Non-equilibrium Electrical Generation of Surface Phonon Polaritons

Christopher R. Gubbin,[1] Stanislas Angebault,[1] Joshua D. Caldwell,[1, 2] and Simone De Liberato[1, 3, 4]

[1] *Sensorium Technological Laboratories, Nashville, TN, USA*
[2] *Vanderbilt University, Nashville, TN, USA*
[3] *School of Physics and Astronomy, University of Southampton, Southampton SO17 1BJ, UK*
[4] *Istituto di Fotonica e Nanotecnologie – Consiglio Nazionale delle Ricerche (CNR), Milano, Italy*
(Dated: October 25, 2024)

Notwithstanding its relevance to many applications in sensing, security, and communications, electrical generation of narrow-band mid-infrared light remains highly challenging. Unlike in the ultraviolet or visible spectral regions few materials possess direct electronic transitions in the mid-infrared and most that do are created through complex band-engineering schemes. An alternative mechanism, independent of dipole active material transitions, relies instead on energy lost to the polar lattice through the Coulomb interaction. Longitudinal phonons radiated in this way can be spectrally tuned through the engineering of polar nanostructures and coupled to localized optical modes in the material, allowing them to radiate mid-infrared photons into the far-field. A recent theoretical work explored this process providing for the first time an indication of its technological relevance when compared to standard thermal emitters. In order to do so it nevertheless used an equilibrium model of the electron gas, making this model difficult to inform the design of an optimal device to experimentally observe the effect. The present paper removes this limitation, describing the electron gas using a rigorous, self-consistent, non-equilibrium Green's function model, accounting for variations in material properties across the device, and electron-electron interactions. Although the instability of the Schrodinger-Poisson iteration limits our studies to the low-bias regime, our results demonstrate emission rates comparable to that of room-temperature thermal emission despite such low biases. These results provide a pathway to design a confirmatory experiment of this new emission channel, that could power a new generation of mid-infrared optoelectronic devices.

I. INTRODUCTION

As a consequence of strong, long-range Coulombic interactions between charged electrons and the crystal lattice, electrical transport in polar materials is dominated by longitudinal optical (LO) phonon scattering, which ultimately determines the electron saturation velocity [1, 2]. Recent works have proposed that LO phonons generated in this way can be put to use: in nanoscopic layers they are able to couple directly to light through shared boundary conditions [3–7], leading to the far-field emission of mid-infrared photons [8].

Polar dielectric crystals support localized excitations termed surface phonon polaritons (SPhPs). These modes are a consequence of strong coupling between photons and the polar optical phonon modes of the lattice [9, 10], existing in the Reststrahlen region separating the lattice longitudinal and transverse optical phonon frequencies in which the dielectric function turns negative. It is within this spectral range that the real part of the dielectric permittivity tensor becomes negative. As hybrid photon-phonon quasi-particles, they enable deep sub-diffractional confinement of energy [11, 12] and fine spectral tunablity [13–18] with applications in nonlinear optics [19–22], sensing [23, 24] and near-field imaging [25, 26]. The phonon-like nature of SPhPs has wide consequences for mid-infrared optoelectronics [27], particularly in the design of tailored narrowband thermal emitters enabled by the mode's underlying spectral tunablity and narrow linewidth [28–30].

Longitudinal and transverse optical phonons are also tunable. In macroscopic systems the dispersion of each is determined by the composition and symmetries of the crystal lattice, but when the lengthscale approaches the phonon propagation length they can become localized. In this regime the continuous phonon dispersion evolves into a discrete one, analogous to the shift in dispersion experienced by photons in a Fabry-Pérot resonator. The SPhPs in such a system exhibit spectral shifts [3] and the emergence of new modes observable from the far-field [31]. In this regime they acquire a new moniker: longitudinal-transverse polaritons (LTPs) [3–5, 7] reflecting the mixture of longitudinal and transverse degrees of freedom and the consequent necessity of considering the true microscopic nature of the underlying phonons.

Longitudinal-transverse polaritons are able to couple directly to electrons through the Coulomb interaction and radiate into the far-field like SPhPs. This enables them to act as the active mode in an electrically driven emitter without the need for dipole-active electronic transitions. Optoelectronic devices designed in this way could provide a simple and flexible pathway to generating radiation: polar dielectric materials have a wide range of phonon frequencies spanning the mid-infrared spectral region and for each material LTP modes can be fine-tuned across the Reststrahlen spectral window. Electrical emission via LTPs is thus a complex multi-particle process, in which non-equilibrium electrons couple to LO phonons, which in turn couple to SPhPs, which themselves couple to free-space radiation. To understand the practicality of LTP-based electrical emitters the previous study of LTP emission in polar nanosystems treated the electronic sub-system using a simple, equilibrium free-

electron model characterized by a uniform electron temperature [8]. Although such a model was sufficient to provide an estimation of the underlying efficiency of the LTP emission channel, it cannot account for local variations of the electron density, the effect of a multilayer structure with varying effective mass and conduction band offsets, or the effect of electron-electron repulsion. These shortcomings prevent calculation of the electrical current across the device, making it difficult to design and optimize a device to observe such an effect.

This Paper aims to overcome these issues, assessing LTP generation rates using self-consistently calculated electron distributions, determined by iterative solution of the coupled Schrodinger and Poisson equations. Our results allow us to assess the changes induced through a rigorous model of the electron gas. Due to the inherent instability in solving the Schrodinger-Poisson equation [32], in this Paper we were able to obtain converged results only in the small bias regime, where emission is expected to be weak. This regime is most interesting for an experiment aiming to provide a first observation of LTP-based electrical emission, as at small bias the device would exhibits less resistive heating and thus less spurious thermal emission. While the problem of simulation of the large-bias regime will be addressed in future works, the results here presented provide strong evidence for the feasibility of LTP based optoelectronic devices, and a pathway toward design of future experiments.

II. THEORETICAL OVERVIEW

A. Longitudinal-Transverse Polaritons in Thin Films

Surface phonon polaritons are hybrid excitations, resulting from the strong coupling of optical phonons to light at the surface of a polar dielectric. Both components exhibit dispersion, meaning their frequency is a function of momentum, however the lengthscales for this dispersion are very different. A photon appreciably shifts frequency for wavevector changes comparable to it's inverse wavelength, while similar shifts for a phonon happen for wavevectors comparable to the inverse lattice spacing. This broad disparity often allows the optical response of an SPhP system to be accurately reconstructed neglecting the phonon dispersion. This is a local-response approximation, termed because the material response at a point $\mathbf{r}$ depends only on the local electric field at $\mathbf{r}$: a consequence of neglecting phonon dispersion is that all phonons are assumed to have zero momentum, so energy is solely transported in the photon field. In systems with length-scales approaching the phonon propagation length a local-response approximation can fail. Energy can be transported through optical phonon excitations, and the material response at $\mathbf{r}$ becomes nonlocal, depending on the driving field in a finite region enclosing $\mathbf{r}$ [3]. In this regime the hybrid nature of the mode becomes increas-
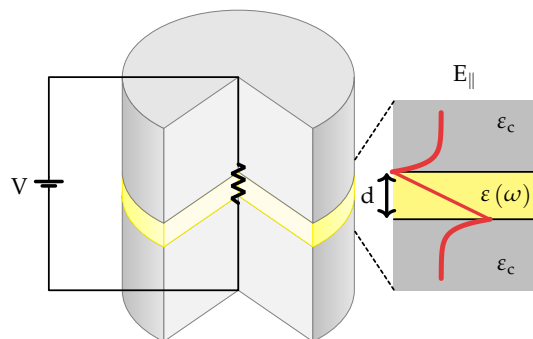

FIG. 1. Sketch of the photonic system under study. A polar dielectric film of thickness d is sandwiched between symmetric high-index cladding layers characterized by non-dispersive dielectric constant ε_c. In the Reststrahlen spectral region of the polar layer the system supports surface phonon polariton excitations, the mode profile for the epsilon-near-zero excitation is sketched in red. The device can emit light by applying a voltage perpendicular to the polar film as indicated.

ingly important. While in the local-response approximation SPhPs can be considered pure transverse excitations, with polarization fields $\mathbf{P}$ orthogonal to wavevector $\mathbf{Q}$ ($\mathbf{Q} \times \mathbf{P} = 0$), in the nonlocal regime this is no longer the case leading to the existence of LTP excitations [4, 5, 7].

In this Paper we focus on a nanoscale polar film of thickness d, sandwiched between non-polar cladding layers with a non-dispersive dielectric constant $\epsilon_c > 0$ illustrated in Fig.1. It has previously been demonstrated [7] that LTPs in this system can be accurately described as linear superpositions of localised LO phonons and long-range epsilon-near-zero (ENZ) SPhP modes of the polar film, with dispersion $\omega_{\mathbf{q}}^{\mathrm{SP}}$, whose mode profile of is sketched in Fig. 1. In this Paper we focus on the case in which a single localized phonon branch, with frequency $\omega_{\mathbf{q}}^{\mathrm{L}}$, is quasi-resonant with the ENZ SPhP, leading to two LTP branches which we will refer to as lower (-) and upper (+) polaritons, with frequencies $\omega_{\mathbf{q}}^{\pm}$. We use the symbols $\mathrm{X}_{\mathbf{q}}^{\pm}$ and $\mathrm{Y}_{\mathbf{q}}^{\pm}$ to indicate the Hopfield coefficients for the longitudinal and the transverse components of the polaritons, obeying the normalization condition $|\mathrm{X}_{\mathbf{q}}^{\pm}|^2 + |\mathrm{Y}_{\mathbf{q}}^{\pm}|^2 = 1$. Details on the derivation of these quantities as a function of the system parameters can be found in Appendix A.

B. Non equilibrium Green's functions

Transport in open systems can be described using a Landauer-Büttiker formalism [33]. In an open-dissipative quantum system electrons transition between energy and momentum states by interacting with external reservoirs, requiring knowledge of two-time correlation functions for the electron gas. Correlation functions can be calcu-

lated in a non-equilibrium Green's function (NEGF) formalism [34–38]. Green's functions generalize the time-dependent electron distribution $n(\mathbf{k}, t)$ to a two-time correlation function between creation and annihilation operators $\hat{c}_{\mathbf{k}}^{\dagger}(t)$ and $\hat{c}_{\mathbf{k}}(t)$ for an electron with wavevector $\mathbf{k}$ at time t. The NEGF formalism describes the electronic population utilizing the distributional lesser and greater Green's functions [39]

$$G^{<}\left(\mathbf{k}, \mathbf{k}'; t, t'\right) = i\langle c_{\mathbf{k}'}^{\dagger}(t'), c_{\mathbf{k}}(t)\rangle, \qquad (1a)$$

$$G^{>}\left(\mathbf{k}, \mathbf{k}'; t, t'\right) = -i\langle c_{\mathbf{k}}(t), c_{\mathbf{k}'}^{\dagger}(t')\rangle, \qquad (1b)$$

which in the steady-state regime considered in this Paper solely depend on the time difference $\tau = t - t'$. Fourier transforming with respect to space and time we obtain the distributional Green's functions $G^{\lessgtr}(\mathbf{r}, \mathbf{r}; E)$ describing correlations between the amplitudes at positions $\mathbf{r}, \mathbf{r}'$ with energy E, whose diagonal element yields the real-space carrier density in the device

$$n(\mathbf{r}) = -\frac{2i}{\Delta A} \int \frac{1}{2\pi} G^{<}(\mathbf{r}, \mathbf{r}; E) \, \mathrm{dE}, \qquad (2)$$

where we wrote the element volume as the product of quantization area A and mesh-spacing Δ [39].

Calculating $G^{\lessgtr}$ requires knowledge of the associated dynamic Green's functions: the retarded Green's function G^{R} and the advanced Green's function $G^{A} = \left[G^{R}\right]^{\dagger}$. These can be derived from a generalization of the Schrodinger equation to an open system, contacted to external reservoirs using leads characterized by self-energies quantifying the complex frequency shift associated with the coupling

$$\left[E - \mathcal{H}_0 - \Sigma^{R}\right] G^{R} = \delta\left(\mathbf{r} - \mathbf{r}'\right), \qquad (3)$$

where $\mathcal{H}_0$ is the closed system Hamiltonian discussed in Appendix B 1 and Σ^{R} is the retarded self-energy whose derivation for coupling to semi-infinite leads is described in Appendix B 2. The distributional Green's functions $G^{\lessgtr}$ are then given by the kinetic equation

$$G^{\lessgtr} = G^{R} \Sigma^{\lessgtr} G^{A}, \qquad (4)$$

where $\Sigma^{\lessgtr}$ are lesser and greater self-energies. Self-energies refer to inflows and outflows of charge from the distributions described by $G^{\lessgtr}$ and their difference relates to the level broadening Γ

$$\Gamma = i\left[\Sigma^{R} - \Sigma^{A}\right] = i\left[\Sigma^{>} - \Sigma^{<}\right]. \qquad (5)$$

Note that as $\mathcal{H}_0$ is dependent on the electrostatic potential ϕ through the Poisson equation, whose solution is outlined in Appendix B 3, which in turn depends on the carrier density calculated through Eq. 2, an iterative solution is necessary, as outlined in Appendix B 4.

C. Figures of Merit

In this Paper we aim to compare the efficiency of the longitudinal phonon driven electrical LTP emission channel with those of SPhP thermal emitters that are already well established in the literature [13, 28, 30]. For this reason we follow the discussion in Ref.[8], defining two figures of merit for the performance of an electrical emitter that rely on comparison of electrical and thermal populations at the same wavevector, rather than net emitted fluxes. The first is the ratio between the steady-state electrically excited LTP population $N_{\mathbf{q}}^{\mathrm{ss}}$ in a lattice at ambient temperature $T_a = 300$ K and thermal population N_0 at the same temperature

$$F^{(1)}\left(\omega_{\mathbf{q}}, T_a, T_e\right) = \frac{N^{\mathrm{ss}}\left(\omega_{\mathbf{q}}, T_a\right)}{N^{0}\left(\omega_{\mathbf{q}}, T_a\right)}. \qquad (6)$$

This represents the emission enhancement compared to the thermal case, which may be observed in an experiment. The thermal population N_0 is given by the Bose-Einstein distribution

$$N^{0}(\omega, T) = \frac{1}{e^{\hbar\omega/k_{\mathrm{B}}T} - 1}. \qquad (7)$$

The steady-state population of excitations in the LTP reservoir under interaction with the electron gas derived in Ref. [8] is found using the Fermi golden rule to be

$$N^{\mathrm{ss}}\left(\omega_{\mathbf{q}}, T\right) = \frac{N^{0}\left(\omega_{\mathbf{q}}, T\right)\gamma_{\mathbf{q}} + \Gamma_{\mathbf{q}}^{\mathrm{in}}}{\gamma_{\mathbf{q}} + \Gamma_{\mathbf{q}}^{\mathrm{out}} - \Gamma_{\mathbf{q}}^{\mathrm{in}}}, \qquad (8)$$

where T is the temperature, $\omega_{\mathbf{q}}$ is the reservoir excitation frequency at in-plane wavevector $\mathbf{q}$, $\gamma_{\mathbf{q}}$ is the wavevector-dependent damping rate for excitations in the reservoir, and $\Gamma_{\mathbf{q}}^{\mathrm{in/out}}$ are the rates of in/out-scattering from the reservoir due to interactions with the electron gas. Note that Eq. 8 is equally valid for excitation of localized LO phonons and LTPs provided the appropriate dispersion relation and electron-phonon interaction rates.

Although the figure of merit is the same as that considered in Ref. [8], the key difference is that in this work the electron distributions entering $\Gamma^{\mathrm{in/out}}$ are calculated self-consistently using a NEGF formalism. Additionally, compared to the free-electron model used in Ref. [8] our theory requires a new parameter. In a free-electron model the carrier distribution is characterized by a single temperature that defines both the shape of the Fermi surface and the drift velocity. In this work the two properties are decoupled: the electron gas exists at ambient temperature regardless of the bias across the device. This improves the accuracy of our results. We take the electronic temperature to be ambient $T_a = 300$ K in all results in this Paper.

The second figure of merit quantifies the number of photons generated by a single electron passing through the system. The photon-fraction of an LTP can be approximated as ratio of it's group-velocity to the speed of light

$$\xi_{\mathbf{q}}^{\pm} = \frac{1}{c}\left\|\frac{\mathrm{d}\omega_{\mathbf{q}}^{\mathrm{SP}}}{\mathrm{d}\mathbf{q}}\right\| |\mathrm{Y}_{\mathbf{q}}^{\pm}|^2. \qquad (9)$$
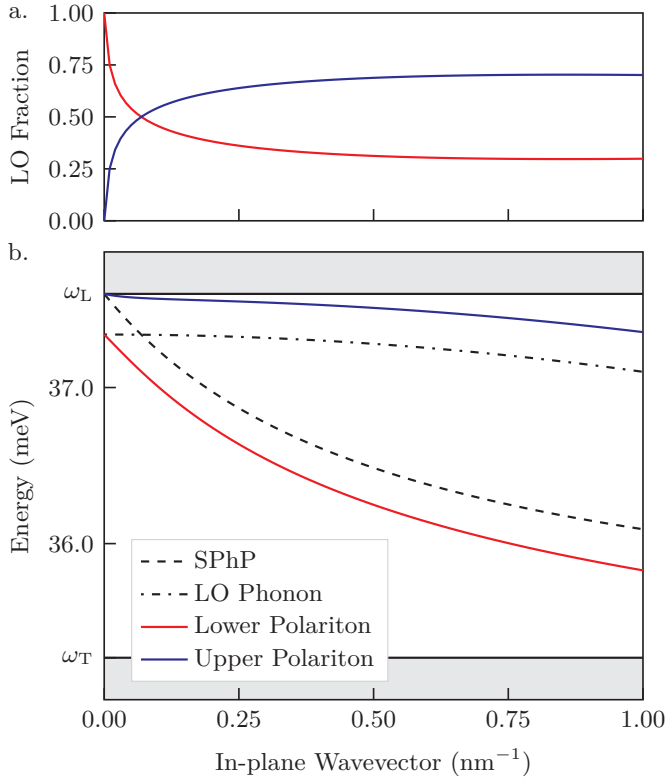
FIG. 2. Longitudinal-transverse polaritons for a 3 nm AlGaAs film. a. The red and blue lines indicate the longitudinal fraction $|X_{\mathbf{q}}^{\pm}|^2$ for the two polariton branches. b. The dashed and dash-dotted lines indicate respectively the dispersions of the ENZ SPhP mode $\omega_{\mathbf{q}}^{\text{SP}}$ and of the localised longitudinal phonon $\omega_{\mathbf{q}}^{\text{L}}$. The red and blue solid lines indicate the LTP frequencies $\omega_{\mathbf{q}}^{\pm}$.

Modes with flat dispersion are matter-like, with low group velocity and vanishing $\xi_{\mathbf{q}}$, while fast dispersing excitations have a sizeable photonic component and acquire a finite $\xi_{\mathbf{q}}$. The electrically generated *photonic* power per unit surface is then given by

$$F^{(2)} = \frac{1}{2\pi} \int \mathrm{d}q \, q \, \hbar\omega_{\mathbf{q}} \xi_{\mathbf{q}} \Gamma_{\mathbf{q}}^{\text{net}}, \qquad (10)$$

where $\Gamma_{\mathbf{q}}^{\text{net}} = [1 + N(\omega_{\mathbf{q}}, T_l)] \Gamma_{\mathbf{q}}^{\text{in}} - N(\omega_{\mathbf{q}}, T_l) \Gamma_{\mathbf{q}}^{\text{out}}$ is the net excitation rate of LTPs.

III.   RESULTS

A.   Electrical characterization

In this Paper we consider a system consisting of an $\text{Al}_{0.24}\text{Ga}_{0.76}\text{As}$ barrier sandwiched between GaAs leads. The polar AlGaAs barrier is a negative dielectric near to it's zone-center LO phonon frequency $\omega_{\text{L}} = 37.6$ meV where it supports an ENZ SPhP excitation [40]. For sim-
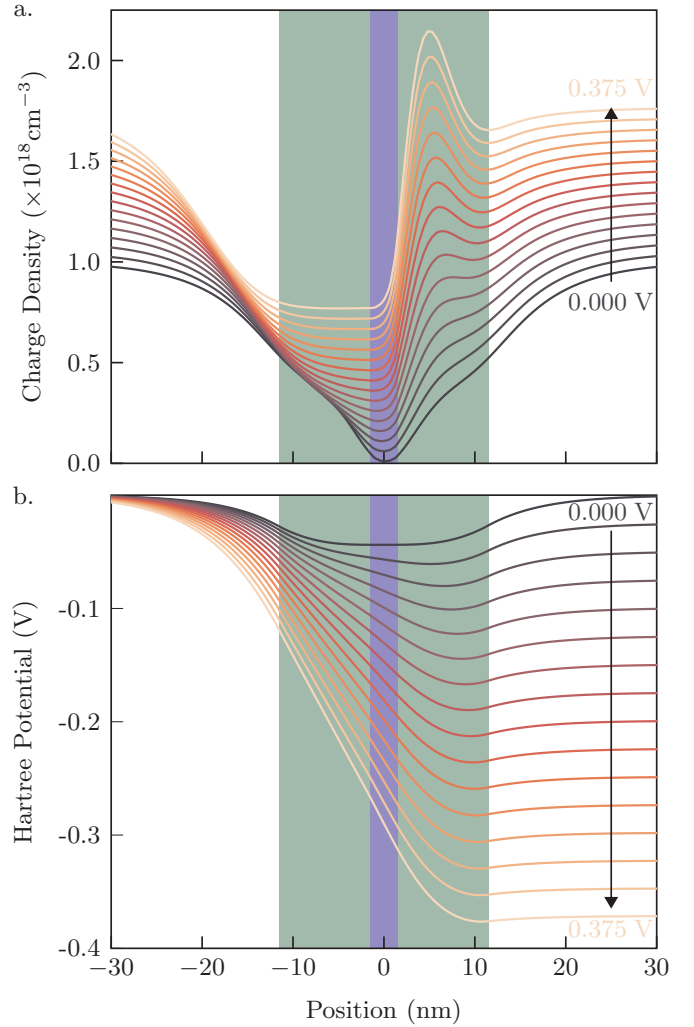

FIG. 3. Characteristics of a 3 nm AlGaAs barrier, sandwiched between GaAs leads as the applied bias is altered. The darkest curve represents the zero bias result, while lighter colors illustrate larger applied bias. The background shading illustrates different material regions of the device: blue indicates the 3 nm AlGaAs barrier, green the 10 nm lightly doped GaAs access regions and whithe the strongly doped GaAs leads. Panels show a. The self-consistent charge distribution. For clarity results for non-zero bias are offset vertically. b. The self-consistent Hartree potential.

plicity we consider the GaAs access regions to be positive dielectrics characterized by high-frequency dielectric function $\epsilon_{\infty} = 10.89$. This approximation allows use of the analytical LTP theory presented in Section. II A. Dispersion in the true dielectric function of the GaAs will shift the dispersion of the SPhP mode reducing the bandwidth, which can be accounted for in a full numerical calculation of the nonlocal optical response [5].

Using formulas from Appendix A, in Fig.2a we plot the longitudinal content of the two LTP branches $|X_{\mathbf{q}}^{\pm}|^2$ and in Fig. 2b their modal dispersions. At zero wavevector
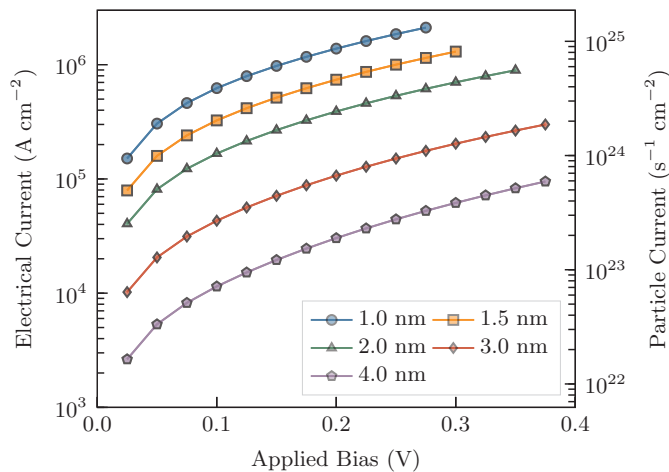
FIG. 4. Comparison of ballistic charge current as a function of barrier thickness. The left axis shows the electrical current density J from Eq. 11 while the right shows the particle current density J/e.

polariton frequencies (solid lines) coalesce with those of the bare modes (black lines): the upper-polariton is entirely ENZ SPhP-like, while the lower is entirely LO-phonon-like. This is a consequence of the fact that, as shown in Appendix A, the longitudinal-transverse coupling tends to zero as the ENZ SPhP mode frequency $\omega_{\mathbf{q}}^{\mathrm{SP}} \rightarrow \omega_{\mathrm{L}}$. At larger wavevectors both modes acquire substantial longitudinal character [7] as they deviate from the bare mode frequencies.

The photonic system is contacted by sandwiching between highly n-doped (doping density $N_{\mathrm{D}} = 10^{18}$ cm^{-3}) semi-infinite GaAs leads. GaAs in the internal photonic system is considered to be lightly doped ($N_{\mathrm{D}} = 10^{12}$ cm^{-3}). A buffer region thickness of 10 nm is used in this Paper, sufficiently thick that the highly confined LTP does not extend into the strongly doped leads. These lightly doped buffer regions ensure the cladding dielectric function ϵ_{c} remains positive, reduce maximal charge density in the central region, and reduces charge accumulation at the emitter. This improves current stability [41] and convergence of the self-consistent calculation. The structure is characterized by effective mass $m^* = 0.063\ (0.083)m_e$ with m_e the bare electron mass, and static dielectric constant $\epsilon_{\mathrm{st}} = 12.90\ (12.21)$ in the GaAs (AlGaAs) region. The system can be seen in the colors of Fig. 3: blue indicates the AlGaAs film, green the lightly doped buffer region and white the heavily doped leads.

The conduction band offset of the AlGaAs barrier (0.3 eV) expels charge from the device central region. This can be observed in Fig. 3a, which shows self-consistent calculations of the coherent charge distributions in the device for a 3 nm barrier at ambient temperature ($T_a = 300$ K) using Eq. 2. At zero bias there is a deep minimum in the charge density at the center of the

AlGaAs barrier. As the bias ramps up, electrons are able to tunnel through the barrier: this increases the charge density in the barrier somewhat but predominantly leads to a strong increase in charge density at the barrier edge as electrons with insufficient energy to tunnel across are pushed up against it. The corresponding Hartree potential in each case is shown in Fig. 3b. In the leads the gradient of the potential tends to zero and the charge density approaches the donor density, while in the device center, where the potential gradient is large there is a strong deviation from the background charge distribution.

Ballistic current conduction across the barrier is given in terms of the dynamic Green's functions and self-energies of the device contacts as

$$J = \frac{2e}{\hbar \mathrm{A}} \sum_{\mathbf{k}} \int \frac{\mathrm{dE}}{2\pi} \mathrm{Tr} \left\{ \Gamma^e G^R \Gamma^c G^A \right\}$$
$$\times \left[f_e \left(T_a, E \right) - f_c \left(T_a, E \right) \right], \quad (11)$$

where f_e (f_c) is the Fermi function in the emitter (collector) contact and Γ^e (Γ^c) is the broadening associated with the contact, calculated from the contact self-energies derived in Appendix B 2 through Eq. 5. This is plotted for a selection of barrier thicknesses in Fig. 4. As the applied bias is ramped up, inter-terminal current increases rapidly as more electrons acquire sufficient energy to tunnel across the barrier. Thinner barriers exhibit higher current at equal applied bias as the evanescent tails of wavefunctions belonging to electrons with energies below the barrier height are more able to penetrate through. As mentioned in the Introduction, and discussed in more details in Appendix B 5, in this Paper we limit ourselves to the small-bias regime, with a maximal bias of 0.4V. Such a value is larger than the conduction band offset, to show the effect of above-barrier transport, but lower than the operating condition of an ideal real-world device. On the large bias side of Fig. 4 the maximum voltage presented decreases for thinner film thickness because the level of instability grows as more current pours over the barrier, which happens at lower bias for thinner barrier layers [32].

B. LTP emission

In this section we consider incoherent transport in the device. We do not calculate a full-self consistent solution of both LTP and electron populations in the device, instead considering the perturbative regime where the equilibrium electron distribution radiates LTPs but is not appreciably altered by doing so. Firstly we formulate some useful figures of merit for assessing incoherent transport in the NEGF formalism. The total rate of out-scattering experienced by a state in the electron gas with in-plane momentum $\mathbf{k}$ and energy E is proportional to
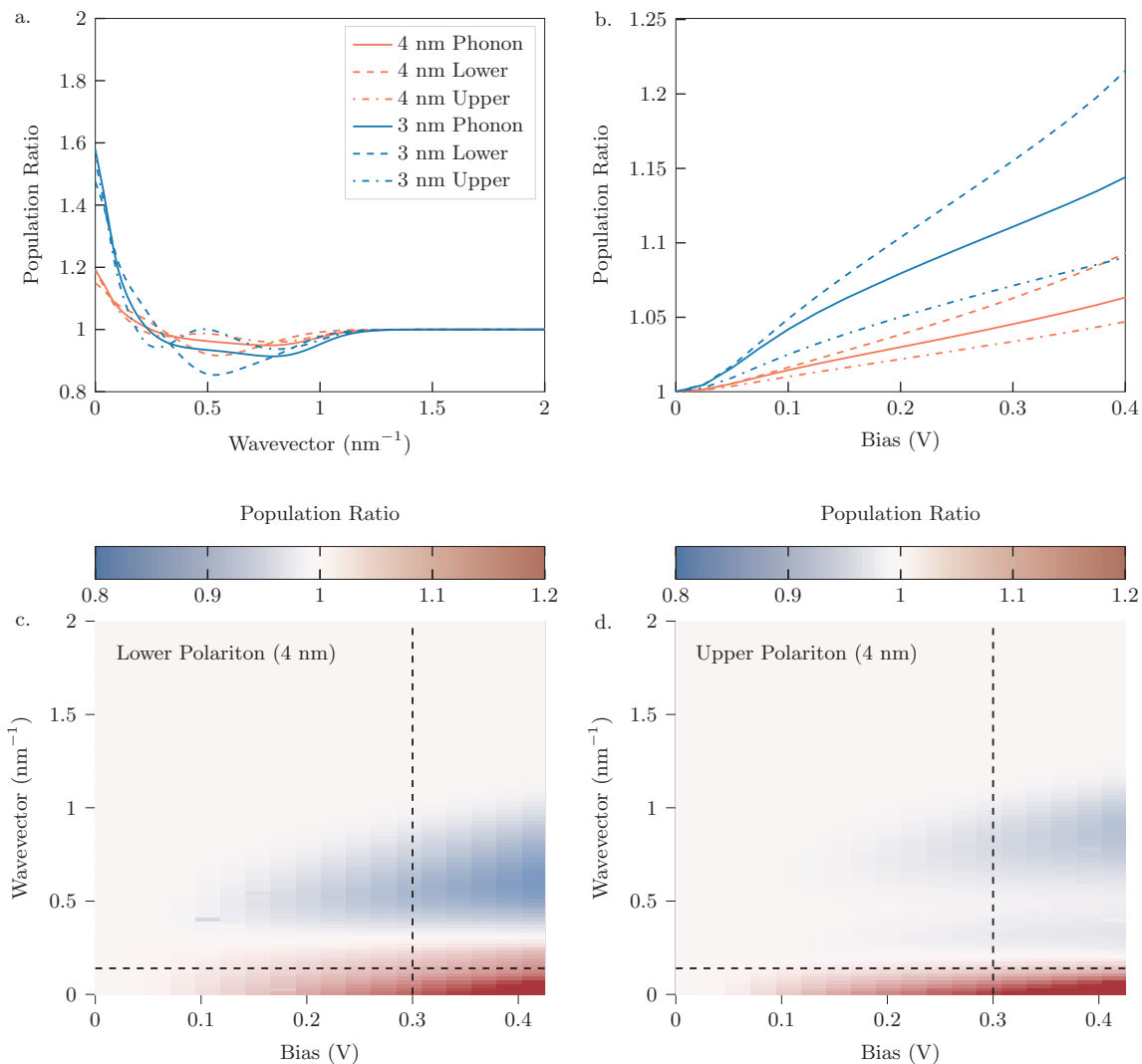
FIG. 5. Figure of Merit $F^{(1)}$ from Eq. 6 evaluated for film thickness 3 nm (red) and 4 nm (blue). Results for localized phonons in the absence of longitudinal-transverse coupling are shown with solid lines, while those for the coupled solutions are shown with dashed (lower polariton) and dash-dotted (upper polariton) lines respectively. Data is presented a. As a function of wavevector for fixed bias 0.3 V. b. As a function of bias for fixed wavevector 0.02 nm^{-1} c. As a colormap of both variables for the 4 nm sample's lower polariton branch. d. As a colormap of both variables for the 4 nm sample's upper polariton branch. Dashed lines in c. and d. indicate the cuts plotted in a. and b.

the distributional greater Green's function

$$\Gamma_i\left(\mathbf{k}, E\right) = \frac{1}{2\pi\hbar} G_{ii}^{>}\left(\mathbf{k}, E\right) \Sigma_{ii}^{<}\left(\mathbf{k}, E\right), \qquad (12)$$

where quantities indexed $\mathbf{k}$ are those defined in Section. II B at $\mathbf{k}' = \mathbf{k}$. For the device under consideration in Fig. 1, the in-plane dimension is translationally invariant, which allows us to pass to Fourier space in the in-plane wavevector. The out-of-plane growth axis cannot be transformed in this way due to the spatial inhomogeneity introduced by the multilayer structure. Along this out-of-plane dimension quantities are defined on a discrete spatial mesh, rather than in the continuous coordinate system used in Section. II B. Subscript indices i

refer to elements in the corresponding real-space matrix or vector, so $G_{ii}^{>}$ refers to the ith element of the greater Green's function diagonal, and Γ_i refers to the ith element of the vector Γ. Note that in Eq. 12 we make a diagonal self-energy approximation $\Sigma_{ij}^{<} = 0$ $(i \neq j)$. Insodoing we implicitly assume the scattering process is local, this means our results give a conservative lower bound on the true scattering rate in an LTP based emitter. In a full model, electrons would be able to interact with phonons in the barrier while outside it [42–44].

The out-scattering rate has contributions from both emission and absorption of longitudinal excitations through self-energy $\Sigma^{<}$. The diagonal component of the
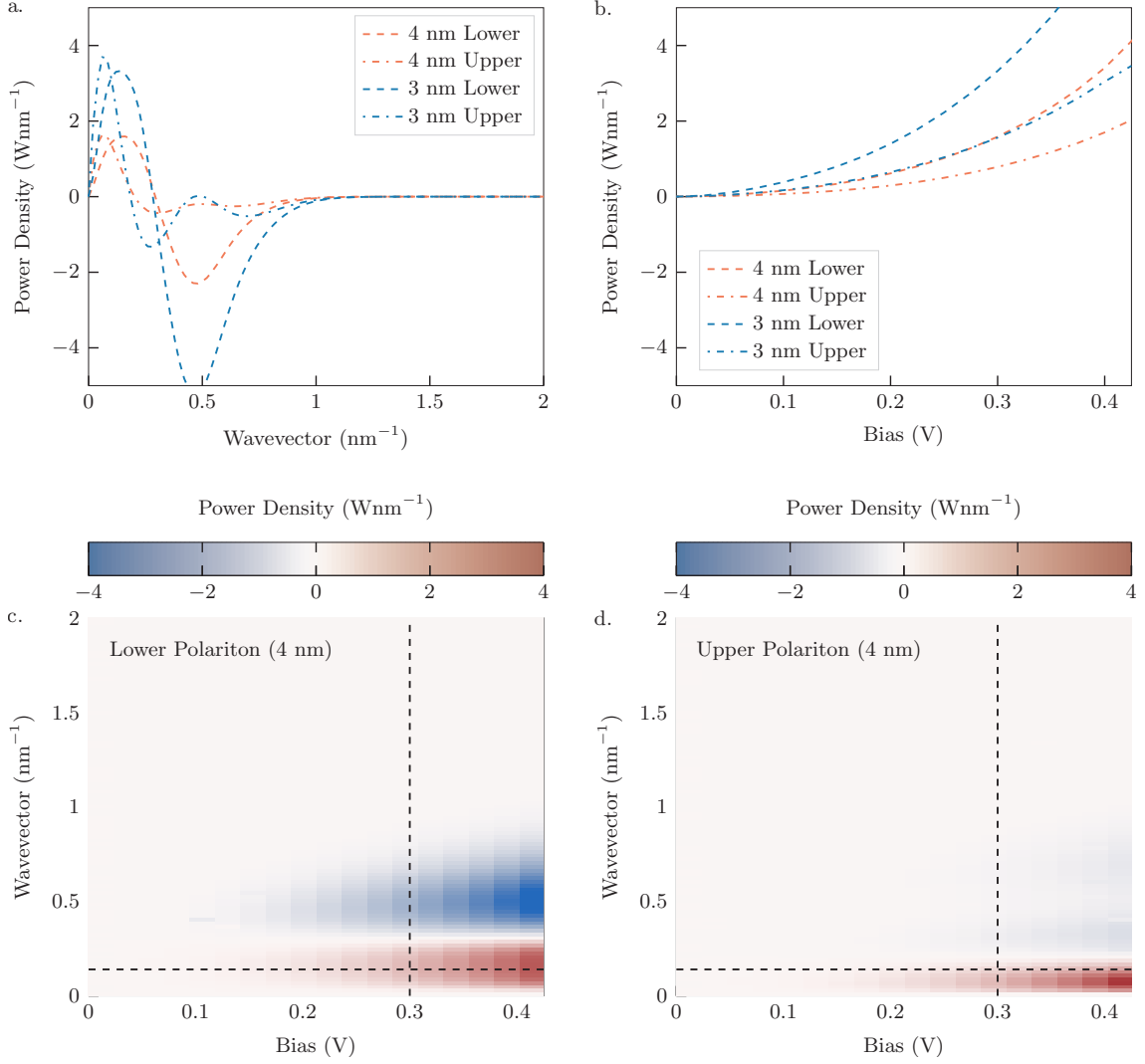
FIG. 6. Integrand for Figure of Merit $F^{(2)}$ from Eq. 10 evaluated for film thicknesses 3 nm (red) and 4 nm (blue). Dashed and dash-dotted lines correspond to lower and upper polaritons, respectively. Data is presented as a. A function of wavevector for fixed bias 0.3 V. b. As a function of bias for fixed wavevector 0.02 nm^{-1}. c. As a colormap of both variables for the 4 nm sample's lower polariton branch. d. As a colormap of both variables for the 4 nm sample's upper polariton branch. Dashed lines in c. and d. indicate the cuts plotted in a. and b.

lesser self-energy can be written in general form as

$$\Sigma_{ii}^{<}(\mathbf{k}, E) = \sum_{\mathbf{q}} \mathcal{K}_i(\mathbf{q})$$

$$\left\{ N^0(\omega_{\mathbf{q}}, T_a) G_{ii}^{<}(\mathbf{k} - \mathbf{q}, E - \hbar\omega_{\mathbf{q}}) \right.$$

$$\left. + \left[N^0(\omega_{\mathbf{q}}, T_a) + 1 \right] G_{ii}^{<}(\mathbf{k} - \mathbf{q}, E + \hbar\omega_{\mathbf{q}}) \right\}, \quad (13)$$

where $\omega_{\mathbf{q}}$ is the frequency of the reservoir mode with in-plane wavevector $\mathbf{q}$, $\mathcal{K}$ is a process-specific scattering kernel and i labels the discrete spatial indices in the mesh. The equilibrium Bose-Einstein population is given by N^0. Calculation of the self-energies for localized phonons is

discussed in Appendix C.

To compute the figures of merit defined in Section. II C it is necessary to derive the rates of in-scattering and out-scattering from the reservoir, which can be written taking appropriate components of $\Sigma^{<}$ and summing over the electronic degrees of freedom

$$\Gamma_{\mathbf{q}}^{\text{in}} = \sum_{i} \mathcal{K}_i(\mathbf{q}) \sum_{\mathbf{k}, E} G_{ii}^{>}(\mathbf{k}, E) G_{ii}^{<}(\mathbf{k} - \mathbf{q}, E - \hbar\omega_{\mathbf{q}}),$$

$$(14a)$$

$$\Gamma_{\mathbf{q}}^{\text{out}} = \sum_{i} \mathcal{K}_i(\mathbf{q}) \sum_{\mathbf{k}, E} G_{ii}^{>}(\mathbf{k}, E) G_{ii}^{<}(\mathbf{k} - \mathbf{q}, E + \hbar\omega_{\mathbf{q}}),$$
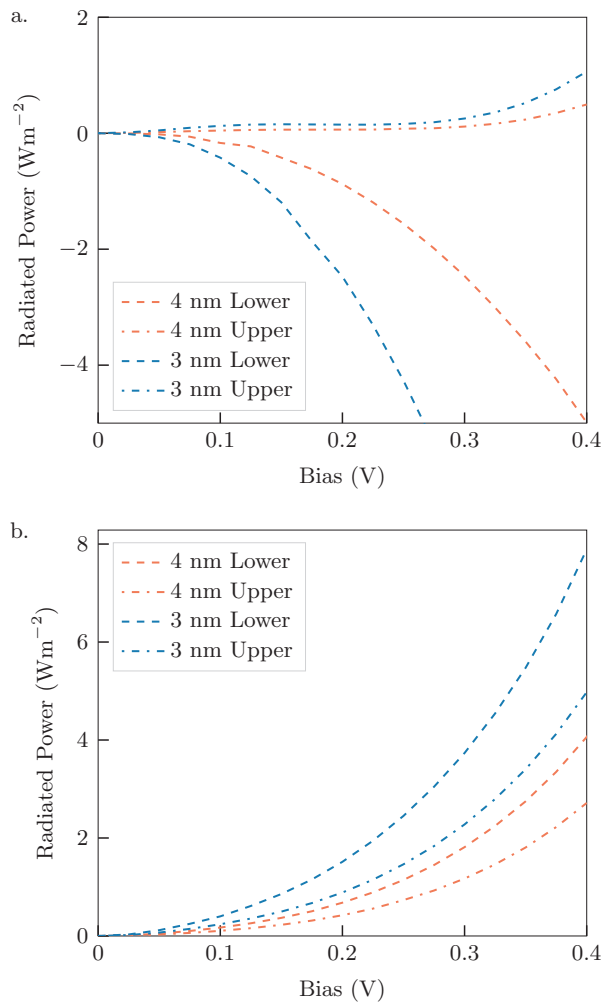
$$(14b)$$

FIG. 7. Figure of merit $F^{(2)}$ evaluated for film thicknesses 3 nm (red) and 4 nm (blue) integrated over the in-plane wavevector of the outgoing LTP for a. All wavevectors. b. Wavevectors less than 0.25 nm^{-1}. Dashed and dash-dotted lines correspond to lower and upper polaritons, respectively.

where terms labeled "in" refer to processes that increase the number of excitations in the reservoir, and those labeled "out" refer to those that reduce the number of excitations in the reservoir.

As LTPs couple to electrons through their longitudinal phonon component, the theory used to calculate the localized phonon scattering rates in Appendix C can be extended to describe LTP scattering by substituting the localized phonon energies with the coupled eigenfrequencies $\omega_{\mathbf{q}}^{\pm}$, and weighting the contribution from each po-

lariton branch by its longitudinal fraction

$$\Sigma_{ii}^{<\,\pm}(\mathbf{k}, E) = \sum_{\mathbf{q}} |U_{\mathbf{q}}|^2 |u(z_i)|^2 |\mathrm{X}_{\mathbf{q}}^{\pm}|^2$$

$$\left\{ N\left(\omega_{\mathbf{q}}^{\pm}, T_a\right) G_{ii}^{<}\left(\mathbf{k} - \mathbf{q}, E - \hbar\omega_{\mathbf{q}}^{\pm}\right) \right.$$

$$\left. \left[N\left(\omega_{\mathbf{q}}^{\pm}, T_a\right) + 1 \right] G_{ii}^{<}\left(\mathbf{k} - \mathbf{q}, E + \hbar\omega_{\mathbf{q}}^{\pm}\right) \right\}, \quad (15)$$

where u is the out-of-plane LO phonon wavefunction and $U_{\mathbf{q}}$ is the matrix element for the scattering process, both defined in Appendix C. The longitudinal phonon Hopfield coefficient $\mathrm{X}_{\mathbf{q}}^{\pm}$ weights the sum by the longitudinal content of the mode, reducing the self-energy in regions of the dispersion where the LTP is photon-like. We can identify the scattering kernel for the LTP as

$$\mathcal{K}_i^{\pm}(\mathbf{q}) = |U_{\mathbf{q}}|^2 |u(z_i)|^2 |\mathrm{X}_{\mathbf{q}}^{\pm}|^2, \quad (16)$$

which is just that for the localized phonon $\mathcal{K}_i^{\mathrm{L}}$ calculated in Appendix A, weighted by the squared longitudinal Hopfield coefficient.

We compute the first figure of merit $F^{(1)}$ using Eq. 6. The steady-state population in the denominator is taken to be that of the device operating under ambient conditions at 300 K. Results are shown as a function of LTP wavevector for fixed applied bias 0.3 V in Figure.5a for two barrier thicknesses 3 nm and 4 nm. Solid lines show the figure of merit in the absence of longitudinal-transverse coupling, where the scattering kernel is $\mathcal{K}_i^{\mathrm{L}}$ and the mode frequency is $\omega_{\mathbf{q}}^{\mathrm{L}}$. The strongest enhancements are observed for wavevectors less than 0.25 nm^{-1}. In this region the figure of merit converges for all branches as the scattering rate significantly exceeds the damping rate. The dominant scattering channel varies as the wavevector increases, a result of the unique modal dispersion Fig. 2 as otherwise reduced phonon content in the polaritonic case would result in a lower scattering rate. The peak at small momentum is a significant new result and suggests that outcoupling of these excitations may be possible using grating features in excess of 100nm. The strongest enhancements are observed for wavevectors less than 0.25 nm^{-1}. In this region the figure of merit converges for all branches as the scattering rate significantly exceeds the damping rate. The dominant scattering channel varies as the wavevector increases, a result of the unique modal dispersion in Fig. 2, as otherwise reduced phonon content in the polaritonic case would result in a lower scattering rate. The peak at small momentum is a significant new result and suggests that outcoupling of these excitations may be possible using grating features in excess of 100nm. Note, however, that the theory we are using is only accurate at wavevectors large enough for the ENZ approximation of the SPhP dispersion to be correct, and it breaks down too close to zero wavevector. All wavevectors in Fig. 5 are large compared to the free-photon at this wavelength ($\approx 1\ \mu\mathrm{m}^{-1}$). Thinner films exhibit nonlocality closer to the lightline,

and will be most promising for optoelectronic applications [6], allowing LTPs to be outcoupled to the far-field with larger grating features.

The same figure of merit at fixed in-plane wavevector 0.14 nm^{-1} is shown in Fig. 5b The quantity increases in all cases as a function of applied bias at this wavevector. At larger bias more electrons are in states that are able to penetrate the barrier (Figure. 4), leading to an increase in the steady-state LTP population. Compared to previous calculations of this quantity [8] we predict weaker growth in $F^{(1)}$ as the bias increases. This is a consequence of the simple free-electron model presented previously, in which larger bias led to a linear increase in electron tunneling through the barrier, and a nearly linear increase in scattering. In the regime explored in this Paper, while increasing bias does increase the current (see Fig. 4), most electrons still do not occupy states capable of tunneling through the barrier, rendering the figure of merit less sensitive to changes in the applied bias. An electron with 0.4 V energy has an effective temperature $\approx$ 4000 K, so despite the low bias considered here we nonetheless extend beyond the parameter range considered in [8]. This reduction could be countered by considering a quantum well active region, or a device such as a resonant tunneling diode where phonon-electron overlap could be greatly enhanced [41]. Note that at zero bias the scattering process deactivates and the thermal population is returned. At this point the device is thermalized. The electronic population continues to increase with bias, and in an experiment could be more readily observed by application of large bias in excess of that explored in this work. In Fig. 5c and d we plot $F^{(1)}$ as a function of both wavevector and bias for the lower and upper polariton branches of the 4 nm film respectively. In both cases the strongest emission occurs close to zero wavevector, increasing monotonously with the applied bias. Emission at larger wavevectors is suppressed as the bias increases, at these wavevectors charge carriers absorb phonons more quickly than they emit reducing the steady-state population. Interaction with the electron gas shifts excitations in the reservoir from large wavevector states to ones closer to the light-line.

The second figure of merit describes the photonic power spectral density generated within the device. Figure. 6a shows the integrand

$$I_{\mathbf{q}} = \frac{|\mathbf{q}|}{2\pi}\hbar\omega_{\mathbf{q}}\xi_{\mathbf{q}}\Gamma_{\mathbf{q}}^{\text{net}}, \qquad (17)$$

which describes a power density Wnm^{-1}. Results for the two film thicknesses as a function of excitation wavevector for fixed bias 0.3 V are shown in Fig. 6. Contrasting Fig. 5 the pure phonon result is omitted as it has no photonic component. At zero wavevector power emitted into both branches tends to zero, which occurs because the power density is linearly proportional to the modal wavevector. In Figure. 6b we plot the same quantity as a function of bias at small wavevector 0.14 nm^{-1}. At zero bias the system is thermalized and there is no net

emission. When the bias is increased, emission grows as more electrons penetrate the barrier with sufficient energy to emit LTPs. Emission is stronger into the lower polariton branch at this wavevector, a consequence of it's larger photonic component (Fig. 2).

In the lower panels of Fig. 6 c and d we plot $I_{\mathbf{q}}$ as a function of both bias and wavevector for the lower and upper branch in the 4 nm film respectively. As the bias increases low-wavevector photon emission increases, particularly for the upper polariton branch. As for $F^{(1)}$, the integrand $I_{\mathbf{q}}$ becomes negative at larger wavevector as the electron gas absorbs phonons from the reservoir. Finally we plot the integrated photonic power density $F^{(2)}$. Results are shown in Fig. 7a. The resulting value is positive for the upper-polariton branch for both thicknesses, but negative for the lower-polariton branch. This can be understood referring to Fig. 6c-d. At larger wavevectors the spectral density becomes negative in both polariton branches, but this is enhanced in the lower branch. When integrating over the in-plane wavevector of the phonon this large wavevector absorption is enhanced compared to the small-wavevector emission. Taken over all wavevectors phonon absorption reduces the total photonic power generated compared to the thermal case.

Fortunately, a real device will not outcouple LTPs from every in-plane wavevector. By restricting the outcoupling mechanism to in-plane wavevectors less than 0.25 nm^{-1} the photonic power density emitted from the device can be greatly increased as shown in Fig. 7b. Longitudinal-transverse polaritons generated in this emission region could be addressed using a grating or similar. As previously mentioned, our simulations are restricted to voltages below 0.4V due to the instability of the outer Schrödinger-Poisson iterations. While results at larger bias will be investigated in furture works, we can predict that, as the average electronic energy increases, the impact of the barrier on electron motion will decrease and the superlinear power increases with bias shown in Fig. 6b and Fig. 7b will cease. In this regime we expect the device performances will more closely match the results obtained with a free electron model in Ref. [8], before eventually saturating as the electronic saturation velocity is approached.

IV. CONCLUSION

Generation of narrowband mid-infrared light is an active problem in mid-infrared optoelectronics [27]. This Paper presented a novel method of designing narrowband mid-infrared electrical emitters without the need for complex band-engineering. Harnessing the natural Coulomb interaction between electrons and the optical phonon modes of polar crystal lattices we demonstrated the generation of photons, capable of coupling to the far-field through intermediary longitudinal-transverse polaritons (LTPs). Our results confirm previous works proposing LTP systems as electrically driven mid-infrared emit-

ters [8], extending to account for the electronic structure of the device in addition to the LTP dynamics. Although we only study small applied bias in this work, we show significant changes in the LTP population which could be further increased in a device operating under larger voltage. Our theory can be enhanced using established techniques to calculate the charge distribution in the presence of scattering in the self-consistent Born approximation [39], and to calculate gain in such a system [45, 46]. Our initial results can be built on to design experimental LTP-based emitters, and the presented figures of merit can be used to assess their likely performance. They underline that this new channel of emission could underpin a new generation of mid-infrared optoelectronic emitters powered by longitudinal-transverse polaritons.

V. ACKNOWLEDGEMENTS

Appendix A: Derivation of Longitudinal-Transverse Polariton dispersion

The polar layer is characterized by an isotropic dielectric function

$$\epsilon(\omega) = \epsilon_\infty \frac{\omega_L^2 - \omega(\omega + i\gamma)}{\omega_T^2 - \omega(\omega + i\gamma)}, \qquad (A1)$$

where ϵ_∞ is the high-frequency dielectric constant, ω_T (ω_L) is the zone-center transverse (longitudinal) optical phonon frequency and γ is the damping constant, assumed here to be independent from frequency and polarization. In the local-response approximation the film supports long-range and short-range SPhPs as a consequence of hybridization between the charge oscillations at each of its interfaces. In the thin-film limit the ENZ SPhP mode dispersion as a function of the in-plane wavevector $\mathbf{q}$ is

$$\omega_{\mathbf{q}}^{SP} = \sqrt{\frac{\omega_L^2 + |\mathbf{q}| x \omega_T^2}{1 + |\mathbf{q}| x}}, \qquad (A2)$$

where $x = d\epsilon_c / 2\epsilon_\infty$ [7], while localized LO phonons supported by the film satisfy the quadratic approximation [6]

$$\omega_{\mathbf{q},n}^{L} = \sqrt{\omega_L^2 - \beta_L^2 (|\mathbf{q}|^2 + \zeta_n^2)}, \qquad (A3)$$

in which β_L is the phonon velocity in the limit of quadratic dispersion, and $\zeta_n = \frac{2(n-1)\pi}{d}$, ($n \in \mathbb{Z}^+$) is the quantized out-of-plane wavevector of the localized mode. In this Paper we focus on the case where only one $n = 1$ localized phonon is near resonance with the ENZ, dropping the index from $\omega_{\mathbf{q}}^L$ and ζ. The LTP spectrum in this system is well characterized using a two-oscillator Hopfield model, given in the rotating-wave approximation by

$$\hat{\mathcal{H}}^{LTP} = \hbar \sum_{\mathbf{q}} \left\{ \omega_{\mathbf{q}}^{SP} \hat{a}_{\mathbf{q}}^\dagger \hat{a}_{\mathbf{q}} + \omega_{\mathbf{q}}^{L} \hat{b}_{\mathbf{q}}^\dagger \hat{b}_{\mathbf{q}} + \Omega_{\mathbf{q}} \left(\hat{a}_{\mathbf{q}} \hat{b}_{\mathbf{q}}^\dagger + \hat{a}_{\mathbf{q}}^\dagger \hat{b}_{\mathbf{q}} \right) \right\}, \qquad (A4)$$

where $\hat{a}_{\mathbf{q}}^\dagger$ and $\hat{b}_{\mathbf{q}}^\dagger$ are bosonic creation operators for the ENZ SPhP and the $n = 1$ localized LO phonon mode, and the Rabi frequency, quantifying the longitudinal-transverse coupling, is given by [7]

$$|\Omega_{\mathbf{q}}|^2 = \frac{2\beta_L^2}{d^2} \frac{(\omega_L - \omega_{\mathbf{q}}^{SP})(\omega_L + \omega_{\mathbf{q}}^{SP})}{\omega_{\mathbf{q}}^{SP} \omega_{\mathbf{q}}^{L}}. \qquad (A5)$$

The frequencies of the two resulting LTP branches, which we will refer to as lower (-) and upper (+) polaritons, can be found by calculating the eigenvalues of the Hamiltonian in Eq. A4

$$\omega_{\mathbf{q}}^{\pm} = \left[\frac{\omega_{\mathbf{q}}^{SP\,2} + \omega_{\mathbf{q}}^{L\,2}}{2} \right.$$
$$\left. \pm \frac{1}{2} \sqrt{\left[\omega_{\mathbf{q}}^{SP\,2} - \omega_{\mathbf{q}}^{L\,2} \right]^2 + 4\Omega_{\mathbf{q}}^2 \omega_{\mathbf{q}}^{SP} \omega_{\mathbf{q}}^{L}} \right]^{1/2}. \qquad (A6)$$

Polaritons can be understood considering the eigenvectors of the Hamiltonian, which describe the mixture of underlying modes in the coupled excitation. The fraction of each bare mode in a given polariton is termed the Hopfield coefficient, which is given by the square of the relevant element of the eigenvector. The Hopfield coefficient describing the longitudinal content of each polariton branch can be written as

$$|\mathrm{X}_{\mathbf{q}}^{\pm}|^2 = \frac{1}{2}\left[1 \pm \frac{\Delta\omega_{\mathbf{q}}}{\sqrt{\Delta\omega_{\mathbf{q}}^2 + 4\Omega_{\mathbf{q}}^2}}\right], \qquad (\mathrm{A7})$$

while the one describing the transverse content is $|\mathrm{Y}_{\mathbf{q}}^{\pm}|^2 = 1 - |\mathrm{X}_{\mathbf{q}}^{\pm}|^2$. Here $\Delta\omega_{\mathbf{q}} = \omega_{\mathbf{q}}^{\mathrm{SP}} - \omega_{\mathbf{q}}^{\mathrm{L}}$ is the frequency difference between the bare mode frequencies.

Appendix B: Self-consistent Charge Calculation

This Appendix summarizes the ingredients necessary to calculate the self-consistent electronic charge distribution in the device in the absence of scattering between energy and momentum states within the device.

1. Calculation of the Closed System Hamiltonian

The dynamic retarded Green's functions are calculated using Eq. 3. This equation requires prior knowledge of the Hamiltonian of the closed electronic system $\mathcal{H}_0$ and the self-energies associated with coupling to the external leads of the device Σ^R.
We consider transport in a system homogeneous in the xy plane, but inhomogeneous along the transport direction z. The closed system, defined in the absence of coupling to the leads, is described by one-dimensional Hamiltonian

$$\mathcal{H}_0 = -\frac{\hbar^2}{2}\left[\frac{\mathrm{d}}{\mathrm{d}z}\frac{1}{m^*(z)}\frac{\mathrm{d}}{\mathrm{d}z}\right] + V_{k_\parallel}(z) + \frac{\hbar^2 k_\parallel^2}{2m_{\mathrm{L}}^*}, \quad (\mathrm{B1})$$

where $m^*(z)$ is the effective mass at position z, m_{L}^* is the effective mass in the semi-infinite left-lead, $k_\parallel$ is the in-plane wavevector and

$$V_{k_\parallel}(z) = V(z) + \frac{\hbar^2 k_\parallel^2}{2m_{\mathrm{L}}^*}\left[\frac{m_{\mathrm{L}}^*}{m^*(z)} - 1\right], \qquad (\mathrm{B2})$$

where $V(z)$ is the on-site potential due to conduction band offsets and the Hartree potential.
We discretize the Hamiltonian in the nearest-neighbour tight-binding approximation, whose matrix elements are given by

$$\langle n, \mathbf{k}_\parallel | \mathcal{H}_c | n', \mathbf{k}_\parallel \rangle = D_n(\mathbf{k})\,\delta_{n,n'} - t_{n,n'}(\mathbf{k}_\parallel)\,\delta_{n,n'\pm1}, \tag{B3}$$

where

$$D_i(\mathbf{k}_\parallel) = \frac{\hbar^2}{2\Delta}\left[\frac{1}{m^+} + \frac{1}{m^-}\right] + V_i(\mathbf{k}_\parallel), \qquad (\mathrm{B4a})$$

$$t_{i,j} = \frac{\hbar^2}{(m_i + m_j)\,\Delta^2}. \qquad (\mathrm{B4b})$$

Note that here $m^+ = (m_i + m_{i+1})/2$ and $m^- = (m_i + m_{i-1})/2$, m_i is the effective mass at mesh site i and Δ is the (uniform) mesh-spacing.

2. Self Energies for Coupling to Semi-Infinite Leads

Coupling to external reservoirs is accounted for considering single-band electrons in a semi-infinite lead. At the device edge the discrete system represented by the closed system Hamiltonian needs to be truncated. The relevant component of the Schrodinger equation is of form

$$E\eta_1 = -t_{0,1}\eta_0 + D_1\eta_1 - t_{1,2}\eta_2, \qquad (\mathrm{B5})$$

where η_i is the wavefunction at lattice site i, E is the energy and $t_{i,j}, D_i$ are the components of the Hamiltonian defined in Eq. B3- B4. To truncate the system to the region $i > 1$ it is necessary to eliminate η_0 from this equation. As the *retarded* Green's function describes the response to an excitation generated within the device, the appropriate boundary condition is

$$\eta_0 = \eta_1 e^{ik_1\Delta}, \qquad (\mathrm{B6})$$

where Δ is the grid spacing at the device edge and k_1 is the electron wavevector at lattice site 1. With this we can write

$$E\eta_1 = -t_{0,1}e^{ik_1 a}\eta_1 + D_1\eta_1 - t_{1,2}\eta_2. \qquad (\mathrm{B7})$$

Note that in the contact D encodes the boundary condition, and therefore the chemical potential in each lead.
A similar result can be found at the rightmost contact, leading to retarded self-energy

$$\Sigma^R = \begin{pmatrix} -t_{0,1}e^{ik_1\Delta} & 0 & \dots & & 0 \\ 0 & \ddots & & & \vdots \\ \vdots & & \ddots & & \\ 0 & 0 & \dots & & -t_{N,N+1}e^{ik_N\Delta} \end{pmatrix}. \qquad (\mathrm{B8})$$

Finally the kinetic equation Eq. 4 can be used to calculate the distributional Green's function $G^<$, which in turn gives the electron density through Eq. 2. Note that we achieve swift convergence in carrier density throughout the device by employing an adaptive Gauss-Kronrod integration scheme to discretize the energy space. Note that the distributional self-energies in the leads necessary to calculate the lesser and greater Green's functions can be calculated from the level broadening associated with the retarded Green's function, and the lead Fermi function as

$$\Sigma^< = if\Gamma. \qquad (\mathrm{B9})$$

3. Calculation of the Hartree potential

Once the charge distribution n has been calculated using Eq. 2 the Hartree potential throughout the device can be found solving the Poisson equation subject to Neumann boundary conditions. For a system homogeneous in the xy-plane but inhomogeneous in z this is given by

$$\frac{\mathrm{d}}{\mathrm{d}z}\left[\epsilon\left(z\right)\frac{\mathrm{d}\phi\left(z\right)}{\mathrm{d}z}\right] + q\left[N_\mathrm{D}^+\left(z\right) - N_\mathrm{A}^-\left(z\right) - n\left(z\right)\right] = F\left(\phi\right),$$
(B10)

where ϵ is the static dielectric constant, N_D^- $\left[N_\mathrm{A}^+\right]$ is the background donor (acceptor) density and F is the residuum to be minimized.

Discretising on a nearest-neighbour tight-binding basis we can recover the non-zero matrix elements of the Poisson operator $\mathbf{A}$

$$A_{ii} = -\frac{\epsilon^+ + \epsilon^-}{\Delta^2} + q\left[N_{\mathrm{D},i}^+ - N_{\mathrm{A},i}^- - n_i\right],$$
(B11a)

$$A_{ij} = -\delta_{j,i\pm1}\frac{\epsilon_i + \epsilon_j}{2\Delta^2},$$
(B11b)

where $\epsilon^\pm = (\epsilon_i + \epsilon_{i\pm1})/2$, ϵ_i is the static dielectric constant at mesh point i and Δ is the (uniform) mesh spacing. In matrix form we can write Eq. B10 as

$$F\left[\phi\right] = \mathbf{A}\left[\phi\right] + q\left[N_\mathrm{D}^+ - N_\mathrm{A}^- - n\right].$$
(B12)

Using the current carrier density n the quasi-Fermi level throughout the device is calculated by minimising

$$n_i = N_\mathrm{C}\mathcal{F}_{1/2}\left[\frac{E_{\mathrm{F},i} - E_{\mathrm{C},i} + q\phi_i}{k_\mathrm{B}T}\right].$$
(B13)

in which T is the device temperature, N_C is the conduction band density of states, $\mathcal{F}_m$ is a Fermi-Dirac integral of order m, $E_{\mathrm{C},i}$ is the conduction band offset at mesh site i, $E_{\mathrm{F},i}$ is the quasi-Fermi level at mesh site i and n_i is the free charge calculated by the NEGF step. Changes in the free charge can then be *predicted* for modest changes in potential $\delta\phi$

$$n_{\mathrm{pred},i}\left(\delta\phi_i\right) = N_\mathrm{C}\mathcal{F}_{1/2}\left[\frac{E_{\mathrm{F},i} - E_{\mathrm{C},i} + q\left(\phi_i + \delta\phi_i\right)}{k_\mathrm{B}T}\right].$$
(B14)

At each iteration we solve for the *change* in the electric potential $\delta\phi = \phi_\mathrm{out} - \phi_\mathrm{in}$ where ϕ_in is the electric potential at the previous outer iteration

$$F_\mathrm{pred}\left[\delta\phi + \phi_\mathrm{in}\right] = \mathbf{A}\left[\delta\phi + \phi_\mathrm{in}\right]$$
$$+ q\left[N_\mathrm{D}^+ - N_\mathrm{A}^- - n_\mathrm{pred}\left(\delta\phi\right)\right] = 0. \quad \text{(B15)}$$

using a pre-conditioned Gauss-Newton predictor corrector scheme. The Jacobian can be calculated analytically as

$$\mathbf{J}_{ij} = \frac{\mathrm{d}F_{\mathrm{pred},i}}{\mathrm{d}\delta\phi_j} = A_{ij} + \frac{\mathrm{d}n_{\mathrm{pred},i}}{\mathrm{d}\delta\phi_j},$$
(B16)

where the differential of the predictor charge density is given by

$$\frac{\mathrm{d}n_{\mathrm{pred},i}}{\mathrm{d}\delta\phi_j} = \delta_{i,j}\frac{q}{k_\mathrm{B}T}N_\mathrm{C}$$
$$\times \mathcal{F}_{-1/2}\left[\frac{E_{\mathrm{F},i} - E_{\mathrm{C},i} + q\left[\phi_i + \delta\phi_i\right]}{k_\mathrm{B}T}\right]. \quad \text{(B17)}$$

4. Self-Consistent Loop

Convergence of an iterative fixed point calculation $x = f\left(x\right)$ is strongly dependent on how the input quantity is computed between iterations. A linear procedure, in which the previous input and output are combined as a simple mixture

$$x^{k+1} = \left(1 - \alpha\right)x^k + \alpha f\left(x^k\right),$$
(B18)

takes the weighting α as a model parameter. Such a method can be unstable, particularly in systems with narrow electronic resonances unless the weighting α is sufficiently small. Minimising the weighting too strongly however leads to unnecessarily slow convergence. This Paper employs an adaptive Type I Anderson mixing procedure to calculate subsequent inputs. In a generic Anderson mixing method inputs are calculated as linear mixtures of the outputs of previous iterations

$$x^{k+1} = \sum_{j=0}^{m_k} \alpha_j f\left(x^{k-m_k+j}\right),$$
(B19)

in which m_k is the memory capacity of the mixing scheme and α_k the mixing weights satisfying $\sum_{j=0}^{m_k}\alpha_j = 1$. This can be recast as a minimisation

$$\text{minimize} \qquad ||\sum_{j=0}^{m_k}\alpha_j g\left(x^{k-m_k+j}\right)||_2^2, \quad \text{(B20a)}$$

$$\text{subject to} \qquad \sum_{j=0}^{m_k}\alpha_j = 1, \quad \text{(B20b)}$$

where $g\left(x\right) = x - f\left(x\right)$ is the residual of the fixed-point problem and the solution is the weight variable $\boldsymbol{\alpha}$. This Paper utilizes the stabilized Type I Anderson acceleration scheme outlined by Zhang *et al.* [47].

5. Stability of the numerical solver

The full iteration of Schrödinger and Poisson equations is nonlinear, and the iterative procedure can diverge, yielding a non-physical solution. This can be mitigated by choosing a good initial guess for the Hartree potential, ensuring that changes over the course of a single iteration are small. This is especially important in this Paper, which employs a predictor corrector scheme when solving

the Poisson equation. This approximates the Jacobian, by assuming explicitly that changes in the charge density are small [48]. Nonlinearities can also be mitigated utilizing a mixing algorithm to avoid rapid changes in the electric potential [49]. Another common source of error is the shift in electronic energy levels as the calculation progresses. If the electron density in Eq. 2 is calculated on a fixed grid the error estimates can fluctuate between iterations, leading again to a diverging result. We employ an adaptive integration mechanism and use a converged solution from a lower bias point as a starting solution to mitigate the issue and increase the range of parameters in which convergence is achieved. All these issues worsen as the bias across the system increases. As the leads are assumed to be held at equilibrium potential and density, larger biases increase the size of the area to be modeled in order to recover a physical solution [50]. Increasing the length of the modeling region enhances the effect of nonlinearities in the Schrödinger-Poisson iterations, making convergence increasingly difficult.

Appendix C: localized Phonon Self-Energies

With a self-consistent potential ϕ it is possible to calculate the greater and lesser Green's functions at all energies and wavevectors. As discussed in Sec. II B these Green's functions are related to correlation functions for electrons and unoccupied states in the conduction band. They can be utilized to calculate coupling between electronic states, mediated by an external reservoir. We can write the lesser self-energy for the reservoir to the lowest order perturbation as

$$\Sigma^{\lessgtr}(\mathbf{r}, \mathbf{r}'; E) = \int \mathcal{D}(\mathbf{r}, \mathbf{r}'; \hbar\omega) G^{\lessgtr}(\mathbf{r}, \mathbf{r}'; E \mp \hbar\omega) \, \mathrm{d}\hbar\omega, \tag{C1}$$

where $\mathcal{D}$ is a function describing the spatial correlation and energies of the scatterers

$$\mathcal{D}(\mathbf{r}, \mathbf{r}'; \hbar\omega) = \sum_{\mathbf{Q}} |U_{\mathbf{Q}}|^2$$
$$\left[e^{-i\mathbf{q}\cdot(\mathbf{r}-\mathbf{r}')} f_{q_z}^*(z) f_{q_z}(z') N^0(\omega_{\mathbf{Q}}, T) \delta(\omega - \omega_{\mathbf{Q}}) \right.$$
$$\left. + e^{i\mathbf{q}\cdot(\mathbf{r}-\mathbf{r}')} f_{q_z}(z) f_{q_z}^*(z') \left[N^0(\omega_{\mathbf{Q}}, T) + 1 \right] \delta(\omega + \omega_{\mathbf{Q}}) \right], \tag{C2}$$

in which $\mathbf{Q} = [\mathbf{q}, q_z]$ is the three-dimensional phonon wavevector, $U_{\mathbf{Q}}$ is the phonon potential felt by a single electron, f_q is the out-of-plane mode profile and the Bose-Einstein distribution function as defined in Eq. 7.

localized longitudinal optic phonon modes have a discrete out-of-plane wavevector space $\zeta_n = \frac{2(n-1)\pi}{d}$, ($n \in \mathbb{Z}^+$ where d is the thickness of the nonlocal layer. Each branch n follows dispersion Eq. A3. The quantized form

of the localized phonon potential is given for the lowest order branch $n = 1$ by [6]

$$\hat{\phi}_{\mathrm{L}}(\mathbf{r}, z) = \frac{1}{e} \sum_{\mathbf{q}} U_{\mathbf{q}} u(z) e^{i\mathbf{q}\cdot\mathbf{r}} \left[\hat{b}_{\mathbf{q}} + \hat{b}_{-\mathbf{q}}^{\dagger} \right]. \tag{C3}$$

Here $\hat{b}_{\mathbf{q}}^{\dagger}$ ($\hat{b}_{\mathbf{q}}$) are creation (annihilation) operators for a phonon of transverse momentum $\mathbf{q}$ in the discrete mode and

$$U_{\mathbf{q}} = \sqrt{\frac{e^2 \hbar \omega_{\mathbf{q}}^{\mathrm{L}}}{\epsilon_{\mathbf{q}}^{\rho} \mathrm{A} d}} \frac{\sqrt{|\mathbf{q}|^2 + \zeta^2}}{|\mathbf{q}|^2 + \zeta^2 + q_0^2}, \tag{C4}$$

where q_0 is the screening length, A is the quantization surface and the dispersive Fröhlich constant is given by

$$\epsilon_{\mathbf{q}}^{\rho} = \epsilon_0 \epsilon_{\infty} \frac{\left(\omega_{\mathbf{q}}^{\mathrm{L}}\right)^2}{\omega_{\mathrm{L}}^2 - \omega_{\mathrm{T}}^2}, \tag{C5}$$

and the out-of-plane mode function is given by

$$u(z) = \sin\left[\zeta(z - z_0 + d/2)\right]. \tag{C6}$$

in which z_0 defines the nonlocal layer center. The leading diagonal of the self-energies are given by

$$\Sigma_{ii}^{\lessgtr}(\mathbf{k}, E) = \sum_{\mathbf{l}} |U_{\mathbf{k}-\mathbf{l}}|^2 |u(z_i)|^2$$
$$\left\{ N^0(\omega_{\mathbf{k}-\mathbf{l}}^{\mathrm{L}}, T_a) G_{ii}^{\lessgtr}\left(\mathbf{l}, E \mp \hbar\omega_{\mathbf{k}-\mathbf{l}}^{\mathrm{L}}\right) \right.$$
$$\left. \left[N(\omega_{\mathbf{k}-\mathbf{l}}^{\mathrm{L}}, T_a) + 1 \right] G_{ii}^{\lessgtr}\left(\mathbf{l}, E \pm \hbar\omega_{\mathbf{k}-\mathbf{l}}^{\mathrm{L}}\right) \right\} \tag{C7}$$

where the subscripts indicate elements of the spatial Green's function or self-energy matrices and we changed wavevector to $\mathbf{l} = \mathbf{k} - \mathbf{q}$. Using Eq. C7 we can define the kernel for localized phonon scattering as

$$\mathcal{K}_i^{\mathrm{L}}(\mathbf{q}) = |U_{\mathbf{q}}|^2 |u(z_i)|^2. \tag{C8}$$

On passing to an integral over wavevector we recover

$$\Sigma_{ii}^{\lessgtr}(\mathbf{k}, E) = \frac{e^2 \hbar}{2\pi d} \int l \, \mathrm{d}l$$
$$\frac{\omega_{\mathbf{k}-\mathbf{l}}^{\mathrm{L}}}{\epsilon_{\mathbf{k}-\mathbf{l}}^{\rho}} \mathcal{G}(l, k, z_i) \left\{ N^0(\omega_{\mathbf{k}-\mathbf{l}}^{\mathrm{L}}, T_a) G^{\lessgtr}\left(\mathbf{l}, E \mp \hbar\omega_{\mathbf{k}-\mathbf{l}}^{\mathrm{L}}\right) \right.$$
$$\left. \left[N^0(\omega_{\mathbf{k}-\mathbf{l}}^{\mathrm{L}}, T_a) + 1 \right] G^{\lessgtr}\left(\mathbf{l}, E \pm \hbar\omega_{\mathbf{k}-\mathbf{l}}^{\mathrm{L}}\right) \right\} \tag{C9}$$

where we applied

$$\int_0^{2\pi} \mathrm{d}\theta \frac{k^2 + l^2 + 2kl\cos\theta + \zeta^2}{\left[k^2 + l^2 + 2kl\cos\theta + \zeta^2 + q_0^2\right]^2}$$
$$= 2\pi \left[\frac{1}{\sqrt{\left(k^2 + l^2 + \zeta^2 + q_0^2\right)^2 - 4k^2l^2}} \right.$$
$$\left. - q_0^2 \frac{k^2 + l^2 + \zeta^2 + q_0^2}{\left[\left(k^2 + l^2 + \zeta^2 + q_0^2\right)^2 - 4k^2l^2\right]^{3/2}} \right], \tag{C10}$$

and defined the following

$$\mathcal{G}\left(l,k,z\right) = u^2\left(z\right)$$

$$\times \left[\frac{1}{\sqrt{\left(k^2+l^2+\zeta^2+q_0^2\right)^2 - 4k^2l^2}} \right.$$

$$\left. - q_0^2 \frac{k^2+l^2+\zeta^2+q_0^2}{\left[\left(k^2+l^2+\zeta^2+q_0^2\right)^2 - 4k^2l^2\right]^{3/2}} \right]. \quad \text{(C11)}$$

[1] B. K. Ridley, W. J. Schaff, and L. F. Eastman, Journal of Applied Physics **96**, 1499 (2004).

[2] J. B. Khurgin, S. Bajaj, and S. Rajan, Applied Physics Express **9**, 094101 (2016), publisher: IOP Publishing.

[3] C. R. Gubbin and S. De Liberato, Physical Review X **10**, 021027 (2020), publisher: American Physical Society.

[4] C. R. Gubbin and S. De Liberato, Physical Review B **102**, 201302 (2020), publisher: American Physical Society.

[5] C. R. Gubbin and S. De Liberato, Physical Review B **102**, 235301 (2020), publisher: American Physical Society.

[6] C. R. Gubbin and S. De Liberato, The Journal of Chemical Physics **156**, 024111 (2022).

[7] C. R. Gubbin and S. De Liberato, Physical Review Applied **17**, 014037 (2022), publisher: American Physical Society.

[8] C. R. Gubbin and S. D. Liberato, Nanophotonics **12**, 2849 (2023), publisher: De Gruyter.

[9] J. D. Caldwell, L. Lindsay, V. Giannini, I. Vurgaftman, T. L. Reinecke, S. A. Maier, and O. J. Glembocki, Nanophotonics **4**, 44 (2015), publisher: De Gruyter.

[10] R. Hillenbrand, T. Taubner, and F. Keilmann, Nature **418**, 159 (2002), number: 6894 Publisher: Nature Publishing Group.

[11] H. Sumikura, T. Wang, P. Li, A.-K. U. Michel, A. Heßler, L. Jung, M. Lewin, M. Wuttig, D. N. Chigrin, and T. Taubner, Nano Letters **19**, 2549 (2019).

[12] A. Woessner, M. B. Lundeberg, Y. Gao, A. Principi, P. Alonso-González, M. Carrega, K. Watanabe, T. Taniguchi, G. Vignale, M. Polini, J. Hone, R. Hillenbrand, and F. H. L. Koppens, Nature Materials **14**, 421 (2015).

[13] G. Lu, C. R. Gubbin, J. R. Nolen, T. Folland, M. J. Tadjer, S. De Liberato, and J. D. Caldwell, Nano Letters **21**, 1831 (2021), publisher: American Chemical Society.

[14] T. Wang, P. Li, B. Hauer, D. N. Chigrin, and T. Taubner, Nano Letters **13**, 5051 (2013), publisher: American Chemical Society.

[15] A. M. Dubrovkin, B. Qiang, T. Salim, D. Nam, N. I. Zheludev, and Q. J. Wang, Nature Communications **11**, 1863 (2020), number: 1 Publisher: Nature Publishing Group.

[16] C. R. Gubbin, S. A. Maier, and S. De Liberato, Physical Review B **95**, 035313 (2017), publisher: American Physical Society.

[17] C. T. Ellis, J. G. Tischler, O. J. Glembocki, F. J. Bezares, A. J. Giles, R. Kasica, L. Shirey, J. C. Owrutsky, D. N. Chigrin, and J. D. Caldwell, Scientific Reports **6**, 32959 (2016), number: 1 Publisher: Nature Publishing Group.

[18] C. R. Gubbin, F. Martini, A. Politi, S. A. Maier, and S. De Liberato, Physical Review Letters **116**, 246402 (2016), publisher: American Physical Society.

[19] C. R. Gubbin and S. De Liberato, ACS Photonics **4**, 1381 (2017), publisher: American Chemical Society.

[20] C. R. Gubbin and S. De Liberato, ACS Photonics **5**, 284 (2018), publisher: American Chemical Society.

[21] S. Kitade, A. Yamada, I. Morichika, K. Yabana, and S. Ashihara, ACS Photonics **8**, 152 (2021), publisher: American Chemical Society.

[22] I. Razdolski, N. C. Passler, C. R. Gubbin, C. J. Winta, R. Cernansky, F. Martini, A. Politi, S. A. Maier, M. Wolf, A. Paarmann, and S. De Liberato, Physical Review B **98**, 125425 (2018), publisher: American Physical Society.

[23] M. Autore, P. Li, I. Dolado, F. J. Alfaro-Mozaz, R. Esteban, A. Atxabal, F. Casanova, L. E. Hueso, P. Alonso-González, J. Aizpurua, A. Y. Nikitin, S. Vélez, and R. Hillenbrand, Light: Science & Applications **7**, 17172 (2018), number: 4 Publisher: Nature Publishing Group.

[24] R. Berte, C. R. Gubbin, V. D. Wheeler, A. J. Giles, V. Giannini, S. A. Maier, S. De Liberato, and J. D. Caldwell, ACS Photonics **5**, 2807 (2018), publisher: American Chemical Society.

[25] R. Kiessling, Y. Tong, A. J. Giles, S. Gewinner, W. Schöllkopf, J. D. Caldwell, M. Wolf, and A. Paarmann, ACS Photonics **6**, 3017 (2019), publisher: American Chemical Society.

[26] T. Taubner, D. Korobkin, Y. Urzhumov, G. Shvets, and R. Hillenbrand, Science (New York, N.Y.) **313**, 1595 (2006).

[27] C. R. Gubbin, S. De Liberato, and T. G. Folland, Journal of Applied Physics **131**, 030901 (2022).

[28] J.-J. Greffet, R. Carminati, K. Joulain, J.-P. Mulet, S. Mainguy, and Y. Chen, Nature **416**, 61 (2002), number: 6876 Publisher: Nature Publishing Group.

[29] J. A. Schuller, T. Taubner, and M. L. Brongersma, Nature Photonics **3**, 658 (2009), number: 11 Publisher: Nature Publishing Group.

[30] C. Arnold, F. Marquier, M. Garin, F. Pardo, S. Collin, N. Bardou, J.-L. Pelouard, and J.-J. Greffet, Physical Review B **86**, 035316 (2012), publisher: American Physical Society.

[31] C. R. Gubbin, R. Berte, M. A. Meeker, A. J. Giles, C. T. Ellis, J. G. Tischler, V. D. Wheeler, S. A. Maier, J. D. Caldwell, and S. De Liberato, Nature Communications **10**, 1682 (2019), number: 1 Publisher: Nature Publishing Group.

[32] L. Zhu, J. Cao, C. Song, B. Li, and H. Z, Nanotechnology

35, 315001 (2024).

[33] M. Büttiker, Physical Review Letters **57**, 1761 (1986), publisher: American Physical Society.

[34] J. Schwinger, Journal of Mathematical Physics **2**, 407 (1961).

[35] L. V. Keldysh, in *Selected Papers of Leonid V Keldysh* (WORLD SCIENTIFIC, 2023) pp. 47–55.

[36] C. Caroli, R. Combescot, P. Nozieres, and D. Saint-James, Journal of Physics C: Solid State Physics **4**, 916 (1971).

[37] Y. Meir and N. S. Wingreen, Physical Review Letters **68**, 2512 (1992), publisher: American Physical Society.

[38] S. Datta, Superlattices and Microstructures **28**, 253 (2000).

[39] R. Lake, G. Klimeck, R. C. Bowen, and D. Jovanovic, Journal of Applied Physics **81**, 7845 (1997).

[40] P. Lugli, P. Bordone, E. Molinari, H. Rucker, A. M. d. Paula, A. C. Maciel, J. F. Ryan, and M. Shayegan, Semiconductor Science and Technology **7**, B116 (1992).

[41] V. N. Do, P. Dollfus, and V. Lien Nguyen, Journal of Applied Physics **100**, 093705 (2006).

[42] T. Kubis and P. Vogl, Journal of Computational Electronics **6**, 183 (2007).

[43] B. K. Ridley, Journal of Physics C: Solid State Physics **15**, 5899 (1982).

[44] B. K. Ridley, Physical Review B **39**, 5282 (1989), publisher: American Physical Society.

[45] A. Wacker, Physical Review B **66**, 085326 (2002), publisher: American Physical Society.

[46] S.-C. Lee and A. Wacker, Physical Review B **66**, 245314 (2002), publisher: American Physical Society.

[47] J. Zhang, B. O'Donoghue, and S. Boyd, SIAM Journal on Optimization **30**, 3170 (2020).

[48] X. Gao, D. Mamaluy, E. Nielsen, R. W. Young, A. Shirkhorshidian, M. P. Lilly, N. C. Bishop, M. S. Carroll, and R. P. Muller, Journal of Applied Physics **115**, 133707 (2014), https://pubs.aip.org/aip/jap/article-pdf/doi/10.1063/1.4870288/15131415/133707_1_online.pdf.

[49] C. Duarte, Computer Physics Communications **181**, 1501 (2010).

[50] N. Papior, N. Lorente, T. Frederiksen, A. García, and M. Brandbyge, Computer Physics Communications **212**, 8 (2017).