Kinetic Monte Carlo approach to nonequilibrium bosonic systems

T. C. H. Liew,[1] H. Flayac,[2] D. Poletti,[3] I. G. Savenko,[4,5,6] and F. P. Laussy[7,8]
[1]*Division of Physics and Applied Physics, School of Physical and Mathematical Sciences, Nanyang Technological University, 21 Nanyang Link, Singapore 637371*
[2]*Institute of Theoretical Physics, Ecole Polytechnique Fédérale de Lausanne, CH-1015 Lausanne, Switzerland*
[3]*Singapore University of Technology and Design, 8 Somapah Road, 487372 Singapore*
[4]*Center for Theoretical Physics of Complex Systems, Institute for Basic Science, Daejeon 34051, Republic of Korea*
[5]*Nonlinear Physics Centre, Research School of Physics and Engineering, The Australian National University, Canberra ACT 2601, Australia*
[6]*ITMO University, Saint Petersburg 197101, Russia*
[7]*Russian Quantum Center, Novaya 100, 143025 Skolkovo, Moscow Region, Russia*
[8]*Faculty of Science and Engineering, University of Wolverhampton, Wulfruna St, Wolverhampton WV1 1LY, United Kingdom*
(Received 8 May 2017; revised manuscript received 1 September 2017; published 18 September 2017)

We consider the use of a kinetic Monte Carlo approach for the description of nonequilibrium bosonic systems, taking nonresonantly excited exciton-polariton condensates and bosonic cascade lasers as examples. In the former case, the considered approach allows the study of the cross-over between incoherent and coherent regimes, which represents the formation of a quasicondensate that forms purely from the action of energy relaxation processes rather than interactions between the condensing particles themselves. In the latter case, we show theoretically that a bosonic cascade can develop an output coherent state.

DOI: 10.1103/PhysRevB.96.125423

I. INTRODUCTION

Although Bose-Einstein condensation (BEC) was originally defined as an effect taking place in thermal equilibrium, it is striking to see that the concept has been generalized to nonequilibrium systems. For example, the BEC of photons in a cavity [1] has been reported and several groups have studied the BEC of exciton-polaritons (hybrid light-matter quasiparticles) appearing in semiconductor microcavities [2–4]. Here BEC is characterized [5] by the spontaneous formation of coherence, typically measured by the transition of the second-order coherence function with increasing particle density, as reported by several experimental groups [6–12].

The physics of nonequilibrium condensates has been shown to be radically different from that of equilibrium systems, where condensates may form in nonground states [13] or multiple states [14] and have distributions undescribable by a single temperature.

The theoretical description of nonequilibrium condensates typically requires an explicit treatment of energy relaxation processes. Such processes compete with dissipative processes, which cause particles to be lost from the system before reaching the ground state. The dynamical interplay of relaxation and dissipation ultimately determines the steady state of the system. Energy relaxation mechanisms have been handled previously in exciton-polariton systems using semiclassical Boltzmann equations [15–18] or introduced phenomenologically into mean-field equations [19–22]. Methods treating energy relaxation from first principles have also been developed based on stochastic sampling of mean-field equations [23] or their hybridization with the Boltzmann equations [24]. However, these methods do not account for quantum fluctuations, which are needed for the unified treatment of nonequilibrium condensation below and above threshold. In principle, density-matrix approaches [25] (possibly supplemented with Monte Carlo techniques [26]) are applicable to this task; however, in practice they are only feasible for systems with small numbers of particles and modes [27]. Bosonic cascade lasers [28,29], which may operate with millions of particles, have been treated with stochastic sampling of the positive-P distribution [30,31]; however, such a method is only accurate in the case of an initial coherent state.

In the present paper we employ a kinetic Monte Carlo approach based on quantum Boltzmann equations for the description of nonequilibrium multimode open quantum systems. The kinetic Monte Carlo approach has been developed under different names in different fields, including vacancy migration in binary ordered alloys [32], the Ising model [33], and chemical reactions [34]. A good overview of the method can be found in Ref. [35]. The approach allows stochastic sampling of the quantum particle distribution function and allows the treatment of systems with up to hundreds of modes with possibly thousands of particles each. From the particle distribution functions we have full access to the coherence statistics, as characterized by the second-order correlation function. We apply the technique to two specific examples: polariton condensation in one-dimensional microwires and terahertz lasing in bosonic cascade lasers.

In the former case we are able to describe the gradual cross-over from incoherent population of excited states to partial coherence in nonground states and the formation of a fully coherent BEC with increasing particle density. The responsible energy relaxation processes are described from first principles, accounting for polariton-phonon scattering and the scattering of polaritons with hot exciton states [15]. Aside from these interaction processes, it is notable that additional interactions between the condensing particles themselves are not required for the formation of a condensate (which is consistent with the original equilibrium theory of BEC of the ideal gas).

In the case of bosonic cascade lasers, we access theoretically the coherence of the lasing mode and show that it can be useful for terahertz lasing with high quantum efficiency.

II. GENERIC KINETIC MONTE CARLO APPROACH

We start with the consideration of a set of M discrete modes with populations $n_1, n_2, \ldots, n_M$. The probability of the system being in any particular state at time t is $P_{n_1,n_2,\ldots,n_M}(t)$. The probability distribution contains sufficient information to calculate the quantum expectation values of a variety of quantities, in particular those with operators that commute with the number operator. For example, one can calculate

$$\langle n_i(t)\rangle = \sum P_{n_1,n_2,\ldots,n_i,\ldots,n_M}(t) n_i, \tag{1}$$

$$\langle n_i^2(t)\rangle = \sum P_{n_1,n_2,\ldots,n_i,\ldots,n_M}(t) n_i^2. \tag{2}$$

Clearly the probability distribution $P_{n_1,n_2,\ldots,n_M}(t)$ does not contain all the information on the state of the system, which would require the full quantum density matrix. However, from the above we can gain access to the second-order correlation function $g_{2,n_i}(t) = \langle n_i^2(t)\rangle / \langle n_i(t)\rangle^2$, which is the parameter typically used to measure the coherence of a given mode. The above prescription can also be easily generalized to the case of nonzero time delay and cross-correlations between different modes.

The calculation of the quantum probability distribution can be based on the quantum Boltzmann master equation, with generic form

$$\begin{aligned}\frac{dP_{n_1,n_2,\ldots,n_M}(t)}{dt} &= \sum_{ij} W_{i\to j} P_{n_1,n_2,\ldots,n_j-1,n_i+1,\ldots,n_M}(t)(n_i+1)n_j \\ &+ \sum_i \frac{1}{\tau_i}[P_{n_1,n_2,\ldots,n_i+1,\ldots,n_M}(t)(n_i+1) \\ &- P_{n_1,n_2,\ldots,n_i,\ldots,n_M}(t) n_i] \\ &+ \sum_i \Gamma_i [P_{n_1,n_2,\ldots,n_i-1,\ldots,n_M}(t) n_i \\ &- P_{n_1,n_2,\ldots,n_i,\ldots,n_M}(t)(n_i+1)].\end{aligned} \tag{3}$$

The first term represents stimulated scattering processes between modes, where $W_{i\to j}$ is the bare (spontaneous) scattering rate from mode i to mode j. In an exciton-polariton system, this term would include phonon emission (or absorption) processes as well as scattering processes involving hot excitons [15]. These processes introduce a temperature dependence of the system via the temperatures of phonon or exciton baths.

The second and third terms represent decay and incoherent or nonresonant pumping of the modes, at rates τ_i and Γ_i, respectively. Their form is consistent with the Liouvillian operator for the full quantum density matrix, written in Ref. [36] for the case of incoherent or nonresonant pumping. In principle other scattering processes (e.g., parametric scattering processes [37]) can also be included, where the generic form is the coupling of one probability in the distribution to another.

Equation (3) can be simulated numerically given $W_{i\to j}$, τ_i, and Γ_i using the kinetic Monte Carlo approach. First, we define an initial state:

$$(n_1, n_2, \ldots, n_M). \tag{4}$$

Equation (3) defines the scattering rates to other possible states to which the above state can jump. A probability distribution of possible jumps to other states is associated with the scattering rates and a random quantum jump is selected from the probability distribution. The jump time defines the amount of time the system spends in the original state, from which the calculation of expectation values can be updated. Then the process is repeated until the end of the time range of the calculation. The process is then further repeated sampling over different quantum trajectories characterized by different stochastic quantum jumps. The system is able to attain a steady state, characterized by constant average expectation values. This is because nonlinear loss processes [38] have effectively been accounted for in the quantum Boltzmann equations; when a given mode becomes highly occupied, the probability for it to lose particles increases such that its occupation is bounded.

Since we neglect off-diagonal elements in the density matrix, we note that our approach is, strictly speaking, valid only when the system is not too far above the condensation threshold and the various scattering processes can be obtained accurately from single-particle wave functions. Far above threshold, polaritons are typically modeled with the mean-field Gross-Pitaevskii equation, where first-order coherence is assumed. Here, we are interested in the behavior crossing the threshold. In principle, polariton-polariton scattering could renormalize the energy dispersion and alter the various energy relaxation rates; however, this does not affect the general trend of relaxing to the lowest available state, which is why we will obtain results consistent with experiments even above threshold.

Given that the range of validity of the quantum Boltzmann approach is the same as that for standard classical Boltzmann equations, the quantum Boltzmann equations can always be reduced to classical ones, depending on the quantities of interest. Our motivation for working with quantum Boltzmann equations is that they give access to second-order correlations.

III. NONEQUILIBRIUM CONDENSATION IN POLARITON MICROWIRES

Exciton-polariton systems are short-lived bosonic quantum systems that are subject to weak energy relaxation processes. Consequently they are exemplary nonequilibrium quantum systems. They have been experimentally shown to form Bose-Einstein condensates [2–4], yet they may also become trapped in nonground states [39] or form non-ground-state condensates [13].

For simplicity, we will consider a one-dimensional exciton-polariton system or microwire [40]. The study of partial energy relaxation processes in such systems is particularly relevant to the study of polariton condensate transistors [41,42] and the control of spin currents for spintronics [43,44]. The main energy relaxation mechanisms in this system arise from polariton scattering with acoustic phonons [45] and the scattering of polaritons with high momentum exciton states that can be considered as a reservoir [15] (provided that we are not interested in the coherence statistics of these excitons). The scattering processes are illustrated in Fig. 1.

The calculation of the polariton-phonon scattering rates is shown in Appendix A. For typical parameters we find the

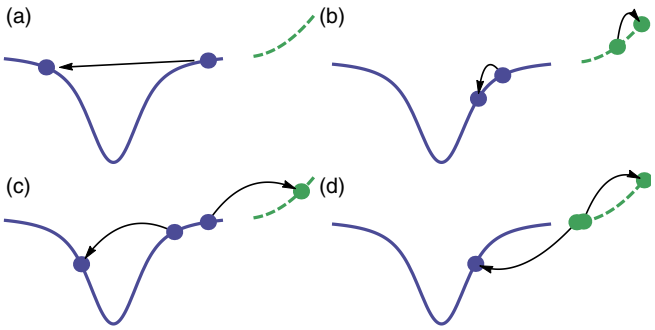

FIG. 1. Polariton energy relaxation mechanisms illustrated on the low-momentum polariton dispersion (blue/solid). The green/dashed curve illustrates the high-momentum exciton dispersion. (a) Polariton-phonon scattering. (b) Polariton-exciton scattering. (c) Polariton-polariton to polariton-exciton scattering. (d) Polariton pumping.

result shown in Fig. 2 for a temperature of 5 K. Due to the reduced density of states in a one-dimensional system, as compared to planar two-dimensional microcavities, we find that the polariton-phonon scattering rates are small, below neV. Given that typical polariton decay rates are at least 10 μeV in typical microcavities, *polariton-phonon scattering alone is insufficient to describe the relaxation of polaritons in one-dimensional systems*. Even accounting for bosonic stimulation, very large polariton occupation numbers or very high lifetime microcavities [46] would be needed to make polariton-phonon scattering dominant.

For this reason, it is important to account for the polariton-exciton scattering processes illustrated in Figs. 1(b)–1(d). The calculation of these rates is outlined in Appendix B. The process in Fig. 1(b) adds to the scattering rates W_{ij} introduced in Eq. (3). The processes in Fig. 1(c) require the addition of new terms in the quantum Boltzmann equation [Eq. (3)] of the form

$$W_{ij\to l} P_{n_1,\ldots,n_l-1,\ldots,n_i+1,\ldots,n_j+1,\ldots,n_M}(t)(n_i+1)(n_j+1)n_l. \quad (5)$$

We assume that the incoherent pumping processes of the system can be derived primarily from the process in Fig. 1(d) to provide Γ_i. It should be noted that the rates of the pumping

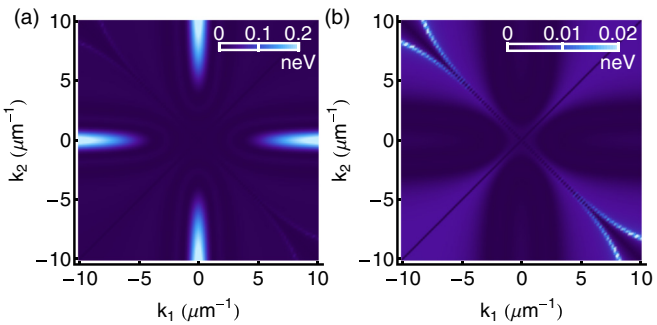

FIG. 2. Polariton-phonon scattering rates from mode k_1 to k_2. (a) Polariton relaxation processes. The scattering to $k = 0$ modes is larger than that of other modes due to the larger density of states at $k = 0$ in a one-dimensional system. (b) Polariton excitation processes, requiring nonzero temperature.

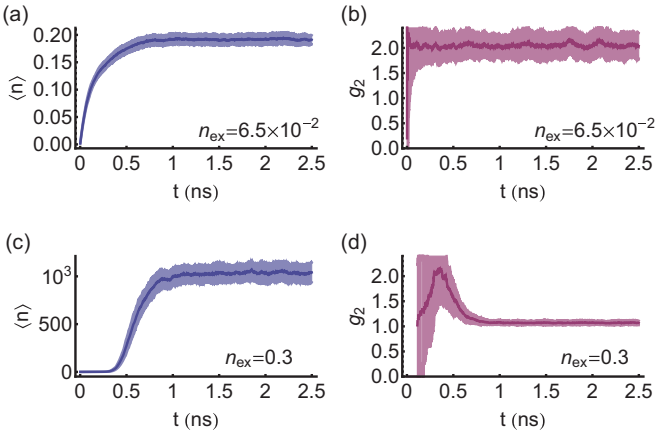

FIG. 3. (a) Time dependence of $\langle n(t)\rangle$ of the ground state below threshold. (b) Corresponding time dependence of $g_2(t)$. (c) Time dependence of $\langle n(t)\rangle$ for the ground state ($k = 0$) above threshold. (d) Corresponding time dependence of $g_2(t)$. In each panel the light shaded region represents the statistical error, corresponding to one standard deviation.

processes and the scattering process illustrated in Fig. 1(b) are proportional to the occupation of exciton reservoir states, which is modeled with a Boltzmann distribution of the form

$$n_{ex,k} = n_{\rm ex} e^{-\hbar^2 k^2/(2m_X k_B T)}, \quad (6)$$

where m_X is the exciton effective mass (taken as 0.22 times the free-electron mass in GaAs based microcavities) and $n_{\rm ex}$ is a parameter representing the maximum occupation of a state in the thermal exciton reservoir. This parameter can be taken as a measure of the strength of incoherent pumping in the system, which would be controlled experimentally via the intensity of a nonresonant laser or current from an electrical injection mechanism.

Given the aforementioned scattering rates for typical microcavity parameters (given in the Appendices) we obtain the time dependence of the average occupation of the ground state below threshold shown in Fig. 3(a). Here the system is evolved from an initial vacuum state. The average occupation remains below unity and the state is incoherent with $g_2 = 2$ [Fig. 3(b)]. The shown quantities are here obtained after averaging over more than 10^9 quantum jumps. By studying the statistical variation over different runs, one can obtain an estimate for the error in the obtained quantities. Since for larger occupations one has to sample a larger number of different states, the statistical error in the occupations is larger.

Above threshold, we find a large occupation of the ground state developing after an initial stabilization time, as shown in Fig. 3(c). This is accompanied by the formation of coherence characterized by $g_2 = 1$, as shown in Fig. 3(d). While the statistical error in g_2 is very large when the occupation numbers are small, which corresponds to a small denominator in calculating g_2 and consequently large effects of small fluctuations in $\langle n\rangle$, the statistical error above threshold becomes small and indistinguishable in the plot.

We stress that while we have accounted for polariton-phonon scattering, this does not affect significantly our results, in which polariton-exciton scattering is the dominant and

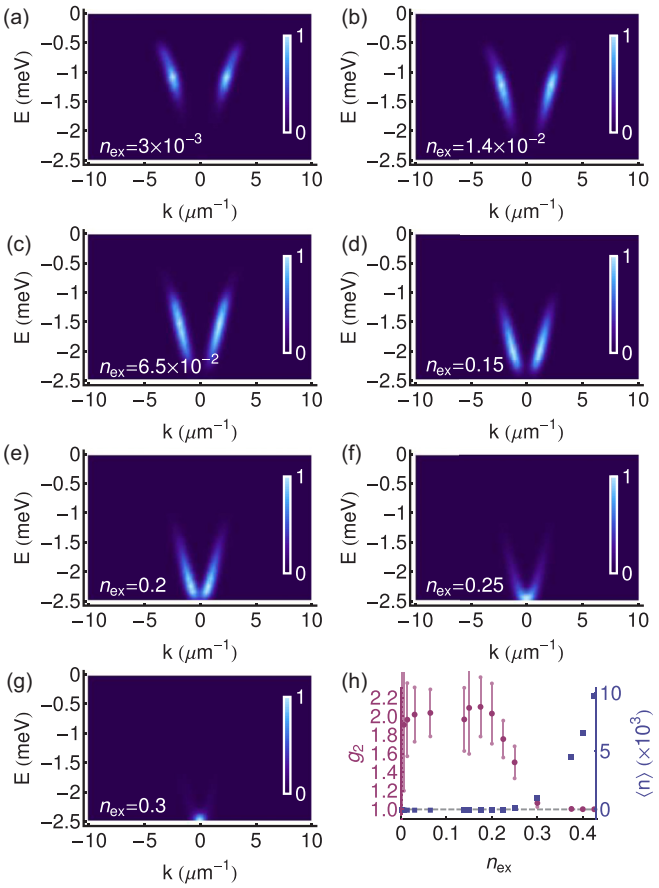

FIG. 4. (a–g) Momentum distribution of polaritons in the long-time (continuous-wave excitation) limit for increasing pumping intensity, represented by the maximum reservoir occupation number, n_{ex}. (h) Dependence of the second-order correlation function, g_2 (circles and left-hand scale), and the average occupation number, $\langle n \rangle$ (squares and right-hand scale), for the highest occupied state as a function of the reservoir occupation number. While there is a significant statistical error below threshold, the error bars narrow upon condensation.

sufficient energy relaxation mechanism. This is consistent with earlier works [15,27].

In addition to describing the behavior of polaritons below and above threshold, the kinetic Monte Carlo theory is able to access the cross-over between the incoherent and condensed regimes. Figure 4 illustrates the change in the momentum distribution of polaritons on the polariton dispersion. As the pumping intensity is increased, the various energy relaxing scattering mechanisms become more and more stimulated. This is captured as a gradual overcoming of the bottleneck region [39], before full condensation is obtained at large pumping strength. Here the large majority of polaritons collect in the system ground state and full coherence is characterized by $g_2 = 1$.

It is worth noting that the obtained phenomenon of condensation is obtained here *without direct interactions between condensing particles*. The interactions that we introduce are only to provide a physical mechanism of energy relaxation, but in principle any mechanism of energy relaxation would generate similar behavior. The considered system is thus a nonequilibrium analog of the noninteracting ideal gas. It should be noted, though, that here we are considering a confined system, in which the quantized modes in k correspond to the levels discretely modeled in our approach.

While we have focused on the second-order coherence, polariton condensates are typically also characterized by the appearance of a first-order spatial coherence, which decays exponentially with distance [47,48]. As our technique neglects off-diagonal terms in the density matrix we are unable to access this property, which could be treated with other techniques [49,50].

IV. COHERENCE FORMATION IN BOSONIC CASCADE LASERS

A bosonic cascade laser is composed of a series of equidistant energy levels and was originally proposed for the high efficiency generation of terahertz-frequency radiation [28]. Potential realizations making use of parabolic quantum wells are in experimental development [29]. When a particle is excited in a particular level of the cascade it is assumed that it can undergo a radiative transition to the next level in the cascade. Thus, in the case that the radiative transition is at terahertz frequency, one can have a high quantum efficiency process where an optical quantum of energy injected into the system undergoes multiple energy relaxing processes resulting in the emission of many terahertz-frequency photons. The terahertz emission processes are typically weak in strength, but they can become enhanced by bosonic final-state stimulation at high occupation numbers. When the system is placed inside a terahertz cavity, it has been assumed that the result will be the generation of a coherent terahertz mode although the theory of such a process has not been attempted.

The bosonic cascade laser can be described by the quantum Boltzmann rate equations:

$$\begin{aligned}
&\frac{dP_{n_1,\ldots,n_M,n_T}}{dt} \\
&= P_0 \sum_{\lambda=1}^{M} [P_{n_1,\ldots,n_\lambda-1,\ldots,n_M,n_T} - P_{n_1,\ldots,n_\lambda,\ldots,n_M,n_T}] \\
&\quad + W \sum_{\lambda=2}^{M} [-P_{n_1,\ldots,n_{\lambda-1},n_\lambda,\ldots,n_M,n_T} n_{\lambda-1}(n_\lambda+1)n_T \\
&\quad + P_{n_1,\ldots,n_{\lambda-1}-1,n_\lambda+1,\ldots,n_M,n_T-1} n_{\lambda-1}(n_\lambda+1)n_T \\
&\quad + P_{n_1,\ldots,n_{\lambda-1}+1,n_\lambda-1,\ldots,n_M,n_T+1}(n_{\lambda-1}+1)n_\lambda(n_T+1) \\
&\quad - P_{n_1,\ldots,n_{\lambda-1},n_\lambda,\ldots,n_M,n_T}(n_{\lambda-1}+1)n_\lambda(n_T+1)] \\
&\quad + \frac{1}{\tau} \sum_{\lambda=1}^{M} [P_{n_1,\ldots,n_\lambda+1,\ldots,n_M,n_T}(n_\lambda+1) \\
&\quad - P_{n_1,\ldots,n_\lambda,\ldots,n_M,n_T} n_\lambda] \\
&\quad + \frac{1}{\tau_T} [P_{n_1,\ldots,n_M,n_T+1}(n_T+1) - n_T P_{n_1,\ldots,n_M,n_T}], \qquad (7)
\end{aligned}$$

where M bosonic levels have populations $n_1, n_2, \ldots, n_M$, and n_T is the number of terahertz photons in the terahertz cavity. We assume for simplicity an equal pumping rate of all levels in

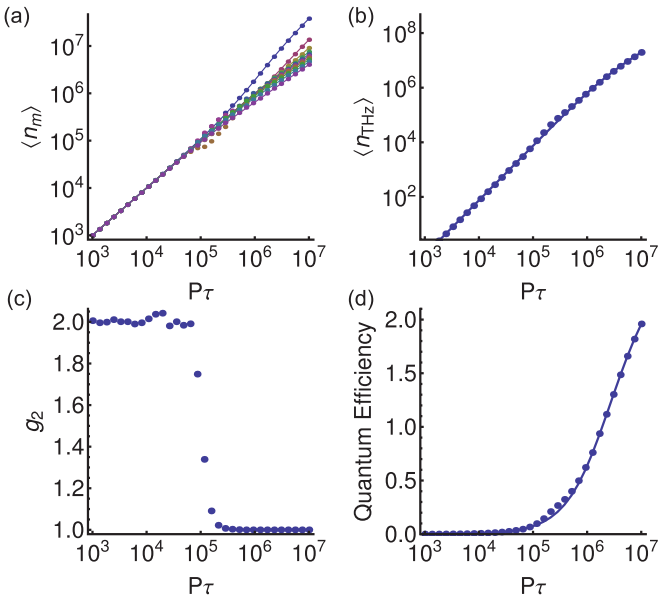

FIG. 5. Kinetic Monte Carlo simulation of a bosonic cascade laser with $M = 10$ levels. (a) Power dependence of the level occupations. (b) Power dependence of the terahertz mode occupation. (c) Power dependence of the second-order correlation function of the terahertz mode. (d) Power dependence of the quantum efficiency of terahertz generation. In all panels the points show results from kinetic Monte Carlo simulation, while the solid curves in (a), (b), and (d) show the result of the solution of the classical Boltzmann equations (see Appendix C).

the cascade by a coherent driving of strength P_0. τ is the decay rate of bosons in each level, τ_T is the decay rate of terahertz photons, and W is the nearest-neighbor level scattering rate, which is assumed here independent of the level index for simplicity. Rather than repeating the detailed calculation of the scattering rate, we take the value of $W\tau = 8.3 \times 10^{-7}$ consistent with Refs. [28,31]. Since our main objective is to derive the formation of coherence, that is, lasing in the terahertz mode, we will consider a fixed value of $\tau_T = 0.1\tau$ rather than presenting a detailed dependence on all parameters.

Figure 5 shows results from the kinetic Monte Carlo modeling of a bosonic cascade of $M = 10$ levels. The mode occupations agree fully with the result from the corresponding classical Boltzmann rate equations; however, the kinetic Monte Carlo approach provides an additional access to the second-order correlation function. This reveals the smooth transition from an incoherent to a coherent state of terahertz photons in the terahertz cavity.

For completeness we also calculated the quantum efficiency of terahertz emission, which we define as the ratio of the number of terahertz photons emitted by the system to the number of optical frequency photons put into the system [28]. This is equivalent to the ratio of the terahertz photon emission rate (n_T/τ_T) to the total system pumping rate (MP_0):

$$Q = \frac{n_T}{M\tau_T P_0}. \tag{8}$$

The cascade geometry allows terahertz lasing to appear with quantum efficiency exceeding unity, in the regime of stimulated scattering despite the weak spontaneous scattering rates in the system.

V. CONCLUSION

The kinetic Monte Carlo approach offers an efficient numerical simulation of quantum Boltzmann equations, suitable for the description of nonequilibrium bosonic systems. Such an approach offers access to the second-order correlation function and allows us to study its dynamics across nonequilibrium phase transitions. As an example, we considered the behavior of the second-order correlation function in one-dimensional exciton-polariton microwires. Here we are able to treat the cross-over from an incoherent state, to populating a nonground state, to a ground quasicondensate. While we required interactions between polaritons and hot excitons to provide a mechanism of energy relaxation from first principles, interactions between the condensing polaritons themselves were not required to generate the condensate. We also studied the formation of coherence in a bosonic cascade laser geometry, verifying the possibility of coherent terahertz emission with quantum efficiency exceeding unity.

We hope that the kinetic Monte Carlo approach can also serve in the description of partial energy relaxation and coherence formation in a variety of other exciton-polariton systems. In particular, we anticipate that the quantum optics of geometries confined with static potentials [4,51–54] and self-induced traps [55,56], where transitions were observed between nonground and ground quasicondensates, would be accessible.

ACKNOWLEDGMENTS

T.C.H.L. was supported by the Singaporean Ministry of Education—Singapore Academic Research Fund Tier-2 project (Project No. 2015-T2-1-055) and Tier-1 project (Project No. 2016-T1-1-084). D.P. was supported by the Singaporean Ministry of Education—Singapore Academic Research Fund Tier-2 project (Project No. MOE2014-T2-2-119). I.G.S. was supported by IBS-R024-D1, the Australian Research Council's Discovery Projects funding scheme (Project No. DE160100167), and the President of Russian Federation (Project No. MK-5903.2016.2). F.P.L. was supported by the POLAFLOW ERC Project No. 308136.

APPENDIX A: POLARITON RELAXATION MEDIATED BY ACOUSTIC PHONONS

The phonon-assisted scattering rate between two polariton states involving excitons of wave vectors $\mathbf{k}_1$ and $\mathbf{k}_2$ is computed following Ref. [45] as

$$W_{\mathbf{k}_1\to\mathbf{k}_2} = \frac{L_z}{\rho u V}\frac{|\Delta\mathbf{k}|^2 + q_z^2}{|\hbar u q_z|}|X(\mathbf{k}_1)X(\mathbf{k}_2)|^2 \times [a_e I_e^{||}(|\Delta\mathbf{k}|)I_e^{\perp}(q_z) - a_h I_h^{||}(|\Delta\mathbf{k}|)I_h^{\perp}(q_z)]^2 \tag{A1}$$

where $\Delta\mathbf{k} = \mathbf{k}_1 - \mathbf{k}_2$ and q_z is the projection of the phonon momentum on the vertical axis. We use here typical GaAs material parameters: $\rho = 5318\,\mathrm{kg/m^3}$ is the material den-

sity, and $L_z = 10$ nm and $V = \pi R^2 L_z$ are the quantum well thickness and volume (for a microcavity radius R), respectively. $u = 3350$ m/s is the speed of sound, and $a_e = -7$ eV and $a_h = 2.7$ eV are the lattice deformation potentials induced by phonons at the locations of electrons and holes, respectively. $I^{||}_{e,h}$ and $I^{\perp}_{e,h}$ are the overlap integrals of the phonon wave functions with the electron and hole wave functions, respectively, in the in-plane and growth directions. They are expressed as

$$I^{||}_{e,h}(|\Delta\mathbf{k}|) = \left[1 + \left(\frac{m_{e,h}}{m_e + m_h}|\Delta\mathbf{k}|a_B\right)^2\right]^{-3/2}, \quad \text{(A2)}$$

$$I^{\perp}_{e,h}(q_z) = \frac{\pi^2}{\frac{q_z L_z}{2}\left[\pi^2 - \left(\frac{q_z L_z}{2}\right)^2\right]}\sin\left(\frac{q_z L_z}{2}\right) \quad \text{(A3)}$$

where $m_e = 0.067 m_0$ and $m_h = 0.18 m_0$ are the effective masses of electrons and holes in terms of the free-electron mass m_0 and $a_B = 10$ nm is the exciton Bohr radius. Finally, $X(\mathbf{k})$ is the excitonic fraction defined as

$$X(\mathbf{k}) = \frac{2}{\sqrt{4 + \left|E_p(\mathbf{k})/\Omega_R\right|^2}} \quad \text{(A4)}$$

where $E_p(\mathbf{k})$ is the polariton dispersion relation and $\Omega_R = 10$ meV is the Rabi splitting.

APPENDIX B: POLARITON RELAXATION MEDIATED BY HOT EXCITONS

The matrix elements of scattering between polariton and hot-exciton states are obtained from Fermi's "golden rule" following Ref. [15]. The scattering rate of the process illustrated in Fig. 1(b), from a polariton state of wave vector k_1 and exciton state of wave vector k_3 to a polariton state of wave vector k_2 and an exciton state of wave vector k_4, is given by

$$W_{k_1,k_2} = \frac{2\pi}{\hbar}\left(\frac{L m_X}{\pi\hbar^2\sqrt{|k_3 k_4|}}\right)\left(\frac{6E_B a_B^2}{S}\right)^2 n_{ex,k_3}. \quad \text{(B1)}$$

Here the factor in the first parentheses on the right-hand side is an average of the initial and final exciton density of states in one dimension (assuming a parabolic exciton dispersion). The factor in the second parentheses is the matrix element of exciton-exciton scattering [57], with E_B the exciton binding energy, a_B the exciton Bohr radius, and S a normalization area. n_{ex,k_3} is the occupation of excitons in initial state k_3. For each combination of wave vectors k_1 and k_2, k_3 and k_4 are obtained from energy and momentum conservation:

$$k_3 = \frac{2m_X}{\hbar^2}\frac{E_{k_1} - E_{k_2}}{2(k_1 - k_2)} - \frac{k_1 - k_2}{2}, \quad \text{(B2)}$$

$$k_4 = k_1 - k_2 + k_3 \quad \text{(B3)}$$

where E_k represents the polariton dispersion.

Similar expressions can be used for the processes in Figs. 1(c) and 1(d). We note that in the case of Fig. 1(c) one should sum over a few different processes that can satisfy the energy and momentum (phase-matching) conditions.

APPENDIX C: CLASSICAL BOLTZMANN EQUATIONS FOR THE BOSONIC CASCADE

In the classical regime, the quantum cascade can be modeled by a set of classical rate equations [28] for the mode occupations n_λ,

$$\frac{dn_M}{dt} = P_0 + W[n_{M-1}(n_M + 1)n_T - n_M(n_{M-1} + 1)(n_T + 1)] - \frac{n_M}{\tau}, \quad \text{(C1)}$$

$$\frac{dn_\lambda}{dt} = P_0 + W[n_{\lambda-1}(n_\lambda + 1)n_T - n_\lambda(n_{\lambda-1} + 1)(n_T + 1) + n_{\lambda+1}(n_\lambda + 1)(n_T + 1) - n_\lambda(n_{\lambda+1} + 1)n_T] - \frac{n_\lambda}{\tau}, \quad \text{(C2)}$$

$$\frac{dn_1}{dt} = P_0 + W[n_2(n_1 + 1)(n_T + 1) - n_1(n_2 + 1)n_T] - \frac{n_1}{\tau}, \quad \text{(C3)}$$

where $1 < \lambda < M$, and for the terahertz mode occupation n_T:

$$\frac{dn_T}{dt} = W\sum_{\lambda}^{2}[n_\lambda(n_{\lambda-1} + 1)(n_T + 1) - n_{\lambda-1}(n_\lambda + 1)n_T] - \frac{n_T}{\tau_T}. \quad \text{(C4)}$$

Equations for the steady state are readily obtained by setting the time derivatives to zero:

$$n_M = \frac{P_0\tau + W\tau n_{M-1}n_T}{W\tau(n_T + 1 + n_{M-1}) + 1}, \quad \text{(C5)}$$

$$n_\lambda = \frac{P_0\tau + W\tau[(n_{\lambda-1} + n_{\lambda+1})n_T + n_{\lambda+1}]}{W\tau(2n_T + 1 + n_{\lambda-1} - n_{\lambda+1}) + 1}, \quad \text{(C6)}$$

$$n_1 = \frac{P_0\tau + W\tau(n_T + 1)n_2}{W(n_T - n_2) + 1}, \quad \text{(C7)}$$

$$n_T = \frac{W\tau\sum_{\lambda=2}^{M}(1 + n_{\lambda-1})n_\lambda}{W\tau(n_1 - n_M) + \frac{\tau}{\tau_T}}. \quad \text{(C8)}$$

A simultaneous solution to this set of equations can be easily found by starting from an initially unoccupied state and evaluating the quantities n_λ and n_T iteratively until the equations become consistent. The result of this procedure gives rise to the solid curves in Figs. 5(a), 5(b) and 5(d), which are in agreement with the result of full kinetic Monte Carlo modeling.

[1] J. Klaers, J. Schmitt, F. Vewinger, and M. Weitz, Nature (London) **468**, 545 (2010).

[2] J. Kasprzak, M. Richard, S. Kundermann, A. Baas, P. Jeambrun, J. M. J. Keeling, F. M. Marchetti, M. H. Szymańska, R. André,

J. L. Staehli, V. Savona, P. B. Littlewood, B. Deveaud, and L. S. Dang, Nature (London) **443**, 409 (2006).
[3] R. Balili, V. Hartwell, D. Snoke, L. Pfeiffer, and K. West, Science **316**, 1007 (2007).
[4] C. W. Lai, N. Y. Kim, S. Utsunomiya, G. Roumpos, H. Deng, M. D. Fraser, T. Byrnes, P. Recher, N. Kumada, T. Fujisawa, and Y. Yamamoto, Nature (London) **450**, 529 (2007).
[5] T. Byrnes, N. Y. Kim, and Y. Yamamoto, Nat. Phys. **10**, 803 (2014).
[6] A. P. D. Love, D. N. Krizhanovskii, D. M. Whittaker, R. Bouchekioua, D. Sanvitto, S. Al Rizeiqi, R. Bradley, M. S. Skolnick, P. R. Eastham, R. Andre, and L. S. Dang, Phys. Rev. Lett. **101**, 067404 (2008).
[7] J. Kasprzak, M. Richard, A. Baas, B. Deveaud, R. André, J-Ph. Poizat, and L. S. Dang, Phys. Rev. Lett. **100**, 067402 (2008).
[8] T. Horikiri, P. Schwendimann, A. Quattropani, S. Höfling, A. Forchel, and Y. Yamamoto, Phys. Rev. B **81**, 033307 (2010).
[9] J-S. Tempel, F. Veit, M. Aßmann, L. E. Kreilkamp, A. Rahimi-Iman, A. Löffler, S. Höfling, S. Reitzenstein, L. Worschech, A. Forchel, and M. Bayer, Phys. Rev. B **85**, 075318 (2012).
[10] A. Rahimi-Iman, A. V. Chernenko, J. Fischer, S. Brodbeck, M. Amthor, C. Schneider, A. Forchel, S. Höfling, S. Reitzenstein, and M. Kamp, Phys. Rev. B **86**, 155308 (2012).
[11] M. Amthor, H. Flayac, I. G. Savenko, S. Brodbeck, M. Kamp, T. Ala-Nissila, C. Schneider, and S. Höfling, arXiv:1511.00878.
[12] S. Kim, B. Zhang, Z. Wang, J. Fischer, S. Brodbeck, M. Kamp, C. Schneider, S. Hofling, and H. Deng, Phys. Rev. X **6**, 011026 (2016).
[13] M. Maragkou, A. J. D. Grundy, E. Wertz, A. Lemaître, I. Sagnes, P. Senellart, J. Bloch, and P. G. Lagoudakis, Phys. Rev. B **81**, 081307(R) (2010).
[14] D. Vorberg, W. Wustmann, R. Ketzmerick, and A. Eckardt, Phys. Rev. Lett. **111**, 240405 (2013).
[15] D. Porras, C. Ciuti, J. J. Baumberg, and C. Tejedor, Phys. Rev. B **66**, 085304 (2002).
[16] T. D. Doan, H. T. Cao, D. B. T. Thoai, and H. Haug, Phys. Rev. B **72**, 085301 (2005).
[17] J. Kasprzak, D. D. Solnyshkov, R. André, L. S. Dang, and G. Malpuech, Phys. Rev. Lett. **101**, 146404 (2008).
[18] H. T. Cao, T. D. Doan, D. B. T. Thoai, and H. Haug, Phys. Rev. B **77**, 075320 (2008).
[19] D. Read, T. C. H. Liew, Y. G. Rubo, and A. V. Kavokin, Phys. Rev. B **80**, 195309 (2009).
[20] M. Wouters, T. C. H. Liew, and V. Savona, Phys. Rev. B **82**, 245315 (2010).
[21] M. Wouters, New J. Phys. **14**, 075020 (2012).
[22] L. M. Sieberer, S. D. Huber, E. Altman, and S. Diehl, Phys. Rev. Lett. **110**, 195301 (2013).
[23] I. G. Savenko, T. C. H. Liew, and I. A. Shelykh, Phys. Rev. Lett. **110**, 127402 (2013).
[24] D. D. Solnyshkov, H. Terças, K. Dini, and G. Malpuech, Phys. Rev. A **89**, 033626 (2014).
[25] D. Racine and P. R. Eastham, Phys. Rev. B **90**, 085308 (2014).
[26] K. Molmer, Y. Castin, and J. Dalibard, J. Opt. Soc. Am. B **10**, 524 (1993).
[27] H. Flayac, I. G. Savenko, M. Möttönen, and T. Ala-Nissila, Phys. Rev. B **92**, 115117 (2015).
[28] T. C. H. Liew, M. M. Glazov, K. V. Kavokin, I. A. Shelykh, M. A. Kaliteevski, and A. V. Kavokin, Phys. Rev. Lett. **110**, 047402 (2013).
[29] A. Tzimis, A. V. Trifonov, G. Christmann, S. I. Tsintzos, Z. Hatzopoulos, I. V. Ignatiev, A. V. Kavokin, and P. G. Savvidis, Appl. Phys. Lett. **107**, 101101 (2015).
[30] S. Chaturvedi, C. W. Gardiner, I. S. Matheson, and D. F. Walls, J. Stat. Phys. **17**, 469 (1977).
[31] T. C. H. Liew, Y. G. Rubo, A. S. Sheremet, S. De Liberato, I. A. Shelykh, F. P. Laussy, and A. V. Kavokin, New J. Phys. **18**, 023041 (2016).
[32] W. M. Young and E. W. Elcock, P. Phys. Soc. **89**, 735 (1966).
[33] A. B. Bortz, M. H. Kalos, and J. L. Lebowitz, J. Comput. Phys. **17**, 10 (1975).
[34] D. T. Gillespie, J. Comput. Phys. **22**, 403 (1976).
[35] K. A. Fichthorn and W. H. Weinberg, J. Chem. Phys. **95**, 1090 (1991).
[36] F. P. Laussy, E. del Valle, and C. Tejedor, Phys. Rev. Lett. **101**, 083601 (2008).
[37] P. G. Savvidis, J. J. Baumberg, R. M. Stevenson, M. S. Skolnick, D. M. Whittaker, and J. S. Roberts, Phys. Rev. Lett. **84**, 1547 (2000).
[38] J. Keeling and N. G. Berloff, Phys. Rev. Lett. **100**, 250401 (2008).
[39] M. Richard, J. Kasprzak, R. André, R. Romestain, L. S. Dang, G. Malpuech, and A. Kavokin, Phys. Rev. B **72**, 201301 (2005).
[40] E. Wertz, L. Ferrier, D. Solnyshkov, R. Johne, D. Sanvitto, A. Lemaître, I. Sagnes, R. Grousson, A. V. Kavokin, P. Senellart, G. Malpuech, and J. Bloch, Nat. Phys. **6**, 860 (2010).
[41] C. Anton, T. C. H. Liew, G. Tosi, M. D. Martín, T. Gao, Z. Hatzopoulos, P. S. Eldridge, P. G. Savvidis, and L. Viña, Appl. Phys. Lett. **101**, 261116 (2012).
[42] C. Anton, T. C. H. Liew, G. Tosi, M. D. Martín, T. Gao, Z. Hatzopoulos, P. S. Eldridge, P. G. Savvidis, and L. Viña, Phys. Rev. B **88**, 245307 (2013)
[43] C. Antón, S. Morina, T. Gao, P. S. Eldrdige, T. C. H. Liew, M. D. Martín, Z. Hatzopoulos, P. G. Savvidis, I. A. Shelykh, and L. Viña, Phys. Rev. B **91**, 075305 (2015).
[44] T. Gao, C. Antón, T. C. H. Liew, M. D. Martín, Z. Hatzopoulos, L. Viña, P. S. Eldridge, and P. G. Savvidis, Appl. Phys. Lett. **107**, 011106 (2015).
[45] C. Piermarocchi, F. Tassone, V. Savona, A. Quattropani, and P. Schwendimann, Phys. Rev. B **53**, 15834 (1996).
[46] Y. Sun, P. Wen, Y. Yoon, G. Liu, M. Steger, L. N. Pfeiffer, K. West, D. W. Snoke, and K. A. Nelson, Phys. Rev. Lett. **118**, 016602 (2017).
[47] L. Mouchliadis and A. L. Ivanov, Phys. Rev. B **78**, 033306 (2008).
[48] V. V. Belykh, N. N. Sibeldin, V. D. Kulakovskii, M. M. Glazov, M. A. Semina, C. Schneider, S. Höfling, M. Kamp, and A. Forchel, Phys. Rev. Lett. **110**, 137402 (2013).
[49] D. Sarchi and V. Savona, Phys. Rev. B **75**, 115326 (2007).
[50] T. D. Doan, H. T. Cao, D. B. T. Thoai, and H. Haug, Phys. Rev. B **78**, 205306 (2008).
[51] N. Y. Kim, K. Kusudo, A. Loffler, S. Hofling, A. Forchel, and Y. Yamamoto, New J. Phys. **15**, 035032 (2013).
[52] E. A. Ostrovskaya, J. Abdullaev, M. D. Fraser, A. S. Desyatnikov, and Y. S. Kivshar, Phys. Rev. Lett. **110**, 170407 (2013).
[53] L. Zhang, W. Xie, J. Wang, A. Poddubny, J. Lu, Y. Wang, J. Gu, W. Liu, D. Xu, X. Shen, Y. G. Rubo, B. L. Altshuler, A. V. Kavokin, and Z. Chen, Proc. Natl. Acad. Sci. USA **112**, E1516 (2015).

[54] K. Winkler, O. A. Egorov, I. G. Savenko, X. Ma, E. Estecho, T. Gao, S. Müller, M. Kamp, T. C. H. Liew, E. A. Ostrovskaya, S. öfling, and C. Schneider, Phys. Rev. B **93**, 121303(R) (2016).

[55] A. Askitopoulos, T. C. H. Liew, H. Ohadi, Z. Hatzopoulos, P. G. Savvidis, and P. G. Lagoudakis, Phys. Rev. B **92**, 035305 (2015).

[56] T. Gao, E. Estrecho, K. Y. Bliokh, T. C. H. Liew, M. D. Fraser, S. Brodbeck, M. Kamp, C. Schneider, S. Höfling, Y. Yamamoto, F. Nori, Y. S. Kivshar, A. G. Truscott, R. G. Dall, and E. A. Ostrovskaya, Nature (London) **526**, 554 (2015).

[57] F. Tassone and Y. Yamamoto, Phys. Rev. B **59**, 10830 (1999).