Laser cooling of semiconductor quantum wells: Theoretical framework and strategy for deep optical refrigeration by luminescence upconversion

Jianzhong Li*

NanoScience Solutions, Inc., Cupertino, California 95014, USA

(Received 24 November 2005; revised manuscript received 8 February 2007; published 12 April 2007)

Optical refrigeration has great potential as a viable solution to thermal management for semiconductor devices and microsystems. We have developed a first-principles-based theory that describes the evolution of thermodynamics—i.e., thermokinetics—of a semiconductor quantum well under laser pumping. This thermokinetic theory partitions a well into three subsystems: interacting electron-hole pairs (carriers) within the well, the lattice (thermal phonons), and the ambient (a thermal reservoir). We start from the Boltzmann kinetic equations and derive the equations of motion for carrier density and temperature, and lattice temperature, under the adiabatic approximation. A simplification is possible as a result of ultrafast energy exchange between the carriers and phonons in semiconductors: a single-temperature equation is sufficient for them, whereas the lattice cooling is ultimately driven by the much slower radiative recombination (upconverted luminescence) process. Our theory microscopically incorporates photogeneration and radiative recombination of the interacting electron-hole pairs. We verify that Kubo-Martin-Schwinger relation holds for our treatment, as a necessary condition for consistency in treatment. The current theory supports steady-state solutions and allows studies of cooling strategies and thermodynamics. We show by numerical investigation of an exemplary GaAs quantum well that higher power cools better when the laser is detuned from the band edge between a *critical* negative value and the ambient thermal energy. We argue for the existence of such a counterintuitive lower bound. Most importantly, we show that there exists an *actual* detuning, 3 meV above the band edge in the simulated free-carrier case and expected to be pinned at the excitonlike absorption peak owing to Coulomb many-body effects, for optimal laser cooling. Significant improvement in cooling efficacy and theoretical possibility of deep refrigeration are verified with such a fixed optimal *actual* detuning. In essence, this work provides a consistent microscopic framework and an optimization strategy for achieving net deep cooling of semiconductor quantum wells and related microsystems.

DOI: 10.1103/PhysRevB.75.155315 PACS number(s): 78.55.Cr, 32.80.Pj, 78.67.De, 78.20.Bh

I. INTRODUCTION

Recent advances in laser cooling of rare-earth-doped glasses and dye-added fluids[1–3] foreshadows the current pursuit with semiconductors. The cooling mechanism is luminescence upconversion, which was proposed for optical refrigeration of matter decades ago.[4] It removes heat from material under refrigeration in three steps, as schematically shown in Fig. 1: (i) A laser beam illuminates the material and creates cold electron-hole (e-h) pairs, (ii) the pairs experience energy exchange through inelastic collision and thereby absorb thermal lattice vibrations, and then (iii) the pairs convert into blueshifted, incoherent photons through radiative recombination and carry heat away from the material. As a result, the material is cooled down by the laser beam.[5] However, other processes can interfere and degrade the cooling operation. For instance, some carriers could lose their energy through nonradiative recombination, instead of the desired radiative kind, and dump the absorbed laser photon energy as heat inside the material. Also, it is highly likely that the photoluminescence (PL) generated in step (iii) cannot escape from the material because of internal reflection (semiconductors have typical refractive indices around 3.4, relative to air's unit value, correspnding to a critical angle of 17.1° for total internal reflection), but recycles in the material. Eventually, the recycled photons either get away or get absorbed and generate new e-h pairs. This recycling process could repeat itself, if without intervention. Ongoing effort mitigates this recycling problem by introducing an antireflective coating through materials engineering and/or by liquid immersion approach that effectively increases the refractive index of the surrounding medium. Whereas optical refrigeration is expected to offer a viable solution to active thermal management of semiconductor devices and microsystems that is nonmechanical (thus vibration-free and noiseless), reliable, and integratable, the phenomenon has yet to be successfully demonstrated in semiconductors. Photon recycling and carrier trappings that lead to nonradiative recombination are shown to be the reasons.[6,7]

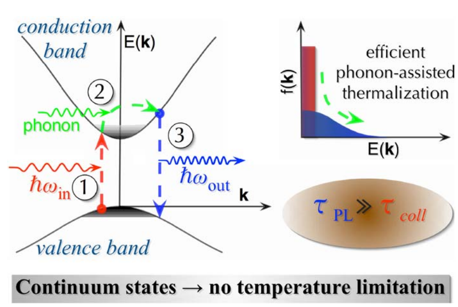

FIG. 1. (Color online) Schematic of luminescence upconversion for laser cooling of semiconductors. Ultrafast carrier-longitudinal optical phonon collision aids efficient thermalization (2) of photogenerated cold carriers (1) and the sequential luminescence upconversion (3). Cooling is achieved as a result of heat extraction by the thermal photoluminescence.

On the theoretical front, effort[8–13] is limited and focuses on bulk materials. Phenomenologically, Bowman[8] analyzed heat generation by an optical pumping and cooling effect by anti-Stokes spontaneous emission (luminescence upconversion) in a solid-state laser system, thus proposed the "radiation balanced laser" to achieve high output power; Mungan[9] quantified the steady-state efficiency of such a doped fiber laser from thermodynamic laws; Sheik-Bahae and Epstein[10] investigated the effects of external efficiency and photon recycling on net optical cooling in GaAs. In addition, a rate-equation approach[13] for carrier and photon densities including photon recycling was used recently to quantify the cooling efficiency, but without temperature dependence. On the other hand, a microscopic description is rare: the results of Huang *et al.*[11,12] were the only known results to the author at the time of submission of this contribution. The studies adopted a nonlocal energy-balance equation for carrier temperatures, without consideration of carrier density. These early studies addressed such critical issues as experimental conditions, efficiency, and limitations, but not to a satisfactory level as no understanding regarding the intrinsic and dynamic behaviors is achieved. The main reason is that a complete and consistent theoretical framework is lacking. Such a framework would allow for a critical examination of the issues, both intrinsic and extrinsic, that prevent us from an ultimate demonstration of laser cooling of semiconductors. Particularly, such a framework would facilitate simulations of the cooling operations, generate sufficient and necessary understanding and insights, and provide indispensable guidance for achieving net cooling of semiconductors. To this end, we present a first-principles-based, consistent thermokinetic theory for the description of laser cooling of quantum wells (QWs) in this work. We focus on intrinsic processes and thus the fundamental limit of laser cooling, and ignore photon recycling and carrier trappings, as well as Auger process. Incorporation of these physical processes into the current framework is straightforward and shall be explored in further studies.

The remainder of the article is organized as follows: First, we set up the physics foundation for the thermokinetic theory and lay out the starting equations in Sec. II. This is followed in Sec. III by a derivation of the main results of this work: the equations of motion for carrier density, carrier and phonon thermal energy, and a conversion to the equation for lattice temperature. Then we treat photogeneration and radiative recombination microscopically in Sec. IV and produce the necessary radiative contributions to the equations of motion, including noninteracting results for illustration. A derivation of the Kubo-Martin-Schwinger relation is conducted in Sec. V to demonstrate the consistency of the theoretical framework, as well as analytical expressions for photogeneration and radiative recombination with many-body effects. To illustrate the implications and utility of the developed framework we present in Sec. VI simulation results and explore the cooling strategy. In Sec. VII technical issues and open questions pertinent to the theoretical framework are discussed, as well as Coulomb many-body effects. This contribution is concluded with a summary in Sec. VIII.

II. MODEL AND BASIC EQUATIONS

The thermokinetic theory is developed from the kinetic equations of the statistical distributions of elementary excitations in a semiconductor quantum well system. We partition a semiconductor quantum well under laser pumping into three subsystems: (i) carriers that are generated by a laser beam, (ii) thermal longitudinal optical (LO) phonons that are the elementary excitations associated with the underlying lattice, and (iii) the ambient. The carriers are electrons and holes, and are confined in the quantum well. Phonons consist of acoustic and optical kinds, of both longitudinal and transverse modes. In particular, the LO phonons are the most effective ones that exchange energy with the carriers through inelastic collision. Due to their vast difference from the others, in terms of the energy exchange time scale with the carriers, the LO phonons (in subpicoseconds) are singled out as the only constituent of the second subsystem, whereas the acoustic ones and transverse optical modes (with time scales in nanoseconds to microseconds) are relegated to be part of the ambient. The ambient, treated as a thermal reservoir, basically includes all forms of elementary excitations that interact with the first two subsystems on a slower time scale and render them incoherent ultimately. To account for the energetics that drives the system into a steady state under laser pumping, our theory microscopically incorporates the photogeneration and radiative recombination of e-h pairs. The steady state is a consequence of thermokinetic balance among the subsystems. First, the laser beam generates nonthermal e-h pairs in the QW through optical excitation. For a monochromatic wave, as considered here, these pairs are distributed along an isoenergetic contour in their joint phase space because of energy and momentum conservation. Next, carriers and LO phonons thermalize as a result of carrier-carrier and carrier–LO-phonon collisions. We assume that *quasiequilibrium* is sustained for carriers and LO phonons by the ultrafast carrier-carrier and carrier–LO-phonon collisions so that the *carriers are well described with Fermi-Dirac distribution*[14,15] and the *phonons are well described by Bose-Einstein distribution throughout the cooling process.* Finally, thermal exchange with the ambient, termed thermal loading, by the lattice pumps heat into the lattice and checks the runaway cooling of the carriers and the lattice.

We start from the Boltzmann kinetic equations for carrier and phonon distributions $(f_{\boldsymbol{k}}^c, n_{\boldsymbol{Q}})$ in momentum space [in-plane wave vector $\boldsymbol{k}$ for carriers and three-dimensional (3D) wave vector $\boldsymbol{Q}$ for phonons]:

$$\dot{f}_{\boldsymbol{k}}^c(t) = \dot{f}_{\boldsymbol{k}}^c(t)|_{abs} - \dot{f}_{\boldsymbol{k}}^c(t)|_{rad} + \dot{f}_{\boldsymbol{k}}^c(t)|_{c\text{-}p}, \quad (1)$$

$$\dot{n}_{\boldsymbol{Q}}(t) = \dot{n}_{\boldsymbol{Q}}(t)|_{e\text{-}p} + \dot{n}_{\boldsymbol{Q}}(t)|_{h\text{-}p} + \dot{n}_{\boldsymbol{Q}}(t)|_{amb}, \quad (2)$$

where the evolution of carrier $(c=e,h)$ distributions, represented by its time derivative denoted by $\dot{f}_{\boldsymbol{k}}^c(t)$, is driven by optical absorption $(|_{abs})$, which generates carriers, radiative recombination $(|_{rad})$, which annihilates carriers, and carrier-phonon collision $(|_{c\text{-}p})$, which redistributes carriers. The change in phonon population is caused by carrier-phonon collision, which emits or absorbs phonons, and thermal exchange with the ambient $(|_{amb})$, which alters lattice temperature, and thus phonon population. The source terms on the right-hand side in the above equations, except the ambient contribution, are systematically obtained by a quantum kinetic treatment from a Hamiltonian that takes into account carriers, LO phonons, photons, and their interactions as the Boltzmann kinetic equations themselves are derived when spatial gradients are neglected for a uniform system. The derivation is nontrivially involving and invokes the well-known Markovian approximation and semiclassical limit of the operator equations. Interested readers can consult Refs. 16–19 for technical details. For completeness, we list the results used in this work as follows:

$$\dot{f}^h_{\boldsymbol{k}}(t)|_{abs} \equiv \frac{2}{\hbar}\,\mathrm{Im}\{d^*_{\boldsymbol{\kappa}_\parallel}\mathcal{E}(t)p_{\boldsymbol{k}}(\boldsymbol{\kappa}_\parallel)\}, \tag{3}$$

$$\dot{f}^e_{\boldsymbol{k}+\boldsymbol{\kappa}_\parallel}(t)|_{abs} = \dot{f}^h_{\boldsymbol{k}}(t)|_{abs}, \tag{4}$$

$$\dot{f}^e_{\boldsymbol{k}}(t)|_{rad} \equiv \frac{2}{\hbar}\,\mathrm{Re}\left\{\sum_{\boldsymbol{q}} \mathcal{E}_{\boldsymbol{q}} u^*_{QW,\boldsymbol{q}} d^*_{\boldsymbol{q}_\parallel}\langle b^\dagger_{\boldsymbol{q}}\hat{p}_{\boldsymbol{k}-\boldsymbol{q}_\parallel}(\boldsymbol{q}_\parallel)\rangle\right\}, \tag{5}$$

$$\dot{f}^h_{\boldsymbol{k}}(t)|_{rad} \equiv \frac{2}{\hbar}\,\mathrm{Re}\left\{\sum_{\boldsymbol{q}} \mathcal{E}_{\boldsymbol{q}} u^*_{QW,\boldsymbol{q}} d^*_{\boldsymbol{q}_\parallel}\langle b^\dagger_{\boldsymbol{q}}\hat{p}_{\boldsymbol{k}}(\boldsymbol{q}_\parallel)\rangle\right\}, \tag{6}$$

where $d_{\boldsymbol{\kappa}_\parallel}$ is the interband dipole matrix element, $\mathcal{E}(t)$ is the pumping laser electric field with in-plane wave vector $\boldsymbol{\kappa}_\parallel$, $V^s_{\boldsymbol{k}-\boldsymbol{k}'}$ is the screened Coulomb potential, and $p_{\boldsymbol{k}}(\boldsymbol{\kappa}_\parallel)$ is the interband polarization; $\mathcal{E}_{\boldsymbol{q}}=\sqrt{\hbar\omega_{\boldsymbol{q}}/2\epsilon_0\epsilon_b}$ is the vacuum-field amplitude for an optical mode with circular frequency $\omega_{\boldsymbol{q}} = c|\boldsymbol{q}|/n$ and wave vector $\boldsymbol{q}\equiv(\boldsymbol{q}_\parallel, q_\perp)$ in the background medium (here the barrier material) with dielectric constant ϵ_b and refractive index $n=\sqrt{\epsilon_b}$, $u_{QW,\boldsymbol{q}}$ is the overlap integral of the QW confinement wave functions and the optical mode function with wave vector $\boldsymbol{q}$, $b^\dagger_{\boldsymbol{q}}$ is the photon creation operator, and $\hat{p}_{\boldsymbol{k}}(\boldsymbol{q}_\parallel)$ is the interband polarization operator between hole state $|v-\boldsymbol{k}\rangle$ and electron state $|c\boldsymbol{k}+\boldsymbol{q}_\parallel\rangle$. The parallel ($_\parallel$), or in-plane, component of a vector lies in the QW plane, and the vertical ($_\perp$) one is perpendicular to the plane. The quantity $\langle b^\dagger_{\boldsymbol{q}}\hat{p}_{\boldsymbol{k}}(\boldsymbol{q}_\parallel)\rangle$, termed the three-point correlation function, measures the creation amplitude of a photon with an in-plane momentum $\hbar\boldsymbol{q}_\parallel$ when an electron-hole pair with the same center-of-mass momentum recombines. The remaining notations have the standard mathematical and physics meaning. Note that the equalities $\Sigma_{\boldsymbol{k}}\dot{f}^e_{\boldsymbol{k}}(t)|_{\mathrm{abs}} = \Sigma_{\boldsymbol{k}}\dot{f}^h_{\boldsymbol{k}}(t)|_{\mathrm{abs}}$ and $\Sigma_{\boldsymbol{k}}\dot{f}^e_{\boldsymbol{k}}(t)|_{\mathrm{rad}} = \Sigma_{\boldsymbol{k}}\dot{f}^h_{\boldsymbol{k}}(t)|_{\mathrm{rad}}$ are always valid, as expected for pairwise processes, where the sum over $\boldsymbol{k}$ implicitly accounts for spin degree of freedom of the carriers.

Photogeneration and luminescence are treated independently in this work. As known, photons need to be second quantized in order to describe luminescence. As a result, both the photon operator and interband polarization operator appear in the theoretical description of luminescence, and standard quantum statistical ensemble averaging $\langle\cdots\rangle$ has to be taken, whereas for absorption, the laser field is semiclassical, and only the polarization operator undergoes the averaging process, which results in its value $p_{\boldsymbol{k}}(\boldsymbol{\kappa}_\parallel)\equiv\langle\hat{p}_{\boldsymbol{k}}(\boldsymbol{\kappa}_\parallel)\rangle$ in Eq. (3). Specifically, the interband polarization is described by the semiconductor Bloch equations,[17,18] and the three-point correlation function for spontaneous emission is described by the semiconductor luminescence equations,[18] as briefly described in Sec. IV. Furthermore, the carrier–LO-phonon collision is not explicitly used in the final thermokinetic theory, as a result of the merger of the carrier subsystem and phonon subsystem, which is delineated in the upcoming section. Therefore, we do not list the related expressions here, but arguments in support of the merger are given. For completeness, interested readers should consult the author's earlier work, Refs. 15 and 20. Finally, our theory considers thermal exchange between the ambient and the QW through thermal radiation. Doubtlessly, thermal transport occurs as a result of local cooling of the QW with respect to the background medium. Whereas transport is driven by the temperature gradient, thermal radiation is proportional to the difference in temperatures with a fourth power, as well known for the blackbody radiation law. Without much compromise in rigor, we neglect thermal transport processes with a weaker dependence on temperature difference and assume that thermal exchange through radiation dominates. Interested readers can refer to the latest work of Huang *et al.*[12] where thermal transport across the QW is considered.

III. THERMOKINETIC EQUATIONS

Computationally, it is daunting and unrealistic to simulate laser cooling thermokinetics with Eqs. (1) and (2). Further simplification is necessary. Fortunately, such a task is possible and straightforward. The underlying reason is that there exist several ultrafast collision processes within the subsystems and between the subsystems. In particular, we have mentioned carrier-carrier and carrier–LO-phonon collisions, which all occur on a subpicosecond time scale. Because of these ultrafast collision processes, each individual subsystem sustains a quasithermal distribution such that a thermokinetic description is possible and sufficient. In other words, Eqs. (1) and (2) overdescribe the laser cooling process. This line of physical analysis is in general termed the *adiabatic approximation*. Thanks to this approximation and a few observations, it is algebraically straightforward to obtain the equations of motion for the carrier density (N), carrier thermal energy (E_c), and LO-phonon thermal energy (E_p), by integrating Eq. (1) over $\boldsymbol{k}$, Eq. (1) over $\boldsymbol{k}$ weighted with e-h kinetic energy, and Eq. (2) over $\boldsymbol{Q}$ weighted with LO-phonon energy $\hbar\omega_{\boldsymbol{Q}}$, respectively, and the results are as follows:

$$\dot{N} = aF(t) - BN^2, \tag{7}$$

$$\dot{E}_c = a_E F(t) - B_E N^2 + \dot{E}_c|_{e\text{-}p} + \dot{E}_c|_{h\text{-}p}, \tag{8}$$

$$\dot{E}_p = \sigma(T^4_{amb} - T^4_p) - \dot{E}_c|_{e\text{-}p} - \dot{E}_c|_{h\text{-}p}. \tag{9}$$

Whereas a is the QW absorbance, a_E describes the carrier kinetic (thermal) energy gain per unit photon flux. B is the radiative recombination coefficient, while B_E is its counterpart for carrier thermal energy. $\dot{E}_c|_{e\text{-}p}$ and $\dot{E}_c|_{h\text{-}p}$ are the energy exchange rates due to electron–LO-phonon and hole–

LO-phonon collisions. Thermal loading by the ambient at temperature T_{amb} is approximated by the blackbody radiation law:[11] σ is the Stefan-Boltzmann constant. The photon flux $F(t)$ relates to laser power and electric field by $P(t) = \hbar\omega F(t) = cn\epsilon_0|\mathcal{E}(t)|^2/2$. Note that a and a_E are functions of detuning (difference between photon energy $\hbar\omega$ and QW band gap, as defined later), N, and T_p (lattice temperature); B and B_E depend on N and T_p explicitly. Their expressions are given in the next two sections. The observations made when deriving the above equations are that (i) electrons and holes are fermions so that independent equations are needed for their number (density) and energy (temperature), whereas phonons are bosons so that only one equation is needed, here for their energy (temperature); (ii) a carrier-phonon collision only facilitates energy exchange between these two subsystems, and the carrier density is conserved during the process; and (iii) photogeneration depends linearly on photon flux when the laser power is reasonably low so that fundamental nonlinear optical processes can be neglected (thus the linear form).

At this point, it is possible to close the thermokinetic equations. However, two insights are instrumental that result in further simplification, in terms of computational budget and algebra. First, equilibration between the e-h subsystem and the LO-phonon subsystem is characterized by a subpicosecond temperature relaxation time scale thanks to efficient energy exchange between the carriers and the LO phonons.[20] In comparison, a pumping-dependent but much slower density relaxation rate, proportional to BN and typically in nanoseconds, characterizes the carrier density kinetics. Second and equally important, the heat capacity of the e-h subsystem is orders of magnitude smaller than that of the phonon subsystem, even at liquid nitrogen temperature. These two insights dictate that Eqs. (8) and (9) can be combined into a single one:

$$\dot{E}_p = \sigma(T_{amb}^4 - T_p^4) + a_E F(t) - B_E N^2. \quad (10)$$

As a result, the subpicosecond thermokinetics due to carrier–LO-phonon collisions does not show in the final form of our theory, represented by Eqs. (7) and (10). We note that Eq. (10) here is formally analogous to Eq. (5) in Ref. 11. *But only* in combination with Eq. (7) does a consistent description of the thermokinetics of laser cooling become complete. Finally, we close Eq. (10) for LO phonons using $E_p = \Sigma_Q n_Q \hbar\omega_Q$ to relate the LO-phonon thermal energy E_p to the phonon temperature T_p. As mentioned, the other phonon modes are treated as part of the ambient and thermokinetically ignored in the current theory. In essence, Eqs. (7) and (10) allow us to simulate laser cooling thermokinetics with the knowledge of photogeneration and radiative recombination rates.

Before proceeding to the derivation of the explicit form of the rates for formal closure, it would be auxiliary to emphasize that the final form of our theory cannot be characterized as phenomenological rate equations for the thermokinetic variables N and T_p. Even though the equations are cast deceivingly in a familiar form in order to bring out some transparency in the relevant physics, all the coefficients (a, B, a_E, and B_E) are variable dependent and incorporate relevant microscopic physics (refer to the following two sections for details). In particular, since the direct process of radiative recombination is microscopically binary and has been conventionally described by a square power law, its rates are written in the same form here. For these reasons, it would be erroneous to take the theory as a phenomenological one.

IV. TREATMENT OF PHOTOGENERATION AND PHOTOLUMINESCENCE

We now supplement our theory with a microscopic treatment of photogeneration and photoluminescence in semiconductor quantum wells so as to derive the photogeneration and radiative recombination rates, thus the a, B, a_E, and B_E coefficients formally. The treatment is at the screened Hartree-Fock level, and correlation effects are phenomenologically modeled under the dephasing rate (γ_p below) approximation. For brevity, we drop the explicit time dependence of carrier distributions hereafter. First of all, optical absorption via interband process is described by the semiconductor Bloch equation[17,18]

$$(\varepsilon^e_{\boldsymbol{k}+\boldsymbol{\kappa}_\parallel} + \varepsilon^h_{-\boldsymbol{k}} - i\gamma_p)p_{\boldsymbol{k}}(\boldsymbol{\kappa}_\parallel) - i\hbar\dot{p}_{\boldsymbol{k}}(\boldsymbol{\kappa}_\parallel) = (1 - f^e_{\boldsymbol{k}+\boldsymbol{\kappa}_\parallel} - f^h_{-\boldsymbol{k}})\left(d_{\boldsymbol{\kappa}_\parallel}\mathcal{E}(t) + \sum_{\boldsymbol{k}'} V^s_{\boldsymbol{k}-\boldsymbol{k}'}p_{\boldsymbol{k}'}(\boldsymbol{\kappa}_\parallel)\right), \quad (11)$$

where $\varepsilon^c_{\boldsymbol{k}}$ $(c=e,h)$ is the renormalized carrier energy[17,18] relative to the conduction band edge. For a monochromatic pumping laser beam $\mathcal{E}(t) = \mathcal{E}_0[\exp(-i\omega t) + \text{c.c.}]$ with a circular frequency $\omega = c|\boldsymbol{\kappa}|/n$ and an in-plane wave vector component $\boldsymbol{\kappa}_\parallel$, a driven (resonant) solution for the polarization can be sought by invoking the ansatz $p_{\boldsymbol{k}}(\boldsymbol{\kappa}_\parallel) = \tilde{p}_{\boldsymbol{k}}(\boldsymbol{\kappa}_\parallel)e^{-i\omega t}$, while ignoring the counterrotating portion (the *rotating-wave approximation*) and an envelope part associated with the slow change in the carrier distributions. Note that for a laser beam with a slowly varying power profile, the same treatment applies. By defining the $\boldsymbol{k}$-resolved susceptibility according to $\tilde{p}_{\boldsymbol{k}}(\boldsymbol{\kappa}_\parallel) \equiv \chi_{\boldsymbol{k}}(\boldsymbol{\kappa}_\parallel)\mathcal{E}_0$, the QW optical susceptibility is obtained via $\chi(\omega) \equiv \Sigma_{\boldsymbol{k}} d^*_{\boldsymbol{k}}\chi_{\boldsymbol{k}}(\boldsymbol{\kappa}_\parallel)/\epsilon_0$ and the QW absorbance can be found using $a = \omega\,\text{Im}\{\chi(\omega)\}/cn$. The thermal energy gain can be calculated with a corresponding summation weighted by the carrier kinetic energies. Thereby, formal expressions for the QW absorbance and the thermal energy gain due to optical absorption can be written as follows:

$$a = \frac{\omega}{cn\epsilon_0}\sum_{\boldsymbol{k}} \text{Im}\{d^*_{\boldsymbol{\kappa}_\parallel}\chi_{\boldsymbol{k}}(\boldsymbol{\kappa}_\parallel)\}, \quad (12)$$

$$a_E = \frac{\omega}{cn\epsilon_0}\sum_{\boldsymbol{k}} (\epsilon^e_{\boldsymbol{k}+\boldsymbol{\kappa}_\parallel} + \epsilon^h_{\boldsymbol{k}})\text{Im}\{d^*_{\boldsymbol{\kappa}_\parallel}\chi_{\boldsymbol{k}}(\boldsymbol{\kappa}_\parallel)\}, \quad (13)$$

where $\epsilon^c_{\boldsymbol{k}} = \hbar^2\boldsymbol{k}^2/2m_c$ is the kinetic energy of a carrier with effective mass m_c. If Coulomb many-body effects (refer to Sec. V for many-body results) are neglected, the free-carrier result (denoted by the superscript 0 throughout this work) is given as

$$\chi_k^0(\boldsymbol{\kappa}_\parallel) = \frac{(1 - f^e_{k+\kappa_\parallel} - f^h_{-k})d_{\kappa_\parallel}}{\epsilon^e_{k+\kappa_\parallel} + \epsilon^h_{-k} - \delta - i\gamma_p}, \tag{14}$$

where the detuning of the laser beam $\delta = \hbar\omega - E_g^{QW}(T_p)$ has been used, with $E_g^{QW}(T_p) = E_g(T_p) + \epsilon_z^e + \epsilon_z^h$ being the temperature-dependent QW band gap, $E_g(T_p)$ being the temperature-dependent bulk band gap of the well semiconductor, and ϵ_z^c being the carrier $(c=e,h)$ confinement energy perpendicular to the QW.

We now turn to the direct radiative recombination process, or photoluminescence. To determine the radiative recombination rates, a three-point correlation function is needed that is described by the semiconductor luminescence equation[18]

$$\begin{aligned}(\varepsilon^e_{k+q_\parallel} + \varepsilon^h_{-k} - \hbar\omega_q - i\gamma_p)\langle b_q^\dagger \hat{p}_k(q_\parallel)\rangle - i\hbar\partial_t\langle b_q^\dagger \hat{p}_k(q_\parallel)\rangle \\ = (1 - f^e_{k+q_\parallel} - f^h_{-k})\sum_{k'} V^s_{k-k'}\langle b_q^\dagger \hat{p}_{k'}(q_\parallel)\rangle \\ - i\mathcal{E}_q u_{QW,q} d_{q_\parallel} f^e_{k+q_\parallel} f^h_{-k},\end{aligned} \tag{15}$$

where ∂_t denotes the time derivative. We have assumed the same dephasing rate for the radiative recombination (photoluminescence) process as for the optical absorption (photogeneration) process. In the same vein as with the absorbance case, we ignore the slow change in carrier distributions and seek an adiabatic solution to the above equation for the three-point correlation function. Because of the nonlocal nature of the Coulomb interaction, only a formal solution exists, which can be found by defining $\xi_k(\boldsymbol{q})$, according to $\langle b_q^\dagger \hat{p}_k(q_\parallel)\rangle \equiv \xi_k(\boldsymbol{q})\mathcal{E}_q$, termed $\boldsymbol{k}$-resolved luminescence susceptibility. Given this quantity, an explicit form of the recombination coefficients in Eqs. (7) and (10) is readily obtained as

$$B = \frac{2}{\hbar N^2}\sum_{k,q} \mathrm{Re}\{\mathcal{E}_q^2 u^*_{QW,q} d^*_{q_\parallel}\xi_k(\boldsymbol{q})\}, \tag{16}$$

$$B_E = \frac{2}{\hbar N^2}\sum_{k,q} (\epsilon^e_{k+q_\parallel} + \epsilon^h_k)\mathrm{Re}\{\mathcal{E}_q^2 u^*_{QW,q} d^*_{q_\parallel}\xi_k(\boldsymbol{q})\}. \tag{17}$$

Similar to the absorbance case, if Coulomb many-body effects are neglected, the free-carrier $\boldsymbol{k}$-resolved luminescence susceptibility is readily available from Eq. (15):

$$\xi_k^0(\boldsymbol{q}) = \frac{- i u_{QW,q} d_{q_\parallel} f^e_{k+q_\parallel} f^h_{-k}}{\epsilon^e_{k+q_\parallel} + \epsilon^h_{-k} - \delta_q - i\gamma_p}, \tag{18}$$

where we have used $\delta_q = \hbar\omega_q - E_g^{QW}(T_p)$. Then, the free-carrier recombination coefficient is

$$B = \frac{1}{\epsilon_0\epsilon_b N^2}\sum_{k,q} \mathrm{Im}\left\{\frac{\omega_q |u_{QW,q}|^2 |d_{q_\parallel}|^2 f^e_{k+q_\parallel} f^h_{-k}}{\epsilon^e_{k+q_\parallel} + \epsilon^h_{-k} - \delta_q - i\gamma_p}\right\}, \tag{19}$$

and the counterpart coefficient for the carrier thermal energy, B_E, can be found correspondingly with a weighted summation by the carrier kinetic energies.

In general, the time-resolved PL spectrum with photon energy $\hbar\Omega$ can be calculated from the $\boldsymbol{k}$-resolved luminescence susceptibility according to

$$I_{PL}(\Omega) = 2\sum_{k,q} \delta(\hbar\Omega - \hbar\omega_q)\mathrm{Re}\{\mathcal{E}_q^2 u^*_{QW,q} d^*_{q_\parallel}\xi_k(\boldsymbol{q})\}, \tag{20}$$

where $\delta(x)$ is the Dirac delta function. Note that an integration over the photon energy of the time-resolved photoluminescence spectrum equals BN^2, as each e-h pair recombines to emit one PL photon. Explicitly, we have $BN^2 = \int d\Omega I_{PL}(\Omega)$. The free-carrier PL result can be found by plugging Eq. (18) into the above equation. Furthermore, the angle-resolved PL spectrum can be obtained using Eq. (20) by integrating over the radial coordinate of $\boldsymbol{q}$ only, whereas the time-integrated PL spectrum integrating over time, of course.

V. KUBO-MARTIN-SCHWINGER RELATION

Incoherent radiative recombination—i.e., thermal photoluminescence—as treated above, can also be obtained via the Kubo-Martin-Schwinger relation[21,22] from the optical absorption spectrum. On the other hand, the relation exhibits one of the connections between stimulated and spontaneous optical processes, which are required for thermodynamic consistency. We now prove that our theory agrees with the famous relation by following Refs. 17 and 18. It is insightful to note that Eq. (11) is analogous to Eq. (15) after invoking the *rotating-wave* ansatz that ignores the slowly varying envelope part. They take the form of a hydrogen-atom-like problem in Fourier space:

$$\sum_{k'} \mathcal{M}_{kk'}(q_\parallel)\mathcal{A}_{k'}(\boldsymbol{q}) - [E(\boldsymbol{q}) + i\gamma_p]\mathcal{A}_k(\boldsymbol{q}) = \mathcal{S}_k(\boldsymbol{q}), \tag{21}$$

where the nondiagonal matrix elements are given by

$$\mathcal{M}_{kk'}(q_\parallel) \equiv (\varepsilon^e_{k+q_\parallel} + \varepsilon^h_{-k})\delta_{k,k'} + (f^e_{k+q_\parallel} + f^h_{-k} - 1)V^s_{k-k'}, \tag{22}$$

with $\delta_{k,k'}$ being the Levi-Cività delta notation. For photogeneration, $\mathcal{A}_k(\boldsymbol{q})$, $E(\boldsymbol{q})$, and $\mathcal{S}_k(\boldsymbol{q})$ correspond to $\tilde{p}_k(q_\parallel)$, $\hbar\omega$, and $(1 - f^e_{k+q_\parallel} - f^h_{-k})d_{q_\parallel}\mathcal{E}_0$. For photoluminescence, they correspond to $\langle b_q^\dagger \hat{p}_k(q_\parallel)\rangle$, $\hbar\omega_q$, and $-i\mathcal{E}_q u_{QW,q} d_{q\parallel} f^e_{k+q_\parallel} f^h_{-k}$. The standard Green's function method applies to Eq. (21), which produces the following formal solution:

$$\mathcal{A}_k(\boldsymbol{q}) = \sum_\nu \psi_{\nu k}\frac{\sum_{k'} \psi^*_{\nu k'}\mathcal{S}_{k'}(\boldsymbol{q})}{E_\nu - [E(\boldsymbol{q}) + i\gamma_p]}, \tag{23}$$

where E_ν and $\psi_{\nu k}$ satisfy $\Sigma_{k'}\mathcal{M}_{kk'}(q_\parallel)\psi_{\nu k'} = E_\nu\psi_{\nu k}$. Note that the eigenvalue set $\{E_\nu\}$ includes discrete (excitonic) and continuous (continuum) entries. Plugging in the corresponding expressions, we have

$$\chi_k(\boldsymbol{\kappa}_\parallel) = \sum_{\nu k'} \psi_{\nu k}\frac{\psi^*_{\nu k'}(1 - f^e_{k'+\kappa_\parallel} - f^h_{-k'})d_{\kappa_\parallel}}{E_\nu - (\hbar\omega + i\gamma_p)}, \tag{24}$$

for the $\boldsymbol{k}$-resolved susceptibility, and

$$\xi_{k}(\mathbf{q}) = -i\sum_{\nu \mathbf{k}'} \psi_{\nu k} \frac{\psi^*_{\nu k'} u_{QW,q} d_{q_\parallel} f^e_{k'+q_\parallel} f^h_{-k'}}{E_\nu - (\hbar\omega_q + i\gamma_p)}, \quad (25)$$

for the $\mathbf{k}$-resolved luminescence susceptibility. By taking the limit $\gamma_p \to 0$ such that $\mathrm{Im}\{1/[E_\nu-(\hbar\omega+i\gamma_p)]\}=\pi\delta(E_\nu-\hbar\omega)$ and ignoring the weak $\mathbf{k}$ dependence in $d_{\mathbf{k}}$ (denoted as d_{cv} hereafter), the QW absorbance is found to be

$$a = \frac{\pi\omega}{cn\epsilon_0}|d_{cv}|^2 \psi_{\bar\nu}(0) \sum_{k'} \psi^*_{\bar\nu k'}(1 - f^e_{k'+\kappa_\parallel} - f^h_{-k'}), \quad (26)$$

where $\bar\nu$ denotes the subset of eigenstates that satisfy $E_{\bar\nu} = \hbar\omega$ and $\psi_{\bar\nu}(\mathbf{0})=\Sigma_{\mathbf{k}}\psi_{\bar\nu \mathbf{k}}$ is the Fourier transform at position $\mathbf{r}=\mathbf{0}$, the origin of the QW. In the same fashion, the PL spectrum is derived as follows:

$$I_{PL}(\Omega) = \frac{\pi\Omega}{\epsilon_0\epsilon_b} D(\Omega)|\bar u_{QW,Q} d_{cv}|^2 \psi_{\bar\tau}(0) \sum_{k'} \psi^*_{\bar\tau k'} f^e_{k'+Q_\parallel} f^h_{-k'}, \quad (27)$$

where $\Omega=c|\mathbf{Q}|/n$, $D(\Omega)=\Sigma_{\mathbf{q}}\delta(\Omega-\omega_{\mathbf{q}})=n^3\Omega^2/2\pi^2c^3$ is the photon mode density, $\bar u_{QW,\mathbf{Q}}$ is the averaged overlap integral $u_{QW,\mathbf{Q}}$ over a 4π solid angle, and $\bar\tau$ denotes the subset of eigenstates that satisfy $E_{\bar\tau}=\hbar\Omega$. Given that the carriers follow the Fermi-Dirac distribution, we have

$$f^e_{k+Q_\parallel} f^h_{-k} = g(\varepsilon_k(\mathbf{Q}))(1 - f^e_{k+Q_\parallel} - f^h_{-k}), \quad (28)$$

$$g(\varepsilon_k(\mathbf{Q})) = \frac{1}{\exp\left(\frac{\varepsilon^e_{k+Q_\parallel} + \varepsilon^h_{-k} - (\mu_e + \mu_h)}{k_B T_p}\right) + 1}, \quad (29)$$

where function $g(\varepsilon_{\mathbf{k}}(\mathbf{Q}))$ denotes a Bose-Einstein distribution with an aggregate carrier energy $\varepsilon_{\mathbf{k}}(\mathbf{Q})=\varepsilon^e_{\mathbf{k}+\mathbf{Q}_\parallel}+\varepsilon^h_{-\mathbf{k}}$ and an aggregate chemical potential for electrons and holes. Finally, by realizing that the resonant condition[23] dictates $\varepsilon_{\mathbf{k}}(\mathbf{Q}) = \hbar\omega$, we arrive at the desired Kubo-Martin-Schwinger relation

$$I_{PL}(\omega) = \frac{c}{n} D(\omega)|\bar u_{QW,Q}|^2 g(\hbar\omega) a, \quad (30)$$

where $\mathbf{Q}$ satisfies dispersion relation $\hbar\omega=c|\mathbf{Q}|/n$.

As a side note, it is easy to check that the generalized 2D Elliot formula[17] is recovered with phase-space filling effect from Eq. (24) when the weak $\mathbf{k}$ dependence in $d_{\mathbf{k}}$ is ignored. To recover from these many-body expressions the free-carrier results presented in the preceding section, it suffices to use two observations: (i) the free-carrier eigenvalues E_ν are all degenerate, and (ii) $\Sigma_\nu \psi_{\nu \mathbf{k}}\psi^*_{\nu \mathbf{k}'}=\delta_{\mathbf{k},\mathbf{k}'}$ is the condition of completeness. In addition, it is well known that an Einstein relation holds between the stimulated and spontaneous emissions for two-level systems, which applies to the present case as an extended example. We do not intend to verify that in this article, but we mention that using the same approach in this section, a formal verification can be achieved.

TABLE I. GaAs parameters used: m_0, free-electron mass; ϵ's, dielectric constants; a_L, lattice constant; P, Kane's parameter; γ_p, dephasing rate; g, mass density.

m_e	ϵ_0	$E_g(0)$	P	γ_p
$0.067m_0$	13.1	1519 meV	25 eV	5 meV
m_h	ϵ_∞	a_L	$\hbar\omega_{LO}$	g
$0.45m_0$	10.9	5.6533 Å	35 meV	5317.6 kg/m^3

VI. NUMERICAL RESULTS

In this section, we demonstrate the utility of the developed theoretical framework by studying laser cooling thermokinetics of an exemplary quantum well system numerically. Most importantly, we put forth an adaptive cooling strategy for deep optical refrigeration of semiconductor QW systems. To this end, we select a 15-nm GaAs QW at ambient temperature T_{amb}=295 K. At such an elevated temperature the Coulomb many-body effects still play a significant role, as discussed in Sec. VII together with aspects of their importance at low temperature (Ref. 24 appeared as this contribution is under review). Nevertheless, we ignore them in our simulations for the sake of numerical efficiency. Table I lists the material parameters used. The bulk GaAs band gap[25] follows $E_g(T_p)=E_g(0)-0.5405T_p^2/(T_p+204)$, and the quantum confinement energies for electrons and holes are $\epsilon^e_z = 16.8$ and $\epsilon^h_z=2.5$ meV, respectively. An adaptive Runge-Kutta algorithm is used to integrate Eqs. (7) and (10). Photogeneration terms (i.e., a and a_E) are calculated on the fly from the semiconductor Bloch equations (SBEs) with band structure determined by $\mathbf{k}\cdot\mathbf{p}$ method.[26] Recombination terms use fitted expressions of N and T_p to calculated emission rates from the semiconductor luminescence equations (SLEs) (see Fig. 2). As expected, $B \propto T_p^{-1}$ (see, for example, Ref. 27) and B_E is T_p independent, nearly; the average energy extracted per photon (B_E/B) is k_BT_p for nondegenerate electrons. The initial condition for integration is an unexcited intrinsic QW. The cw photon flux is ramped up according to $F(t)=F_l\{1-\exp[-(t/\tau)^2/2]\}$ with a judicious time constant

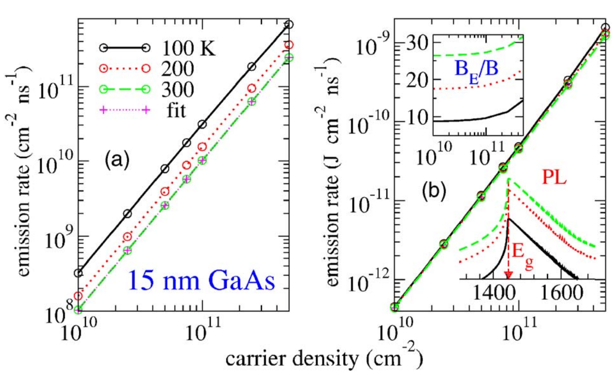

FIG. 2. (Color online) Microscopic radiative recombination rates for carrier density (a) and thermal energy (b) used in thermokinetic study of a 15-nm GaAs QW. Upper (b) inset shows their ratio in meV. Lower inset shows semilog arithmic PL spectra for 1, 5, 10×10^{10} cm^{-2} at 295 K with energy in meV.

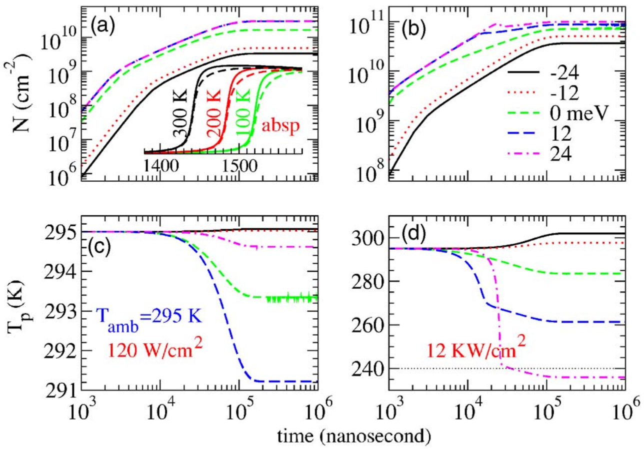

FIG. 3. (Color online) Laser cooling thermokinetics of the GaAs QW at ambient temperature T_{amb}=295K for indicated *nominal* detunings at low (a),(c) and high (b),(d) power. (a) Inset: absorbances at 100, 200, 300 K for 1, 10×10^{10} cm^{-2}. High power cools better and its behavior reveals an optimal *actual* detuning for laser cooling. See text for simulation details.

τ=50 ns. The laser power is given by $P_l = h\nu F_l$.

Figure 3 presents pumping thermokinetics at a series of *nominal* detunings $\tilde{\delta} = h\nu - [E_g(T_{amb}) + \epsilon_z^e + \epsilon_z^h]$ for low-power [120 W/cm^2, (a) and (c)] and high-power [12 KW/cm^2, (b) and (d)] cases. Note that (i) the change in band gap during cooling effectively alters the *actual* detuning $\delta = h\nu - [E_g(T_p) + \epsilon_z^e + \epsilon_z^h]$ (thus *nominal*) and (ii) it takes about $3\tau = 150$ ns to reach peak laser pumping value. Clearly, our model supports a conditional cooled steady state (we delay discussion about heating). The low-power case features two phases: (i) a generation-dominated ramp-up phase and (ii) a recombination-compensating cooling phase. Carrier generation collapses onto a single curve at positive detuning owing to the constant QW density of states. Optical nonlinearity from the phase space filling effect is negligible in the studied pumping range. Maximum cooling is achieved around $\tilde{\delta} = 12$ meV. (The optimal detuning is actually at 3 meV, as shown later.) At high power, the cooling thermokinetics depends much on detuning. Near the band gap, the behavior is similar to the low-power case, except that faster, deeper cooling is observed, obviously due to power increase. Then occurs a transition around $\tilde{\delta} = 12$ meV. As can be seen, (i) the detuning for maximum cooling is blueshifted and (ii) cooling accelerates before the transition. The blueshift is due to the band gap increase as cooling continues. As for the transition, when correlating data in (b) and (d), we see that it starts when the carrier density approaches saturation. More revealingly, the cooling behavior at positive detuning strikingly resembles the near-gap case ($\tilde{\delta} = 0$ meV curve) after the transition, which indicates that the *actual* detuning after the transiton tends to lock into a near-gap value. Indeed, take the $\tilde{\delta} = 24$ meV case for example: When cooled to 240 K where the transition kicks in, $E_g(T_p) = 1451.37$ meV; the gap has increased by 24.11 meV from 1427.26 meV at 295 K. Correspondingly, the carrier density ceases to grow. Therefore, the initial phase relates to a positive *actual* detuning, which settles into a near-gap value after transition. In essence, *cooling is the most effective right before the transition.*

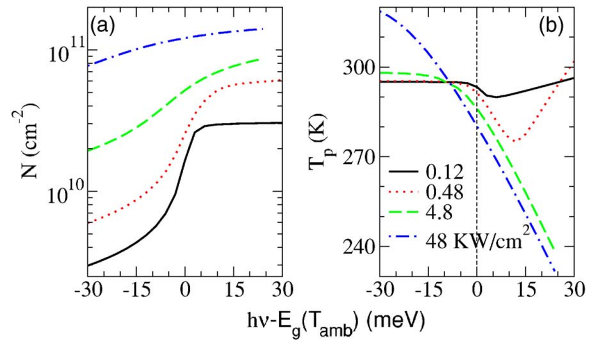

FIG. 4. (Color online) Thermokinetic steady-state solutions vs laser power and *nominal* detuning. Laser cooling is observed from a *critical* detuning value of 9 meV below to $k_B T_{amb}$ above the QW band gap (x=0 dashed line). Higher power cools deeper and blueshifts the optimal cooling detuning.

As the cooling process is driven by thermal energy extraction via luminescence upconversion, the amount of thermal energy taken away is usually determined by $BN^2(k_B T_p - \delta)$, where k_B is the Boltzmann constant, without Coulomb many-body effects. Whereas the initial cooling is dominated by an increase in carrier density, the second phase appears thanks to the slower growth in N (near balance) and in the magnitude of δ due to the further blueshift of the band gap. In addition, the cooling acceleration in the first phase is an indication that *there exists an optimal actual detuning for cooling*. Furthermore, the crossovers for different detuning cases in (d) result from an interplay of N and δ: Pumping above the gap increases photogeneration (BN^2), but reduces the average net energy extracted per photon ($k_B T_p - \delta$) vice versa. However, the increase in carrier density pays off eventually (thus acceleration) owing to the quadratic dependence. We turn now to discuss heating at large negative detuning. This counterintuitive observation deserves a close reexamination of the photogeneration process in our model, since the recombination process leaves no ambiguity. We argue that the *actual* detuning is *not* a proper measure for the average carrier energy in the present case. Pumping below the edge creates carriers via absorption in the Lorentzian tail. In contrast with a resonant pumping case where carriers are created with energies of a spread characterized by the linewidth (γ_p=1 meV here), far-off-resonant cases with a Lorentzian line shape, as considered now, feature a divergent spread. When the detuning is too far off the band gap, this energy spread will surpass the thermal energy scale $k_B T_p$. As a result, there exists a *critical* negative detuning, which is apparently independent of pumping level, where the average energy of the photogenerated carriers equals that of the recombined carriers. Passing it, instead of generating cold carriers that cool the QW, the laser beam heats up the QW.

Next, we show steady-state solutions to Eqs. (7) and (10) as a function of *nominal* detuning and laser power in Fig. 4. First of all, laser cooling of the GaAs QW is observed between a *critical* negative detuning and $k_B T_{amb}$. Second, the higher the power, the lower the lattice temperature. Third, the *critical* negative detuning is around −9 meV and indepen-

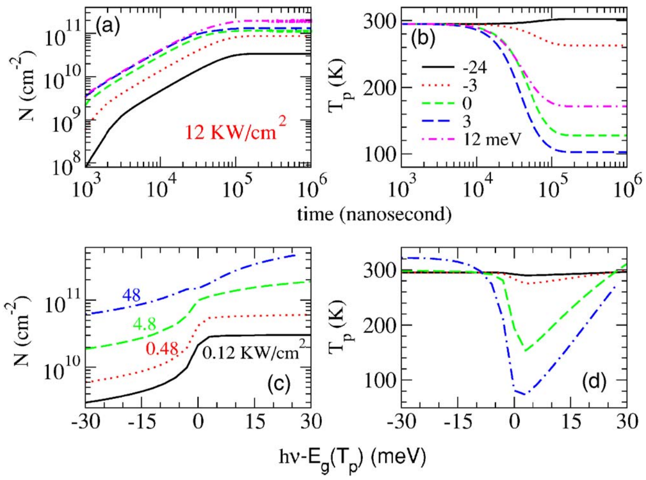

FIG. 5. (Color online) Laser cooling thermokinetics (a),(b) and efficacy (c),(d) vs laser power and *actual* detuning. (b) Far deeper cooling is achieved at 12 kW cm^{-2} with fixed *actual* detuning than fixed *nominal* detuning [Fig. 3(d)]. (d) Optimal *actual* detuning is 3 meV above QW band gap.

dent of pumping level, indeed. At low power, the carrier density approximately follows the absorbance spectrum and the optimal detuning for cooling is near the band edge. As power increases, the steady state supports a larger carrier density and deeper cooling. Moreover, the optimal *nominal* detuning for cooling moves toward $k_B T_{amb}$, as the QW band gap moves up with further cooling. On the other hand, heating below the *critical* detuning increases with laser power as well.

As shown, an increase in laser power leads to deeper refrigeration with the optimal *nominal* detuning for cooling shifting with the band edge. However, the obvious drawback is the high demand on laser pumping power. A weaker power level would help mitigate issues that are not addressed in this work, such as carrier trapping and photon recycling, which will become worse as the pumping level (carrier density) increases. A strategy to reduce the pumping level arises from the following insight: *There is an optimal actual detuning for cooling*, as mentioned while discussing the behavior transition in the high-power case in Fig. 4. Direct evidence is shown in Fig. 5(d) where the laser frequency is tuned to fix the *actual* detuning during cooling. As can be seen, the optimal detuning lies 3 meV above the band edge and does not depend on laser power. Whereas there seemed to be a lower bound in Fig. 4(b), sub-100-K cooling is now observed. Furthermore, we show in (a) and (b) the thermokinetics for fixed *actual* detuning at 12 kW cm^{-2}. The distinction is clear: (i) By trailing the band edge, carrier generation and cooling are the most efficient; (ii) the second phase vanishes; and (iii) the coolest temperature jumps from 235 down to 100 K. The strategy works wonders: It creates much more carriers and reduces the power level by orders of magnitude to cool the QW below 225 K.

VII. TECHNICAL REMARKS

Some comments on the technical limitations and issues with the current theoretical framework are in order. First of all, our theory only considers fundamental linear optical process, and optical nonlinearity arises as a consequence of the pumping-induced phase-space filling effect. Therefore, it is limited to a moderate laser power level. Second, we have used the dephasing rate approximation in the description of the optical processes. The resulting Lorentzian line shape has been known to largely overestimate the absorption below the band edge, which is supposed to follow the well-known "Urbach tail" behavior at low temperature. It is possible to overcome this shortcoming within the framework of our theory by considering the actual dephasing processes, which requires a systematic treatment of the Coulomb correlation effects within the second or self-consistent Born approximation and is mathematically involving and computationally costly. In view of the present transparent and insightful formulation, the approximation is a "necessary sin." Third, inclusion of the other phonon branches, as well as their corresponding interactions with carriers in the theory, is straightforward and amounts to increasing the QW heat capacity at elevated temperature when the LO phonons play a dominant role, which would accordingly reduce the cooling speed and adversely increase the computing time. Nevertheless, such a treatment does not affect the steady state solution and much of the thermokinetic characteristics. Fourth, our theory serves as a local description for the laser cooling of semiconductor QWs. Thermal transport and carrier diffusion shall modify the thermokinetics to a limited degree at the quantitative level. Last, we note that cooling into the cryogenic regime (below 77 K) requires careful treatment of Coulomb many-body effects, as they largely alter the characteristics of photogeneration and radiative recombination (see, for example, Refs. 17–19, 22, and 24). Inclusion of Coulomb many-body effects normally renders Eqs. (11) and (15) intractable analytically. However, both equations can be numerically solved by the matrix inversion method. Nevertheless, note that Ref. 22 provides some approximate analytical results on many-body effects for optical absorption and photoluminescence, as well as Refs. 17 and 18. Besides Coulomb many-body effects, acoustic phonons play a more important role, as the number of optical phonons dwindles. In other words, the optical refrigeration physics at cryogenic temperature becomes fundamentally different from that of ambient cooling. The theoretical framework presented in this article serves as a good starting point for such studies.

We now take the liberty to reckon with possible quantitative modifications by Coulomb many-body effects, at the Hartree-Fock level, to the simulated free-carrier results presented in the preceding section before the conclusion of this article. In contrast to their fundamental qualitative impact at low temperature, such as the existence of discrete excitonic states, Coulomb many-body effects are known for two major reasons at elevated temperature: (i) band gap renormalization (BGR), which shifts the spectroscopic response of the many-body system, and (ii) Coulomb enhancement (excitonic effect), which alters spectroscopic features around the van Hove singularities, such as the band edge. The former increases roughly proportional to the carrier density while the latter is weakened. Depicted in Fig. 6 is a telltale demonstration of the effects in the 15-nm GaAs QW. Clearly, the effects are drastic, even at room temperature, and lead to a

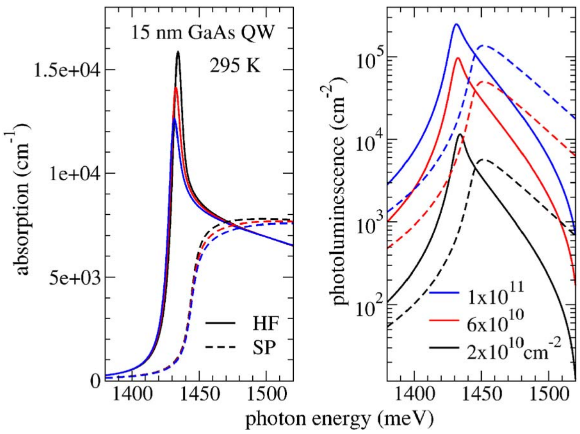

FIG. 6. (Color online) Coulomb many-body effects in a 15-nm GaAs QW at room temperature (295 K). Coulomb enhancement and band-gap renormalization lead to spectral shifts and excitonlike features in optical absorption (left panel) and photoluminescence (right panel) within the Hartree-Fock (HF) treatment at a constant dephasing rate of 5 meV. Free-carrier results are shown in comparison with dashed lines (denoted as SP). Also visible is the phase-space filling effect in the absorption spectra as carrier density increases.

large condensation (spectral narrowing) of the oscillator strength onto the excitonlike peak which is redshifted from the free-carrier band edge. Within the relevant density range, the excitonic effect dominates over that of BGR, as evidenced from the insensitive dependence of the peaks on density. Because of the spectral narrowing and maximal feature, optical pumping becomes more efficient and faster around the band edge, which will lead to a higher steady-state carrier density as well as a lower steady-state lattice temperature. Therefore, one expects that the excitonlike absorption peak pins down the optimal *actual* detuning near the peak. On the other hand, the detuning range for cooling is affected insignificantly for the lower bound is determined by the Lorentzian tail and the upper bound by the LO phonon energy. As such, Coulomb many-body effects are expected to (i) draw the optimal *actual* detuning to the band-edge absorption peak and (ii) redshift the spectral response of the QW relative to the free-carrier case. As for the adaptive cooling strategy, it is expected to work as before, subject to a minor spectral shift and slight change in the optimal *actual* detuning thanks to the aforediscussed Coulomb many-body effects, and to follow the much more significant band-gap shrinkage resulting from lattice cooling. Finally, we note that Coulomb many-body effects become more pronounced for more polar semiconductors, such as II-VI compounds, and narrower QWs or nanostructured materials at a lower dimension. Hence, it is imperative to take into account the effects and even exploit them for more efficient cooling schemes and engineering.

VIII. SUMMARY

In summary, we have presented a first-principles-based thermokinetic theory for laser cooling of semiconductor quantum wells. The theory reduces the Boltzmann kinetic equations under *the adiabatic approximation* to a set of thermokinetic equations for carrier density and lattice temperature. Photogeneration and radiative recombination of carriers that are in quasiequilibrium with thermal LO phonons are microscopically taken into account. In particular, semiconductor Bloch equations and semiconductor luminescence equations are used to derive the generation and recombination rates. We verify that the Kubo-Martin-Schwinger relations are valid within the current theory. On the numerical end, we show that stronger laser pumping leads to deeper optical refrigeration within the detuning range demarcated by a *critial* value and the ambient thermal energy. Of particular interest, we prove for the free-carrier case the existence of a pumping-independent optimal *actual* detuning for cooling above the band gap (Coulomb many-body effects are expected to pin the optimal detuning at the excitonlike absorption peak). This revelation leads to a cooling strategy that reduces the laser power drastically. This work provides a consistent theoretical framework to model laser cooling operations, may be invaluable to achieving laser cooling of semiconductors as a thermal management solution, and serves to inspire further work in this emerging technologically important research area.

ACKNOWLEDGMENTS

The author acknowledges helpful discussions with Yong-Hang Zhang (Arizona State University) who suggested the adaptive cooling scheme, Danhong Huang (Air Force Research Lab) for introduction to optical refrigeration of semiconductors, and Cun-Zheng Ning (NASA Ames Research Center) for proofreading part of the manuscript.

*Electronic address: jianzhng@netzero.com

[1] R. I. Epstein, M. I. Buchwald, B. C. Edwards, T. R. Gosnell, and C. E. Mungan, Nature (London) **377**, 500 (1995).
[2] J. L. Clark and G. Rumbles, Phys. Rev. Lett. **76**, 2037 (1996).
[3] C. W. Hoyt, M. Sheik-Bahae, R. I. Epstein, B. C. Edwards, and J. E. Anderson, Phys. Rev. Lett. **85**, 3600 (2000).
[4] P. Pringsheim, Z. Phys. **57**, 739 (1929).
[5] L. Landau, J. Phys. (Moscow) **10**, 503 (1946).
[6] H. Gauck, T. H. Gfoerer, M. J. Renn, E. A. Cornell, and K. A. Bertness, Appl. Phys. A: Mater. Sci. Process. **64**, 143 (1997).
[7] E. Finkeissen, M. Potemski, P. Wyder, L. Viña, and G. Weimann, Appl. Phys. Lett. **75**, 1258 (1999).
[8] S. R. Bowman, IEEE J. Quantum Electron. **35**, 115 (1999).
[9] C. E. Mungan, J. Opt. Soc. Am. B **20**, 1075 (2003).
[10] M. Sheik-Bahae and R. I. Epstein, Phys. Rev. Lett. **92**, 247403 (2004).
[11] D. Huang, T. Apostolova, P. M. Alsing, and D. A. Cardimona, Phys. Rev. B **70**, 033203 (2004).

[12] D. Huang, T. Apostolova, P. M. Alsing, and D. A. Cardimona, Phys. Rev. B **72**, 195308 (2005).
[13] J.-B. Wang, S. R. Johnson, D. Ding, S.-Q. Yu, and Y.-H. Zhang, J. Appl. Phys. **100**, 043502 (2006).
[14] H. Haug and S. W. Koch, Phys. Rev. A **39**, 1887 (1989).
[15] J. Li and C. Z. Ning, Phys. Rev. A **66**, 023802 (2002).
[16] M. Bonitz, *Quantum Kinetic Theory* (Teubner-Verlag, Stuttgart, 1998).
[17] H. Haug and S. W. Koch, *Quantum Theory of the Electrical and Optical Properties of Semiconductors* (World Scientific, Singapore, 1994).
[18] M. Kira, F. Jahnke, W. Hoyer, and S. W. Koch, Prog. Quantum Electron. **23**, 189 (1999).
[19] T. Kuhn, in *Theory of Transport Properties of Semiconductor Nanostructures*, edited by E. Schöll (Chapman & Hall, London, 1998).
[20] J. Li and C. Z. Ning, Phys. Rev. A **66**, 023803 (2002b).
[21] H. Haug and S. Schmitt-Rink, Prog. Quantum Electron. **9**, 3 (1984).
[22] S. Schmitt-Rink, C. Ell, and H. Haug, Phys. Rev. B **33**, 1183 (1986).
[23] Note that Coulomb many-body effects are treated at the Hartree-Fock level. As only the linear polarization induced by the pump laser and the subsequent e-h pair generation are considered in the kinetic description, no many-body effects are directly taken into account in the distribution functions, except inherently those originated from dielectric response. Therefore, it is consistent within the Hartree-Fock approximation and linear response theory to replace the aggregate carrier energy with the photon energy.
[24] G. Rupper, N. H. Kwong, and R. Binder, Phys. Rev. Lett. **97**, 117401 (2006).
[25] Landolt-Börnstein, *Numerical Data and Functional Relationships in Science and Technology* (Springer-Verlag, Berlin, 1982).
[26] S. L. Chuang, *Physics of Optoelectronic Devices* (Wiley, New York, 1995).
[27] W. W. Chow, S. W. Koch, and M. Sargent, III, *Semiconductor Lasers Physcis* (Springer-Verlag, Berlin, 1994).