TOPOCAL REVIEW

Next-generation mid-infrared sources

To cite this article: D Jung et al 2017 J. Opt. 19 123001

View the article online for updates and enhancements.

Related content
- Novel InP- and GaSb-based light sources for the near to far infrared
  Sprengel Stephan, Demmerle Frederic and Amann Markus-Christian

- Quantum dot optoelectronic devices: lasers, photodetectors and solar cells
  Jiang Wu, Siming Chen, Alwyn Seeds et al.

- Optoelectronic device physics and technology of nitride semiconductors from the UV to the terahertz
  Theodore D Moustakas and Roberto Paiella
Topical Review

Next-generation mid-infrared sources

D Jung¹, S Bank², M L Lee³ and D Wasserman²,⁴

¹ Department of Electrical and Computer Engineering, University of California Santa Barbara, Santa Barbara, CA, 93106, United States
² Department of Electrical and Computer Engineering, University of Texas Austin, Austin, TX, 78758, United States
³ Department of Electrical and Computer Engineering, University of Illinois Urbana-Champaign, Urbana, IL, 61801, United States

E-mail: dw@utexas.edu

Received 15 November 2016, revised 26 September 2017
Accepted for publication 16 October 2017
Published 14 November 2017

Abstract

The mid-infrared (mid-IR) is a wavelength range with a variety of technologically vital applications in molecular sensing, security and defense, energy conservation, and potentially in free-space communication. The recent development and rapid commercialization of new coherent mid-infrared sources have spurred significant interest in the development of mid-infrared optical systems for the above applications. However, optical systems designers still do not have the extensive optical infrastructure available to them that exists at shorter wavelengths (for instance, in the visible and near-IR/telecom wavelengths). Even in the field of optoelectronic sources, which has largely driven the growing interest in the mid-infrared, the inherent limitations of state-of-the-art sources and the gaps in spectral coverage offer opportunities for the development of new classes of lasers, light emitting diodes and emitters for a range of potential applications. In this topical review, we will first present an overview of the current state-of-the-art mid-IR sources, in particular thermal emitters, which have long been utilized, and the relatively new quantum- and interband-cascade lasers, as well as the applications served by these sources. Subsequently, we will discuss potential mid-infrared applications and wavelength ranges which are poorly served by the current stable of mid-IR sources, with an emphasis on understanding the fundamental limitations of the current source technology. The bulk of the manuscript will then explore both past and recent developments in mid-infrared source technology, including narrow bandgap quantum well lasers, type-I and type-II quantum dot materials, type-II superlattices, highly mismatched alloys, lead-salts and transition-metal-doped II-VI materials. We will discuss both the advantages and limitations of each of the above material systems, as well as the potential new applications which they might serve. All in all, this topical review does not aim to provide a survey of the current state of the art for mid-IR sources, but instead looks primarily to provide a picture of potential next-generation optical and optoelectronic materials systems for mid-IR light generation.

Keywords: mid-infrared, optoelectronics, lasers, light emitting diodes

(Some figures may appear in colour only in the online journal)

1. Introduction

The mid-infrared (mid-IR) wavelength range can most generally be described as the portion of the electromagnetic (EM) spectrum that extends from the long-wavelength edge of the telecom/near-IR wavelength range (~2 μm) to the short wavelength edge of the THz range (30 μm). Thus defined, the mid-IR encompasses the wavelength range associated with the fundamental rotational and vibrational resonances of a large number of technologically vital molecules. For this
reason, the mid-IR is oftentimes referred to as the ‘molecular fingerprint’ region of the EM spectrum, where the presence of these molecules can be detected via their unique spectral signatures (figure 1), enabling a range of environmental, health and medical, industrial as well as security and defense sensing applications [1–3]. At the same time, the mid-IR is also home to the spectral peak of the thermal emission signatures of most biological and mechanical objects with temperatures ranging from 200 K–1400 K, and for this reason is of significant importance for security and defense applications related to thermal targeting and night vision, and also for energy auditing/conservation applications [4].

While technologically vital for its numerous potential sensing and imaging applications, the mid-IR is also important as a wavelength range for fundamental scientific investigation. First, and perhaps most mundanely, the longer wavelengths of the mid-IR (when compared to the visible or near-IR portions of the EM spectrum), scale many wavelength- and subwavelength-scale optical components and structures to dimensions achievable with standard photolithography, making the mid-IR an ideal testbed for next-generation optical materials such as metamaterials, optical antennas and sub-diffraction features based on plasmonic or designer effective media [5–7]. More fundamentally, the mid-IR is a wavelength range where a wide variety of light–matter interactions can not only be studied, but also engineered. Interaction with free carriers, the foundation of plasmonics and many metamaterial structures, is not only controlled by defining the geometry of metallic optical structures, but using ‘designer’ materials such as doped semiconductors, by engineering the optical properties of the ‘metals’ themselves [8]. The low energies of mid-IR photons allow traditional optoelectronic semiconductor materials to be used, whose bandgap energies are well above the energy of most mid-IR photons, as high index and low loss dielectrics. With careful band structure engineering, these materials can also provide designed optical transitions (the subject of much of this review) and enhanced optical nonlinearities [9, 10], from quantum-engineered states in superlattices, and quantum wells, wires and dots (QWs, QWRs, and QDs). At the same time, many III–V, II–VI and lead-salt semiconductor alloys can provide bandgaps well into the mid-IR, offering opportunities to engineer semiconducting materials able to serve as the foundation for mid-IR devices mimicking well-established visible, near-IR and telecom wavelength optoelectronic devices [11]. Interestingly, the mid-IR even offers opportunities for investigating the interaction of light with lattice vibrations (optical phonons), most frequently in polar semiconducting materials such as III-nitrides or SiC [12, 13].

Historically, despite the many applications and fundamental implications associated with the mid-IR, research and development in this wavelength range has been stymied by a sparsity of optoelectronic devices (both coherent and incoherent sources, modulators, detectors), optical materials, and the compact optical systems that would result from these materials and devices. Most mid-IR characterization is performed using a Fourier transform infrared (FTIR) spectrometer, which has long been the workhorse of mid-IR optics research. The FTIR uses a broadband thermal source (glo-bar) and Michelson interferometer to generate high-resolution reflection and transmission spectra across the mid-IR and THz portions of the EM spectrum. The basic operation of the FTIR has changed little over the past decades, and despite the widespread use and effectiveness of the FTIR, these systems are poorly suited to the compact and portable optical systems desired for many mid-IR applications. The limitations of mid-IR optical systems, the FTIR being only one example, can arguably be traced back to a common denominator: the lack of suitable mid-IR optical sources. Historically, the thermal emitters used in FTIR spectroscopy have been the sole light source available for mid-IR applications. However, such sources are wildly inefficient, and certainly not compatible with applications requiring either low power consumption, or highly directional, narrow-band or coherent mid-IR light. For such applications, one of the few potential commercial sources, historically, has been the CO2 laser, which can offer extremely high power coherent radiation, but only at certain distinct wavelengths (in the 9–11 μm range). Moreover, despite offering high power and coherent mid-IR emission, the CO2 laser, like most gas lasers (including other mid-IR gas lasers [14]), is poorly suited for applications requiring mobility or those that have size and power use limitations. Such applications at shorter wavelengths have long been enabled by the remarkable development of semiconductor lasers and light emitters, which have served as a foundation for the technological behemoths that are telecommunications, solid state lighting and displays, and visible/near-IR imaging and sensor fields. Without longer wavelength versions of these semiconductor optoelectronic devices, the mid-IR has long faced a severe technological bottleneck.

In this topical review, we will present the current state of mid-IR source technology and review both new and old material systems, as well as potential applications for next generation mid-IR source technology. In section 2, we discuss the current state-of-the-art for mid-IR sources, in particular (a) thermal emitters, (b) quantum cascade lasers and (c) interband cascade lasers. We will discuss the significant recent advances encompassed in these classes of emitters, and also their limitations and the challenges faced for further development. Section 3 is a short discussion of the features and capabilities that might be desired for next-generation mid-IR sources, as well as potential applications and technologies that are...
Figure 2. Thermal emission (spectral emittance) as a function of temperature and wavelength across the mid-IR for (a) a perfect blackbody emitter \( \epsilon (\lambda) = 1 \), and spectrally selective emitters with the emission centered at (b) \( \lambda = 10 \, \mu m \) and (c) \( \lambda = 1.5 \, \mu m \).

currently underserved or unavailable due to the limited optical infrastructure in the mid-IR. In sections 4 and 5, we will present a number of potential semiconductor material systems that could serve as the basis for next-generation mid-IR sources. In section 4, we focus on interband emitters based on (a) type-I active regions on GaSb, (b) type-I and type-II active regions on InP, and (c) type-II superlattice active regions on both InAs and GaSb, (d) highly mismatched alloys and (e) lead-salt and transition metal doped II-VI materials. The majority of these structures leverage traditional band structure engineering along the epitaxial growth direction, including separate confinement heterostructures, QWs, strain-balancing, digital alloys and superlattices. In section 5, we discuss the potential for mid-IR emitters utilizing three-dimensional (quantum dot) confinement for (a) intersubband transitions and (b) type-II transitions, as well as (c) lead-salt epitaxial dots and new classes of nanocrystal and colloidal quantum dots with mid-IR transition energies. Finally, in section 6, we summarize the outlook for the next generation sources discussed in our review, and present both the challenges and opportunities of the sources discussed.

2. The current mid-IR source landscape

2.1. Thermal emitters, benefits and limitations

The mid-IR, as mentioned above, is the spectral home of peak thermal—or blackbody—emission from objects with temperatures ranging from 200 K–1400 K. The thermal emittance from the surface of a heated object is given, in W-m\(^{-2}\), by Planck’s law of thermal radiation:

\[
E(\lambda, T) = \epsilon(\lambda, T) \frac{4\pi c^2}{h^3} \frac{1}{e^{hc/\lambda kT} - 1}
\]

where \( h \) is Planck’s constant, \( \lambda \) is the wavelength of the emitted light, \( T \) is the object’s temperature, \( c \) is the speed of light in a vacuum and \( k \) is the Boltzmann constant. The emissivity, \( \epsilon(\lambda, T) \), modulates the spectral emittance of a perfect blackbody, as a function of wavelength and temperature, with the limit \( \epsilon(\lambda, T) = 1 \) being the case of an ideal blackbody emitter.

A typical thermal emitter, such as the ‘glo-bars’ used as the default light source in an FTIR, generally consists of a ceramic element heated to a high temperature via resistive heating, providing the broadband emission suitable for many spectroscopic applications. The advantage of the thermal emitter, in addition to its broadband nature, lies primarily in its cost-effectiveness.

However, the thermal emitter suffers from some rather significant drawbacks. Figure 2(a) shows the emission of a perfect blackbody emitter as a function of temperature and wavelength. For such an emitter, the peak thermal emission can only be controlled by temperature, which also controls the total emitter power. Applications only requiring light in a limited bandwidth of the mid-IR, say the 8–12 \( \mu m \) ‘molecular fingerprint’ band, would be able to utilize a thermal emitter centered in this wavelength band. However, such a source would basically be operating at or near room temperature, providing emission that is barely distinguishable from the thermal background. An increase in emission intensity would only come about by increasing the temperature, shifting the peak emission to shorter wavelengths and decreasing the effective efficiency of the emitter. Moreover, the bulk of the additional emission provided by the increasing temperature would come at shorter wavelengths that are outside the band of interest. Thus, traditional thermal emitters, with near uniform emissivity, while useful for broadband FTIR spectroscopy, become increasingly less useful and less efficient, when the wavelength range of interest narrows, such as for sensing applications focused on distinct spectral features or bands.

However, control over the thermal emission spectrum can be achieved by engineering the emissivity of the surface. Kirchhoff’s law states that the emissivity of a surface can be written as \( \epsilon(\lambda, T; \theta) = a(\lambda, T; \theta) = 1 - R(\lambda, T, \theta) - T(\lambda, T; \theta) \), where \( a \) is the material’s absorptivity, and \( R \) and \( T \) are the reflectivity and transmissivity of the surface (including all diffraction), recognizing that all of the above depend not only on the wavelength and temperature, but also on the angle of incidence. Thus, by engineering spectrally selective losses onto a structure’s surface, using for instance plasmonic or micro/nano-antenna resonators, one can control
the emissivity and thus the thermal emission of that surface upon heating. Alternatively, phase-change materials offer the potential for designing thermal emitters with the dynamic control of emissivity [15]. Selective thermal emitters have been the subject of some recent interest due partly to their potential utility as mid-IR sources with relatively narrow band emission [16–20], but even more for potential applications regarding thermophotovoltaic (TPV) energy conversion [21, 22], where the thermal emission from a hot surface is absorbed by a photovoltaic cell and converted to electrical power. Both of the above applications require a high temperature, in the former case to maximize the emitted power and in the latter to align the emission peak with the bandgap of efficient photovoltaic devices. Figures 2(b) and (c) show the simulated thermal emission as a function of wavelength and temperature for two selective thermal emitters, one designed for selective emission at \( \lambda = 10 \, \mu m \), and the other at \( \lambda = 1.5 \, \mu m \). In both cases, we have set the emissivity to be \( \epsilon_{\text{max}} = 0.95 \) on-resonance and \( \epsilon_{\text{off}} = 0.05 \) off-resonance, which is an idealized but not entirely unrealistic approximation of a good spectrally selective surface.

As can be seen in figure 2(b), for longer wavelength selective emitters, while the off-resonant emission can be quenched somewhat by the introduction of a structured surface with engineered emissivity, at higher temperatures, the shift of the peak thermal emission to shorter wavelengths quickly results in a strong, off-resonance component of thermal emission. For the shorter wavelength selective emitter, significant on-resonance emission is only seen at high temperatures (\( \gtrsim 700 \, K \)), when the center of the blackbody emission spectrum lines up with the designed resonance. For this reason, there has been significant interest in the development of high-temperature materials for selective thermal emitters [23, 24]. However, increases in material losses at higher temperatures impose limitations on the selectivity of structured surfaces, on top of the inherent limitations imposed by Planck’s law of thermal radiation. Thus while such emitters are attractive for their relative cost-effectiveness, the efficiency and bandwidth limitations of thermal emitters leave ample motivation for the development of semiconductor-based emitters, long-wavelength analogs of near-IR and visible LEDs and lasers that provide the foundation for shorter wavelength optical infrastructure.

2.2. The quantum cascade laser—benefits and limitations

The landscape of mid-IR sources changed dramatically in the mid-1990s with the demonstration and rapid development of the quantum cascade laser (QCL) [25]. The operation of the QCL is, in many ways, fundamentally different from that of the more-established interband semiconductor lasers, which operate at shorter wavelengths. Instead of electrons recombining with holes across the semiconductor bandgap (figure 3(a)), the QCL relies on intersubband (ISB) transitions between engineered conduction band states in complex heterostructures (as shown in figure 3(b)). Such an approach offers significant design flexibility, limited on the short-wavelength side by the conduction band offset of the QCL material system, and on the long-wavelength side (in the mid-IR) by semiconductor phonon resonances (their Reststrahlen bands), though THz QCLs, at energies below those of the semiconductor optical phonons, are also now well-established technology [26, 27]. The most common QCL material systems are the InGaAs/AlInAs lattice-matched to InP and the GaAs/AlGaAs system, which is more frequently used for longer mid-IR wavelengths (10–100 \( \mu m \)), though less so for the shorter wavelength side of the mid-IR due to its limited conduction band offset. Since the first low-temperature (10 K), pulsed operation demonstration of the QCL [25], remarkable strides in QCL growth, design and operation have been demonstrated, with room-temperature (and above), continuous wave (CW) and high power QCLs operating across much of the mid-IR [28]. The rapid development of the QCL enables compact and high-performance mid-IR lasers suitable for mid-IR optical systems, which previous to the development of the QCL, were largely inconceivable. The commercial availability of these lasers attests to the significant demand for such sources, and has spurred significant efforts to integrate QCLs into all manner of mid-IR sensing, imaging and countermeasure systems.

However, QCLs utilizing the most common InGaAs/InAlAs material system, have generally struggled to achieve shorter wavelength (3–4.5 \( \mu m \)) operation, largely due to the fixed conduction band offset (520 meV) of the lattice-matched ternary alloys. Strain-balancing of the lattice-mismatched InGaAs/InAlAs layers on InP substrates, can allow for greater conduction band offsets, and thus be used to extend the short wavelength range of QCLs on InP substrates [30–33]. New QCL material systems have also been extensively researched, such as InGaAs/AlAsSb [34], InAs/AlSb [35], III-nitrides [36, 37], or alternatively even II–VI material systems, such as ZnCdMgSe alloys [38], in efforts to extend QCL emission to shorter wavelengths. However, these material systems are less developed in both growth and processing than the more common (In/Ga/Al)As materials, and thus QCLs using such novel materials lag well behind more technologically mature material systems.

The QCL’s wavelength flexibility results from the ample design space afforded by the careful quantum design of the ISB transitions in semiconductor heterostructures. However, these ISB transitions are also, in many ways, the origin of its greatest weaknesses. In a planar heterostructure, the dipole matrix element of an ISB transition is in the growth direction, ensuring that only TM-polarized light can be emitted from the active region of the QCL. This dipole matrix orientation prevents surface emission, unless additional outcoupling structures are fabricated onto the device surface. In addition, and perhaps most importantly, phonon scattering between electronic states in the QCL active region results in extremely fast nonradiative recombination lifetimes [39, 40]. The ~ps timescale nonradiative lifetimes associated with QCL ISB optical transitions lead to large threshold current densities (\( J_{\text{th}} \)), which limit the wall-plug efficiency of the QCL [41], and lead to weak sub-threshold emission. Thus, while the QCL is in many ways an ideal emitter for integration with optical systems requiring coherent and compact sources.
across much of the mid-IR, the aforementioned limitations resulting from the ISB transitions at the heart of the QCL constrain its utility for applications requiring narrow bandwidth—if not monochromatic—incoherent light, or alternatively, configurations requiring surface emission.

2.3. The interband cascade laser, benefits and limitations

Many of the shortcomings of the QCL are avoided in the interband cascade laser (ICL), first proposed [42–44] and experimentally realized [45–47] soon after the initial demonstration of the QCL. In 2008, the first room-temperature continuous wave operation of an ICL was demonstrated [48]. The ICL generates light via electron–hole recombination in complex heterostructures, usually leveraging type-II interfaces, such as the ‘W’ structure schematically shown in figure 3(c), though type-I ICLs have also been demonstrated [43, 49, 50]. In the ICL device structure shown in figure 3(c), electrons in the InAs conduction band QWs recombine with the holes in the InGaSb valence band QW. In this case, the broken gap type-II alignment allows emission energies below the bulk bandgap of either constituent material by control of the QW thickness, and thus the position of the electron and hole energy states in the active region. Electrons (holes) are supplied to the active region from a semimetallic interface through the electron (hole) injector states, shown in green (purple) in figure 3(c).

While at longer wavelengths (>7 μm), the ICL design struggles against decreased wavefunction overlap and band-filling effects [51]. ICLs are particularly promising for the 3–7 μm wavelength range, and have a number of advantages over the QCL for operation at these wavelengths. In particular, ICLs generally offer a lower $J_0$ than QCLs, especially in the short wavelength (≤5 μm) range, where the former have been demonstrated with room-temperature current densities below 100 A cm$^{-2}$ in pulsed mode operation [52]. The lower threshold powers are largely a result of the longer upper state lifetimes in the ICLs, which are estimated to be orders of magnitude larger than the phonon-scattering limited lifetimes of the ISB transitions in QCLs [53], resulting in power thresholds over an order of magnitude lower than QCLs for comparable wavelengths. In addition, the interband transitions of the ICL active region result in TE polarized emission, allowing for the possibility of ICL-based light emitting diodes [54] and vertical cavity surface emitting lasers (VCSELs) [55], though for the latter, developing high-quality as well as low loss top and bottom mirrors in the ICL material system offers significant challenges.

ICLS, like their QCL counterparts, have made tremendous strides over the past years [56], and are now commercially available. The low $J_0$ for the ICL offers potential applications in mid-IR systems with low power budgets, perhaps battery-operated, mobile or hand-held systems. However, the device does present certain challenges associated with high-power operation. ICLs, which require careful alignment of both the valence and conduction band states in adjacent semiconductors, as well as careful carrier balancing in the active region [57], are particularly sensitive to temperature changes. While ICLs can be designed for operation in a predetermined temperature range, at high power, heating of the active region will alter the temperature of the laser core and affect performance. In addition, while the high gain per active region allows few-stage ICLs and results in limited voltage defects, it also limits the ICL slope efficiency. As a general rule, QCLs demonstrate higher slope efficiencies and higher peak powers, as well as weaker temperature dependence, than comparable ICLs.

3. Opportunities for new mid-IR source technology

The performance of ICL and QCL emitters has improved continuously; however, efficient mid-IR emission from these devices requires complex bandstructure engineering, making the modeling, design and growth of cascade lasers time-consuming and difficult. Nonetheless, these established mid-IR sources have certainly been a driving technology for the rapid growth of, and interest in, mid-infrared optical systems and applications. Both the QCL and the ICL continue to be optimized, expanding their operational wavelength range, improving wall-plug efficiency and increasing peak power. It is quite likely that both devices will remain a dominant part of IR optical infrastructure for some time to come. However, each device presents challenges associated with certain aspects of performance or operation. While many of these challenges are sure to be overcome with further research and development, others are intrinsic to the physical operation of...
the devices. Short carrier lifetimes and TM polarized emission from QCLs limit efficiency and the development of surface emitting QCLs, while the significant temperature dependence of ICL designs make high temperature, or more importantly, high power operation difficult. Both device architectures require complex and often thick epitaxial growth, as well as significant expertise in the design and modeling.

As a result of this complexity, there is appeal in the prospect of simpler, diode-like emitters, similar to the near-IR and visible devices, both lasers and LEDs, that form the backbone of our modern communication, imaging and lighting technologies. Such devices, which are somewhat more straightforward in their modeling, design, growth and operation, have played a significant role in the increasing cost-effectiveness, energy efficiency and compactness of near-IR, and visible optical and optoelectronic systems. Similar devices for the mid-IR could offer improved power consumption efficiency and cost-effectiveness, and spur the development of compact mid-IR sensing systems, or even free-space communication systems. The mid-IR equivalent of LEDs, if capable of demonstrating reasonable efficiency at lower cost, could see significant integration in infrared scene projection [58], or sensing applications where wavelength selectivity is achieved by optics external to the source [59].

In addition, the current stable of mid-IR sources has no equivalent to the IR and visible quantum emitters such as quantum dots, quantum wires, or nitrogen vacancies in diamonds, which have spurred advances in optical quantum information and communication technologies [60–63]. The dominant mid-IR sources, though dependent on the quantum engineering of electronic states in complex heterostructures, most generally leverage the 1D confinement of carriers (i.e. along the growth direction) in wavelength-scale active regions. Thus, deeper investigation of mid-IR light–matter interaction at the extremes of carrier confinement and sub-diffraction-limited optical confinement are out of reach with current mid-IR sources. This is despite the fact that the mid-IR can reasonably be described as the wavelength range where the greatest number of material-based phenomena and structures (electronic transitions, free-carrier interactions, phononic effects, nonlinear optics and dielectric structures) can not only be exploited, but actively engineered to develop new multi-physics-based devices and structures for fundamental investigations and potential applications in quantum information and communication.

Finally, there is growing interest in the demonstration of, and potential applications for, ultrafast mid-IR sources [64]. Such sources have potential spectroscopy, diagnostic, material processing, and security and defense applications, as well as fundamental applications focused on resolving the temporal behavior of vibrational or rotational excitations in atoms or the solid state. QCLs offer one approach to achieving ultrafast mid-IR sources, with active mode-locking used to generate 5 ps pulses at \( \lambda = 8 \ \mu m \) [65], 3 ps pulses at \( \lambda \approx 6 \ \mu m \) [66], and most recently, \( \lambda \approx 5 \ \mu m \) pulses using an external ring cavity [67]. Passive mode-locking, on the other hand, is essentially impossible in the QCL due to fast gain-recovery [68]. Alternative approaches leverage ultrafast sources typically operating at shorter wavelengths, whose emission is down-converted to the mid-IR via nonlinear processes such as optical parametric amplification, optical parametric generation, or difference frequency generation. In the mid-IR, such processes require nonlinear materials with broadband transparency [69], and can offer ultra-fast (fs) pulses for a range of fundamental and commercial applications.

In this topical review, we will endeavor to present potential next-generation mid-IR sources offering utility across a range of applications, or alternatively, insight into a range of fundamental phenomena. We readily acknowledge the current primacy of the cascade-based sources (QCLs and ICLs) in the mid-IR source landscape, and expect the continued rapid development of these sources to ensure that cascade lasers will have a significant, if not dominant, role to play in the mid-IR optoelectronic infrastructure. A number of recent reviews of quantum cascade and interband cascade lasers have provided substantial insight into the history, current state-of-the-art and future directions for cascade laser research and development [28, 29, 70, 71], and we will not attempt to reproduce these works here. Similarly, we will limit the scope of our discussion to semiconductor sources providing direct emission of mid-IR light, and for the sake of (relative) brevity, omit the vibrant and growing field of mid-IR light generation via nonlinear processes. Instead, we will look to discuss a broad range of mid-IR sources that may offer potential for niche applications where cascade lasers may not be able to compete, or alternatively, for investigations of fundamental phenomena that are difficult to access with cascade laser structures. Our discussion is divided into two primary sections. First, we will discuss both bulk and 1D heterostructures as potential mid-IR sources. We will begin by discussing type-I QW lasers grown on GaSb, followed by type-I and type-II QW lasers grown on InP substrates, both operating in the short-wavelength IR. Next, we discuss emitters using type-II superlattice (T2SL) active regions grown on both InAs and GaSb substrates. Finally, we will discuss both a new class of highly mismatched alloys with intriguing potential at mid-IR wavelengths, and a long-standing material system, lead-salts, that might deserve a fresh look. In the second part of our review, we will consider nanostructured (QD and nanocrystal) mid-IR emitters, discussing both attempts to generate light from ISB transitions in QDs as well as type-II, lead-salt and II-VI QD emitters. We conclude by offering our thoughts on the viability of the presented emitter architectures for a range of applications, as well as the larger questions remaining for the development of alternative mid-IR sources.

4. Next-generation mid-IR sources: quantum well emitters

Cascade lasers, which have had such a dramatic effect on the mid-IR optical infrastructure, rely on complex engineered heterostructures, often with hundreds or thousands of epitaxial layers. As discussed above, their design is a clear departure from the ubiquitous diode emitters that serve as the
backbone of our telecommunications and solid state lighting systems. Cascade lasers require significant expertise, in both design and growth, and interest is expected to remain in more straightforward diode design for the demonstration of mid-IR light emitting diodes (LEDs) and diode lasers. However, such noncascade structures come with their own set of challenges. In particular, using narrow bandgap materials significantly increases Auger recombination rates, making mid-IR sources based on narrow bandgap materials inefficient at the high carrier injection levels required for population inversion, where most efficient emitters operate. In addition, the use of narrow bandgap materials limits the potential alloys available for heterostructure growth. As an example, InSb (which has a bandgap of 170 meV at room temperature) has a lattice constant of \( \sim 6.5 \, \text{Å} \), which is much larger than the next highest binary material (GaSb, 6.1 Å), preventing the growth of quantum wells or heterostructures without significant strain. Thus, while the challenge of using narrow bandgap materials for longer wavelength mid-IR emitters is formidable, significant progress has been made on the shorter wavelength end of the mid-IR using material systems leveraging interband transitions in diode-like structures. In the following sections, we discuss recent advances in the state-of-the-art for mid-IR interband emitters on GaSb, InP and InAs substrates.

### 4.1. GaSb-based emitters

Light emission from type-I QWs occurs via the spatially direct recombination of conduction band electrons and valence band holes, at wavelengths determined by the ground state energies of the carriers in the QWs, as shown in figure 4(a). In the 1980s, GaSb-based double-heterostructure (DH) laser diodes emitting from 1.8 to 2.4 μm were first developed by liquid phase epitaxy (LPE) [72, 73] and molecular beam epitaxy (MBE) [74]. Double heterostructure devices can be thought of as the first important step towards QW-based devices, where optical mode and charge carrier confinement (but not quantization) are achieved simultaneously in a thick, low-\( E_g \) active region surrounded by higher-\( E_g \) barriers that also serve as low-index clads. These early works utilized the InGaAsSb/AlGaAsSb ‘well’/barrier material system and showed promising laser diodes operating at room temperature (RT) with \( J_{th} \) values of \( 2–5 \, \text{kA cm}^{-2} \) [72–74]. The steady improvements achieved in MBE technology allowed the first demonstration of multi-QW (MQW) lasers with InGaAsSb/AlGaAsSb active regions to be made soon thereafter [75]. Additional efforts looked to the development of lower \( J_{th} \) (~200 A cm \(^{-2} \)) GaSb-based interband type-I QW lasers, and at the same time, the extension of the operational wavelength to 3 μm and beyond. The resulting lasers used ~1.5% compressively strained InGaAsSb QWs and AlGaAsSb barriers, demonstrating high RT output power (>500 mW), low \( J_{th} \) (~200 A cm \(^{-2} \)) and relatively high wallplug-efficiency (WPE) (>10%), in the 2–3 μm spectral region [76–81]. However, as shown in figure 4(b), laser diodes based on the InGaAsSb/AlGaAsSb material system show decreasing optical output powers and increasing \( J_{th} \) as the lasing wavelengths increase, especially for wavelengths beyond 3 μm. Additional incorporation of indium into the InGaAsSb QW lowers the bandgap of the QW material, and is thus required to extend the wavelength beyond 3 μm. However, increasing the In fraction in the QW also decreases the valence band offset between the well and barrier [82]. As a result, the InGaAsSb/AlGaAsSb material system suffers from poor hole confinement in the QWs with increasing In concentration in the QWs, which is an effect that has damaging repercussions for laser performance [83].

To mitigate the shallow hole confinement issue, AlInGaAsSb quaternary barriers were first introduced by Grau et al. Adding ~25% of indium to the barrier material increases the valence band offset from ~50 meV to ~150 meV, as shown in figure 5 [82], which enables improved performance at 3 μm with \( J_{th} = 300 \, \text{A cm}^{-2} \) and an output power of 130 mW [84]. Perhaps more importantly, the
increased valence band offset allows the demonstration of lasing out to 3.7 μm, which is the longest emission wavelength observed in GaSb-based type-I QW lasers [85].

The use of highly strained (i.e. >2%) QWs has emerged as a promising alternative approach to improving hole confinement at wavelengths >3 μm [86]. The heavy compressive strain warps the valence band, resulting in both lower hole density of states (DOS) and improved hole confinement. Another recently developed approach involves the use of metamorphic (i.e. strain-relaxed) buffers on GaSb, to shift the lattice constant of the entire device while aiming to maintain low dislocation density [87]. The shift to a 1%–2% larger lattice constant platform (e.g. Ga0.7In0.3Sb) allows a higher In% to be used in the QW without exceeding the critical thickness. Both the high strain and metamorphic approaches have been successful in enabling RT lasing at 3.0–3.2 μm with a peak power of ~150–250 mW. More recently, Shterengas et al demonstrated a 3.4 μm laser at 10°C through the growth of a 2.4%-strained In0.55Ga0.45As0.14Sb0.86/GaSb MQW active region without resorting to metamorphic growth [88].

As is the case for most mid-IR emitter materials, large Auger recombination rates and free carrier absorption in lower-Eg materials, are well-known obstacles to efficient laser operation [75]. Novel cascade pumping has recently been introduced to improve the efficiency of GaSb-based type-I QW lasers. The cascade pumping concept aims to ‘recycle’ injected carriers from one gain stage to the next, in a manner similar to QCLs and ICLs. Shterengas et al employed a GaSb/InAs injector grown on AlGaAsSb graded layers to demonstrate cascade-pumped 3.0 μm GaSb lasers, as shown in figure 6 [89]. Under a forward bias, electrons in the p+GaSb valence band tunnel into the n+InAs conduction band. Then, tunneling electrons are used for type-I QWs in the next stage. Even though the absorptive injector layer hurts optical confinement and increases loss, the benefit of recycling carriers outweighs the disadvantages. Shterengas et al reported output power reaching almost 1 W, a double decrease in Jth, and a peak power conversion efficiency (WPE) of 16% from 3.0 μm emitting GaSb-based lasers [90]. Optimization of the injector region, with light p-doping of the AlGaAsSb graded layers improved the hole transport between the stages, resulting in 500 mW output at 3.2 μm [91]. One major concern of type-I QW cascade pumping lasers is the increased turn-on voltage up to ~2 V as the number of cascaded stages increase.

Electrically pumped vertical cavity surface-emitting-lasers (VCSELs) on GaSb substrates have also been developed for applications such as low-cost and small-footprint gas sensing systems [92–94]. In 2008, the first RT CW 2.3 μm GaSb-based QW VCSELs were reported using a highly doped p-GaSb/n-InAsSb buried tunnel junction and five compressively strained (~1.6%) InGaAsSb/AlGaAsSb MQWs. The VCSELs employed an epitaxial AlAsSb/GaSb distributed Bragg reflector (DBR) for the bottom mirror and an a-Si/SiO2 dielectric DBR for the top mirror; since no current flows through the top DBR, the turn-on voltage and Joule heating are somewhat mitigated compared to monolithic VCSEL designs. Shortly thereafter, VCSELs with longer emission wavelengths at 2.5 μm and 2.63 μm were demonstrated using all-epitaxial n-AlAsSb/GaSb DBRs for both the top and bottom mirrors [95, 96]. These monolithic VCSELs showed relatively high Jth values of 8.8–11.5 kA cm−2 at 300 K under quasi-CW operation; turn-on voltages of ~3 V were higher than expected due to voltage drops across the DBRs. Andrejew et al reported improved performance from their GaSb-based VCSELs emitting at 3 μm using an undoped epitaxial DBR to achieve low optical loss in the cavity, which results in Jth = 4.1 kA cm−2 at ~10°C [97]. The device operated up to 50°C in pulsed mode while persisting only up to 5°C in CW, suggesting that the high turn-on voltage (~3 V) across the device gives rise to heating issues in the VCSELs. Recently, GaSb-based VCSELs with a type-II QW (InGaSb/InAs/AlGaSb) structure enabled 4 μm emission with a relatively low CW Jth of ~3 kA cm−2 near RT (~7°C) [98], allowing GaSb-based mid-IR VCSELs to cover the spectral regime from 2.3 μm to 4.0 μm, as shown in figure 7(b).
4.2. InP-based type-I and type-II emitters

InP-based laser diodes using type-I InGaAsP QW active regions are the main workhorse at the long-haul optical telecommunication wavelengths of 1.3 and 1.55 μm [99]. In this section, we describe efforts to grow non-QCL mid-IR lasers on InP substrates. One of the advantages of developing InP-based QW lasers for the mid-IR is the relatively straightforward active region design and growth compared to the QCLs. Also, such lasers may offer lower costs due to greater compatibility with the existing infrastructure for InP-based lasers and photonic integrated circuits compared to GaSb-based technology.

Highly strained InAs type-I QWs with InGaAs barriers on InP have reached lasing wavelengths up to 2.37 μm [100, 101]. Distributed feedback (DFB) lasers with the same active region materials show excellent performance, with CW operation up to 95 °C, a T0 of 71 K, and a high side-mode suppression ratio (SMSR) of 30 dB [102]. However, the substantial lattice-mismatch between InAs and InP constrains the thickness of the InAs QW to ∼5 nm, and this fundamentally limits type-I QW emission to ∼2.4 μm due to strong quantum confinement-induced shifts [103, 104]. Growth of InAs QWs thicker than ∼5 nm on InP substrates generates a high density of misfit dislocations, which act as nonradiative recombination centers in the QW active region. Owing to the high lattice mismatch between InAs and InP, type-I QW lasers on InP show continuously decreasing RT CW output power as the emission wavelength increases from ∼2.0 μm.

Interestingly, InP-based QCLs show decreasing RT CW output power as the emission wavelength is decreased from ∼5.0 μm. Therefore, neither type-I QW lasers nor QCLs on InP offer efficient operation at RT in the spectral wavelength gap between ∼2.4 and 3.0 μm, as seen in figure 8.
To grow dislocation-free InAs QWs thicker than ~5 nm on InP substrates, InAlAs or InAsP metamorphic graded buffers can be used to serve as a virtual substrate. Kirch et al. grew 16.5 nm thick InAs QWs with tensile strained InAsP barriers on metamorphic InAsP step-graded buffers [104]. While their PL showed 3 μm emission at RT, lasing was observed at 2.45 μm and 77 K from the InGaAs waveguides, not from the InAs QWs. Gu et al. utilized InAlAs continuously graded buffer layers on InP substrates to enable a 2.9 μm laser diode from type-I 15 nm thick InAs QWs, operating up to 180 K and 230 K in CW and pulsed mode, respectively [105, 106].

Graded InAsP buffers grown by MBE have been studied as another platform for high-quality InAs type-I QWs by providing low threading dislocation densities of 2–4 × 10^6 cm^-2 and smooth surface morphology [103]. Furthermore, InAsP alloys on InP have a higher E_g than InGaAs at the same lattice constant and thus offer more favorable carrier confinement in the QWs: Jung et al. showed a barrier-emission-free EL peak from 2.5 μm to 3.0 μm at RT for InAs type-I QWs surrounded by InAs,P_{x-0.5} barriers with x = 0.49–0.73 [108]. More recently, 2.75 μm RT lasing has been achieved in InAs/InAsP MQWs grown on InAs_0.8P_0.2 multi-functional metamorphic buffers (MFMBs), as shown in figure 9 [107]. Jung et al. showed that the InAsP MFMBs serve not only as a virtual substrate to grow InAs QWs on InP substrates, but also function as a bottom graded-index cladding layers for optical confinement. For the top cladding, they used a fully relaxed p-AlGaAs layer that does not generate threading dislocations penetrating back to the InAs QW active region. However, the device turn-on voltage is ~3 V, which indicates an unwanted voltage drop outside the InAsP p-n junction. This high turn-on voltage leads to device heating, limiting the CW operation to 200 K and degrading the WPE. Finally, the lifetime of MFMB mid-IR lasers remains untested, and further study has to be conducted on the role of the 10^5 cm^-2 threading dislocations on device reliability.

Another alternative material system for InP-based interband QW emitters is InGaAs/GaAsSb type-II QWs where electrons are confined in InGaAs and holes in GaAsSb (figure 10). Unlike the spatially direct (vertical) optical transitions of type-I quantum structures, type-II band alignments provide electron and hole confinement in adjacent semiconductor layers, and thus provide spatially indirect (diagonal) optical transitions. Because both the InGaAs and GaSb layers comprising the ‘W-shaped’ QW are both compressively strained, an As-rich, tensile GaSb layer may be inserted between each W-structure to serve as a strain compensation layer [109]. While the indirect recombination rate is smaller than the type-I due to the decreased electron–hole wavefunction overlap, long-wavelength emission is nonetheless achievable [110]. The emission wavelength can be widely tuned by modifying the thickness and composition of the layers in the W-structure, and RT PL emission from 2.56 to 3 μm has been demonstrated from the materials grown by Sprengel et al. and Pan et al. via MBE [111, 112]. In 2012, Pan et al. reported optically pumped type-II QW lasers emitting at 2.56 μm at RT on InP substrates [113], and Sprengel et al. demonstrated 2.55 μm electrically injected type-II lasers using a similar active region material [114]. Sprengel’s electrically pumped lasers operated up to 42 °C in pulsed mode and 0 °C in CW at 2.31 μm; the RT pulsed mode J_0 was ~5 kA cm^-2. Later in 2015, 2.5 μm VCSELs operating up to 10 °C in CW mode with a J_0 of 7.2 kA cm^-2 (figures 10(b)) and 2.7 μm Fabry–Perot lasers operating up to 30 °C in CW mode were demonstrated [114, 115]. However, InP-based type-II lasers typically show a drastic decrease in T_0 from ~45 K to ~24 K as the temperature increases from 0 °C to 70 °C [115]. To overcome this strong temperature dependence, the increasing carrier loss in QWs needs to be addressed [109].

4.3. Type-II superlattice-based mid-IR emitters on InAs and GaSb substrates

This section reviews LEDs and lasers where light is emitted by type-II interband transitions, tuned around the range of ~3–5 μm. Mid-IR lasing from a single type-II heterojunction was observed as early as 1986, based on the theoretical predictions of Kroemer et al. from 1983 [116]. Since then, many type-II active regions have been proposed and demonstrated, most notable of which is the ICL; the reader is referred elsewhere for an excellent review of ICLs [29]. After presenting a brief background on type-II lineups, we focus this section on two classes of type-II superlattice (T2SL) active regions: InAs/GaSb superlattices and InAs/InAsSb superlattices. Both material systems have been heavily developed for infrared detector applications [117–119], and also show great promise for light emission applications.

The canonical type-II heterojunction is the broken-gap InAs/GaSb superlattice (shown in figure 11), where the bulk conduction band edge of InAs lies below the bulk valence band edge of GaSb [120, 121]. A semiconductor superlattice can be thought of as a periodic structure of two or more materials, in this case InAs and GaSb, where the nm scale thickness of the
layers allows coupling between electron (and/or hole) states in adjacent periods, thus resulting in ‘minibands’ of allowed electron and hole energies. Such superlattices offer distinct advantages when compared to bulk narrow bandgap materials, in particular diminished Auger scattering \cite{122}, control over the effective bandgap—and thus emission energy—by control of the superlattice layer thicknesses, allowing access to longer wavelength emission than would be possible with a bulk alloy using any combination of the constituent materials of the superlattice \cite{123}. Quantum confinement shifts the ground state electron energy level in the InAs to above the heavy hole state in the GaSb, leading to tunable bandgap energy for spatially indirect transitions. Further bandgap tuning can be achieved by substituting strained InGaSb layers for the GaSb \cite{124} with 10%–30% In. In contrast, the nature of the InAs/InAsSb offset

Figure 9. A schematic illustration of a processed multi-functional metamorphic buffer laser diode with an optical transverse electric mode 2D profile and bright-field transmission electron microscopy images. (b) Temperature-dependent pulsed mode $L$–$I$ curves. The inset shows a 300 K lasing spectrum. Reprinted from \cite{107}, with the permission of AIP Publishing.

Figure 10. (a) The band structure of the W-shaped type-II laser. (b) CW $L$–$I$–$V$ curves from an InP-based VCSEL with type-II QWs as a function of the driving current at different heat sink temperatures. (a) Reprinted from \cite{109}, with the permission of AIP Publishing. (b) Reprinted from \cite{114}, with the permission of AIP Publishing.

Figure 11. The band structure of a broken gap type-II InAs/GaSb SL. The confinement energy of heavy holes in the GaSb layers is indicated by gray lines, the electron miniband in the InAs conduction band is shown as a grey shaded region. Reprinted from \cite{119}. Copyright 2005, with permission from Elsevier.
is more controversial, which is likely due to the coupled effects of strain, quantum confinement, ordering, and phase separation, with various reports of both type-I and type-II lineups [125]. One recent report found the lineup to be type-II for strain-balanced InAs/InAsSb superlattices on GaSb, with electrons confined to InAs, which is qualitatively similar to InAs/GaSb T2SLs [126].

The tendency for the confinement of electrons in InAs can be qualitatively understood by noting its large electron affinity of 4.9 eV and small bulk bandgap of 0.36 eV [124]. In contrast, the confinement of holes in the Sb-containing layer, e.g. (In) GaSb or InAsSb [117], is understood to arise from the low binding energy of Sb and the valence band character that is strongly derived from Sb atomic orbitals [115]. One notable design feature of such T2SLs is that the quantum confinement effect in the layer containing Sb is relatively weak due to the heavier hole mass compared to electrons. Hence, bandgap control can be affected just by altering the thickness of the InAs layers. [127] Both strain and quantum confinement work together to break the light-hole/heavy-hole degeneracy, while also affording some control over the position of the split-off band. Such valence band engineering is considered to be crucial both in light emission and light detection applications as a means of controlling nonradiative Auger recombination [125].

The emitters described in this section are bipolar devices based on pn-junctions for the simultaneous injection of holes and electrons into a T2SL active region. The active region typically contains an alternating sequence of two layers repeated 10–30 times, resulting in a total thickness of ~100–200 nm. As described below, separate pn junctions with differently designed active regions can be cascaded to enable broadband IR light emission [127]. The pn-junction devices described here typically have a low turn-on voltage (<1 V) due to their relatively low bandgap energies.

Most work to date on T2SLs for light emitters has been by MBE [121], with a few reports of growth by MOCVD [128, 129]. Interface control for InAs/GaSb SLs can be notoriously difficult due to the intermixing of both anions and cations, and it is possible to form InSb-like or GaAs-like interface regions [121]. InAsSb/InAs T2SLs are touted as somewhat simpler from the standpoint of interface control due to the common cation [117]. To date, much of the work on InAs/InAsSb SLs has utilized InAs substrates, while work on InAs/(In)GaSb SLs has primarily been done on GaSb substrates.

4.3.2. InAs/(In)GaSb-based type-II superlattice light emitters.

The first lasers to use an InAs/(In)GaSb superlattice were demonstrated by Hasenberg et al in 1995 [137]. Their active regions contained SLs of 17 Å InAs/35 Å In0.25Ga0.75Sb sputtering at 3.5 μm, while their optical clads consisted of InAs/AlSb digital alloy layers. Pulsed, electrically injected operation was observed up to 160 K. By 1997, Baranov et al had demonstrated room temperature pulsed lasing at 1.98–2.32 μm using 6–12 Å InAs/300 Å GaSb active regions with Al0.4Ga0.6Sb0.05Sb0.95 clads [138]. More recently, researchers have learned to intentionally engineer the formation of the InSb interfacial layer between InAs and GaSb (figure 13) to further control the emission of the active region [139, 140]. The primary benefit has been to dramatically increase the wavelength range of room-temperature operation, which now stands at 3.3 μm.

While the spectral purity and directional output of mid-IR lasers are well suited to a wide range of sensing applications, IR scene illumination could benefit from efficient sources of broad-band, incoherent (i.e. speckle-free) IR light. Thermal sources such as glo-bars are currently dominant, but high-brightness mid-IR LEDs would offer benefits including higher speed and greater tunability of the output spectrum [141–143]. In a recent paper, Ricker et al took advantage of the wide energy tunability of InAs/GaSb SLs on GaSb substrates to demonstrate an LED with an apparent temperature of 100 K in the 3.5–5 μm range [127]. By holding the GaSb thickness at 16 MLs and varying the InAs thickness from 6–9.7 MLs in eight separate SL active regions, Ricker tuned the emission peak from 3.3 to 5.0 μm; in essence, the broadband sources that they demonstrated consist of eight pn-junction LEDs connected electrically in series through tunnel junctions. At low current density, the emission from the
separate active regions can be resolved, but at high current, the individual peaks merge into a broad spectrum (figure 14).

Recent work from the same group on LEDs grown on GaAs shows a higher performance at 77 K than comparable devices on GaSb, despite high dislocation density on the GaAs wafers [144], while the Shockley–Read–Hall recombination was significant at a low injection current for typical operating current densities (100–1000 A cm$^{-2}$). The superior thermal conductivity and lower IR absorbance of GaAs compared to GaSb enabled higher external quantum efficiency for dislocated

Figure 12. (a) The growth and schematic band structure of an InAs/InAsSb SL laser, and (b) the temperature-dependent output at a fixed optical pump power showing lasing up to 95 K. Reprinted from [130], with the permission of AIP Publishing.

Figure 13. (a) The layer structure of an InAs/GaSb T2SL laser, where 1 monolayer (ML) of InSb is taken into account in the design, and (b) the pulsed operation (100 ns, 1 kHz) threshold current density ($J_{th}$) as a function of temperature for a range of laser devices of different dimensions fabricated from the same epitaxial growth, using the structure shown in (a). (a) Reprinted from [139], with the permission of AIP Publishing.
devices on GaAs. This finding is promising for future work integrating cascaded superlattice LEDs on substrates with a lower cost than GaSb, such as GaAs or Si. While InAs/InAsSb T2SLs offer the promise of simpler growth than InAs/GaSb T2SLs, light emitters based on the latter have attained higher performance to date. This situation is closely mirrored in the detector community [118, 145]; however, with growing interest in InAs/InAsSb T2SL detectors, re-examination of their promise as light emitters is in order.

4.4. Highly mismatched alloys for band engineering

Intriguing opportunities for band-engineering appear with the end-points of the group-III and group-V elements (boron, thallium, nitrogen and bismuth), which differ significantly in size and electronegativity from their more traditional counterparts. This often results in significant band bowing, which is typically localized at either the conduction or valence band, affording routes to independently control the band edges. Such control can be enabling for sources based on both type-I and type-II active regions. In type-I diode lasers for example, elements that affect the valence band can be employed to restore the valence band offset, which, as discussed in section 4.1, is reduced when the well and barrier compositions are tuned to extend the emission wavelength. The type-II active regions discussed earlier in section 4 could potentially benefit from independent control of the valence and conduction band edges, affording new flexibility in controlling the emission wavelength and band offsets. Unfortunately, these materials tend to be challenging to grow; this is essentially a fundamental consequence of their highly mismatched nature, which leads to low solid solubility. Fortunately, growth technology has advanced tremendously and high-efficiency emitters have been demonstrated in a number of highly mismatched alloy systems.

Nitrogen is perhaps the best-studied of the highly mismatched elements. It is immensely electronegative, and in dilute amounts, introduces a localized s-like state above the conduction band edge in most III-V materials; this spatially-localized level is broad in momentum space, leading to band-anticrossing interactions with the host matrix [146]. The resulting repulsion in these ‘dilute-nitrides,’ splits the conduction band into two bands, with the lower band being pushed rapidly downward in energy with increasing N% [147], as shown in figure 15 [148]. Moreover, this bandgap reduction is accompanied by a reduction in lattice-constant that is attractive for long-wavelength sources [149]. This enables high-performance and advanced GaAs-based lasers across the key portions of the optical fiber window [150–154]. In the mid-IR, dilute-nitrides have attracted much interest for extending the emission wavelength of type-I GaSb-based diode lasers. Although initial experiments do not yield room-temperature luminescence, due to the comparatively weaker bond-strength of antimonide materials than their arsenide counterparts [155–158], room-temperature luminescence has been achieved by mitigating ion-related damage from the nitrogen plasma source and by spatially-localizing the nitrogen dose during growth [159]. Type-II W-type structures have also been realized, where dilute-nitride layers were used to produce a low-lying conduction band in one layer and interspersed between antimonide layers that produced a high-lying valence band, resulting in promising performance in the near-IR [160] and mid-IR [110].

Converse to the case of nitrogen, we would expect the introduction of a highly electropositive group-V atom to introduce an energy level below the valence band edge, resulting in valence band anticrossing. Indeed, this intuitive observation has been confirmed experimentally for antimony and bismuth in GaAs, with the bismuth and antimony levels lying within the valence band of many of the traditional III-V [161, 162]. Bismuth is much larger than the constituents of the host matrix, making the solid solubility low, leading to a propensity for bismuth adatoms to ride the growth front in a number of dilute bismuthide (sometimes termed dilute bismide) alloy systems [163–165]. This produces a driving force towards bismuth droplet formation, which Ptak and co-workers showed can be suppressed with a combination of low growth temperature and high growth for GaAsBi [166]. Droplet-free growth in similar growth windows was also found for InGaAsBi [167] and InAsBi [88]. An added advantage is that the combination of low growth temperature and high growth rate also enables films to greatly exceed the classic critical thickness limitation by as much as twelve-fold [88, 167]. The addition of bismuth also dramatically increases the spin–orbit splitting [168], offering the exciting prospect of being able to drive the optical transition energy lower than the spin–orbit splitting to mitigate Auger recombination transitions involving the split-off band [169]. This condition has been achieved in GaAsBi on GaAs [170], GaInAsBi on InP [167], and GaSbBi on GaSb [171]. Progress towards mid-IR dilute bismuthide lasers has been rapid, with recent reports of room temperature photoluminescence out to 3.8 μm from GaSbBi with a Bi concentration of 14% [171], as well as
room temperature lasing out to 2.71 μm from GaSbBi/GaSb quantum well lasers shown in figure 16 [172]. A potential risk is that valence band anticrossing in the dilute bismides results in a myriad of valence bands, which would enable a number of additional Auger recombination transitions. More work is needed to assess this possibility, but the $T_0$ of these devices around room temperature was $\sim 50$ K, suggesting Auger recombination is at play, though this could simply be a consequence of the 4.22 kA cm$^{-2}$ threshold current density in these initial exciting devices.

The dilute boride compounds are an intriguing family of alloys, perhaps in the large part because very little about them has been definitively established. For example, even for BGaAs, the bandgap has been reported to both increase [173–176] and decrease [177, 178] with increasing boron composition; the predicted composition at which the direct-to-indirect bandgap transition occurs varies from $\sim$30% [179] to $\sim$80% [180] boron; there are even theoretical predictions that alloys will not obey properties as fundamental as Vegard’s law [179, 180]. The optical transition edge of compounds like BGaAs was not found to be governed by band-anticrossing effects [181], but are instead dominated by boron-related cluster states near the band edge [182], perhaps rendering them less advantageous for emitters. However, the small size of boron atoms leads to a significant reduction in the lattice-constant, making them potentially quite valuable for strain-compensating barriers, which has been demonstrated in the near-IR [183]. Similar to other highly mismatched alloys, the growth of dilute borides has proven challenging, with record boron concentrations limited to $\sim$8% in both MBE [184] and MOCVD [173]. Initial growth attempts by MOCVD were limited to B concentrations <6%, with roughening and structural degradation brought about by the formation of nanocrystals [185] and twinning defects [186]. Improvement in MOCVD growth of dilute borides was achieved using lower temperatures, with a large arsenic precursor pressure and with an appropriate boron precursor [187]. With MBE, boron concentrations were initially limited to $\sim$1% [173], with larger concentrations accessible by growth at lower temperatures and higher arsenic overpressures [184]. Additionally, surfactant-mediated epitaxy has been shown to positively affect boron incorporation in

![Figure 15](image1.png)

Figure 15. (a) The effects on the conduction band structure of GaNAs with the introduction of nitrogen; nitrogen bifurcates the conduction band into two bands, $E_+$ and $E_-$, with the $E_-$ band forming a new conduction band minimum. (b) Splitting between anticrossed bands versus pressure, which is consistent with band-anticrossing theory. (c) The fundamental energy gap of GaNAs with nitrogen content. Band anticrossing theory prediction is shown in red. Reproduced from [148] © IOP Publishing Ltd. All rights reserved.

![Figure 16](image2.png)

Figure 16. (a) Voltage and light output versus current, under pulsed operation from a GaSbBi/GaSb quantum well laser. (b) Laser emission spectra at different temperatures under pulsed operation. Reprinted from [172], with the permission of AIP Publishing.
BGaAs films grown by MBE by reducing anti-site/interstitial boron incorporation [188]. This is consistent with the common trends in highly-mismatched alloy growth, where thermodynamic instability may be overcome with kinetically limited growth. Indeed, boron concentrations well in excess of previous reports are now accessible via MBE growth in a highly kinetically limited regime. Using high growth rates with an e-beam boron deposition source has enabled BGaAs growth with boron concentrations of ~12% in strained epilayers and excellent structural quality, as well as up to 22% boron in relaxed films [189].

Thallium-containing compounds first drew significant interest as potential near-to-long-wavelength infrared photodetector materials [190, 191] and can presumably be applied to emitters as well. For example, a potentially advantageous mid-IR material, zincblende InTlAs has been demonstrated, though concomitant with metallic thallium droplets. This is perhaps quite similar to the aforementioned challenges associated with dilute bismides [192]; indeed, droplet-free GaTlAs films have been achieved with a thallium content of up to 7% [193]. A key remaining challenge appears to be the detrimental effects when both indium and thallium are present in the crystal; despite growth at low temperatures, a thallium content >4% resulted in twinning-type defects [194] and also tended to form diluted (ordered) alloy clusters [195].

4.5. II-VI and lead-salt mid-IR emitters

Perhaps the earliest semiconductor materials used for mid-IR applications were not III-V semiconductors, which are so central to the current optoelectronic infrastructure, but II-VI and IV-VI (lead-salt, or lead-chalcogenide) materials. II-VI semiconductors can be grown with bandgaps ranging from far-IR to UV, depending on the alloy material and composition. Narrow bandgap II-VI semiconductors, such as HgCdTe, have long been the dominant material system for mid-IR detector technology, due to their direct bandgap and the broad wavelength tuning of their bandgap by controlling the alloy composition. However, challenges associated with large area, uniform growth and the large Auger recombination of II-VI materials has hampered the development of direct interband mid-IR optoelectronic devices outside of photodetectors.

However, there currently exist two distinct paths towards the development of mid-IR sources based on II-VI materials. The first utilizes high bandgap II-VI alloys, such as ZnCdSe/ZnCdMgSe, with their large conduction band offsets, as a binary material system for the demonstration of II-VI quantum cascade emitters operating at short wavelengths, though initial demonstrations showed emission at longer (6–8 μm) wavelengths [37, 196]. The second approach leverages transition metal (TM) doped II-VI materials, first suggested in 1996 as having potential as a mid-IR gain medium, in particular for ZnSe doped with Cr$^{2+}$/Cr$^{3+}$:ZnSe, which showed promising emission, and in the following year, lasing out to ~3 μm [197, 198]. Longer wavelength lasing (4–4.5 μm) was soon thereafter observed in Fe:ZnSe at temperatures up to 180 K [199] and then room temperature [200, 201]. The majority of subsequent TM-doped II-VI efforts utilized either Fe or Cr doping, in materials such as ZnS, ZnSe, CdSe or CdMnTe. These materials, effectively being the mid-IR analog of the shorter wavelength Ti:sapphire gain medium, have certain very real advantages over mid-IR optoelectronic sources, in particular low-energy phonons (weak nonradiative decay) and broad gain bandwidths which allow for broadly tunable emitters, or alternatively, ultra-short pulse generation (figure 17). The significant development of TM-doped II-VI gain media, and their potential for mid-IR applications requiring short pulse widths, high CW power or high energy pulses has been chronicled in a series of excellent review articles [202–204]. However, the aforementioned lasers are all optically pumped (though often with great efficiency), and electrical pumping of these materials is yet to be demonstrated. While offering a unique mid-IR gain material, TM-doped II-VI is most likely poorly suited to use in compact, on-chip applications, or mid-IR opto-electronic device development.

Lead-salt materials, on the other hand, have a long history of use in optoelectronic mid-IR emitters, including the demonstration of mid-IR lasing from PbTe (λ = 6.7 μm) and PbSe (λ = 8.5 μm) in 1964 [205, 206], only two years after the demonstration of the first semiconductor lasers. For the next 30 years, the lead-salt laser was effectively the semiconductor laser of choice for the mid-IR [207–209], achieving CW operation at temperatures as high as 223 K at the time of the first QCL demonstration [210, 211]. The IV-VI materials have the advantage of lower Auger recombination rates (by 1–2 orders of magnitude compared to HgCdTe with similar bandgaps) [212] and a broadly tunable direct bandgap (as well as QW barrier materials) using ternary or quaternary Pb-alkyls containing Sn, Se, Sr, and/or Te. However, lead-salt lasers suffer from severe self-heating, and while this is a problem for all semiconductor lasers, it is exacerbated in
lead-salts by the extremely low thermal conductivity of the lead-salts themselves (a factor of 20 smaller than GaAs or InP), and the BaF$_2$ substrates upon which the majority of lead salt lasers have been grown (a factor of $\sim$5 lower than typical III-V substrates). The most recent lead-salt efforts employed optical pumping in vertical external cavity surface emitting laser (VECSEL) or VCSEL configurations [213], demonstrating pulsed emission at 4.5 $\mu$m above room temperature with the former [214]. Optically pumped microdisk cavity lasers have demonstrated CW emission near room temperature (2 °C) at wavelengths longer than 4 $\mu$m [215]. Interestingly, there have been multiple successful demonstrations of lead-salt mid-IR lasers grown directly on Si, using optically pumped VECSEL (as shown in figure 18) [216–218] or photonic crystal [219] configurations, with the latter operating at room temperature.

TM-doped II-VI lasers offer a wavelength-flexible gain medium for ultra-fast or high power/energy mid-IR sources. While there is always interest in moving towards electrically driven sources, it could be argued that the TM-doped II-VI already have a niche position as a broadband gain medium for ultrafast pulse generation: one which may become increasingly important as spectral coverage of the mid-IR saturates, and interest turns to sources with the ability to provide information in the time-domain. Lead-salt lasers may also have a role to play in the future of mid-IR optical systems. Although efforts focusing on lead salts waned with the emergence of the QCL, there remains interest in their continued development for broadly tunable lasers in surface emitting configurations. Lead-salts also hold particular promise for Si integration. Finally, both TM-doped II-VI and lead-salt material systems are intriguing candidates for mid-IR lasers based on quantum dot active regions, which is the subject of the next section.

5. Next-generation mid-IR sources: quantum dot emitters

Virtually all of the emitters discussed in the previous section—with the exception of TM-doped II-VI materials—utilize active regions with either type-I or -II interband transitions in QW or superlattice structures resulting from the 1D modulation of semiconductor band structure achieved by the epitaxial growth of planar layers. Such structures are more similar in design to the traditional and ubiquitous near-IR and visible diode lasers than the cascade lasers currently serving the majority of the mid-IR. Utilizing the diode laser design for long wavelengths requires significant adjustments, such as metamorphic buffer layers, quaternary and quinary materials, and oftentimes the use of less technologically mature substrates such as GaSb, InAs, or even BaF$_2$ in order to accommodate strain and at the same time increase the emitter’s operational wavelength. Key to the performance of these lasers is the ability to strongly confine carriers in 1D quantum well structures (most often type-I, though also type-II) in order to minimize the thermalization of carriers from the active region and to mitigate Auger scattering, which can be suppressed by final state engineering in heterostructure devices [220, 221]. Controlling carrier lifetimes is essential for the development of any semiconductor-based source, but perhaps even more so for emitters at long wavelengths, where nonradiative scattering processes can be significantly stronger than at shorter wavelengths. The importance of both minimizing nonradiative recombination in mid-IR lasers, as well as improving their temperature performance, has driven interest in QD-based active regions.

QDs, loosely defined, are semiconductor nanostructures that provide the three-dimensional (3D) confinement of carriers, leading to strong localization of electron and hole wavefunctions, as well as a delta-function density of states in the valence and conduction bands. These effects for QD-based lasers were predicted to result in a largely temperature-independent, low $J_{th}$ and high differential gain operation, and had been explored theoretically for interband lasers well before the ability to form QDs was experimentally demonstrated [222, 223]. Subsequent realization of QDs was achieved with a variety of approaches, including colloidal...
chemistry (primarily for visible wavelengths, but more recently for mid-IR QDs as well [24, 225]), top-down lithography, and self-assembled epitaxial growth. It was the self-assembled approach, however, leveraging the Stranski–Krstanow growth of QDs in lattice mismatched materials, that provided the experimental realization best-suited for the development of semiconductor-based QD lasers [226]. The projected benefits of QD lasers, predicted assuming uniform, perfect quantum box-like structures, never quite materialized due not only to nonuniformity in dot shape and size, but also to deviations from the perfect delta-function density of states caused by finite potential barriers and wetting layer formation in the epitaxially grown QD layers [227]. Nonetheless, interband QD lasers have been demonstrated to have a lower \( J_{th} \) [228], a nearly flat temperature dependence of \( J_{th} \) [229] and greater insensitivity to defects, making SAQDs ideal for the integration of III-V optoelectronic materials on Si [230].

The development of these QD-based emitters occurred at roughly the same time in which QCLs were drawing significant interest for the mid-IR; thus it is not surprising that increasing interest was given to efforts designed to leverage quantum dot or box-like structures for mid-IR applications.

### 5.1. Intersublevel quantum dot emitters

Interestingly, many of the same arguments trotted out to motivate interband QD devices were used to make the case for mid-IR quantum dot lasers, in particular the strong and narrow gain spectrum and improved temperature performance associated with the \( \delta \)-function density of states. An additional suggested benefit, for emitters based on intersublevel transitions in structures with cascade-like designs, was the so-called phonon bottleneck [231–233]. Here, the elimination of the continuum of states between the conduction band states reduces nonradiative phonon scattering of the excited state electrons, predicted to result in significant increases in carrier lifetimes, and therefore commensurate reductions in \( J_{th} \). The rapid phonon-assisted scattering from excited states in QCLs, as described earlier, results in decreased wall-plug efficiency, and makes the quantum cascade architecture largely unsuitable for subthreshold (LED) emitters. The ability to reduce upper state scattering was the primary motivation in the initial proposals for intersublevel QD emitters, which generally imagined their formation via the epitaxial growth of planar cascade-like devices, followed by top-down lithography and etching to create nanopillars with vertical quantization provided by the epi-grown heterostructure [234, 235].

The fabrication of quantum dots via top-down lithography, using any number of patterning techniques [236–239], offers the exact control of confinement in the growth direction (via epitaxy), and the deterministic control of the lateral positioning of the QDs. However, achieving the high density needed for efficient emission and the small feature size needed to reap the purported benefits of lateral quantization requires a small pitch and small lithographic features (on the order of 10s of nm). For features of this size, small variations in lateral dimension can cause significant changes to the intersublevel energy and the absolute position of energies. In addition, surface effects in such structures can greatly complicate both the design and operation. For this reason, the first demonstration of intersublevel emission from QDs came from a single layer of InAs QDs formed by Stranski–Krstanow self-assembly, positioned in an AlGaAs/GaAs cascade-like heterostructure [240]. Other efforts looked to integrate InAs QD layers into 20 periods of GaAs/AlGaAs superlattices, shown in figures 19(c), (d), with emission argued to come from transitions between excited QD states [241]. Subsequent efforts demonstrated improved temperature performance in a single layer device, with electroluminescence up to room temperature, as shown in figure 19(b), using a simplified graded AlGaAs injector and QW energy filter (figure 19(a)) [242]. There have been additional efforts to integrate QDs into cascade structures leveraging the InGaAs/InAlAs material system lattice matched to InP substrates, which are the most commonly used substrates for quantum cascade lasers, and which offer significant conduction band offset, and thus greater design flexibility. In this system, InAs deposition results in the formation of quantum dashes aligned along the [1–10] direction. When integrated into 50-period QC-like designs, electroluminescence up to room temperature was observed. The resulting emission was observed to be polarization-dependent [243], as would be expected from the elongated dashes (an effect also observed from InAs QDs grown on AlAs in an AlGaAs/GaAs heterostructure [244]). All of the above demonstrations of mid-IR QD emission show TE polarized emission, which serves as strong evidence that light emission is in fact a result of intersublevel transitions in 3D quantized structures (as mentioned earlier, selection rules mandate TM emission from intersubband transitions in QWs).

While initial results using intersublevel transitions in self-assembled quantum dots in QC-like structures showed promise, moving this work forward proved challenging. Ultimately, control of dot uniformity, size and shape posed significant design problems, especially when the QDs were required to be buried in complex heterostructures aligned to energy states in the dots. Thus, the approach described above required the control of dot sizes and shapes (and thus energies) such that the heterostructures could be built around the dots, so as to funnel electrons into excited dot states. However, changes in the heterostructure could also result in changes of the QD energy, making the control of QD energies dependent on the surrounding material and significantly complicating both the design and growth process. For this reason, it is only recently that lasing has been observed in cascade-like structures with embedded QDs [245], though for these devices, the QDs provided only marginal lateral confinement, and emission appears to be TM polarized, as for traditional QCLs. The difficulty in engineering three-dimen-
sional QDs with the layered heterostructures of the injector and extraction structures has proven difficult to overcome. However, alternative approaches to utilizing QDs in less standard optoelectronic material systems for mid-IR emission have shown some real potential. Such approaches will be discussed in the next sections.
5.2. Type-II quantum dot emitters

The development of intersublevel QD emitters for the mid-IR not only requires the growth of SAQDs with the appropriate intersublevel energy spacing and the design and growth of surrounding injection/extraction heterostructures, but also the alignment of energies in each of these regions of the device, in essence, adding a level of difficulty to the already complex QCL/ICL, and then compounding this additional challenge with lack of control inherent in the self-assembled nature of QD formation. As discussed in the introduction, there are obvious benefits to mid-IR emitter devices with straightforward diode-like designs, such as the narrow bandgap QW materials discussed in section 4. Such materials are promising at short wavelengths, but are limited at long wavelengths by lattice mismatch and the long wavelength limit of direct bandgap III-V materials (InAsSb). The low energy bandgap limit of InAsSb can be overcome by using type-II band alignments in InAs/InAsSb (staggered type-II) and InAs/InGaSb (broken gap type-II) material systems, as shown in figure 20 ([246]). In such materials, electrons in the InAs recombine, with holes confined in either the InAsSb or InGaSb layers, with the emitted photon energy controlled by the thickness of the material layers.

Variations of these material systems were presented in section 4, in quantum well and superlattice forms. The devices discussed used strain-balanced or lattice-matched systems for planar material systems with the 1D engineering of carrier wavefunctions. However, as figure 20(a) shows the percentage lattice mismatch between InSb and InAs (∼7%) is remarkably close to that between InAs and GaAs, a material combination whose lattice mismatch is understood to be a primary driver for the formation of SAQDs. The opportunity thus exists for the development of type-II quantum dots for mid-IR emitter applications. Such emitters would benefit from straightforward device architecture mimicking short wavelength interband diode lasers, as well as many of the advantages of 3D quantization presumed and observed for interband QD lasers.

Initial demonstration of InSb nanostructures utilized submonolayer (SML) QDs, formed by anion exchange during exposure of the InAs surface to Sb-overpressure, with 2D fill factors estimated between 0.6–1 ([247]). Such structures
provide weak confinement of holes, due to the ML InSb thickness, and thus emission close to the InAs band edge, with larger SML fill factors approaching wavelengths on the order of $\sim 3.8\ \mu m$ at 77 K, commensurate with the emission observed from ML InSb insertion layers in InAs [248]. SML InSb QD emitters have been used to demonstrate mid-IR lasers at low temperature (60 K) [247]. However, these devices emit within 50 meV of the InAs band edge, and no information regarding $J_{th}$ or temperature performance has been presented. Later work demonstrated room temperature emission from similar structures, with estimated internal quantum efficiencies and output power estimated to be 0.4% and 6 $\mu W$, respectively [249], and lasing, again near the InAs band edge, of up to 120 K [250]. The above work suggests that poor confinement in the thin SML QDs may not provide significant improvement (either in wavelength extension, or emission efficiency) from bulk InAs [251].

Stronger confinement of holes can be achieved by moving from SML to self-assembled QDs, formed by MBE [253], MOCVD droplet heteroepitaxy [254], or liquid phase epitaxy [255, 256]. The initial demonstrations of InSb QD growth focused more on growth parameters and structural characterization than optical properties, though subsequent efforts showed weak emission close to the InAs bandgap [257]. The addition of Ga to the InSb layers, in an effort to form In(Ga)Sb QDs on InAs, successfully increased emission wavelength to as far as $\sim 8\ \mu m$, though direct microstructural evidence for InGaSb QD formation was not shown [258, 259]. Although the demonstrations of In(Ga)Sb QD growth on InAs utilize a variety of epitaxial techniques, all suggest that the formation of these QDs is extremely sensitive to growth parameters, with subsequent growth studies suggesting that the window for InSb QD formation is rather narrow, but that the addition of Ga appears to result in the clear formation of InGaSb QDs with emission wavelengths in the 5–6 $\mu m$ range, demonstrating strong confinement of holes in the InGaSb QD valence band [248, 252]. Perhaps most intriguingly for the future development of QD-based mid-IR emitters is a recent work that suggests significant improvement in nonradiative recombination lifetimes—both Auger [260] and defect-related [248]—for such active regions (figure 21). As a result, type-II InGaSb QD-based emitters have demonstrated significant improvements in high-temperature emission quenching (a decrease in emission intensity by a factor of only 4–5 from 77 K to 250 K) [248]. The encouraging temperature performance and strong mid-IR emission from these QDs offers hope for future sources with not only potential for mid-IR LEDs and lasers, but perhaps devices leveraging the unique QD density of states for quantum information and communication applications in the mid-IR.

5.3. Alternative material quantum dot and nanocrystal emitters

III-V material systems, and in particular cascade laser architecture, have quickly become the dominant approaches in the generation of mid-IR radiation, replacing lead-salts, as mentioned earlier, which up until the mid-1990s, had been the material system of choice for mid-IR emitters. The rapid emergence of the QCL and the challenges associated with lead-salts as optoelectronic materials, quickly shifted the balance of mid-IR emitter research towards III-V materials. However, in addition to compatibility with Si, the demonstration of QD formation in lead-salts did provide some impetus for the continued development of lead-salt emitters. In particular, it was demonstrated that the deposition of PbSe/PbEuTe superlattices resulted in highly organized and reasonably controllable PbSe quantum dot crystals with strikingly uniform size distribution [261, 262]. The narrow size distribution of PbSe QDs, combined with the direct bandgap of PbSe, opened the door to the development of mid-IR QD lasers, in particular surface-emitting lasers, where QDs have an advantage over edge-emitting QCLs, which at the time were quickly becoming the clear device of choice for the mid-IR. Optically pumped lasing was demonstrated from PbSe QDs in the VCSEL configuration [263]. However, despite the projected advantages associated with QD-based emitters, and the high uniformity of PbSe QDs, lead-salt mid-IR QD lasers struggled to outperform even their bulk or QW material lead-salt counterparts [264], the latter of which has been shown to lase under optical excitation at temperatures as high as 2 °C [265]. The poor performance of lower dimensional QD emitters in this work was attributed to non-radiative...
recombination, which was most likely the result of defects associated with the QD layers [265].

The appeal of direct narrow bandgap lead-salt materials was such that efforts to utilize QDs from these materials continued, and shifted towards a new approach for QD formation, known as epitaxial precipitation. Here, coherent PbTe islands are formed by the low-temperature epitaxy of planar PbTe layers in CdTe matrices, followed by a postgrowth anneal, during which the PbTe precipitates into an array of coherent embedded QDs. The identical lattice constants and very different crystal structures of PbTe and CdTe (rocksalt and zincblende, respectively) result in strong segregation, and highly symmetric PbTe QD formation [266, 267]. The narrow bandgap of the PbTe results in mid-IR emission in the 2–3 μm range, tunable to shorter wavelengths with the addition of Sr, or longer with the addition of Sn, to the PbTe QDs [268, 269]. The symmetric, embedded and strongly segregated nanocrystalline nature of the precipitated QDs, allows room temperature mid-IR emission from LEDs using PbTe QDs as the active region [270], though the temperature dependence of the LED $I–V$ and emission suggests that charge transport and low-resistance contacts in the less mature IV-VI material system may prove to be challenging moving forward.

Interestingly, the same material systems are particularly appealing for the development of colloidal QDs emitting in the mid-IR. Colloidal QDs, as opposed to the self-assembled QDs discussed thus far in this section, are in fact the most well-developed, technologically mature variant of QDs, emitting most frequently at visible wavelengths, though sometimes extending to the near-IR [271]. Due to the relative ease of growth (in solution, as opposed to epitaxially at ultra-high vacuum), and their strong optical properties, colloidal QDs have been integrated into display technology, medical imaging applications, and even solar photovoltaics. However, very little work had been done to extend the center wavelengths of colloidal QDs out to the mid-IR.

**Figure 21.** The AFMs of In$_{0.6}$Ga$_{0.4}$Sb QD samples with a thickness of $t_{\text{InSb}} = 2.25$ ML grown at (a) high and (b) low Sb overpressure. (c) Temperature-dependent PL for the InGaSb QD sample (AFM shown in (b)). (d) Temperature-dependent integrated PL from samples (i) 1.75 ML InSb QW (black), (ii) 1.75 ML of In$_{0.6}$Ga$_{0.4}$Sb (green), (iii) 2.25 ML of In$_{0.6}$Ga$_{0.4}$Sb (blue), and (iv) 2.25 ML of In$_{0.6}$Ga$_{0.4}$Sb (red). (e) Experimental TRPL (solid) and modeled TRPL (dashed) for samples in (d) (with the exception of the 1.75 ML InSb sample). Reprinted from [252], with the permission of AIP Publishing.
The demonstration of mid-IR emission from lead-salt semiconductors, however, opened the door to a new class of long-wavelength colloidal QDs, first realized from PbSe QDs, and shown to be tunable from ∼2 to 4 μm [272]. Additional efforts demonstrated mid-IR emitting QDs not with lead-salts, but with the II-VI semiconductors HgTe [224], CdS and CdSe [273]. The HgTe nanoparticles, and the associated absorption and emission, are shown in figure 22. Control of the nanoparticle size allows the interband transition to be broadly tuned across the short wavelength mid-IR. Interestingly, the CdS and CdSe mid-IR emission comes not from interband transitions (both materials have large bandgaps) but from intraband transitions within the QD conduction band. These QDs showed strong degradation in emission intensity with exposure to an ambient atmosphere, argued to be a result of photo-oxidation effects. More recently, core–shell colloidal QDs of PbTe/CdTe have demonstrated mid-IR emission at 3 μm wavelengths, with little to no degradation in emission with exposure to atmosphere, which is presumably the result of the protective shell surrounding the PbTe QD core [274]. Core–shell QDs in the HgSe/CdS material system have also demonstrated emission out to 5 μm [275].

As they do for the bulk and QW mid-IR emitters previously discussed, TM-doped II-VI also provides an additional viable alternative to the III-V for the development of QD-based mid-IR sources. TM-doped II-VI micro- (Cr:ZnSe) and nanocrystals (Cr:ZnS) have been used to demonstrate room-temperature random lasing in the mid-IR [276, 277]. QDs (Fe:ZnSe), fabricated by microemulsion hydrothermal synthesis, with sizes in the 4–8 nm range, have been shown to emit in the 3.5–4.5 μm range, but with severely degraded performance approaching room temperature. Unlike III-V and lead-salt QDs, where emission energies are largely determined by QD size, transitions in TM-doped II-VI come from the states in the transition metal dopants, effectively encased in a mid-IR transparent glass matrix, and appear to be less susceptible to quantum size effects, though this does not appear to have been rigorously explored in the literature.

These materials thus offer an intriguing micro- or nano-scale gain medium at mid-IR frequencies, with the potential for fundamental investigation, or perhaps as optically pumped, highly localized sources of mid-IR light.

The development of mid-IR emitting nanoparticles (NPs) and colloidal QDs opens the door to intriguing new research.
directions in the mid-infrared. At shorter wavelengths, QDs have had significant technological impact on the development of consumer electronics and energy harvesting applications, but they have also been central to the development of a better understanding of quantum optics phenomena, providing the near ideal replication of a quantized two-level system. Mid-IR QDs could follow in much the same path. Already, mid-IR colloidal QDs have been used as light-absorbing material for mid-IR photodetectors, demonstrating the potential for optoelectronic devices based on QDs formed by chemistry and integrated into planar, electrically contacted, devices [225]. Future efforts may look to leverage individual NPs or QDs to act as point sources of mid-IR light for a range of quantum- and sub-diffraction optics experiments.

6. Summary and conclusions

The mid-IR, which is home to a wide range of molecular absorption signatures, multiple atmospheric transmission windows, and the peak thermal emission of biological and mechanical objects, has significant technological importance for a range of sensing, imaging and communication applications. At the same time, the mid-IR can serve as a test-bed for engineering a range of light–matter interactions at length scales achievable with current lithographic techniques, making it an intriguing space for studying fundamental optical phenomena. Key to both the fundamental and applied aspects of mid-IR research and development are sources of electromagnetic radiation in this vital wavelength range. In this review, we have attempted to present the history of mid-IR sources, starting with simple blackbody sources and moving through a wide range of optical and optoelectronic sources developed over the past decades. The steady narrative of mid-IR source development went through an inflection with the demonstration of the QCL, and later the ICL, which, it is now clear, were paradigm-changing developments in mid-IR photonics and optics. These novel sources permanently altered the mid-IR optical and optoelectronic landscape. Now, over 20 years since the first demonstration of the QCL, we are well-positioned to survey this landscape, and make a reasoned account of what is currently available and what is needed moving forward. Our goal in this review was not to suggest that new mid-IR sources are imminently set to overtake cascade lasers, but to ask what niche applications or capabilities these sources do not currently serve, and what technologies and materials might be capable of serving in their stead.

In some ways, this review is the intellectual equivalent of searching through scientific rubble for usable material systems and device designs buried by the disruptive force of cascade lasers, dusting off technologies which are decades-old, and which might warrant a second look now that we have a better sense of the capabilities and limitations of cascade lasers. So, the story of what is new for mid-IR sources is then partly a story of what was old. To tell this story we started with the very simplest mid-IR source there is: the blackbody. We discussed the obvious limitations of these sources and the motivation for efficient, narrowband sources of radiation in the mid-IR. The QCL and ICL are introduced, and every effort is made to acknowledge the tremendous impact of these sources, not only in the field of IR source development, but for IR optical systems as a whole. At the time that cascade lasers burst onto the scene, there was a range of ongoing efforts to develop and improve the existing mid-IR sources. The shift of research funding and effort towards cascade lasers had the effect of slowing—or even freezing—these parallel efforts in their tracks. These efforts included interband emitters using a double heterostructure or QW active regions with narrow bandgap materials grown on GaSb and InAs substrates, type-II interfaces, QWs, and superlattices also on GaSb and InAs substrates, as well as lead-salt materials, typically grown on BaF$_2$ substrates. Interestingly, efficient and reliable InGaN blue LEDs were announced around the same time as the first QCLs and ICLs, which abruptly ended most research on ZnSe-based green/blue light emitters.

Each noncascade mid-IR source faced serious shortcomings when compared to both QCLs and ICLs. The poor thermal properties of lead-salt devices—both in the active region and substrate—limited operation to low temperatures, while the narrow-gap type-I and type-II emitters suffered from significant Auger recombination, as well as challenges associated with growth on InAs and GaSb substrates. The QCL offered the opportunity to develop wavelength-flexible, engineered mid-IR light emitters, on InP or GaAs substrates. While the intersubband phonon scattering rates in QCLs certainly resulted in a high $J_{th}$ once lasing was achieved, slope efficiency was quite good. In addition, while the temperature performance of the cascade lasers took years to improve, these devices certainly did not face the inherent limitations of lead-salts. However, this is not to say that cascade lasers do not face inherent challenges to continued development. The ISB transition in QCLs results in TM-polarized emission and the aforementioned high $J_{th}$ values, while shorter wavelength QCL operation is difficult to achieve without the highly strained systems required to give large conduction band offsets. Both ICLs and QCLs require significant design and growth complexity, and are poorly suited to the sort of point-source emitters required for many sub-diffraction and/or quantum optics applications.

We then returned to our stable of pre-cascade laser sources for a fresh look, but with the benefit not only of improvements in epitaxial growth and materials science, but also a more complete understanding of the limitations of cascade laser technology. The short wavelength edge of the mid-IR (2.0–3.5 $\mu$m), which QCLs struggle to reach, is quickly being filled by interband QW light emitters either grown on GaSb substrates, or leveraging significant improvements in metamorphic buffer layer growth, on more technologically mature substrates, such as InP. These sources also offer opportunities for the development of mid-IR surface emitting devices (LEDS, VECSELs or VCSELs). Significant investment in type-II detectors has invigorated early work on superlattice emitters, which are now being developed as wavelength flexible superlattice emitting diode (SLED) sources for thermal scene projection applications. Optically pumped gain media, such as TM-doped II-VI, while probably...
not suitable for compact, on-chip applications, are gaining purchase for ultrafast, high pulse energy lasers, which are effectively the mid-IR equivalent of the shorter wavelength Ti:sapphire workhorse. Even veteran lead-salt materials show promise as a mid-IR gain medium, and can be grown directly on Si, thus making them a potential mid-IR source for long wavelength silicon-integrated photonics.

Quantum dots, which are nanoscale structures that confine charge carriers in all three dimensions, also offer a path forward towards new types of mid-IR sources. While early efforts to integrate epitaxial dots into cascade-like structures stalled, type-II quantum dots exhibit promising optical properties and improved temperature performance when compared to QWs in the same material system. Random lasing at mid-IR wavelengths in TM-doped II-VI nanocrystal powders, and colloidal QDs formed by wet chemistry, may signify a new front in mid-IR optoelectronics for spin-on, flexible mid-IR optoelectronics. New classes of epitaxial and colloidal QDs offer additional promise for the development of sub-diffraction mid-IR optoelectronics and as point sources of light for fundamental investigations and quantum optics efforts at long wavelengths.

Although it would be foolish to argue that cascade lasers will not be playing a dominant role in mid-IR source technology in the foreseeable future, it is also the case that there exist a wide range of materials and device geometries that are either currently able to, or have the potential to serve as mid-IR sources for a number of applications either outside the range of cascade lasers, or for which cascade lasers are poorly suited. In this review, we have provided an overview of such sources, presenting not only the state of the art for each source, but the challenges to further development and the potential applications for which these sources are suited. The continued growth of the mid-IR, fueled in large part by QCLs and ICLs, ensures that the sources discussed herein will likely continue to play a significant and expanding role as a mid-IR continua. Moreover, continued research into these materials may open new doors for the development of nanoscale, subdiffraction mid-IR optoelectronics, or ultrafast, high-power pulsed mid-IR systems. The mid-IR continues to play a significant and expanding role as a wavelength range for both applied technologies and fundamental investigations. The continued growth of the mid-IR ensures that there is room for a wide range of mid-IR materials, and the exciting promise of new sources for perhaps yet unknown applications.

Acknowledgments

DW gratefully acknowledges funding from NSF ECCS-1711858. MLL acknowledges partial support from NSF DMR-1713068. SB acknowledges funding from ARO W911NF-15-1-0612 and NSF ECCS-1408302.

ORCID iDs

D Wasserman @ https://orcid.org/0000-0003-3234-0803

References


[27] Williams B S 2007 Terahertz quantum cascade lasers Nat. Photonics 1 517


[31] Semtis M P, Ziegler M, Dressler S, Masselink W T, Georgiev N, Dekorsy T and Helm M 2004 Above room temperature operation of short wavelength (λ = 3.8 μm) strain-compensated In0.53Ga0.47As/InAs quantum-cascade lasers Appl. Phys. Lett. 85 1478


[33] Bandypadhyay N, Silviken S, Bai Y and Razeghi M 2012 High power, continuous wave, room temperature operation of λ ∼ 3.4 μm and λ ∼ 3.55 μm InP-based quantum cascade lasers Appl. Phys. Lett. 100 212104

[34] Revin D G, Cookburn J W, Steer M J, Airy R J, Hopkinson M, Krysa A B, Wilson L R and Menzel S 2007 InGaAs/AlAsSb/InPInGaAs/AlAsSb/InP quantum cascade lasers operating at wavelengths close to 3 μm Appl. Phys. Lett. 90 021108


[49] Allermand A A, Biefeld R M and Kurtz S R 1996 InAsSb-based mid-infrared lasers (3.8–3.9 μm) and light-emitting diodes with AlAsSb claddings and semimetal electron injection, grown by metalorganic chemical vapor deposition Appl. Phys. Lett. 69 465


[51] Tian Z, Li L, Ye H, Yang R Q, Mishima T D, Santos M B and Johnson M B 2012 InAs-based interband cascade lasers with emission wavelength at 10.4 μm Electron. Lett. 48 113


Garbuzov D Z, Martineili R U, Belenky G L, Garbuzov D Z and QW W C 2005 2.5 μm InGaAsSb/AlGaAsSb diode lasers emitting 1 W continuous waves Appl. Phys. Lett. 81 3146

Lin C, Grau M, Dier O and Amann M–C 2004 Low threshold room-temperature continuous-wave operation of 2.24–3.04 μm GaInAsSb/AlGaAsSb quantum-well lasers Appl. Phys. Lett. 84 5088

Shterengas L, Belenky G, Kisin M V and Donetsky D 2009 High power 2.4 μm heavily strained type-I quantum well GaSb-based diode lasers with more than 1 W of continuous wave output power and a maximum power-conversion efficiency of 17.5% Appl. Phys. Lett. 90 011119

Belenky G, Shterengas L, Kipshidze G and Hosoda T 2001 Type-I diode lasers for spectral region above 3 μm IEEE J. Sel. Top. Quant. 17 1420


Hosoda T, Belenky G, Shterengas L, Kipshidze G and Kisin M V 2008 Continuous-wave room temperature operated 3.0 μm type GaSb-based lasers with quantumnary AlInGaAsSb barriers Appl. Phys. Lett. 92 091106

Vizbaras K and Amann M C 2012 Room-temperature 3.73 μm GaSb-based type-I quantum-well lasers with quantumnary barriers Semicond. Sci. Tech. 27 032001

Kaspi R, Lu C A, Yang C and Luong S 2015 GaSb-based >3 μm laser diodes grown with up to 2.4% compressive strain in the quantum wells using strain compensation J. Cryst. Growth 424 24


Ortsiefer M, Neumeyr C, Rosskopf J, Arafis S, Bohm G, Hangauer A, Chen J, Strzoda R and Amann M–C 2011 GaSb and InP-based VCSELs at 2.3 μm emission wavelength for tunable diode laser spectroscopy of carbon monoxide Quantum Sensing and Nanophotonic Devices VIII p 7945


Sato T, Mitsuhara M, Nunoya N, Fujisawa T, Kasaya K, Kan F and Kondo Y 2008 2.33 μm wavelength distributed feedback lasers with InAs-In0.53Ga0.47As multiple-quantum wells on InP substrates IEEE Photonics Technol. Lett. 20 1045

Jung D, Song Y C, Yu L, Wasserman D and Lee M L 2012 2.8 μm emission from type-I quantum wells grown on InAsP-InP metamorphic graded buffers Appl. Phys. Lett. 101 251107

Kirch J et al 2010 InAs0.34P0.66/1–3 μm metamorphic buffer layers on InP substrates for mid-IR diode lasers J. Crystal Growth 312 1165

Cao Y Y, Zhang Y G, Gu Y, Chen X Y, Zhou L and Li H 2013 2.7 μm InAs quantum well lasers on InP-based InAlAs metamorphic buffer layers Appl. Phys. Lett. 102 201111


Jung D, Yu L, Wasserman D and Lee M L 2015 Mid-infrared electroluminescence from InAs type-I quantum wells grown on InAsP/InP metamorphic buffers J. Appl. Phys. 118 183101


Pan C H, Lin S D and Lee C P 2010 2–3 μm mid infrared light sources using InGaAs/GaAsSb ‘W’ type quantum wells on InP substrates J. Appl. Phys. 108 103105

Sprengel S, Grasse C, Vizbaras K, Gruendl T and Amann M C 2011 Up to 3 μm light emission on InP substrate using GaInAs/GaAsSb type-II quantum wells Appl. Phys. Lett. 99 221109

Pan C H, Chang C H and Lee C P 2012 Room temperature optically pumped 2.56 μm lasers with ‘W’ type InGaAs/GaAsSb quantum wells on InP substrates IEEE Photon. Tech. Lett. 24 1145


Sakaki H, Chang L L, Ludeke R, Chang C-A, SAI-HALASZ G A and Easki L 1997 In0.53Ga0.47As/InAsSb1−μ heterojunctions by molecular beam epitaxy Appl. Phys. Lett. 31 213

Kroemer H 2004 The 6.1 angstrom family (InAs, GaSb, AlSb) and its heterostructures: a selective review Physica E 20 196


Smith D L and Mailhiot C 1987 Proposals for strained type II superlattice infrared detectors J. Appl. Phys. 62 2545

Mailhiot C and Smith D L 1987 Electronic structure of (001) and (111) growth axis InAs–Ga1−μInSb strained-layer superlattices J. Vac. Sci. Technol. B 5 1268

Wei S H and Zunger A 1995 InSb/InAs: a type-I or a type-II band alignment Phys. Rev. B 52 12039


Zhang Y H 1995 Continuous wave operation of InAs/InAsSb1−μ, midinfrared lasers Appl. Phys. Lett. 66 118


[133] Li B et al 1997 Band alignments and offsets in In(As, Sb)/ InAs superlattice Electron. Lett. 33 4589

[134] Wilk A et al 2001 MBE growth of InAs/InAsSb/AlAsSb structures for mid-infrared lasers J. Cryst. Growth 232 341


[136] Lackner D, Pitts O J, Sieger M, Yang A, Thewalt M L W and Watkins S P 2009 Strain balanced InAs/InAsSb superlattice structures with optical emission to 10 µm Appl. Phys. Lett. 95 081906


[215] Eibelhuber M, Schwarz T, Pichler S, Heiss W and Springholz G 2010 Near room temperature continuous-wave laser operation from type-I interband transitions at wavelengths beyond 4 μm Appl. Phys. Lett. 97 061103


[228] Liu G T, Stintz A, Li H, Malloy K J and Lester L F 1999 Extremely low room-temperature threshold current density diode lasers using InAs dots in In0.53Ga0.47As quantum well Electron. Lett. 35 1163

[229] Sugawara M and Usami M 2009 Quantum dot devices: handling the heat Nat. Photonics 3 30


[240] Wasserman D and Lyon S A 2002 Midinfrared luminescence from InAs quantum dots in unipolar structures Appl. Phys. Lett. 81 2648


[244] Wasserman D, Gimachi C, Lyon S A and Shaner E A 2006 Multiple wavelength anisotropically polarized mid-infrared emission from InAs quantum dots Appl. Phys. Lett. 88 191118


[247] Ivanov S V et al 2005 Molecular beam epitaxy of type II InSb/InAs nanostructures with InSb sub-monolayers J. Cryst. Growth 278 72


[265] Eibelhuber M, Schwarzl T, Pichler S, Heiss W and Springholz G 2010 Near room temperature continuous-wave laser operation from type-1 interband transitions at wavelengths beyond 4 μm Appl. Phys. Lett. 97 061103


[267] Groiss H et al 2007 Size control and midinfrared emission of epitaxial PbTe/CdTe/PbTe/CdTe quantum dot precipitates grown by molecular beam epitaxy Appl. Phys. Lett. 91 222106


[275] Deng Z and Guyot-Sionnest P 2016 Intraband luminescence from HgSe/CdS core–shell quantum dots ACS Nano 10 2121
