar vertical transport in GaN/AlGaN/GaN heterostructures. *Applied Physics Letters*, 103(2):022102, 2013.
41. J. Piprek. Efficiency droop in nitride-based light-emitting diodes. *Physica Status Solidi (A)*, 207(10):2217–2225, 2010.
42. H.-Y. Ryu and J.-I. Shim. Effect of current spreading on the efficiency droop of InGaN light-emitting diodes. *Optics Express*, 19(4):2886–2894, February 2011.
43. M. Calciati, M. Goano, F. Bertazzi, M. Vallone, X. Zhou, G. Ghione, M. Meneghini, G. Meneghesso, E. Zanoni, E. Bellotti, G. Verzellesi, D. Zhu, and C. Humphreys. Correlating electroluminescence characterization and physics-based models of InGaN/GaN LEDs: Pitfalls and open issues. *AIP Advances*, 4(6):067118, 2014.
44. C. Kang Li, M. Rosmeulen, E. Simoen, and Y.-R. Wu. Study on the optimization for current spreading effect of lateral GaN/InGaN LEDs. *IEEE Transactions on Electron Devices*, 61(2):511–517, February 2014.
45. C.-H. Lu, Y.-C. Li, Y.-H. Chen, S.-C. Tsai, Y.-L. Lai, Y.-L. Li, and C.-P. Liu. Out-put power enhancement of InGaN/GaN based green light-emitting diodes with high-density ultra-small In-rich quantum dots. *Journal of Alloys and Compounds*, 555:250–254, 2013.

46. W. Lv, L. Wang, L. Wang, Y. Xing, D. Yang, Z. Hao, and Y. Luo. InGaN quantum dot green light-emitting diodes with negligible blue shift of electroluminescence peak wavelength. *Applied Physics Express*, 7(2):025203, 2014.
47. N. Van der Laak, R. Oliver, M. Kappers, and C. Humphreys. Characterization of InGaN quantum wells with gross fluctuations in width. *Journal of Applied Physics*, 102(1):013513, 2007.
48. N. K. van der Laak, R. A. Oliver, M. J. Kappers, and C. J. Humphreys. Role of gross well-width fluctuations in bright, green-emitting single InGaN/GaN quantum well structures. *Applied Physics Letters*, 90(12):121911–121911, 2007.

19

Superluminescent Light-Emitting Diodes

Nicolai Matuschek
and
Marcus Duelk

19.1 Introduction

More than 30 years ago, the concept of a superluminescent light-emitting diode (SLED) was proposed for the first time (Kaminow and Marcuse 1983). It can be briefly described as an edge-emitting semiconductor light source that operates in the so-called superluminescence regime, also known as the amplified spontaneous emission (ASE) regime. This means that the gain medium is pumped to a level beyond transparency but below the threshold for starting lasing activity. As a consequence of this particular operating range, SLEDs combine some aspects of the electro-optical performance from standard light-emitting diodes (LEDs) and laser diodes (LDs). They are able to produce high-output powers with high brightness similar to high-power LDs. On the other hand, they show a broadband emission spectrum similar to LEDs. With respect to their coherence properties, this translates into high spatial coherence similar to LDs and low temporal coherence like LEDs. A more detailed comparison is given in Rossetti et al. (2012).

Due to these intermediate characteristics, SLEDs are preferred light sources for many applications, including fiber-optic gyroscopes (FOGs) (Burns et al. 1983), fiber-optic current sensors (FOCSs) (Bohnert et al. 2002), optical coherence tomography (OCT) (Schmitt 1999; Drexler and Fujimoto 2008), structural health monitoring with optical fiber sensors (Wild and Hinckley 2009), speckle-free illumination (Rossetti et al. 2012), metrology systems (Dufour et al. 2005), or optical test equipment for fiber-optic networks (Senior 2009). In order to meet the different requirements of these applications, broadband SLEDs at various emission wavelengths with different spectral bandwidths and shapes are used. SLEDs can be designed for wavelengths ranging from 390 nm up to 2700 nm. They can be realized in GaN (390–570 nm), GaAs (570–1150 nm), InP (1150–2000 nm), or GaSb (2000–2700 nm). Figure 19.1 shows a selection of typical ASE spectra obtained for various commercially available SLED modules operating at center wavelengths from around 400 to 1550 nm (Exalos 2017).

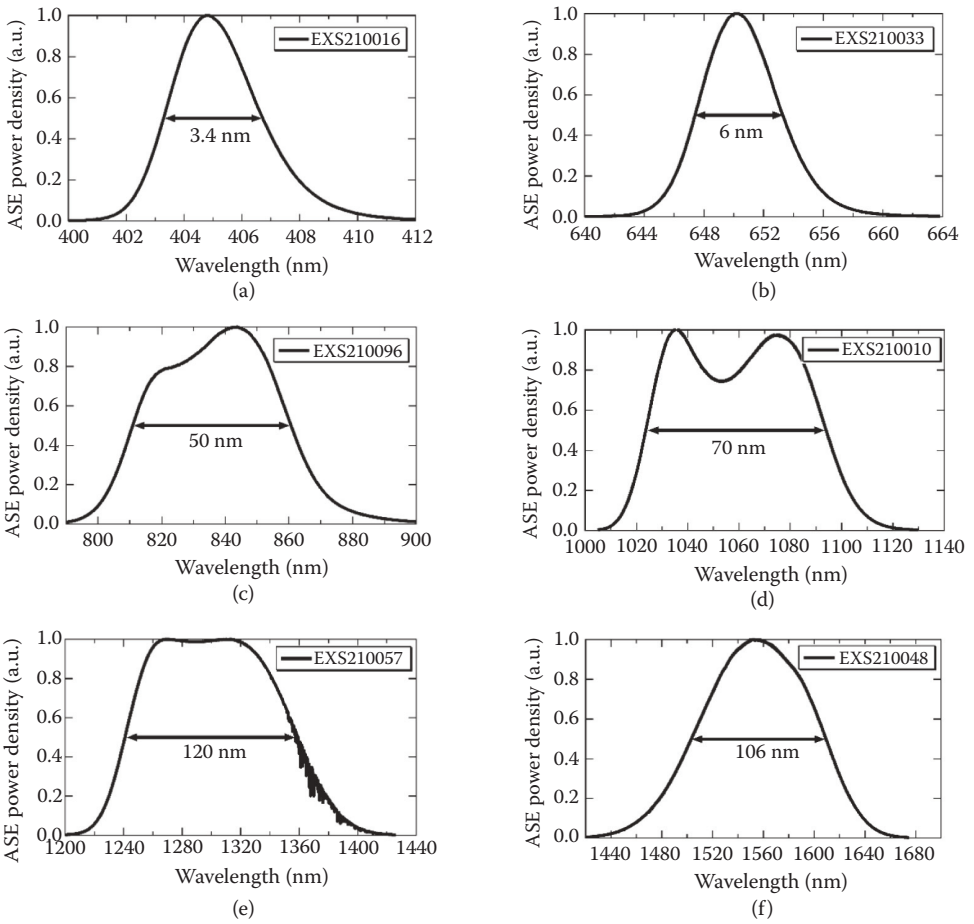

FIGURE 19.1 Measured amplified spontaneous emission (ASE) spectra (on a linear scale) obtained for six different superluminescent light-emitting diode (SLED) modules operating at relevant wavelength ranges from about 400 to 1550 nm. The full-width at half maximum (FWHM) or 3-dB bandwidth is given in the diagrams. The Gaussian-shaped spectra shown in (a), (b), and (f) are based on a single transition epitaxial design, whereas the spectra with two humps shown in (c), (d), and (e) are based on a multitransition epitaxial design, as explained in Section 19.3.1.

Traditionally, the vast majority of SLEDs are employed in systems with light emission in the near infrared from 750 to 1600 nm (see Figure 19.1c through f). Over the last decade, the emerging interest in speckle-free red–green–blue (RGB) applications forced the development of SLEDs operating in the visible wavelength range (see Figure 19.1a and b). Based on the technology for red LDs, which has been well established over the last 30 years, the demonstration of a red SLED realized in GaAs was a straightforward task (Semenov et al. 1993a). In contrast, the development of blue and green SLEDs has been confronted with a couple of intrinsic technological difficulties (Rossetti et al. 2012). One of them is the choice of the substrate material, which can be GaN, SiC, sapphire, or others. For SLEDs and LDs, high-quality substrate materials with good thermal conductivity and low dislocation densities are considered to be of utmost importance. Therefore, past development activities have focused on the growth of the epitaxial layer structure on free-standing (mainly *c*-plane) GaN substrates. The first GaN-based SLEDs emitting in the blue violet spectral region could be demonstrated just a few years ago (Feltin et al. 2009; Rossetti et al. 2010), whereas research and development for the demonstration of green SLEDs is still ongoing. Very recently, the first long-wavelength SLED realized in GaSb for optical sensing at 2.4 μm has been demonstrated (Wootten et al. 2014).

19.2 Design and Modeling of SLED Devices

19.2.1 SLED Geometry and Epitaxial Design

Typically, SLEDs are realized as epitaxially grown layers on a substrate in an index-guided ridge-waveguide geometry similar to Fabry–Pérot-type LDs. The latter consists of a straight waveguide with optical dielectric coatings applied at the facets of the chip, typically a high-reflective (HR) at the back facet and a partial-reflective (PR) coating at the front facet (see, e.g., Macleod 2010). The straight waveguide in combination with reflecting facets defines a resonant cavity that allows for the build-up of longitudinal cavity modes and laser oscillations. In contrast, several tricks are applied to SLEDs for the suppression of resonant cavity modes in order to minimize modulations in the ASE spectrum (so-called spectral ripples) or, even worse, to prevent the starting of lasing. The most popular and very efficient way to achieve this is the tilting of the waveguide by a few degrees with respect to the normal of the facets (Alphonse et al. 1988) in combination with the application of antireflection (AR) coatings (Macleod 2010) at the facet sides. The residual net modal reflectivity of such a design can be calculated using the analytical model presented in Marcuse (1989). Exemplary reflectivity calculations as a function of the tilt angle are given in Matuschek and Duelk (2013). Other methods are based on the incorporation of an absorber section at the back-facet side of the waveguide (Patterson et al. 1994; Kwong et al. 2008), the use of a bent (Semenov et al. 1993b), or tapered (Koyama et al. 1993) waveguide structure, or a combination of various methods (see, e.g., Lee et al. 1973; Nagai et al. 1989; Semenov et al. 1993a; Middlemast et al. 1997).

The ASE process occurs in the active region along the SLEDs waveguide structure. The most common approach to realize a positive material gain is based on a single-quantum well (SQW) or multiquantum well (MQW) active region design (see, e.g., Chuang 1995). For some applications, a bulk layer approach might be advantageous, as it is the case, for example, for SLEDs that require a low polarization extinction ratio (PER),[†] that is, the amount of transverse electric (TE)- and transverse magnetic (TM)-polarized output power should be as equal as possible (Heo et al. 2011). Quantum-dot epi structures for SLEDs have been proposed (Sun et al. 1999) and demonstrated a few times as well. However, this technology has various severe limitations, like a limited wavelength range (so far 1050–1200 nm), low differential gain requiring long SLED chips and large drive currents, manufacturing challenges, and others (Rossetti et al. 2008).

Figure 19.2 shows the schematic front view of a typical SLED structure for an SQW epitaxial layer design. Generally, quantum well (QW) active-region designs allow for the use of strained layers. Compressively strained QW layers are required for SLED designs showing a high PER greater than ∼10 dB, which is desired for many applications. Maximum strain values up to 3% (with respect to the lattice constant of the substrate material) may be realized in practice. Higher strain values may lead to the formation of dislocations and are therefore detrimental to achieving a good long-term reliability. In addition to the relative strain of an individual layer, the total integrated net strain for the entire active region must be taken into account too. Hence, strain compensation by incorporation of tensile strained barrier layers is a useful option for strain relaxation over the active region (Tansu and Mawst 2001).

The QW and barrier layers are sandwiched between (undoped or partially doped) waveguide layers that are surrounded by highly doped *n*- and *p*-cladding layers. The high refractive index of the active-region and waveguide layers leads to the confinement of the optical mode in vertical direction. On the other hand, the ridge-waveguide geometry ensures the lateral mode confinement. As a result, the epitaxial layer stack realized in a ridge-waveguide geometry allows for the build-up of transverse optical modes with a lateral extension roughly given by the ridge width and a vertical extension as given by the thickness of the waveguide layers. For almost all applications, single-mode operation on the fundamental optical mode is preferred. Hence, care has to be taken to avoid the occurrence of higher-order modes or leaky substrate

[†] The PER is defined as $\mathrm{PER} = 10 \cdot \lg\left(\frac{P_{\mathrm{TE}}}{P_{\mathrm{TM}}}\right)$ dB, where P_{TE} and P_{TM} is the SLED's total output power from the front facet in TE and TM polarization, respectively.

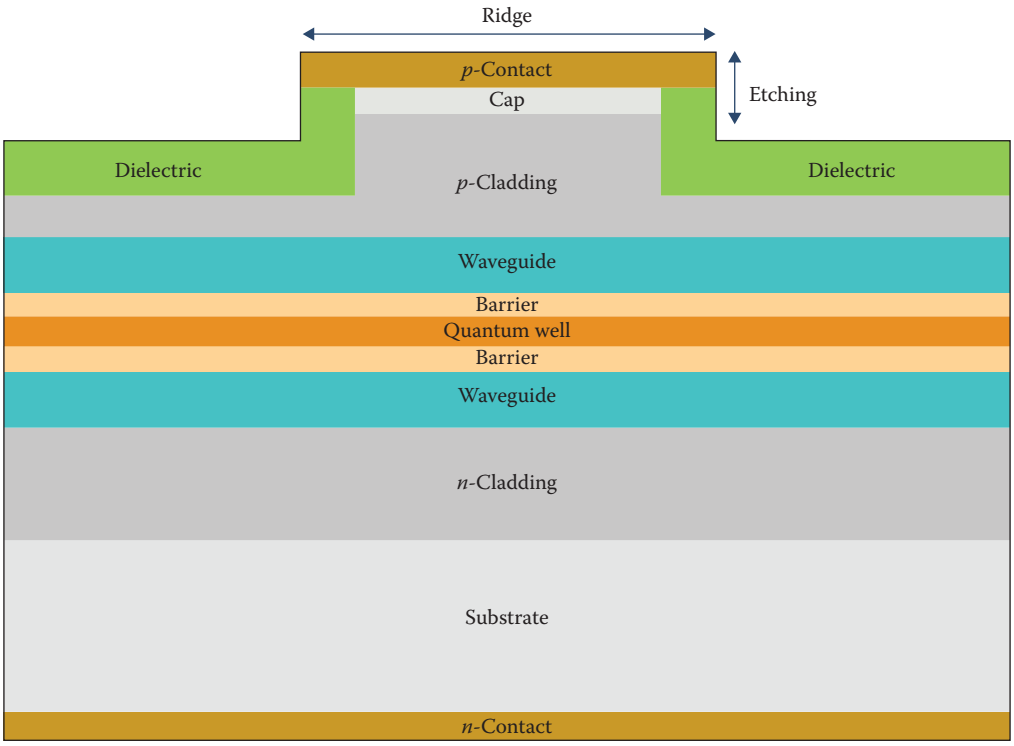

FIGURE 19.2 Basic layer structure of a single-quantum well (SQW) epitaxy. The QW and barrier layers are sandwiched between waveguide layers followed by the cladding layers on both sides. The ridge-waveguide geometry ensures the lateral guiding of the optical mode, whereas the vertical guiding is a consequence of the refractive-index profile defined by the material composition of the different layers.

modes. The latter are particularly a problem for GaN-based epi structures in the visible due to the high refractive index of the substrate material (Laino et al. 2007; Matuschek and Duelk 2013). Moreover, some output power might be lost into higher-order lateral modes if such modes are supported by the lateral waveguide (see, e.g., Coldren et al. 2012). The number of supported lateral modes, and thus, the corresponding power loss scales with the width of the ridge waveguide. Therefore, the stripe width is typically chosen to be smaller than 10 μm.

As mentioned earlier, the appropriate choice of semiconductor materials depends primarily on the wavelength range of interest, and thus, on the bandgap energy of available semiconductor compound material. In Table 19.1, the substrate materials and the most commonly used compound materials for the functional layers of the epitaxial structure, as shown in Figure 19.2, are summarized for the different wavelength regimes.

19.2.2 Modeling and Simulation Approaches

A couple of different approaches exist for the modeling and simulation of SLEDs. Generally, the analytical insight into the electro-optical performance decreases with the complexity of the set of coupled equations taken under consideration. On the other hand, the accuracy increases when more effects are taken into account.

Commercially available packages (Synopsis 2017; Crosslight Software 2017; Photon Design 2017) allow for full three-dimensional (3D) simulations. They are based on the coupling of solvers for the various problems involved in the simulation of an SLED (so-called multiphysics treatment) (Li and Li 2010). Typically,

TABLE 19.1 Materials Used for the Epitaxial Layers of an SLED Grown on a Substrate Depending on the Wavelength Range of Interest

Wavelength Range	390–570 nm	570–1150 nm	1150–2000 nm	2000–2700 nm
Substrate	GaN	GaAs	InP	GaSb
Cladding	$Al_xGa_{1-x}N$	$Al_xGa_{1-x}As$ $Al_xGa_yIn_{1-x-y}P$	InP $Al_xIn_{1-x}As$	$Al_xGa_{1-x}As_{1-y}Sb_y$
Waveguide	GaN	$Al_xGa_{1-x}As$ $Al_xGa_yIn_{1-x-y}P$	$Al_xGa_yIn_{1-x-y}As$ $In_xGa_{1-x}As_{1-y}P_y$	$Al_xGa_{1-x}As_{1-y}Sb_y$
Active region (QWs + barriers)	$In_xGa_{1-x}N$	$Al_xGa_yIn_{1-x-y}P$ $Al_xGa_yIn_{1-x-y}As$ $In_xGa_{1-x}As_{1-y}P_y$	$Al_xGa_yIn_{1-x-y}As$ $In_xGa_{1-x}As_{1-y}P_y$	$Al_xGa_{1-x}As_{1-y}Sb_y$ $In_xGa_{1-x}As_{1-y}Sb_y$

Note: The general definition of the quaternary materials includes the special case that the materials may reduce to ternary or binary materials if the material fractions x and/or y are equal to 0 or 1.

the quantum-mechanical problem is described using the **k·p** method and the optical modes are found from a Helmholtz equation (Chuang 1995). These equations are coupled to the semiconductor equations, for example, as Poisson equation and drift-diffusion equation and to a heat-transfer equation if the thermal properties are also taken into account (Loeser and Witzigmann 2008). All equations are then solved simultaneously and self-consistently for each iteration step, yielding a solution for all local and global variables of the system.

In analytical approaches both the physical dimensionality of the problem and the number of coupled equations are reduced. They are normally based on a traveling-wave equation for the electric field (Park and Li 2006) or the optical power (or photon density) (Matuschek and Duelk 2013; Milani et al. 2015) and on rate equations for the carrier density. The material gain is either calculated from quantum-mechanical equations or using simplified semianalytical equations assuming a logarithmic or linear gain dependence on the carrier density (Coldren et al. 2012). In order to calculate the net modal gain, both the optical confinement factor and the internal loss coefficient are required. The first can be calculated from the overlap integral of the optical modes with the active layers. The latter can be used, for instance, as a free fitting parameter or it can be extracted from an *inverse efficiency versus chip length* plot, which is obtained from the *L–I* characteristics of LDs measured for various chip lengths. The LD design for such measurements differs from the SLED design only in the straight ridge waveguide and its facets are typically uncoated.

19.2.3 *L–I* Characteristics of Reflecting SLEDs

SLEDs operate in the ASE regime, that is, spontaneously emitted photons induce the process of stimulated emission of photons while traveling along the waveguide. Hence, the number of forward- and backward-propagating photons grows exponentially in direction to both facets of the chip. On the other hand, the SLED's *L–I* characteristic follows to good approximation a power law with chip length–dependent exponent because of the logarithmic gain dependence on current density as shown in Matuschek and Duelk (2013). The analytic derivation of this power law was based on the assumption of an ideal SLED without any residual reflectivity at both facet sides.

Here, we want to extend this approach to the general case of SLEDs with arbitrary effective modal reflectivities at the chip's front- and back-facet sides. This derivation includes special cases of standard SLEDs and of so-called reflecting SLEDs (R-SLEDs). Typical SLEDs feature a tilted ridge-waveguide geometry with tilt angles smaller than 15° and high-quality AR coatings deposited at both facet sides in order to avoid any resonant cavity effect. In contrast, R-SLEDs consist of a tilted waveguide section in direction to the front facet with deposited AR coating and a straight waveguide section toward the chip's back facet with a deposited HR coating (Matuschek and Duelk 2014), as shown in Figure 19.3. As explained below in more detail, R-SLEDs are able to produce much higher output powers due to the double-pass geometry

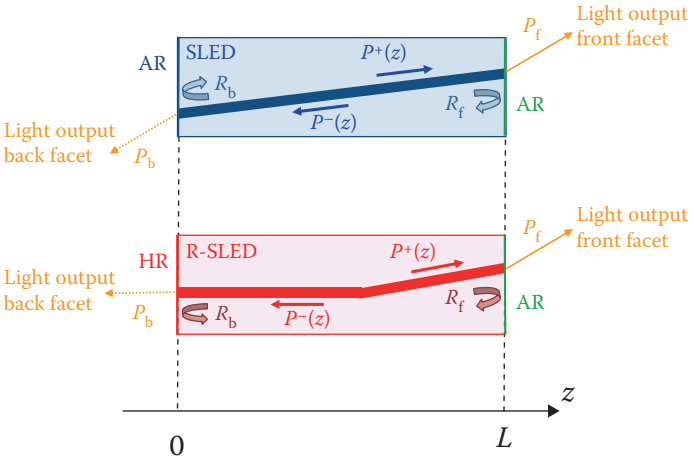

FIGURE 19.3 Schematic top view of an SLED and an R-SLED with effective modal reflectivities R_f and R_b. The power is exponentially amplified in both directions along the (partially) tilted ridge-waveguide structure, where P^+ denotes the power distribution for forward-propagating photons and P^- for backward-propagating photons, respectively. P_f and P_b are the powers, which are coupled out and can be measured at the front- and back-facet sides of the SLED or R-SLED. The z-axis is defined by the propagation direction of photons along the ridge-waveguide structure, so that the single-path length is slightly greater than the vertical distance between the facets due to the tilt.

compared to standard SLEDs under same operating conditions. However, the drawback is that such devices are much more susceptible to the appearance of ripples in the ASE spectrum and/or starting of undesired lasing operation.

In the following, the longitudinal power distribution for two counter-propagating waves is derived under the condition of steady-state continuous wave (CW) operation. Based on a one-dimensional (1D) traveling-wave amplifier approach similar to (Marcuse 1983; Matuschek et al. 2008), the differential equations

$$\frac{\mathrm{d}P^+(z)}{\mathrm{d}z} = \left(g_{\mathrm{mod}} - \alpha_{\mathrm{i}}\right) \cdot P^+(z) + \frac{n_{\mathrm{sp}}\hbar\omega v_{\mathrm{gr}}}{2L} g_{\mathrm{mod}}\,, \tag{19.1}$$

$$-\frac{\mathrm{d}P^-(z)}{\mathrm{d}z} = \left(g_{\mathrm{mod}} - \alpha_{\mathrm{i}}\right) \cdot P^-(z) + \frac{n_{\mathrm{sp}}\hbar\omega v_{\mathrm{gr}}}{2L} g_{\mathrm{mod}}\,, \tag{19.2}$$

determine the power distribution $P^{\pm}(z)$ for forward- and backward-propagating photons of the same longitudinal mode with angular frequency ω. $g_{\mathrm{mod}} = \Gamma g_{\mathrm{mat}}$ is the modal gain with g_{mat} being the material gain and Γ the mode confinement factor. α_{i} is the internal loss coefficient, n_{sp} the population inversion factor, L the effective chip length of the tilted ridge-waveguide structure, and $v_{\mathrm{gr}} = c/n_{\mathrm{gr}}$ the group velocity with the speed of light in vacuum, c, and the effective group index, n_{gr}. The independent variable (z-coordinate) is defined in the range $z \in [0;\ L]$, where the boundaries are defined by the chip's front and back facet (see Figure 19.3). The second term on the right-hand side describes the generation of light by spontaneously emitted photons, whereas the first term describes the stimulated amplification process by already existing photons.

It should be noted that, generally, there is one set of equations for each longitudinal mode. Here, we have omitted an index for the number of the longitudinal mode. We solve the equations for the mode closest to the gain maximum. Finally, the total power values are obtained by summing over an effective number of longitudinal modes, which is proportional to the full-width at half maximum (FWHM) or 3-dB bandwidth of the emitted ASE power spectrum (Matuschek and Duelk 2013). As explained in

detail in Marcuse (1983), *it is not necessary to treat this ASE process as a coherent signal since it distributes itself continuously over a relatively wide band of wavelengths with random phases between adjacent wavelength components*. Hence, a description using powers instead of fields with definite amplitude and phase is sufficient.

Equations 19.1 and 19.2 can be easily solved if we assume that the gain is homogeneous in longitudinal direction, that is, we neglect any gain saturation effects. This approximation is justified for low-power SLEDs but might be questionable for SLEDs and R-SLEDs operating at high injection currents, and thus, at high-output powers, as will be shown at the end of this section. Then the general solution is given by

$$P^{+}(z) = P_0^{+} \cdot G(z) + a \cdot \{G(z) - 1\}\,, \tag{19.3}$$

$$P^{-}(z) = P_L^{-} \cdot G(L - z) + a \cdot \{G(L - z) - 1\}\,, \tag{19.4}$$

with

$$a = \frac{n_{\mathrm{sp}} \hbar \omega v_{\mathrm{gr}}}{2L} \frac{g_{\mathrm{mod}}}{g_{\mathrm{mod}} - \alpha_{\mathrm{i}}}, \tag{19.5}$$

where we have introduced the gain amplification factor

$$G(z) = \exp\left\{(g_{\mathrm{mod}} - \alpha_{\mathrm{i}})z\right\}. \tag{19.6}$$

The subscripts 0 and L denote power values at positions $z = 0$ and $z = L$ inside the chip.

Equations 19.3 and 19.4 are coupled by the boundary conditions

$$P_L^{-} = R_{\mathrm{f}} \cdot P_L^{+}, \tag{19.7}$$

$$P_0^{+} = R_{\mathrm{b}} \cdot P_0^{-}, \tag{19.8}$$

where R_{f} and R_{b} are the effective modal reflectivities at the front and back sides, respectively. Using these boundary conditions, we can derive the expressions

$$P_L^{+} = a \cdot \frac{1 + G(L)R_{\mathrm{b}}}{1 - G^2(L)R_{\mathrm{b}}R_{\mathrm{f}}} \cdot (G(L) - 1)\,, \tag{19.9}$$

$$P_0^{-} = a \cdot \frac{1 + G(L)R_{\mathrm{f}}}{1 - G^2(L)R_{\mathrm{b}}R_{\mathrm{f}}} \cdot (G(L) - 1)\,, \tag{19.10}$$

for the light powers that hit the facets from inside the chip. It should be noted that Equations 19.9 and 19.10 are symmetric with respect to an exchange of the chip's front- and back-facet sides.

The power coupled out of the chip is obtained by multiplying (Equations 19.9 and 19.10) with the effective transmittance of the front- and back-facet coatings, T_{f} and T_{b}.[†] Thus, taking the values of all variables at the spectral position of the gain maximum, which we indicate by adding a subscript index 0 at each

[†] Generally, the inequality $R_{\mathrm{f/b}} + T_{\mathrm{f/b}} < 1$ holds because of slight absorption that might occur in the applied dielectric coatings. Moreover, for tilted SLED sections light that is not reflected back at the facet into the waveguide but elsewhere, is lost and not covered by the reflectivities $R_{\mathrm{f/b}}$.

relevant variable and summing over the number of effectively contributing modes finally leads to

$$P_f = \tilde{P} \cdot \frac{1 + G_0(L)R_\mathrm{b}}{1 - G_0(2L)\,R_\mathrm{b}R_\mathrm{f}} \cdot \left(G_0(L) - 1\right) \cdot T_\mathrm{f}, \tag{19.11}$$

$$P_\mathrm{b} = \tilde{P} \cdot \frac{1 + G_0(L)R_\mathrm{f}}{1 - G_0(2L)\,R_\mathrm{b}R_\mathrm{f}} \cdot \left(G_0(L) - 1\right) \cdot T_\mathrm{b}. \tag{19.12}$$

The prefactor $\tilde{P}$ is given by

$$\begin{aligned} \tilde{P} &= n_\mathrm{eff} a_0 = n_\mathrm{eff} \cdot \frac{n_\mathrm{sp}\hbar\omega_0 v_\mathrm{gr,0}}{2L} \frac{g_\mathrm{mod,0}}{g_\mathrm{mod,0} - \alpha_\mathrm{i,0}} \\ &= \Delta\lambda_\mathrm{3dB} \cdot \frac{n_\mathrm{sp} hc^2}{\lambda_0^3} \frac{g_\mathrm{mod,0}}{g_\mathrm{mod,0} - \alpha_\mathrm{i,0}}, \end{aligned} \tag{19.13}$$

where we have estimated the effective mode number, n_eff, by the 3-dB bandwidth of the ASE output spectrum, $\Delta\lambda_\mathrm{3dB}$, and the longitudinal mode spacing, $\lambda_0^2/(2n_\mathrm{gr,0}L)$. Strictly speaking, the replacement of the sum by an effective mode number requires that the dispersion of the effective modal facet reflectivities and coating transmittance can be neglected over the relevant spectral range.

Equations 19.11 and 19.12 together with Equation 19.13 represent the general solution for the front- and back-facet output power of an SLED with arbitrary front- and back-side reflectivity. They are particularly useful for analyzing differences observed in the output power from both facets caused by residual facet reflections. In Matuschek et al. (2008), very similar expressions were derived and a short discussion of the consequences on the front-to-back output power ratio was given. In the following, we apply our general results to two special cases of an SLED, namely an ideal standard SLED and an ideal R-SLED.

19.2.3.1 Ideal Standard SLED

An SLED device is assumed where both facets do not have any residual reflection. This condition is nearly fulfilled for real devices with sufficiently great tilt angle and AR-coated facets. In this limit, we can set $R_\mathrm{f} = R_\mathrm{b} = 0$ and $T_\mathrm{f} = T_\mathrm{b} = 1$ and obtain

$$P_f = P_\mathrm{b} = \tilde{P} \cdot \left(G_0(L) - 1\right) = \Delta\lambda_\mathrm{FWHM} \cdot \frac{n_\mathrm{sp} hc^2}{\lambda_0^3} \frac{g_\mathrm{mod\,,0}}{g_\mathrm{mod\,,0} - \alpha_\mathrm{i,0}} \cdot \left(e^{(g_\mathrm{mod\,,0} - \alpha_\mathrm{i,0})\cdot L} - 1\right). \tag{19.14}$$

Inserting a logarithmic gain model in Equation 19.14 yields a power law for the SLED's *L–I* characteristic with chip length–dependent exponent. The consequences of this power law have been discussed and experimentally proven in detail in Matuschek and Duelk (2013) and are not repeated here.

19.2.3.2 Ideal R-SLED

An SLED device is assumed with a perfect front facet having zero residual reflection, similar to an ideal standard SLED, but with a back facet being a perfect reflector that reflects all incident light back into the waveguide. For real devices, this condition is fairly well realized with effective modal reflectivity values of about 90% or even higher for the HR coating. Moreover, the waveguide at the back facet must be straight or perpendicular with respect to the facet in order to avoid any geometrical reduction of the effective modal reflectivity. In the limit of an ideal R-SLED we can set $R_\mathrm{f} = T_\mathrm{b} = 0$ and $R_\mathrm{b} = T_\mathrm{f} = 1$ and obtain

$$P_f = \tilde{P} \cdot \left(G_0(2L) - 1\right) = \Delta\lambda_\mathrm{FWHM} \cdot \frac{n_\mathrm{sp} hc^2}{\lambda_0^3} \frac{g_\mathrm{mod\,,0}}{g_\mathrm{mod\,,0} - \alpha_\mathrm{i,0}} \cdot \left(e^{2\cdot(g_\mathrm{mod\,,0} - \alpha_\mathrm{i,0})\cdot L} - 1\right) \tag{19.15}$$

for the total power out of the front facet. It should be obvious that under those assumptions no power will exit the back facet of the chip, that is, $P_b = 0$.

Comparing Equation 19.15 with Equation 19.14, it follows directly that with respect to output power from the front facet an R-SLED of length L behaves like a standard SLED of twice the length $2L$. This result seems intuitive since the amplification process for photons is extended over a distance, which corresponds to one round-trip in the chip with perfect reflection at the back facet. However, it is important to realize that with respect to current injection the chip is still a short chip of length L. This means that an R-SLED of length L emits the same power out of the front facet as a standard SLED of length $2L$ at same current density values J or, in other words, at half values of the injection current I. Also, under these conditions the average carrier densities are approximately the same and, thus, the emitted ASE output spectra with their corresponding FWHM bandwidth. Consequently, the prefactor of Equation 19.13 has the same value in both cases.

19.2.4 Experimental and Numerical Verification of the Analytical Model

In this section, we want to verify the results derived from the analytical treatment by comparing them with experimental results as well as results obtained from a full 3D simulation of the SLED and R-SLED under consideration. The simulation software is based on a finite-element discretization of the device. Using the **k·p** method, the quantum mechanical system is described by a 6×6 Luttinger–Kohn Hamiltonian and the eigenvalue problem is solved as well as a Helmholtz equation for the optical modes (Chuang 1995). These equations are coupled to the semiconductor equations and thermal equations as described earlier. All equations are then solved self-consistently on a discrete mesh.

As an object for our study, we use SLED and R-SLED chips from the same epi wafer with light emission in the wavelength region around 860 nm. All chips are operated at a heat-sink temperature of 25°C. The MQW epi structure of these chips is based on (Al)GaAs/(In)GaAs layers grown on a GaAs substrate (see Table 19.1). For the simulation, the effective modal reflectivity values have not been set to the ideal values 0 and 1 but to values of 10^{-9} for the AR-coated facets with tilted waveguide section and 0.95 for the back facet with straight waveguide section. These values are quite realistic for well-designed SLED and R-SLED structures. The simulator has been calibrated by adjusting the L–I characteristics and ASE output spectra for short standard SLEDs with chip lengths ranging from 500 to 850 μm.

Figure 19.4 shows the results obtained for the 850-μm-long SLED chip after the calibration procedure. Obviously, the measured L–I curve is almost exactly reproduced by the simulation. Moreover, the detailed shape of the measured ASE spectra is also simulated very accurately for different levels of injection currents. The quality of the analytical model is shown by the dotted L–I characteristic, which is obtained from Equation 19.14 by following the procedure described in Matuschek and Duelk (2013). The center wavelength and 3-dB bandwidth have been taken from the ASE spectrum at an injection current of 100 mA. The modal gain as a function of the current density has been derived from the full 3D simulation by averaging over the z-direction. Apparently, the output power at high-injection currents is somewhat overestimated by the analytical model because the relatively simple model does not take into account self-heating and gain saturation effects of the chip.

After successful calibration, we are able to compare the electro-optical properties of a 1750-μm standard SLED with an 840-μm R-SLED. The chip length ratio is 2.08, which is close to the ideal value of two. Moreover, the back-facet coating of the R-SLED chip is not a real HR coating but only a PR coating with a reflectance value of about 20%. Figure 19.5 shows the measured and simulated front-facet output power as a function of the current density for both types of SLEDs. For our considerations the effective width of current injection is not relevant. Therefore, we scale the current values only by the chip length so that the density values are defined by $J = I/L$. As can be seen, the agreement between all curves is excellent. Only the measured power of the R-SLED (dash-dotted line) is slightly smaller because of the nonideal reflectance of the PR coating. First of all, this proves that an R-SLED behaves like a standard SLED of twice the length

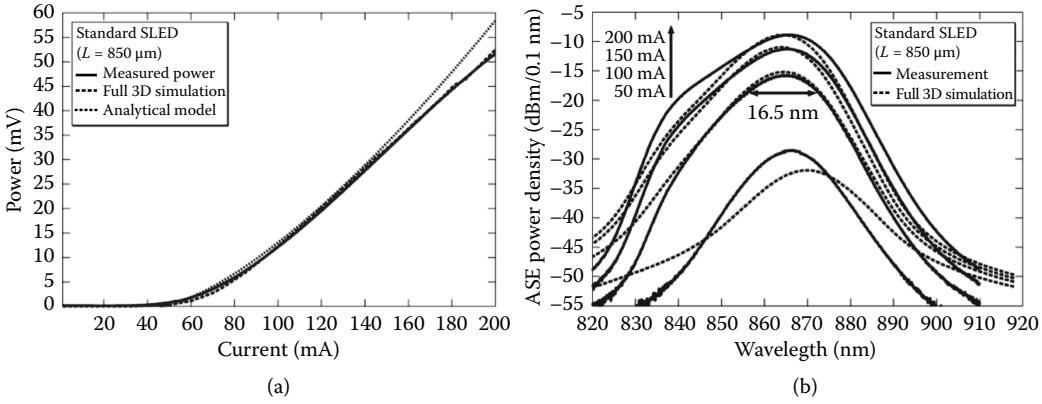

FIGURE 19.4 (a) *L–I* characteristics and (b) ASE output spectra (on an absolute logarithmic scale) for an 850-μm-long standard SLED chip. Solid lines represent measurement data and dashed lines simulation results obtained from a full 3D simulation after the calibration procedure. The dotted *L–I* characteristic follows from the analytical model according to Equation 19.14, as explained in the text. The ASE spectra are shown for four different injection currents: 50, 100, 150, and 200 mA. At 100 mA the 3-dB spectral bandwidth is about 16.5 nm. Note that all simulated ASE spectra have been shifted slightly by the same amount in wavelength so that they coincide at their maximum values with the measured spectra. This allows a better comparison of the detailed behavior of the spectral shapes.

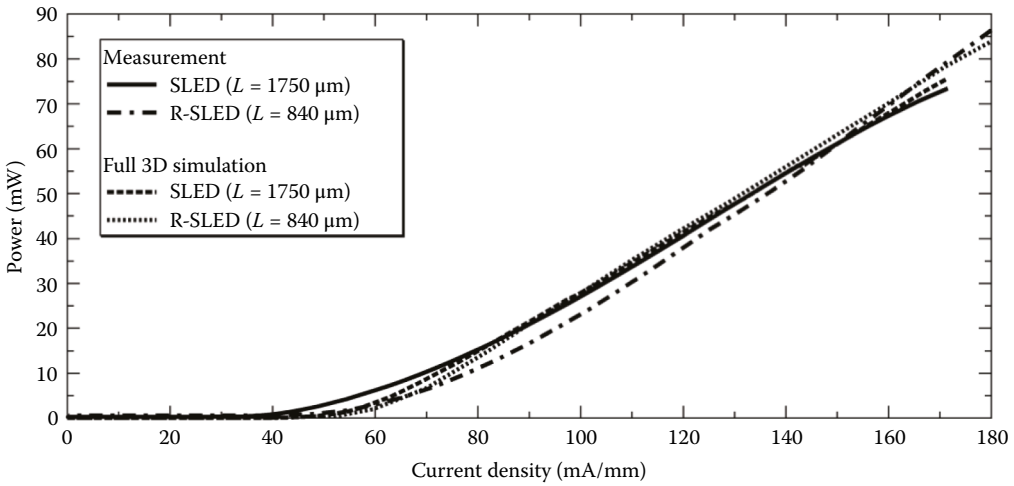

FIGURE 19.5 Measured (solid and dash–dotted line) and simulated (dashed and dotted line) front-facet output power of an SLED chip with length 1750 μm and an R-SLED chip with length 840 μm as function of current density, as described in the text. All curves agree almost perfectly.

at the same current density. Second, experimental results can be accurately reproduced and extrapolated once the simulator is well calibrated.

This is further demonstrated by analyzing the ASE output spectra. Figure 19.6a shows measured ASE spectra and Figure 19.6b shows simulated ASE spectra for both chips at two different levels of current density. In the first case, the current density with values slightly below 60 mA/mm is just above the threshold for ASE operation, which is why the output power is low. In the second case, the current density is beyond 170 mA/mm, resulting in the chips operating at high front-facet output powers of about 75 mW. The simulated curves coincide more or less over the entire spectral range for both current densities, whereas a

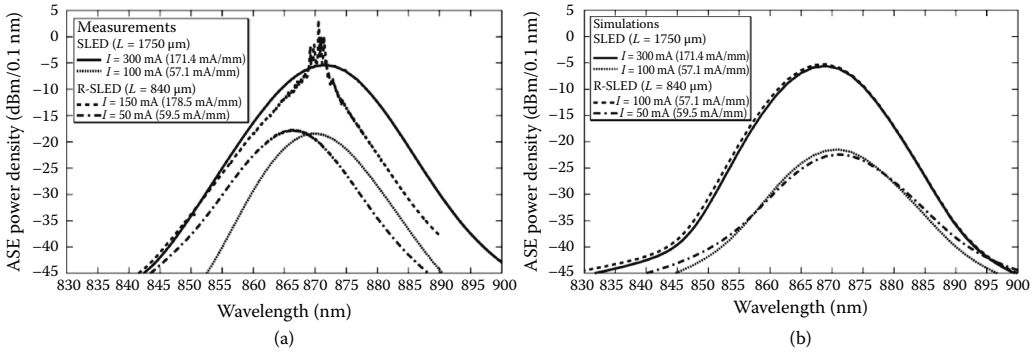

FIGURE 19.6 (a) Measured and (b) simulated ASE output spectra of an SLED chip with length 1750 µm and an R-SLED chip with length 840 µm at two different levels for the current density. The dotted and dash-dotted lines correspond to a current density slightly below 60 mA/mm and the solid and dashed lines to a value slightly above 170 mA/mm. In contrast to Figure 19.4b, no additional wavelength shift has been applied to the ASE spectra.

small wavelength shift can be observed for the measured curves. This is explained by the fact that even nominally identical chips from the same epi wafer have slightly different center wavelengths of the emitted ASE spectra. Furthermore, the thermal heat sinking toward the submount might not be identical for all chips. A more important difference, which can be observed for the measurements at the high current densities, is that the ASE spectrum of the R-SLED is somewhat narrower and shows strong periodic modulations (spectral ripples) in the center part. The periodicity corresponds to the longitudinal mode spacing defined by the length of the linear cavity. This is a clear indication of a resonant cavity effect (Fabry–Pérot modulation) caused by the high back-side reflectivity and demonstrates the tendency of an R-SLED to start lasing at high levels of current injection. In Section 19.3.2, the influence of spectral ripples on the coherence function will be discussed.

To gain further insight into the physical properties of an SLED and R-SLED, respectively, we analyze the longitudinal distribution of the optical intensity and the carrier density along the z-axis inside the chip based on the results obtained from the full 3D simulation (see Figure 19.7). The intensity profiles have been extracted at the center of the optical mode, and the maximum value at the front facet has been normalized to unity. The carrier density profiles have been extracted at the center of one of the QWs. The z-axis has been defined so that the back facet of the R-SLED coincides with the middle of the standard SLED and the front facets almost coincide, too. With this definition, all chip properties of the standard SLED are symmetric with respect to the origin. As shown in the upper plot, the shapes of the optical intensity profiles are quite similar for the same current density. However, as can be clearly seen, for low current densities (dotted and dash–dotted lines) the growth of the optical intensity in direction to the front facet is almost exponential, whereas the optical intensity saturates for higher current densities (solid and dashed lines) and, hence, higher optical output powers. This gain saturation behavior due to high optical powers is directly linked to the longitudinal carrier density distribution as shown in the lower plot. At low current densities, the carrier density is almost constant along the chip with a value of about 3.3×10^{-18} cm^{-3}. In contrast, at high current densities the carrier density decreases strongly by more than 40% from the origin toward the front facet with a terminal value of about 2.8×10^{-18} cm^{-3}. Low values of the carrier density mean low material gain, and thus, low gain amplification of the optical power.

Our investigations can be summarized as follows:

- At same current densities an R-SLED behaves very similar like a standard SLED with twice the length.
- An important difference is that R-SLEDs show ripples in the ASE spectrum and may have the tendency to start lasing operation.
- For long SLED chips and for R-SLEDs at high-output powers, gain saturation effects are not negligible anymore.

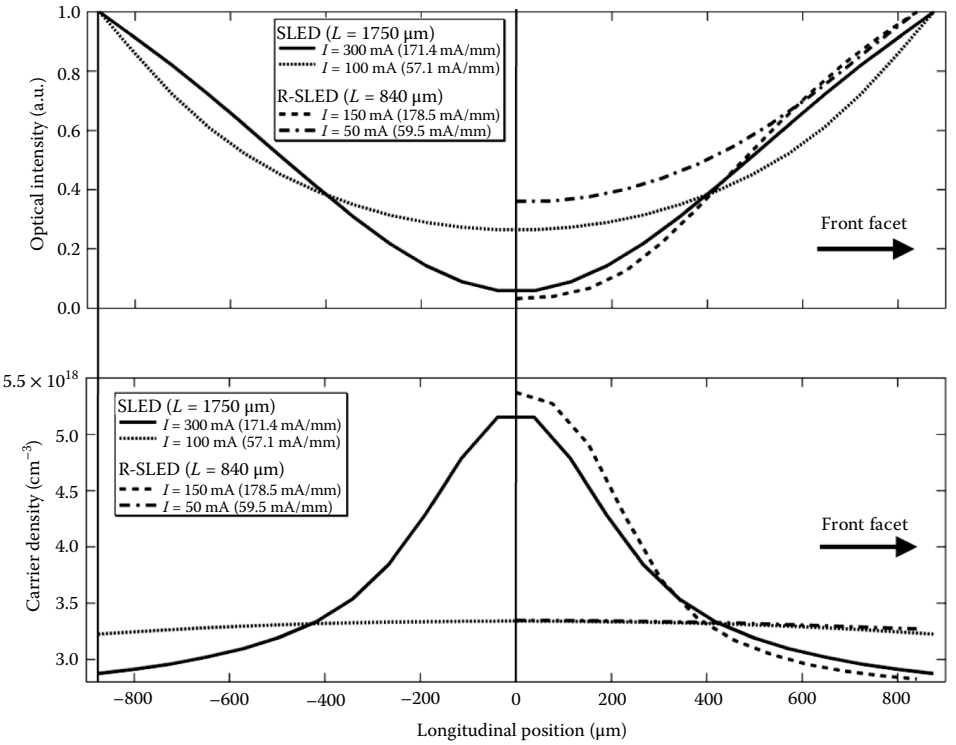

FIGURE 19.7 Longitudinal distribution of the optical intensity (upper plot) and carrier density (lower plot) along the *z*-axis inside the chip of an SLED chip with length 1750 μm and an R-SLED chip with length 840 μm at two different levels for the current density. The dotted and dash-dotted lines correspond to a current density slightly below 60 mA/mm and the solid and dashed lines to a value slightly above 170 mA/mm.

19.3 SLED Designs with Specified Targets

Commercial SLEDs have to fulfill target specifications for the electro-optical performance (output power, wavelength, bandwidth, spectral shape, ripple, PER, etc.) that are defined by the application. Also, their long-term reliability has to meet certain lifetime requirements (e.g., 10% power drop over 5000 hours of operation or 50% power drop over 100,000 hours of operation) that are dependent on the application and the intended use of operation. The shape and width of the ASE spectrum is directly linked to the coherence function, which is the autocorrelation function of an SLED. This is of particular importance for applications like OCT, as explained below. It is well known that increasing the output power is, in general, accompanied by a reduction in the ASE spectral bandwidth. Hence, methods are required to overcome this limitation.

19.3.1 Broad-Bandwidth SLED Designs

Generation of light in an SLED is based on the recombination from an electronic state in the conduction band with a hole state in the valence band (VB). In a QW, the bound states build a discrete set of sub-bands. In the simplest case, the SLED's epitaxial structure is based on an SQW active-region design or an MQW structure consisting of uncoupled identical QWs. Such designs can be reduced to the discussion of the quantized SQW states. Here, we restrict our discussion to the VB for heavy holes because ASE spectra of SLEDs consisting of compressively strained QWs are dominated by transitions to the heavy-hole sub-bands.

19.3.1.1 SLEDs with Bell-Shaped (Gaussian) ASE Spectra

The spectral shape of the ASE spectra is mainly determined by the fundamental sub-band transition from the first electronic state to the first heavy-hole state $(e_1 \rightarrow hh_1)$, as long as the carrier density is sufficiently low, so that the population of other sub-band states can be neglected (Chen et al. 1990). Generally, single transition designs yield bell-shaped (Gaussian) ASE spectra as shown in Figure 19.1a, b, and f. Such spectra are preferred with respect to the avoidance of side lobes in the near optical path length difference (OPD) range of the coherence function (see Section 19.3.2). An increasing carrier density in the QW leads at first to a rising population of high-energy states in the fundamental sub-bands, and thus, to a broadening of the spectrum. If other sub-band states start to become significantly populated, an additional hump tends to appear on the short-wavelength side (as shown in Figure 19.1c). Hence, with increasing population of the sub-bands, the spectral appearance changes continuously from being bell shaped in direction to an M-shaped double-humped spectrum (see Figure 19.1d).

19.3.1.2 SLEDs with M-Shaped (Double-Humped) ASE Spectra

Generally, all approaches for the design of SLED structures showing an ASE spectrum with ultrawide emission bandwidth are based on a multitransition design. This means that more than just one optical transition contributes to the emission spectrum. One has to distinguish between two principally different approaches. The first one is based on the SQW (or MQW) epitaxial design, as just discussed, in which, in addition to the fundamental transition, other transitions fulfilling the quantum-mechanical selection rules are utilized to contribute to the generation of light. If properly designed, such QW structures allow for the enormous enlargement of the spectral bandwidth (Semenov et al. 1993b; Kondo et al. 1992). The second approach is based on the use of nonidentical (so-called chirped) QWs (Lin et al. 2004). Chirping means that at least one QW of an MQW epitaxial layer structure is differently designed compared to the other QW(s), yielding optical transitions at different wavelengths. Chirping can be achieved by a different QW thickness and/or different material compositions for the QW and/or the barrier layers. The basic idea of utilizing chirped QWs is demonstrated in Figure 19.8 for a chirped double-quantum well (DQW) structure. Due to the greater band-gap energy and thinner QW thickness of the right QW, the fundamental transition results in photons with higher energy (shorter wavelength) compared to the left QW. Thus, an ultrawide bandwidth design may be achieved if the spectral separation chosen between both fundamental transitions is sufficiently large.

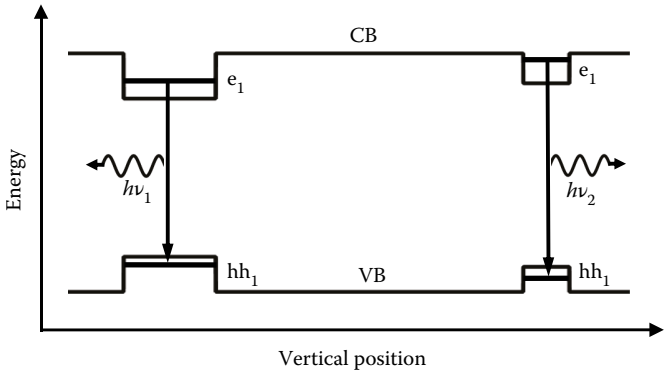

FIGURE 19.8 Energy-band diagram for two uncoupled chirped quantum wells (QWs). The horizontal lines in the wells indicate the first bound electronic state in the conduction band (CB), e_1, and the first bound hole state in the valence band (VB) for each QW separately. For the VB we restrict ourselves to the heavy hole band with its states hh_1. The arrows indicate the fundamental transition from the CB to the VB for both QWs. The electron–hole recombination results in the emission of a long-wavelength photon with energy $h\nu_1$ from the left QW and a short-wavelength photon with energy $h\nu_2$ from the right QW.

As mentioned earlier, ultrawide ASE spectra tend to exhibit a double-humped spectral shape. Such spectra show their maximum spectral bandwidth when both humps are well balanced, i.e., for approximately the same peak strength (the so-called flat-top condition). For a given chip design and a given heat-sink temperature, this situation can be achieved only over a limited range of operation currents. This makes such designs less flexible compared to single-transition designs. Moreover, the non-Gaussian spectral shape is unfavorable with respect to the occurrence of side lobes in the coherence function as mentioned earlier. The stronger the deviation from a Gaussian shape, the higher the side lobes are in the coherence function.

Generally, the width of the ASE spectrum broadens with the energy separation between allowed transitions to the heavy-hole VB. As a consequence of a greater spectral separation, first of all, the dip between both humps obvious in the ASE spectra at flat-top condition increases yielding a negative impact on the coherence function. And second, the SLED has to be driven at higher currents for operation at flat-top condition. The reason is that higher carrier densities in the QWs are required in order to lift the energetically more distant transitions on the short-wavelength side. This might be used as a positive side effect or if it is undesired, this effect can be compensated for by using chips with shorter active segment length. Nevertheless, it is obvious that the approach described here is limited to finding the best compromise between spectral bandwidth and the depth of the spectral dip.

19.3.1.3 Flattening of M-Shaped ASE Spectra

The M-shape approach can be optimized by a flattening of the spectral dip (see Figure 19.1e). This requires the use of at least one additional transition between QW states yielding photons with a wavelength somewhere around the spectral position of the dip. The additional transition fills up the dip between both peaks in the ASE spectrum. It can be realized in different ways, as discussed in the following.

The first possibility is the use of a chirped MQW active region design with at least one of the chirped QWs having its fundamental transition in the spectral range of the dip (Lin et al. 1996). For MQW structures consisting of a large number of QWs an elaborated approach is based on the continuous chirp of the fundamental transition wavelength for each QW over a wide spectral range.

Another possibility is the utilization of coupled QWs, where the coupling is caused by the finite potential walls between the QWs. For a coupled DQW active region design, the coupling leads to a splitting of the fundamental sub-band states, where the coupling strength, and thus, the splitting can be adjusted by the thickness of the barrier layers between the QWs and/or the well depth. The latter can be adjusted by the material composition of the barrier layers. If properly designed, the transition wavelength of the short-wavelength photon appears in the spectral range of the dip.

A third possibility to reduce or even remove the dip is the combination of two or more SLEDs with spectrally shifted ASE spectra. The SLEDs are designed so that the spectral peak of one SLED coincides with the dip of the other SLED. The combined spectrum is flatter and potentially broader if the SLEDs are appropriately designed. The flattening effect is demonstrated by the example of the M-shaped ASE spectrum for the 1050-nm SLED shown in Figure 19.1d. We assume that a second SLED has been designed with similar spectrum just shifted by 24 nm to shorter wavelengths. For simplicity reasons we take the original spectrum (the dashed curve in Figure 19.9) and shift it accordingly in order to obtain the dotted curve. Now, we assume that the output spectra are combined via a dispersion-free 50:50 coupler. This corresponds to a simple linear addition of both spectra with the same weight, resulting in the solid curve shown in Figure 19.9. Obviously, the original dip at around 1050 nm is removed and the resulting spectrum looks rather bell shaped. In the following section, the impact of the SLED combination on the coherence function is discussed.

19.3.2 Coherence Function and Coherence Length

As mentioned earlier, for OCT applications, the coherence function is of major interest. It allows the extraction of the coherence length, which determines the theoretical limit for the axial resolution of an

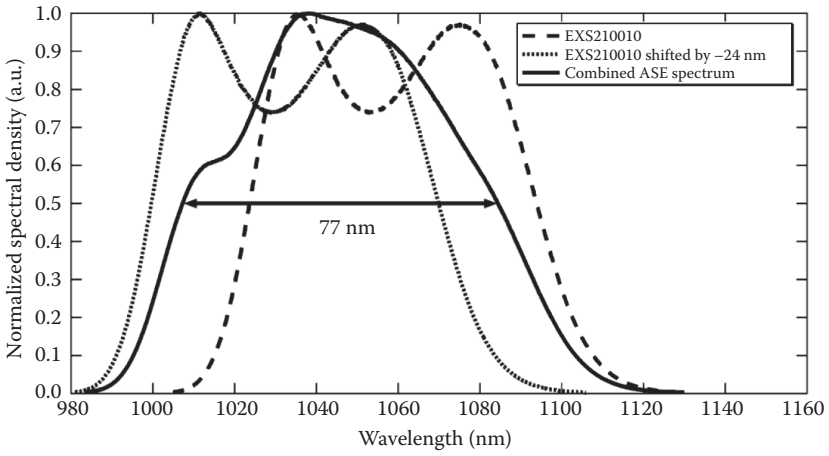

FIGURE 19.9 The dashed line shows the same ASE spectrum as plotted in Figure 19.1d for module EXS210010. The dotted curve shows the same spectrum shifted by 24 nm to shorter wavelengths. The solid curve is obtained as the linear superposition of both spectra. The FWHM bandwidth of the combined spectrum is 7 nm greater compared to the width of each single spectrum.

OCT imaging system (Duelk and Hsu 2015). For a broadband light source, the coherence length can be written as

$$l_{\mathrm{coh}} = \gamma \cdot \frac{2 \cdot \ln(2)}{\pi} \cdot \frac{\lambda_{3\mathrm{dB}}^2}{\Delta\lambda_{3\mathrm{dB}}}, \tag{19.16}$$

with $\lambda_{3\mathrm{dB}}$ being the 3-dB center wavelength and $\Delta\lambda_{3\mathrm{dB}}$ the 3-dB bandwidth of the ASE spectrum. The factor γ depends on the shape of the ASE spectrum. For an ideal Gaussian shape, the factor is unity, $\gamma = 1$, for other shapes, $\gamma > 1$. For SLEDs with double-humped ASE spectra operating at flat-top condition, a value of $\gamma \approx 1.19$ is typically used. This yields a penalty of roughly 20% in coherence length caused by the deviation from an ideal Gaussian shape. As an example, for the ASE spectrum shown in Figure 19.1e, we calculate for the coherence length $l_{\mathrm{coh}} \approx 7.4$ μm using Equation 19.16 with $\lambda_{3\mathrm{dB}} = 1300$ nm, $\Delta\lambda_{3\mathrm{dB}} = 120$ nm, and assuming a flat-top shape with $\gamma = 1.19$.

The coherence length can be directly determined from the coherence function, too. The latter is equivalent to the SLED's autocorrelation function, and thus, according to the Wiener–Khinchin theorem, the Fourier transform of the ASE spectrum (Schmitt 1999). Figure 19.10 shows the coherence function as a function of the OPD obtained for the ASE spectrum just discussed. The coherence function is symmetric with respect to the OPD. Its drop-off from the maximum value at zero OPD to 50% defines the coherence length, that is, the coherence length is the half-width at half maximum (HWHM). From the coherence function shown in Figure 19.10a near zero, we extract a coherence length of 7.9 μm. This value is close to the value estimated from Equation 19.16. Moreover, side lobes are visible at an OPD of 20 μm with a side-lobe suppression ratio (SLSR) of about 10 dB. The side lobes are caused by the non-Gaussian flat-top shape of the ASE spectrum. For applications like OCT, those side lobes may generate imaging artifacts if they are not suppressed by an additional windowing function in the OCT signal processing (Duelk and Hsu 2015). Therefore, SLED designs with bell-shaped ASE spectra are preferred for such applications. Figure 19.11 shows the coherence functions, which correspond to the ASE spectra plotted in Figure 19.9 for module EXS210010 and the combined SLED spectrum. Obviously, going from the M-shaped to the rather bell-shaped ASE spectrum leads to a strong reduction of the side lobes with an increase of the SLSR from 7 to 13 dB. As a side effect, the coherence length decreases by almost 1 μm from 8.6 to 7.7 μm due to the broadening of the combined SLED spectrum by 7 nm.

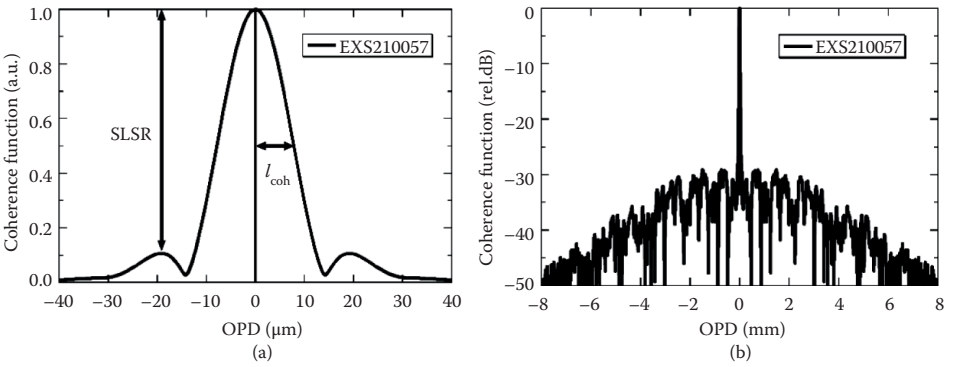

FIGURE 19.10 Coherence function versus optical path length difference (OPD) obtained for the ASE spectrum shown in Figure 19.1e for SLED EXS210057. (a) Linear plot over a small OPD range. The coherence length defined by the half-width at half maximum (HWHM) value is found to be 7.9 μm (in air) and the side-lobe suppression ratio (SLSR) roughly 10 dB. (b) Logarithmic plot over the full OPD range.

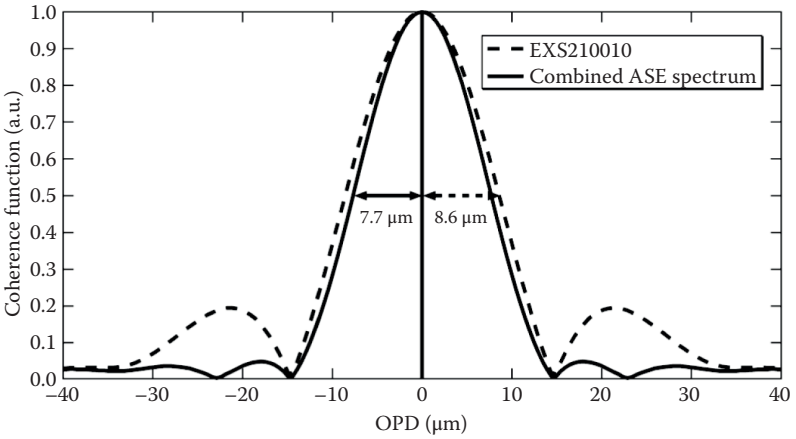

FIGURE 19.11 Coherence function in the near OPD range obtained for the ASE spectra plotted in Figure 19.9 for module EXS210010 (dashed curve) and the combined SLED spectrum (solid curve). From the dashed curve an SLSR of 7 dB and a coherence length of 8.6 μm is extracted, whereas the values found from the solid curve are 13 dB and 7.7 μm, respectively.

Periodic modulations in the ASE spectrum caused by the build-up of a resonant cavity, as discussed in Section 19.2.4, may show up as secondary coherence subpeaks at an OPD of a few millimeters (see, e.g., Figure 18.1 in Duelk and Hsu 2015). Figure 19.10b is free of such secondary coherence peaks as the ASE spectrum is smooth and free of spectral ripple. In order to avoid OCT imaging artifacts (so-called "ghost lines"), a secondary peak suppression ratio (SPSR) of at least 25 dB is needed. The SPSR requirement defines the tolerance for the maximum allowed spectral ripples in the ASE spectrum. As a consequence, this might cause R-SLED designs to be impractical for OCT applications.

19.3.3 High-Power SLED Designs

In addition to the requirements for the ASE output spectrum, SLEDs should typically deliver a minimum amount of output power at a given injection current. A variety of methods exist for increasing the SLED's

front-facet output power, and thus achieve a high-power design. In the following, we describe possible approaches and discuss their advantages and drawbacks.

19.3.3.1 Increasing the Injection Current

The most obvious approach is just to increase the injection current. This is possible as long as the point for the thermal roll-over is not reached. But one should take into account that the increasing carrier density may have a strong influence on the shape of the ASE output spectrum. As discussed in Section 19.3.1, ASE spectra resulting from a single optical transition will typically remain bell shaped with broader bandwidth. However, for M-shaped ASE spectra, which may be well balanced for a lower injection current, as shown in Figure 19.1d and e, the low-wavelength peak will strongly grow with increasing injection current. This leads to an unfavorably strong reduction of the spectral bandwidth compared to the preferred flat-top condition.

19.3.3.2 Increasing the Chip Length

The output power can be increased for a given injection current by increasing the active segment length, that is, the length of the gain medium. This method works for operating currents sufficiently above the ASE threshold level because the latter increases with increasing chip length. Longer chips lead to a reduction of both the thermal resistance and the series resistance leading to lower junction temperatures for SLED chips running in CW mode of operation. Moreover, the carrier density is reduced resulting in the opposite effect on the ASE spectra, as explained in Section 19.3.3.1.

19.3.3.3 Using an R-SLED Design

As discussed in Section 19.2.3, a very efficient way to design an ultimate high-power SLED structure is the R-SLED approach. This allows for very high-output powers using relatively short chips. The main drawback of this approach is the enhancement of spectral ripples or even worse the tendency to start lasing due to the build-up of a resonant cavity. This may lead to an increase of the noise level in general and the appearance of strong secondary coherence subpeaks. Hence, the R-SLED design might be the optimum choice for insensitive applications with respect to resonant-cavity effects. For other applications, this approach might not work or is not recommended.

19.3.3.4 High-Power SLEDs by Epitaxial Design

The epitaxial design of the active-region layers has a strong influence on the SLED's *L–I* characteristic. Typically, it is easier to design SLEDs delivering higher-output powers by increasing the number of QWs and/or by increasing the thickness of the QW layers. Modifying the epitaxial design in this way has diverse and complicated effects on the expected electro-optical performance. For example, the injection current required for reaching ASE threshold increases in general with increasing total thickness of the active region. Moreover, the carrier density is reduced, having a similar effect on the ASE spectrum as described earlier in Section 19.3.3.2 for an extended chip length.

Therefore, the approach that has to be chosen for the realization of a specific SLED design depends on the whole set of target specification. The expected *L–I* characteristics and ASE spectra for a given active region design finally determine the appropriate chip length and injection current for device operation. There are no simple and strictly valid rules for finding the optimal SLED design. However, it can be approached using the discussion presented in this section as a guideline.

19.4 Summary

SLEDs are employed over a wide range of different applications with emitted ASE spectra covering the wavelength range from the visible to the near infrared. The band-gap energy of available semiconductor materials determines the appropriate choice of the substrate and epitaxial layer materials for the design

of a specific SLED structure. Most SLEDs are realized in a tilted ridge-waveguide geometry similar to a Fabry–Pérot LD, which allows for the formation of optical modes. The target specifications mainly on the output power and the ASE spectrum determine the active-region design. SQW and MQW structures are most commonly used to achieve material gain.

The performance of SLED structures can be simulated like other opto-electronic devices using full 3D simulation tools. Such complex tools couple the different problems involved in the physical device description (multiphysics treatment). The simulation results are rather accurate at the expense of analytical insight. Analytical approaches based on traveling-wave equations describing the longitudinal propagation of light reduce the dimensionality of the problem and the number of coupled equations. They are particularly useful to derive an analytical expression for the SLED's *L–I* characteristic. We have presented a simple model that includes the influence of arbitrarily strong modal reflectivities at both facets on the output power. Using our analytical results, we have compared the performance of standard SLEDs with the new class of so-called R-SLEDs. Furthermore, our model has been verified by a comparison against experimental data and results obtained from full 3D simulations.

The design of SLEDs with specific target requirements has been discussed. In particular, the appearance of single-peaked (bell-shaped) and double-humped (M-shaped) ASE spectra for broad bandwidth SLED structures and how it affects the corresponding coherence functions has been analyzed in detail. Various methods for boosting the output power have been illustrated with their implications particularly on the ASE spectra. These findings can be finally used as a guideline for the design of a specific broad-bandwidth high-power SLED structure.

References

Alphonse, G. A., D. B. Gilbert, M. G. Harvey, et al. 1988. High-power superluminescent diodes. *IEEE J. Quant. Electron*. 24, no. 12: 2454–2457.

Bohnert, K., P. Gabus, J. Nehring, et al. 2002. Temperature and vibration insensitive fiber-optic current sensor. *IEEE J. Lightwave Technol.* 20, no. 2: 267–275.

Burns, W. K., C. Chen, and R. P. Moeller. 1983. Fiber-optic gyroscopes with broad-band sources. *IEEE J. Lightwave Technol.* 1, no. 1: 98–105.

Chen, T. R., L. Eng, Y. H. Zhuang et al. 1990. Quantum well superluminescent diode with very wide emission spectrum. *Appl. Phys. Lett.* 56, no. 14: 1345–1346.

Chuang, S. L. 1995. *Physics of Optoelectronic Devices*. New York: John Wiley & Sons.

Coldren, L. A., S. W. Corzine, and M. L. Masanovic. 2012. *Diode Lasers and Photonic Integrated Circuits*, 2nd ed. Hoboken, NJ: John Wiley & Sons.

CrosslightSoftware2017

Crosslight Software, Inc. Crosslight Software, 2017. http://www.crosslight.com/

Drexler, W. and J. G. Fujimoto, eds. 2008. *Optical Coherence Tomography: Technology and Applications*, 2nd ed. Berlin: Springer-Verlag.

Duelk, M. and K. Hsu. 2015. SLEDs and swept source laser technology for OCT. In *Optical Coherence Tomography: Technology and Applications*, 2nd ed., Drexler, W. and J. G. Fujimoto (eds.), 527–562. Berlin: Springer-Verlag.

Dufour, M. L., G. Lamouche, V. Detalle, et al. 2005. Low coherence interferometry—An advanced technique for optical metrology in industry. *Insight–Non-Destructive Testing and Condition Monitoring* 47, no. 4: 216–219.

Exalos, A. G. EXALOS: High-performance broadband SLEDs, swept sources and sub-systems, 2017. http://www.exalos.com/sled-modules/

Feltin, E., A. Castiglia, G. Cosendey, et al. 2009. Broadband blue superluminescent light-emitting diodes based on GaN. *Appl. Phys. Lett*. 95: 081107.

Heo, D., I.-K. Yun, J.-S. Lee, et al. 2011. High-power SLD-based BLS module for WDM-PON applications. *J. Korean Phys. Soc*. 58, no. 3: 429–433.

Kaminow, I. P. and D. Marcuse. 1983. Superluminescent LED with efficient coupling to optical waveguide. U.S. Patent 4 376 946.

Kondo, S., H. Yasaka, Y. Noguchi, et al. 1992. Very wide spectrum multiquantum well superluminiscent diode at 1.5 μm. *Electron. Lett.* 28, no. 2: 132–134.

Koyama, F., K. Y. Liou, A. G. Dentai, et al. 1993. Multiple-quantum-well GaInAs/GaInAsP tapered broad-area amplifiers with monolithically integrated waveguide lens for high-power applications. *IEEE J. Photonics Technol. Lett*. 5, no. 8: 916–919.

Kwong, N. S. K., K.-Y. Lau, and N. Bar-Chaim. 2008. High-power, high-efficiency GaAlas superluminescent diodes with integrated absorber for lasing suppression. *IEEE J. Quant. Electron*. 25, no. 4: 696–704.

Laino, V., F. Roemer, B. Witzigmann, et al. 2007. Substrate modes of (Al,In)GaN semiconductor laser diodes on SiC and GaN substrates. *IEEE J. Quant. Electron*. 43, no. 1: 16–24.

Lee, T. P., C. A. Burrus, and B. I. Miller. 1973. A stripe-geometry double-heterostructure amplified-spontaneous-emission (superluminescent) diode. *IEEE J. Quant. Electron*. 9, no. 8: 820–828.

Li, Z. Q. and Z. M. S. Li. 2010. Comprehensive modeling of superluminescent light-emitting diodes. *IEEE J. Quant. Electron.* 46, no. 4: 454–461.

Lin, C.-F., Y.-S. Su, C.-H. Wu, et al. 1996. Broad-band superluminescent diodes fabricated on a substrate with asymmetric dual quantum wells. *IEEE J. Photonics Technol. Lett.* 8, no. 11: 1456–1458.

Lin, C.-F., Y.-S. Su, C.-H. Wu, et al. 2004. Influence of separate confinement heterostructure on emission bandwidth of InGaAsP superluminescent diodes/semiconductor optical amplifiers with nonidentical multiple quantum wells. *IEEE J. Photonics Technol. Lett.* 16, no. 6: 1441–1443.

Loeser, M. and B. Witzigmann. 2008. Multi-dimensional electro-opto-thermal modeling of broadband optical devices. *IEEE J. Quant. Electron.* 44, no. 6: 505–514.

Macleod, H. A. 2010. *Thin-Film Optical Filters*. Boca Raton, FL: Taylor & Francis.

Marcuse, D. 1983. Computer model of an injection laser amplifier. *IEEE J. Quant. Electron*. 19, no. 1: 63–73.

Marcuse, D. 1989. Reflection loss of laser mode from tilted end mirror. *J. Lightwave Technol*. 7, no. 2: 336–339.

Matuschek, N. and M. Duelk. 2013. Modeling and simulation of superluminescent light emitting diodes (SLEDs). *IEEE J. Sel. Top. Quant. Electron*. 19, no. 5: 7800307.

Matuschek, N., T. Pliska, and N. Lichtenstein. 2008. Properties of pump-laser modules exposed to polarization-dependent and wavelength-selective feedback from fiber Bragg gratings. *IEEE J. Quant. Electron*. 44, no. 3: 262–274.

Matuschek, N. and M. Duelk. 2014. Modeling and simulation of reflecting SLEDs. *14th International Conference on Numerical Simulation of Optoelectronic Devices* (NUSOD '14), Postdeadline Poster MPD43, Mallorca, Spain.

Middlemast, I., J. Sarma, and S. Yunus. 1997. High power tapered superluminescent diodes using novel etched deflectors. *Electron. Lett.* 33, no. 10: 903–904.

Milani, N. M., V. Mohadesi, and A. Asgari. 2015. A novel theoretical model for broadband blue InGaN/GaN superluminescent light emitting diodes. *J. Appl. Phys.* 117: 054502.

Nagai H., Y. Noguchi, and S. Sudo. 1989. High-power, high-efficiency 1.3 μm superluminescent diode with buried bent absorbing guide structure. *Appl. Phys. Lett.* 54, no. 18: 1719–1721.

Park, J. and X. Li. 2006. Theoretical and numerical analysis of superluminescent diodes. *IEEE J. Lightwave Technol*. 24, no. 6: 2473–2480.

Patterson, B. D., J. E. Epler, B. Graf, et al. 1994. A superluminescent diode at 1.3 μm with very low spectral modulation. *IEEE J. Quant. Electron*. 30, no. 3: 703–712.

Photon Design. Photon Design - Your source of photonics CAD tools, 2017. https://www.photond.com/

Rossetti, M., J. Dorsaz, R. Rezzonico, et al. 2010. High power blue-violet superluminescent light emitting diodes with InGaN quantum wells. *Appl. Phys. Express.* 3: 061002.

Rossetti, M., L. H. Li, A. Fiore, et al. 2008. Quantum dot superluminescent diodes. In *Handbook of Self Assembled Semiconductor Nanostructures for Novel Devices in Photonics and Electronics*, ed. Henini, M., 565–599. Oxford: Elsevier.

Rossetti, M., J. Napierala, N. Matuschek, et al. 2012. Superluminescent light emitting diodes—The best out of two worlds. *SPIE Photonics West, MOEMS and Miniaturized Systems XI*, 8252–8206.

Schmitt, J. M., 1999. Optical Coherence Tomography (OCT): A Review. *IEEE J. Sel. Top. Quant. Electron.* 5, no. 4: 1205–1215.

Semenov, A. T., V. R. Shidlovski, S. A. Safin, et al. 1993a. Superluminescent diodes for visible (670 nm) spectral range based on AlGaInP/GaInP heterostructures with tapered grounded absorber. *Electon. Lett.* 29, no. 6: 530–532.

Semenov, A. T., V. R. Shidlovski, and S. A. Safin. 1993b. Wide-spectrum SQW superluminescent diodes at 0.8 μm with bent optical waveguide. *Electon. Lett.* 29, no. 10: 854–856.

Senior, J. M. 2009. *Optical Fiber Communications Principles and Practice*. Harlow: Pearson Education Limited.

Sun, Z. Z., D. Ding, Q. Gong, et al. 1999. Quantum-dot superluminescent diode: A proposal for an ultra-wide output spectrum. *Opt. Quant. Electron*. 31: 1235–1246.

Synopsis, Inc. Synopsys, 2017. http://www.synopsys.com/

Tansu, N. and L. J. Mawst. 2001. High-performance strain-compensated InGaAs–GaAsP–GaAs (λ = 1.17 μm) quantum-well diode lasers. *IEEE J. Photonics Technol. Lett*. 13, no. 3: 179–181.

Wild, G. and S. Hinckley. 2009. Distributed optical fibre smart sensors for structural health monitoring: A smart transducer interface module. *5th International Conference Intelligent Sensors, Sensor Networks and Information Processing*, Melbourne, Australia, 373–378.

Wootten, M. B., J. Tan, Y. J. Chien, et al. 2014. Broadband 2.4 μm superluminescent GaInAsSb/AlGaAsSb quantum well diodes for optical sensing of biomolecules. *Semicond. Sci. Technol*. 29, no. 11: 115014.