This is the accepted manuscript made available via CHORUS. The article has been published as:

Theory of Thermal Relaxation of Electrons in Semiconductors
Sridhar Sadasivam, Maria K. Y. Chan, and Pierre Darancet
Phys. Rev. Lett. 119, 136602 — Published 27 September 2017
DOI: 10.1103/PhysRevLett.119.136602
We compute the transient dynamics of phonons in contact with high energy “hot” charge carriers in 12 polar and non-polar semiconductors, using a first-principles Boltzmann transport framework. For most materials, we find that the decay in electronic temperature departs significantly from a single-exponential model at times ranging from 1 ps to 15 ps after electronic excitation, a phenomenon concomitant with the appearance of non-thermal vibrational modes. We demonstrate that these effects result from the slow thermalization within the phonon subsystem, caused by the large heterogeneity in the timescales of electron-phonon and phonon-phonon interactions in these materials. We propose a generalized 2-temperature model accounting for the phonon thermalization as a limiting step of electron-phonon thermalization, which captures the full thermal relaxation of hot electrons and holes in semiconductors. A direct consequence of our findings is that, for semiconductors, information about the spectral distribution of electron-phonon and phonon-phonon coupling can be extracted from the multi-exponential behavior of the electronic temperature.

Following the seminal works of Kaganov et al. [1] and Allen [2], the thermalization of a system of highly energetic charge carriers with a lattice is frequently understood as an electron-phonon mediated, temperature equilibration process with a single characteristic timescale \( \tau_{el-ph} \). Such description, referred to as the two temperature (2T) model, relies on the central assumption that both electrons and phonons remain in distinct thermal equilibria and can therefore be described by time-dependent temperatures \( T_{el}(t) \) and \( T_{ph}(t) \) during the thermal equilibration process. In metals, due to the relative homogeneity of the electron-phonon interactions and the rates of thermalization within the electronic and phononic subsystems, the hypothesis of subsystem-wide thermal equilibrium is generally accurate, and the 2T model has been successful in modeling ultra-fast laser heating [3–5], despite some notable deviations from the 2T predictions in graphene and aluminum [6–8]. In semiconductors, the highly heterogeneous electron-phonon interactions (e.g. in polar semiconductors with Fröhlich interactions [9]) and, in some cases, the higher lattice thermal conductivity in comparison to metals weaken the hypothesis of a thermalized phononic subsystem [10–11], hence calling for the reexamination of the 2T physical picture in semiconductors.

In this context, the advent of first-principles techniques able to predict the mode- and energy-resolved electron-phonon [12–14] and phonon-phonon interactions [15–16] provides an important opportunity: In their modern implementations [13–16,17], these methods have been able to predict lattice thermal conductivities [18–21], the temperature- and pressure-dependence of the electronic bandgap [22–23], electrical conductivities [29–30], and hot carrier dynamics [31–32]. However, to the best of our knowledge and despite these early successes, these approaches have yet to be applied to the computation of electron-induced, non-equilibrium phonon distributions and their effects on thermal relaxation of electrons.

In this work, we combine first-principles calculations of electron-phonon and third-order phonon-phonon interactions within the semi-classical Boltzmann transport equation (BTE) for predicting the joint time-evolution of electron and phonon populations after hot carrier excitation. For 12 polar and non-polar cubic semiconductors, we show that the resulting phonon and electron dynamics departs qualitatively from the 2T physical picture over timescales of 1-15 ps after excitation. We demonstrate that this disagreement stems from the breakdown of the hypothesis of thermal equilibrium within the lattice subsystem, caused by the wide range of timescales associated with electron-phonon and phonon-phonon interactions in these systems. We generalize the 2T model of Allen to account for the slow phonon thermalization as a limiting step of electron-phonon thermalization, show that our generalized 2T model captures the transient dynamics for all compounds, and discuss its implication for time-resolved spectroscopy experiments. We anticipate our findings to apply to any material with broad spectral distributions of electron-phonon interactions (e.g. polar materials) and weak phonon-phonon interactions (in comparison to bulk metals).

We compute the mode- and time-resolved phonon occupation function \( n_{\mathbf{q},\nu}(t) \) in the presence of an electron occupation function \( f_{\mathbf{n} \mathbf{k}}(t) \) solving a coupled system of equations parametrized using density functional theory (DFT)-based approaches. The time-evolution of the phonon occupation function is obtained by solving the BTE:  

\[
\frac{\partial n_{\mathbf{q},\nu}(t)}{\partial t} = \left( \frac{\partial n_{\mathbf{q},\nu}(t)}{\partial t} \right)_{el} \left[ n_{\mathbf{q},\nu}(t), f_{\mathbf{n} \mathbf{k}}(t) \right] + \left( \frac{\partial n_{\mathbf{q},\nu}(t)}{\partial t} \right)_{ph} \left[ n_{\mathbf{q},\nu}(t) \right],
\]

where the drift term has been neglected due to the lack of spatial temperature gradient, and \([\ldots]\) indicates the functional dependence. The two terms on the right denote the time-dependent scattering potentials due to electron-phonon (EPI) and phonon-phonon interactions (PPI), both computed using first-principles methods, as detailed below. Importantly, we
make the assumption that charge carriers are in thermal equilibrium and that \( f_{nk}(t) \) can be approximated by a time-dependent Fermi-Dirac function centered near the top of the valence band for holes and near the bottom of the conduction band for electrons at the temperature \( T_{el}(t) \). Depending on the material and the nature of charge carriers, the timescale of the phonon-mediated carrier thermalization to the band edges was found to range from 0.1 to 1 ps \([31, 33]\) which also corresponds to limits of validity of the semi-classical description. Hence, we expect our simulation method and the approximation of \( f_{nk}(t) \) to be quantitative at subsequent times.

Specifically, we define the EPI scattering potential as an explicit functional of the phonon and electron occupation functions at time \( t \), and compute it using Fermi’s golden rule: 
\[
\left( \frac{\partial n_{\nu}(t)}{\partial t} \right)_{cp} = \frac{4\pi}{h^2} \sum_{k,m,n} |g_{q\nu}(mk+q,nk)|^2 M_{mn\nu\nu}(t),
\]
in which \( |g_{q\nu}(mk+q,nk)| \) is the time-independent electron-phonon matrix elements involving electronic states \( |n\rangle \) and \( |mk+q\rangle \) and vibrational state \( |\nuq\rangle \) evaluated using Wannier interpolation with the EPW code \([13]\). \( M_{mn\nu\nu}(t) \) is the time-dependent joint density of states computed from \( n_{q\nu}(t) \), \( f_{nk}(t) \), \( f_{nk+q}(t) \), and the electron and phonon spectral densities (detailed formulas are given in Supplemental Material). Similarly, we evaluate the scattering caused by PPI (\( \left( \frac{\partial n_{\nu}(t)}{\partial t} \right)_{pp} \)) from Fermi’s golden rule, using the time-independent 3-phonon scattering matrix elements \( |\Psi_{qq\nu\nu\nu}\rangle^2 \) computed with DFT \([15]\) and the time-dependent density of final states computed from \( n_{q\nu\nu}(t) \), \( n_{q\nu\nu}(t) \), \( n_{q\pm q\nu\nu}(t) \). At each time step, the net energy transfer \( Q_{cp} \) between electrons and phonons is computed and a new electronic temperature is derived as 
\[
T_{el}(t + \Delta t) = T_{el}(t) - \frac{Q_{cp}(t)}{C_{el}(T_{el})} \text{ where } C_{el}(T_{el}) \text{ is the instantaneous electronic heat capacity at temperature } T_{el}.
\]
The BTE is solved for 48000 phonon modes using an explicit time-stepping scheme with a time-step of 0.5 fs and a total simulation time of 25 ps for 12 cubic semiconducting compounds (BN, BP, BAs, BSb, AlP, AlAs, AlSb, GaN, GaP, GaAs, diamond, Si). All the simulations discussed below were initialized with an equilibrium phonon distribution at 300 K and a Fermi-Dirac distribution of electrons at 3000 K with the Fermi level set at 0.3 eV below the valence band maximum (other choices of initial temperatures and Fermi energies are shown to lead to similar conclusions in Supplemental Material).

The electronic structure was computed with DFT in the local density approximation, using norm-conserving pseudopotentials, a \( 10 \times 10 \times 10 \) \( k \)-grid and the Quantum Espresso package \([34]\). The phonon dispersion was computed using density functional perturbation theory \([35]\) and a \( 5 \times 5 \times 5 \) \( q \)-grid. Third-order force constants were computed in real space using finite differences on a \( 6 \times 6 \times 6 \) supercell \([36]\) and Fourier transformed to obtain phonon-phonon interaction matrix elements \( |\Psi_{qq\nu\nu\nu}\rangle^2 \) on a \( 20 \times 20 \times 20 \) \( q \)-grid \([37]\). The present approach neglects the temperature dependence of the third-order force constants \([38, 39]\), as this simplification has been shown to accurately predict the temperature-dependent lattice thermal conductivity for cubic semiconductors \([19, 37, 40, 41]\). Electron-phonon interactions were evaluated on \( 20 \times 20 \times 20 \) and \( 40 \times 40 \times 40 \) grids for phonons and electrons, respectively. Convergence stud-
ies are provided in the Supplemental Material.

In Fig. 1 we show the time-dependent phonon occupations along high symmetry directions of the Brillouin zone for BN and BAs (snapshots for all materials can be found in Supplemental Material). At short times \( t < 1 \) ps, the electronic energy is transferred to long-wavelength optical phonons, an effect originating from the larger electron-phonon scattering phase space associated with low-momentum phonons near the top (bottom) of the valence (conduction) bands in all 12 compounds, and further magnified by the \( 1/q \) divergence in the Fröhlich coupling in polar semiconductors \([9,12]\). Accordingly, we observe that more energy is transferred to the LO and TO modes of BN than to the modes of BAs, as expected from the Born effective charges (1.86 for BN vs 0.56 for BAs) and polarity, which also lead to a larger electron-phonon coupling and LO-TO splitting. Surprisingly, these “hot” phonon modes are found for all compounds to achieve near-thermal equilibrium with the electrons rather than with the rest of the phonons, a strong departure from the hypothesis of local thermal equilibrium within the lattice. At longer times, \( 1 < t < 10 \) ps, long wavelength LO and TO phonons in BN remain in near-thermal equilibrium with electrons, while transferring their energy to acoustic modes via 3-phonon processes, through Klemens (decay to two acoustic phonons) \([13]\) and Ridley (decay to one optical & one acoustic phonons) mechanisms \([44]\). The “hot phonon” cooling in BAs is slower in comparison to BN as the large acoustic-optical phonon band gap \([19,45]\) (originating from the mass mismatch) truncates the Klemens scattering phase space, while the Ridley decay is reduced by the small LO-TO splitting. Near-thermalization within the phonon subsystem (and, concomitantly, between electrons and phonons) is achieved in BN at \( t \approx 25 \) ps, with an electronic temperature 50 K away from the average lattice temperature \( T_{ph} = 380 \) K. In stark contrast to BN, for BAs both electrons (602 K) and hot phonons remain in near equilibrium with each other, but far from the average lattice temperature \( T_{ph} = 344 \) K. Importantly, the same two trends are observed for all simulated materials: (1) electrons first achieve near thermal equilibrium with a small number of high energy phonon modes; (2) full electron-lattice thermalization and intra-phonon thermalization are always achieved simultaneously: both trends implying that electron cooling is limited by thermalization within the phonon-subsystem.

Further illustrating the non-equilibrium between phonon modes, we see in Fig. 2 that the agreement between the BTE simulation and a 2T model parametrized from first-principles (see Supplemental Material for details) is good at times \( t < 0.05 \) ps and \( t < 0.2 \) ps for BN and BAs, respectively, but quickly deteriorates afterwards. For all compounds, the 2T model predicts a thermalization that is at least an order of magnitude faster than that observed in the full BTE simulation, proving that electronic cooling becomes limited by another mechanism, not accounted for in the 2T model \([46]\).

To test our hypothesis of a phonon-thermalization limited process, we perform a constrained simulation of electron cooling in which the phonons are partitioned into multiple subspaces defined by the strength \( \eta_{q\nu} \) of their interactions with electrons and phonons:

\[
\eta_{q\nu} = \hbar \omega_{q\nu} \left\{ \sum_{k,m,n} |g_{q\nu}(mk + q, nk)|^2 \delta_{q\nu,mnk} + \sum_{q'\nu',q\nu} |\Psi_{q\nu,q'\nu'}|^2 \delta_{q\nu,q'\nu'} \right\},
\]

where the terms on the right approximate the scattering due to EPI and PPI for each mode \( |q\nu\rangle \) (the PPI term only includes modes with a larger interaction strength and is computed self-consistently). \( \delta_{q\nu,mnk} \) and \( \delta_{q\nu,q'\nu'} \) are energy conservation delta functions for electron-phonon and phonon-phonon scattering respectively. The phonon modes in the largest \( \eta_{q\nu} \) subspace \([47]\) are a small subset (see Fig. 2c,d) of the total number of phonons and primarily consists of long-wavelength optical phonons with strong electron-phonon interaction \([48]\).
At time $t = 0$, only the modes belonging to the subspace with the largest $\eta_{q\nu}$ are allowed to interact with electrons until thermalization. Upon thermalization of the first subspace, the next subspace is introduced in the simulation along with the thermalized system of electrons and the first subspace. This constrained, “successive thermalization” (ST) process is continued until all modes are included [19]. As shown in Fig. 2 and Supplemental Material, this constrained ST simulation achieves quantitative agreement at all times for all materials considered, validating our central finding: Electron cooling in semiconductors is limited by intra-phonon thermalization, a direct consequence of the order-of-magnitude heterogeneities in the mode-dependent electron-phonon interactions and slow phonon thermalization.

We conclude this work by proposing a simple generalization of the 2T model based on these findings, and discuss its consequences in interpreting materials properties measured by the time-resolved decay of the electronic temperature [50, 51]. We start by partitioning the system in an analogous way to our ST simulation, i.e., with 3 subsystems well described by a subsystem-wide temperature: (I) - a system containing electrons and phonons fully thermalized with each other (by definition, $T_I(t) = T_{el}(t)$), (II) - phonon modes in contact with electrons and/or phonons of system (I) (in the process of being thermalized) at temperature $T_{ph}(t = 0) < T_{II}(t) < T_{el}(t)$, and (III) - “cold” phonons not in contact with (I) ($T_{III}(t) = T_{ph}(t = 0)$). In this generalized partition scheme, the 2T model is strictly recovered by setting (I) = {el}, (II) = {ph} and (III) = ∅. As more modes become thermalized with electrons as a function of time, the long-time electron-phonon thermalization can be understood as system (I) absorbing systems (II) and (III). Hence, the heat capacity of (I) becomes time-dependent with $C_I(t)$ increasing from $C_I(t = 0) = C_{el}$ to $C_I(t \to \infty) = C_{el} + C_{ph}$ (temperature dependences were omitted for simplicity of notation). Such time-dependent heat capacity $C_I(t)$ can be understood as an accumulation of the phonons over the timescales of their interactions. $C_I(t)$ can be computed heuristically by defining an effective mode-dependent thermalization time $\tau_{q\nu}$ (that we set to the relaxation time) and 

$$C_I(t) = \sum_{q\nu} C_{q\nu} \Theta(t - \tau_{q\nu})$$

where $\Theta(t)$ is the Heaviside function. Similarly, the electron-phonon coupling accumulation can be defined as 

$$G_{ep}(t) = \sum_{q\nu} G_{ep,q\nu} \Theta(t - \tau_{q\nu})$$

(see Supplemental Material for definitions of $\tau_{q\nu}$, $C_{q\nu}$, $G_{ep,q\nu}$ and their values for all compounds). Noteworthily, for materials with large heterogeneities in their mode-dependent electron-phonon coupling rates $G_{ep,q\nu}$, these two accumulation functions have very different time-dependences: as seen in Fig. 3 b,d), $G_{ep}(t)$ reaches 50% of its total value 1-1000 times faster than $C_I(t)$.

At short times (comparable to the time of accumulation of $G_{ep}(t)$), the observed electronic temperature decay rate given by this generalized 2T model can be approximated by $G_{II}/C_I \simeq G_{ep}(t \to \infty)/C_{el}$, i.e., the decay rate predicted by a “standard” 2T model. Correspondingly, in Fig. 3 c,e), we observe an excellent correlation between the initial decay rate and the electron-phonon coupling strength predicted directly from first-principles for all compounds considered in this work – indicating that, at short time, the determination of the single-exponential decay of the electronic temperature yields the total electron-phonon coupling. At longer times, the decay rate of the electronic temperature $G_{II}/C_I$ is reduced by the accumulation of heat capacity in I, as $C_I >> C_{el}$. As shown in Fig. 3 d), the reduction of the decay rate for all compounds shows a good correlation with the disparity of timescales between $C_I(t)$ and $G_{ep}(t)$, suggesting that a measurement of the electronic temperature decay across timescales in semiconductors would yield both the total electron-phonon coupling coefficient and information about the distribution of phonon interaction strength (and its heterogeneity) in a given material. Interestingly, as the phonon interaction strength involves both EPI and PPI (see Eq. 11), the time-dependence of the decay rates is particularly im-

FIG. 3: a) Electronic temperature decay in BA$_3$ along with the decay rates at short ($\alpha$) and long ($\beta$) time instants. b) Accumulation of lattice heat capacity $C_I(t)$ at $T_{ph} = 300$ K and electron-phonon coupling coefficient $G_{ep}(t)$ at $T_{el} = 3000$ K, $T_{ph} = 300$ K as a function of phonon thermalization time-scale in BA$_3$. c) Comparison of electron-phonon coupling coefficient obtained from the decay rate of electronic temperature at $t \to 0$ and directly from DFT for all the 12 semiconductors considered in this work. d) Comparison of the ratio between initial and long-time decay rates with the ratio of time-scales of heat capacity and electron-phonon coupling accumulation for all compounds considered in this work.
portant for materials with very heterogeneous EPI (Diamond, BN, GaN), and large phonon-bandgaps (BAs, Bsb), and vanishes for nearly homogeneous EPI (e.g. GaAs [22]).

In conclusion, we have demonstrated that electron cooling in semiconductors is limited by intra-phonon thermalization at timescales on the order of 1 – 20 ps. We have proposed a generalized 2-Temperature model accounting for this effect, and shown that such a model can be used to extract information from the measurement of the electronic temperature about both the total electron-phonon coupling and the distribution of electron-phonon and phonon-phonon interactions. More generally, we expect the phonon-limited thermalization identified in this work to have consequences on both heat and electron transport, fields in which long-lasting non-equilibrium phonon distributions have been shown to impact spectroscopic measurements [6, 7], current-voltage characteristics [52], and hot electron lifetimes [33] [54]. Specifically, our work offers a direct estimate of the timescales at which equilibrium models become quantitative in the presence of hot electrons (and their relationship to materials properties), and, via the tunability of the phonon-interaction strength, new pathways to control the timescales of electronic energy dissipation.

Use of the Center for Nanoscale Materials, an Office of Science user facility, was supported by the U. S. Department of Energy, Office of Science, Office of Basic Energy Sciences, under Contract No. DE-AC02-06CH11357. This material is based upon work supported by Laboratory Directed Research and Development (LDRD) funding from Argonne National Laboratory. We gratefully acknowledge the computing resources provided by the Laboratory Computing Resource Center at Argonne National Laboratory. We thank Stephen Gray, Richard Schaller, and Yi Xia for fruitful discussions.

[36] Decay of real-space third-order force constants for all 12 semiconductors is provided in Supplemental Information.
[46] As shown in Supplemental Material, this overly fast cooling is not corrected by simple higher level descriptions such as the 3-temperature model proposed by Waldecker et al. [8] in which the phonon branches are sub-divided into two categories (for example, optical and acoustic), depending on their coupling to electrons.
[47] Phonon modes with $\eta_{q\nu} > c\eta_{q\nu,\text{max}}$ are chosen to belong to a subspace and the predictions for $c = 0.5$ are presented in the main text. Results for $c = 0.1$ are reported in Supplemental Material with similar temperature decay predictions.
[48] See Supplemental Material (which includes Refs. [55][57]) for distributions of $\eta_{q\nu}$ for all 12 semiconductors considered in this study along with information on the average momentum of phonons within each subset.
[49] The effective coupling coefficient between the interacting systems at each thermalization step is chosen to be proportional to the sum of all interaction strengths $\eta_{q\nu}$ of modes belonging to the subspace undergoing thermalization. We note that the proportionality constant is chosen to be the same for every subspace and is independent of temperature.