Ab initio optimization of phonon drag effect for lower-temperature thermoelectric energy conversion

Jiawei Zhou[a], Bolin Liao[a], Bo Qiu[a], Samuel Huberman[a], Keivan Esfarjani[b,c], Mildred S. Dresselhaus[d,e], and Gang Chen[a,1]

[a]Department of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139; [b]Department of Mechanical and Aerospace Engineering, Rutgers University, Piscataway, NJ 08854; [c]Institute for Advanced Materials, Devices and Nanotechnology, Rutgers University, Piscataway, NJ 08854; [d]Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, MA 02139; and [e]Department of Physics, Massachusetts Institute of Technology, Cambridge, MA 02139

Edited by Robert C. Dynes, University of California, San Diego, La Jolla, CA, and approved October 27, 2015 (received for review June 24, 2015)

Although the thermoelectric figure of merit zT above 300 K has seen significant improvement recently, the progress at lower temperatures has been slow, mainly limited by the relatively low Seebeck coefficient and high thermal conductivity. Here we report, for the first time to our knowledge, success in first-principles computation of the phonon drag effect—a coupling phenomenon between electrons and nonequilibrium phonons—in heavily doped region and its optimization to enhance the Seebeck coefficient while reducing the phonon thermal conductivity by nanostructuring. Our simulation quantitatively identifies the major phonons contributing to the phonon drag, which are spectrally distinct from those carrying heat, and further reveals that although the phonon drag is reduced in heavily doped samples, a significant contribution to Seebeck coefficient still exists. An ideal phonon filter is proposed to enhance zT of silicon at room temperature by a factor of 20 to ∼0.25, and the enhancement can reach 70 times at 100 K. This work opens up a new venue toward better thermoelectrics by harnessing nonequilibrium phonons.

phonon drag | nonequilibrium phonon | electron phonon interaction | thermoelectrics | nanocluster scattering

In metals and semiconductors, lattice vibration-, or phonon-, induced electron scattering has a major influence on electronic transport properties (1). The strength of the electron–phonon interaction (EPI) depends on the distribution of electron and phonon populations. The EPI problem was first studied by Bloch (2), who assumed the phonons to be in equilibrium (so-called Bloch condition) when calculating the scattering rates of electrons caused by EPI, because of the frequent phonon–phonon Umklapp scattering. This assumption is widely adopted for the determination of electronic transport properties at higher temperatures (1, 3), including the electrical conductivity and the normal ("diffusive") Seebeck coefficient. Below the Debye temperature, however, the nonequilibrium phonons become appreciable because the phonon–phonon Umklapp scatterings are largely suppressed, and this assumption becomes questionable. The significance of nonequilibrium phonons on the Seebeck coefficient was first recognized by Gurevich (4). The experimental evidence given later (5, 6) clearly showed an "anomalous" peak of the Seebeck coefficient at around 40 K in germanium. To address this unusual observation, Herring proposed that the nonequilibrium phonons can deliver excessive momenta to the electrons via the EPI (7). This process generates an extra electrical current in the same direction as the heat flow, as if the electrons were dragged along by phonons. Therefore, this effect has been dubbed "phonon drag" (7), which makes itself distinct from the normal (diffusive) Seebeck effect. Subsequent explorations (8–13) revealed that this effect exists in various material systems. In particular, it has been suggested that phonon drag is responsible for the extremely high Seebeck coefficient $S = -4{,}500$ μV/K experimentally found in $FeSb_2$ (14–16).

The efficiency of thermoelectric materials is characterized by the figure of merit zT, defined as $zT = \sigma S^2 T/\kappa$, where σ, S, κ, and T are the electrical conductivity, Seebeck coefficient, thermal conductivity, and absolute temperature, respectively. Thermoelectric energy conversion at lower temperatures (around and below 300 K) can benefit a wide range of applications including refrigeration, air conditioning, and cryogenic cooling (17) but is also challenging mainly because the diffusive Seebeck coefficient drops in magnitude whereas the thermal conductivity increases as the temperature decreases. It is therefore tempting to make use of the phonon drag effect to boost the Seebeck coefficient for better thermoelectrics at lower temperatures. However, controversies exist as to whether zT can be enhanced at all by using the phonon drag effect. Theoretical estimations concluded that the phonon drag will not benefit the thermoelectrics (18), whereas an experiment on silicon nanowires has suggested the possibility of using the phonon drag to reach zT of 1 at 200 K (19). One major concern lies in the fact that a significant phonon drag requires phonons to be far away from their equilibrium under a certain temperature gradient (or equivalently, to have long mean free paths), which usually implies a high thermal conductivity. This is reflected in the experimental fact that the low-temperature peak of the phonon drag Seebeck coefficient usually coincides with the peak of the thermal conductivity (20, 21). Although it has been qualitatively known that phonons involved in phonon drag have longer mean free paths than those that carry heat (7), so far it is not clear to what extent one can decouple the contributions, that is, keep a significant phonon drag effect while reducing the lattice thermal conductivity, the latter of which has become a common strategy in increasing the thermoelectric efficiency (22–24). Furthermore, it was believed that phonon drag will become negligible at high doping levels necessary for thermoelectrics (20, 25), because the increased carrier concentration largely weakens the phonon drag [the so-called saturation effect (7)].

Significance

It has been well known that the phonon drag effect—an extra electrical current induced by phonon heat flow via electron–phonon interaction—can lead to unusually high Seebeck coefficient at low temperatures. However, its use for improving thermoelectric performance has been controversial. Here, using first principles calculations we examine the phonon drag with detailed mode-specific contributions and reveal that even in heavily doped silicon at room temperature, phonon drag can still be significant, which challenges the previous belief that phonon drag vanishes in heavily doped samples. A phonon filter is designed to spectrally decouple the phonon drag from the heat conduction. Our simulation explores the coupled electron phonon transport and uncovers the possibility of optimizing the phonon drag for better thermoelectrics.

Author contributions: J.Z., B.L., B.Q., and S.H. performed research; J.Z., B.L., K.E., M.S.D., and G.C. analyzed data; and J.Z., B.L., and G.C. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

[1]To whom correspondence should be addressed. Email: gchen2@mit.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1512328112/-/DCSupplemental.

The key information for a better understanding here is a mode-by-mode quantification of phonon contributions to the phonon drag and the thermal conductivity, individually. An exact description of the phonon drag effect requires the solution of coupled electron–phonon Boltzmann transport equations (BTEs) with accurate matrix elements of various scattering processes (most importantly, the electron–phonon scattering and phonon–phonon scattering). Early theoretical works (1, 5, 7, 26) attempted to concurrently solve the coupled electron–phonon BTEs mostly using a variational approach (1). Although reasonable agreements with experiments were achieved (27, 28), the calculations were limited to the lowest orders of the trial functions for the variational method and assumed simplified scattering models, which lack predictive power due to the large number of adjustable parameters involved and make the interpretation of the calculated results less intuitive. An alternative approach to solving the coupled BTEs seeks to partially decouple the electron and phonon transport (29). Compared with the variational method, it has the advantage of a more transparent interpretation of the results in terms of contributions from individual phonon modes. This approach has been applied to Si metal–oxide–semiconductor field-effect transistors and GaAs/AlGaAs heterojunctions (30) and to bulk silicon at low carrier concentrations recently by incorporating accurate phonon–phonon scattering rates obtained from first-principles calculation into the model (31). Although the quantitative agreement with experiments at low carrier concentrations was improved, the appearances of adjustable material-dependent parameters (i.e., the deformation potentials) for the crucial electron–phonon scattering processes still limit the predictive power and are not satisfactory for a modern mode-by-mode understanding of the phonon drag effect. In addition, a quantitative evaluation of the phonon drag across a range of carrier concentrations (particularly into the heavily doped region, which is practically relevant to thermoelectrics) was missing. In this context, a fully first-principles understanding of the phonon drag effect becomes highly desirable. The recent development of first-principles simulation tools (32, 33) has made the mode-specific contributions to thermal conductivity accessible. Obtaining the same information for phonon drag, however, can be exceedingly challenging due to the ultradense sampling meshes entailed by the convergence. In this paper, we undertake the challenge of examining the detailed phonon mode contributions to the phonon drag effect in silicon from first principles, by combining the recently invented interpolation scheme based on the maximally localized Wannier functions (34, 35) for the determination of EPI matrix elements with an accurate description of the nonequilibrium phonon distributions (32, 33), including first-principles calculations of the intrinsic phonon–phonon interactions as well as the phonon scatterings by electrons. By generalizing the partially decoupled electron–phonon BTE framework (30) into the heavily doped region, we implemented an ab initio computational approach without any adjustable parameters and justified it by comparing it to experiments across a wide range of temperatures and carrier concentrations. Our result quantifies the previous explanation of the saturation effect in heavily doped samples (7) as caused by the increased phonon scattering by electrons. Based on the information revealed, we show that the phonon drag effect can be engineered to enhance the Seebeck coefficient while largely reducing the thermal conductivity by identifying the "preferable" phonon modes and filtering out the others. Following this strategy, an ideal phonon filter is proposed to increase zT in n-type silicon by a factor of 20 to ∼0.25 at room temperature, and the enhancement can be as large as 70 times at 100 K.

Results

Theoretical Formalism of Phonon Drag. A temperature gradient can concurrently drive a phonon flow and an electron flow, which results in the lattice thermal conduction and the Seebeck effect, respectively. A complete description of the transport requires a full solution of the coupled BTEs for the distribution functions of both electrons and phonons (1):

$$\begin{cases} \mathbf{v}_\alpha(\mathbf{k})\cdot\dfrac{\partial f_\alpha(\mathbf{k})}{\partial T}\nabla \mathrm{T} - e\mathbf{v}_\alpha(\mathbf{k})\cdot\dfrac{\partial f_\alpha(\mathbf{k})}{\partial E}\nabla\varphi = -\dfrac{f_\alpha(\mathbf{k})-f_\alpha^0(\mathbf{k})}{\tau_\alpha^*(\mathbf{k})} + \left(\dfrac{\partial f_\alpha(\mathbf{k})}{\partial t}\right)_{e-ph} \\ \mathbf{v}_\lambda(\mathbf{q})\cdot\dfrac{\partial n_\lambda(\mathbf{q})}{\partial T}\nabla \mathrm{T} = -\dfrac{n_\lambda(\mathbf{q})-n_\lambda^0(\mathbf{q})}{\tau_\lambda^*(\mathbf{q})} + \left(\dfrac{\partial n_\lambda(\mathbf{q})}{\partial t}\right)_{e-ph} \end{cases}, \quad [1]$$

where φ is the electrochemical potential, f and n are the distribution functions for electrons and phonons, respectively (the distribution functions with 0 in the superscript denote the equilibrium statistics: Fermi–Dirac for electrons and Bose–Einstein for phonons), and $\mathbf{v}$ is the group velocity for electrons (with wave vector $\mathbf{k}$) and phonons (with wave vector $\mathbf{q}$). The electron bands and the phonon branches are denoted using $\alpha(\beta)$ and λ, respectively. The collision terms on the right-hand sides of Eq. **1** describe the changes of the distribution functions caused by various scattering mechanisms. We have separated out those due to the electron–phonon coupling (denoted by "*e-ph*") and assumed that all other scattering processes (electron impurity, phonon–phonon, and phonon impurity) can be described by the mode-dependent relaxation time τ^*.

The electron–phonon interactions are three-particle processes and involve distribution functions of three different states: the initial electron state $f_{\mathbf{k}\alpha}$, the final electron state $f_{\mathbf{k}'\beta}$ and the participating phonon $n_{\mathbf{q}\lambda}$. Within the linearized Boltzmann equation framework (1), only first-order deviations of the distribution functions from their equilibrium values are kept, which include terms linear in $\Delta f_{\mathbf{k}\alpha}=f_{\mathbf{k}\alpha}-f^0_{\mathbf{k}\alpha}$, $\Delta f_{\mathbf{k}'\beta}=f_{\mathbf{k}'\beta}-f^0_{\mathbf{k}'\beta}$, or $\Delta n_{\mathbf{q}\lambda}=n_{\mathbf{q}\lambda}-n^0_{\mathbf{q}\lambda}$, each characterizing the nonequilibrium state of electrons or phonons. For the normal electrical property calculations (1) (e.g., electrical conductivity and diffusive Seebeck coefficient), the Bloch condition is usually assumed, which implies $\Delta n_{\mathbf{q}\lambda}=0$ (2). This assumption breaks down when nonequilibrium phonons become appreciable and the appearance of $\Delta n_{\mathbf{q}\lambda}$ in the electron BTE (the first line of Eq. **1**) is essentially responsible for the phonon drag contribution S_{ph} to the Seebeck coefficient. However, the determination of $\Delta n_{\mathbf{q}\lambda}$ in turn requires the knowledge of $\Delta f_{\mathbf{k}\alpha}$, which appears in the phonon BTE (the last term in the second line of Eq. **1**) and makes solving fully coupled BTEs a formidable task. One further step toward solving Eq. **1** is to realize that the influence of nonequilibrium electrons on the phonon drag effect, indirectly through affecting the nonequilibrium phonons, is a higher-order effect (i.e., $\Delta f_{\mathbf{k}\alpha}\approx 0$ can be assumed in the phonon Boltzmann equation; also see discussions in *SI Appendix, Supplementary Note 2*). As a result, the electron–phonon BTEs can be partially decoupled, which leads to a feasible computational approach.

The electron–phonon coupling terms in Eq. **1** also contain parts that characterize electron scattering by equilibrium phonons (3) and phonon scattering by equilibrium electrons (36, 37), which can be separated out and described by additional relaxation times (*SI Appendix, Supplementary Note 1*). The former is usually the dominant scattering mechanism for the electron transport, whereas the latter weakens the phonon drag at high carrier concentrations, as we will show. We define the total relaxation times τ taking into account both of these scatterings and also τ^* using Matthiessen's rule. Because the phonon BTE is decoupled from the electron BTE, the nonequilibrium phonon distribution can be directly written down as $\Delta n_{\mathbf{q}\lambda}=\tau_{\mathbf{q}\lambda}\mathbf{v}_{\mathbf{q}\lambda}\cdot\nabla \mathrm{T}\frac{\partial n^0_{\mathbf{q}\lambda}}{\partial T}$. The electron BTE in Eq. **1** can then be readily solved for the nonequilibrium electron distribution $\Delta f_{\mathbf{k}\alpha}$ given $\Delta n_{\mathbf{q}\lambda}$. Considering the electrical current density given by $\mathbf{j}=e\sum_{\mathbf{k}\alpha}\mathbf{v}_{\mathrm{k}\alpha}\Delta f_{\mathrm{k}\alpha}$, we arrive at the phonon drag contribution to the Seebeck coefficient besides the diffusive contribution

$$S_{ph}=\frac{2e}{3\sigma\Omega N_{\mathbf{k}}N_{\mathbf{q}}k_BT^2}\sum_{\mathbf{q}\lambda}\left[\hbar\omega_{\mathbf{q}\lambda}\tau_{\mathbf{q}\lambda}\mathbf{v}_{\mathbf{q}\lambda}\left(\sum_{\mathbf{k}\alpha,\,\mathbf{k}'\beta}\left(\tau_{\mathbf{k}\alpha}\mathbf{v}_{\mathbf{k}\alpha}-\tau_{\mathbf{k}'\beta}\mathbf{v}_{\mathbf{k}'\beta}\right)\cdot f^0_{\mathbf{k}\alpha}\left(1-f^0_{\mathbf{k}'\beta}\right)n^0_{\mathbf{q}\lambda}\cdot\frac{2\Pi}{\hbar}\right)\right]$$

with

$$\prod=\pi\left|g_{\alpha\beta\lambda}(\mathbf{k},\mathbf{k}',\mathbf{q})\right|^2\cdot\delta\left(E_{k'\beta}-E_{k\alpha}-\hbar\omega_{q\lambda}\right)\delta(\mathbf{k}'-\mathbf{k}-\mathbf{q}),\qquad[2]$$

where e is the electron charge, σ is the electrical conductivity, Ω is the unit cell volume, $N_{\mathbf{k}}$ and $N_{\mathbf{q}}$ are the number of sampling points in the discrete reciprocal space meshes for electrons and phonons, E is the electron energy, $\omega_{\mathbf{q}\lambda}$ is the phonon frequency, $g_{\alpha\beta\lambda}(\mathbf{k}, \mathbf{k}', \mathbf{q})$ is the electron–phonon interaction matrix element (*SI Appendix*, *Supplementary Note 1*), and the two delta functions impose the energy and momentum conservation conditions. The term inside the bracket of Eq. **2** is the phonon drag contribution from each phonon mode. The same formula has been obtained in previous work under the assumption of weak electron–phonon coupling (29), which suggests that the dominant scattering for electrons is impurity scattering and also the phonon scattering by electrons is neglected. A similar model included the phonon scattering by electrons but was largely simplified (38). Here we have derived Eq. **2** from partially decoupled BTEs without making a priori assumptions on the scattering processes, thus extending its applicability particularly to the heavily doped region. We note that a strong coupling between dopants and the host material can alter the band structure at extremely high doping concentrations, where the current formalism will be less accurate without necessary modifications.

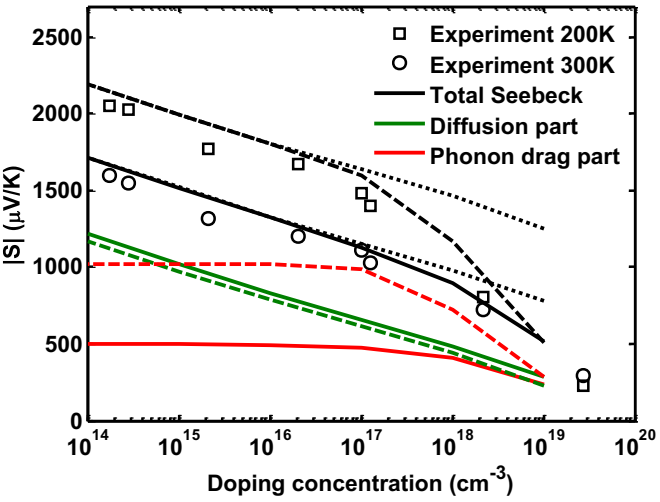

Fig. 1. Calculated Seebeck coefficient with respect to doping concentrations for n-type silicon at 300 K and 200 K on a semilog plot. The solid lines describe the calculated results at 300 K and dashed lines represent those at 200 K. Circles and squares are taken from the experiment (8). At each temperature the total Seebeck coefficient (black) as well as the decomposition into the phonon drag part (red) and diffusion part (green) is shown. Dotted lines are the total Seebeck coefficient calculated using the low doping level value of the phonon drag contribution and assuming this value will not be reduced as the doping concentration increases (i.e., neglecting the "saturation" effect). Compared with the dotted lines, the experimental Seebeck coefficient has a further decrease beyond 10^{17} cm^{-3} doping concentration, which is due to the decrease of the phonon drag contribution. This is captured by considering electron scattering of phonons and the resulting Seebeck coefficient (black solid lines for 300 K and black dashed lines for 200 K) agrees with experiments across the full range of doping concentrations.

Temperature and Carrier Concentration Dependence. To justify our numerical implementation (*Methods*), we first examine the phonon drag Seebeck coefficient in lightly doped silicon. The low doping level implies that the impurity scattering (3) as well as the phonon scattering by electrons (37) plays a negligible role. Similar to previous work (29, 31), good agreement is obtained between the calculation results and the experimental data (8), from 300 K down to 80 K for electrons and to 60 K for holes (*SI Appendix*, Fig. S2). An extremely dense sampling mesh of $100 \times 100 \times 100$ **q**-points in the phonon Brillouin zone is used, which is necessary for convergence at very low temperatures. We have also noticed that, even though the phonon drag effect is only dominant at very low temperatures, it has influences across a wide range of temperatures, extending beyond room temperature. This latter fact is consistent with previous simulation work (31, 39) and quenched thermopower experiments (40).

To optimize zT, the carrier concentration is a common experimental variable and usually sits around 10^{19} to 10^{21} cm^{-3}, for achieving a higher electrical conductivity (41). In Fig. 1, the Seebeck coefficient as a function of carrier concentration is shown for n-type silicon, which will be focused on in the following discussion. At low doping levels, the intrinsic phonon drag Seebeck coefficient should be independent of the carrier concentration (7). This is because the total momentum transfer to the electron system depends on the available electron states that can couple to phonons. At low doping levels, the number of such states is proportional to the carrier concentration, and therefore the momentum gain per electron characterizing the phonon drag Seebeck coefficient does not depend on the carrier concentration (*SI Appendix*, *Supplementary Note 6*). However, as the carrier concentration increases, the total momentum transfer deviates from the increasing trend of the carrier concentration and starts to saturate. As a result, the momentum gain per electron decreases, leading to a reduction of the phonon drag [the saturation effect (7)]. If the phonon drag Seebeck coefficient is assumed to be unchanged, an apparent discrepancy with experiments occurs above 10^{17} cm^{-3} doping concentration (Fig. 1).

As the carrier concentration increases, several scattering processes start to affect the transport properties. Impurity scattering of electrons can largely decrease the mobility but is found from our calculation to have only a small influence on the phonon drag effect (*SI Appendix*, *Supplementary Note 3*). It is the scattering of phonons that shortens the phonon mean free paths, therefore leading to a significant reduction of the phonon drag effect. Our calculation confirms that the phonon scattering by electrons (7, 38) is the major reason for this reduction whereas the phonon-impurity scattering contributes much less. This difference stems from the different phonon frequency dependence of the scattering rates. As will be clear later, phonons that contribute to the phonon drag mostly have low frequencies. For these phonons, the phonon-impurity scattering (scattering rate scales as ω^4) (42) drops much faster with the frequency than the phonon-electron scattering (scattering rate scales as ω) (37) and therefore plays a negligible role in reducing the phonon drag. However, the phonon–electron scattering drops more slowly than the intrinsic phonon–phonon scattering (scattering rate scales as ω^2) and will eventually dominate the scattering of low-frequency phonons as the carrier concentration increases (*SI Appendix*, Fig. S3). It has been inferred that the saturation effect is associated with the interaction of increased electrons on phonons (7, 9), but the interpretation is often vague due to the use of the variational approach (26) and the lack of quantitative evaluation. By performing first-principles simulation, we clarify the reason of the phonon drag reduction as the shortening of phonon mean free paths by electron scatterings and clearly show that the inclusion of this scattering mechanism leads to a good agreement of the phonon drag with the experiment

extending to the heavily doped region (Fig. 1). If the carrier concentration becomes extremely high, the long-wavelength phonons will be predominantly scattered by electrons and we expect an approximate $1/n$ dependence for their averaged relaxation times (36, 37). As a result, the phonon drag will tend to decrease as $1/n$, which was first pointed out by MacDonald (43) for the phonon drag in metals. Despite such reduction, the carrier concentration in thermoelectric materials is not sufficiently high, as in metals, and therefore the reduction of phonon drag with respect to the carrier concentration is typically much weaker than $1/n$ (*SI Appendix*, *Supplementary Note 6* and Fig. S8), implying that the phonon drag contribution in the thermoelectrics-relevant doping region cannot be simply ignored. As is seen in Fig. 1, at 10^{19} cm^{-3} doping concentration, the phonon drag contribution to the Seebeck coefficient is still comparable to the diffusive contribution. This finding is against the previous belief that the phonon drag effect vanishes in heavily doped samples (20, 25, 39) and makes it possible to optimize the phonon drag in heavily doped materials for better thermoelectric performance.

Identification of Preferable Phonon Modes. Having justified our calculation by looking into the temperature and carrier concentration dependence of the phonon drag, we proceed to quantify how one can benefit from the phonon drag effect for better thermoelectrics. First we examine the phonon mode-specific information in lightly doped silicon. The accumulated contributions to the phonon drag Seebeck coefficient and to the thermal conductivity from each phonon mode with respect to their frequencies from 100 K to 300 K are shown in Fig. 2*A*. Compared with the modes that carry heat, the specific phonons that contribute the most to the phonon drag effect have lower frequencies, indicating that they are closer to the zone center and also possess longer wavelengths (*SI Appendix*, Fig. S5), which explains why phonons involved in the phonon drag processes have longer mean free path than those that carry heat (Fig. 2*B*).

These features were qualitatively understood previously (7, 28, 44) and originate from the energy and momentum conservations in the electron–phonon scattering process, for which the phonon wave vector must be small to connect different electron states close to the band edge with the energies differing by a phonon energy. Here with the full knowledge of the spectral contribution, we can quantitatively determine how important each phonon mode is in contributing to the phonon drag. To enhance zT, the factor S^2/κ needs to be maximized. Provided that the phonon drag contribution is nonnegligible in the total Seebeck coefficient, one can ask whether we can reduce the thermal conductivity without sacrificing the Seebeck coefficient much. According to Fig. 2*B*, one can achieve this by designing a mean-free-path-selective phonon filter. For example, at 300 K phonons with mean free paths shorter than 1 μm contribute around 70% to the total thermal conductivity while contributing negligibly to the phonon drag effect, implying that the thermal conductivity can be reduced by 70% without affecting the Seebeck coefficient much by "filtering out" these phonon modes. At lower temperatures, the accumulated contribution to the phonon drag effect has a larger shift toward the long mean free path region compared with the contribution to the thermal conductivity. Therefore, this "decoupling" strategy becomes even more effective at lower temperatures.

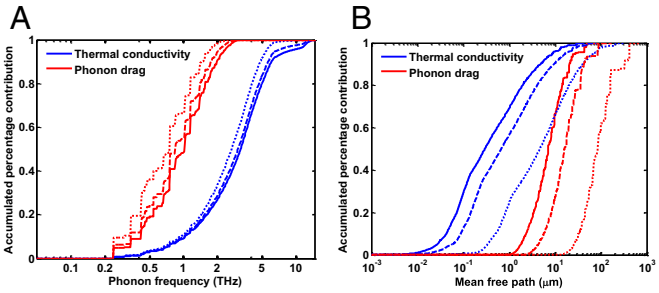

Fig. 2. Phonon mode-specific accumulated contributions to the phonon drag Seebeck coefficient and the thermal conductivity with respect to (*A*) phonon frequency and (*B*) phonon mean free path. Solid lines show the contribution at 300 K, and dashed lines are used at 200 K and dotted lines at 100 K. These results are obtained for lightly doped n-type silicon.

Because the phonon drag, as we have revealed, makes a significant contribution even at high carrier concentrations, the decoupling strategy naturally extends to the heavily doped region. We should note that as the carrier concentration increases, the restriction on the phonon wave vector to fulfill the momentum conservation is relaxed because electrons occupy a larger region in the reciprocal space. This tends to weaken the spectral difference between the phonons that contribute to phonon drag and thermal conduction. However, even in heavily doped semiconductors, we have discovered that the phonon wave vector is still limited to small values, and therefore there is large room for optimizing the phonon drag effect using the decoupling strategy. The resulting optimal zT achievable can be estimated by selecting those phonons that contribute preferably to the phonon drag than to the thermal conductivity, and then "filtering out" all other phonons as much as possible. We can define a mode-specific figure of merit $\zeta_{\mathbf{q}\lambda}$ as the ratio of the mode's contribution to the phonon drag to its thermal conductivity contribution (*Methods*). Modes with larger $\zeta_{\mathbf{q}\lambda}$ are desirable to maximize the Seebeck coefficient given an upper bound of the thermal conductivity. Fig. 3*A* shows the largest possible phonon drag Seebeck coefficient with different upper bounds of the thermal conductivity. In heavily doped silicon, it was previously believed that the phonon drag effect is completely suppressed, especially when one also tries to reduce the thermal conductivity. State-of-the-art material synthesis techniques have shown the capability of reducing the room-temperature thermal conductivity of silicon to below 4 W/mK (19, 45) via introducing phonon-blocking nanostructures. Our results show, however, even at such a low value of the thermal conductivity in a 10^{19} cm^{-3} doped n-type silicon sample, there still can be a phonon drag contribution that is about 25% of the diffusive contribution, if preferable modes are chosen carefully. Fig. 3*B* compares the zT when selecting preferable modes with that when neglecting the phonon drag effect, assuming that the thermal conductivity is reduced to 4 W/mK. The optimized zT for normal bulk silicon is ~0.01 at 300 K around 4×10^{19} cm^{-3} doping concentration (20), which can be boosted to ~0.25 by combining the optimized phonon drag effect and the reduced thermal conductivity. However, the conventional means of nanostructuring to reduce the thermal conductivity is to introduce nanoscale grain boundaries or precipitates that strongly scatter long-wavelength phonons, by which the phonon drag effect will also be largely weakened, diminishing the possible enhancement of zT by half (Fig. 3*B*). Furthermore, this enhancement increases as the temperature decreases, reaching a value as large as 70 at 100 K (*SI Appendix*, Fig. S6). This striking result indicates the large potential of exploiting the phonon drag effect for lower-temperature thermoelectrics.

Fig. 4*A* shows the distribution of the preferable phonon modes as a function of wavelength and frequency, where it is clearly seen that more preferable modes typically have longer wavelengths and lower frequencies. Although the discussions above set the upper bound for the enhancement of S^2/κ, simple selective mechanisms can be devised based on either one of the variables. Here we show the possibility of using nanoclusters as impurities to selectively scatter phonons with different frequencies. The impurity scattering is generally stronger for phonons with higher frequencies and therefore serves as a low-pass filter. Besides, in the volume fraction range we have explored, the electrical conductivity is barely affected (*SI Appendix*, *Supplementary Note 5*).

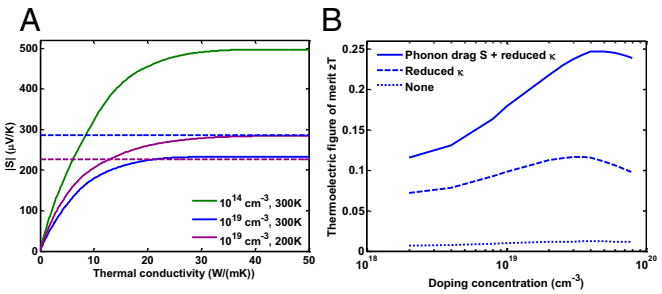

Fig. 3. (*A*) Contribution of the most preferable modes to the phonon drag Seebeck coefficient at different reduced thermal conductivity values and (*B*) the enhancement of zT achieved by selecting preferable modes at 300 K for n-type silicon. In *A*, for lightly doped silicon, the thermal conductivity can be reduced to 30 W/(mK) before observing significant diminishment of the phonon drag effect. Dashed lines represent the diffusive Seebeck coefficient for heavily doped silicon at different temperatures. For heavily doped silicon, the phonon drag part is still nonnegligible and becomes larger compared with the diffusion part when the temperature is decreased. In calculating zT, the experimental data are used for the electrical conductivity as a function of doping concentration for n-type silicon (52).

Fig. 4*B* shows the thermoelectric figure of merit zT with different volume fractions and nanocluster sizes at 300 K. A large enhancement of zT can be achieved, due to the fact that the thermal conductivity is largely reduced whereas the Seebeck coefficient is less affected. An optimal nanocluster size of 1 nm in diameter (*SI Appendix*, *Supplementary Note 5*) with a volume fraction of 0.2%, is found to enhance zT by a factor of 5, for which a significant portion of the Seebeck coefficient comes from the phonon drag. At lower temperatures, this enhancement becomes more pronounced (*SI Appendix*, Fig. S7).

Discussion

In this work the phonon drag effect is investigated in unprecedented detail from first principles to study the possibility of enhancing the thermoelectric performance around and below 300 K. Although it is widely believed that the phonon drag effect is only dominant at very low temperatures and will be largely suppressed when the carrier concentration becomes higher, especially after the thermal conductivity is reduced, we clearly show that even for silicon with a 10^{19} cm^{-3} carrier concentration at room temperature and a low thermal conductivity of 4 W/mK, the phonon drag contribution to the Seebeck coefficient is still appreciable if preferable phonon modes with long wavelength are carefully selected. We emphasize the importance of recognizing these long-wavelength (low frequency) phonons in the phonon drag effect: When optimizing the phonon drag, one should avoid nanostructure techniques that will destroy the long-wavelength phonons; otherwise, the phonon drag will also be reduced, leading to a significant drop in the possible enhancement of zT. The benefit of the optimized phonon drag effect becomes greater when the temperature is decreased. Moreover, we have proposed practical phonon filtering mechanisms based on the identification of preferable phonon modes. A phonon frequency filtering approach based on nanocluster scattering is shown to have the potential of enhancing zT significantly due to the combined effect of a reduced thermal conductivity and the optimized phonon drag by preferentially scattering high-frequency phonons.

Here we discuss some implications of our results in the context of the previous experimental work. The widely applied way of using grain boundaries as phonon scatterers (22), though efficient in reducing the thermal conductivity, may also destroy the phonon drag effect in certain materials at low temperatures due to the scattering of long-wavelength phonons (15). Similarly, the defects (extended defects such as micrometer-sized pores or cracks or line defects such as dislocations) that have detrimental effects on long-wavelength phonons should be avoided. However, there have been extensive experiments on SiGe alloys (46, 47), which preferentially scatter high-frequency phonons. However, these alloys typically contain large fractions of both Si and Ge, where the perturbation theory framework breaks down and the phonon drag effect needs to be reconsidered. We should also mention that the experiment showing zT of 1 in nanowires (19) cannot be simply explained by our results, because only bulk properties are examined in this work whereas phonons in their experiment are claimed to have 1D-like features. Nonetheless, our results present a rational principle to maintain the phonon drag effect while reducing the thermal conductivity by the appropriate use of a phonon filter.

The computational formalism we implemented to solve the coupled electron–phonon BTEs for the phonon drag effect should be generally applicable beyond silicon. We envisage that along this path more material systems can be systematically studied for quantitative understandings of their coupled electron–phonon transport. We also noticed recent experimental findings (21) of enhanced Seebeck coefficients of a thin conductive layer deposited on an insulating substrate, and the possible explanation that the phonons generated in the substrate with long mean free paths "leak" through the interface and drag the electrons. Generalizing the current formalism to understanding this interfacial phonon drag phenomenon will provide more space for tuning the phonon drag property of materials. Furthermore, other filtering mechanisms can be designed and engineered to better select the preferable phonon modes, thereby bringing benefits to thermoelectric applications, particularly at lower temperatures.

Methods

Electron–Phonon Interaction. The equilibrium properties of electrons and phonons are calculated using the density functional theory and density functional perturbation theory as implemented in the QUANTUM ESPRESSO package (48). We use the norm-conserving pseudopotential with the Perdew and Zunger (49) local density approximation for the exchange-correlation functional and a cutoff energy of 60 Ryd. A $12 \times 12 \times 12$ **k**-mesh is used for the electronic band structure and a $6 \times 6 \times 6$ **q**-mesh is used for the phonon dispersion. On these coarse meshes we also obtain electronic wave functions, phonon mode-specific perturbing potentials, as well as the electron–phonon coupling matrix elements. We use the EPW package (50) with the Wannier interpolation scheme (34) to map this information to a much denser mesh and

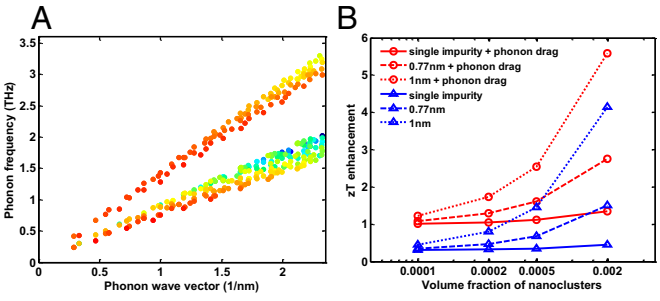

Fig. 4. (*A*) Distribution of preferable phonon modes in wave vector and phonon frequency, along with (*B*) the enhancement of the thermoelectric figure of merit zT compared with that of bulk doped silicon (~0.1 at 300 K) with respect to the volume fraction of nanoclusters. Data in *A* are obtained on a $70 \times 70 \times 70$ mesh and only long-wavelength phonons are shown. Red colors represent a higher figure of merit $\zeta_{q\lambda}$ than blue colors. In *B*, the calculation is done for n-type silicon with a doping concentration of 10^{19} cm^{-3} at 300 K. Curves labeled with phonon drag include the phonon drag contribution to the Seebeck coefficient, whereas others neglect the phonon drag effect. The single impurity case is an extreme case of nanocluster scattering, where one nanocluster only contains one impurity atom. For nanoclusters with more than one impurity atom, we calculate the equivalent diameter corresponding to the total volume of unit cells contained in that nanocluster. We note that the 0.77-nm size contains six unit cells, whereas the 1-nm size contains 14 unit cells. Details of how the nanocluster scattering is treated can be found in *SI Appendix*, *Supplementary Note 6*.

then carry out the integration over the Brillouin zone to obtain scattering rates due to the electron–phonon interaction (these include electron scattering by phonons and phonon scattering by electrons; for electron-impurity scattering see *SI Appendix, Supplementary Note 3*). More simulation details can be found in *SI Appendix*.

Lattice Conduction Calculation. The phonon relaxation time $\tau_{\mathbf{q}\lambda}$ is required for the calculation of both the phonon drag Seebeck coefficient and the thermal conductivity and characterizes the nonequilibrium phonons. Here we briefly summarize the method, and more details are available in the literature (32, 33). Because phonon relaxation times are essentially characterized by the anharmonic force constants (third-order force constants in this calculation), a minimal set of anharmonic force constants is first determined using the symmetry of silicon. They are fitted, based on first-principles data regarding forces acting on different atoms and their displacements in a large supercell ($2 \times 2 \times 2$ conventional unit cells, 64 atoms), with imposed translational and rotational invariances (51). These third-order force constants are then used to calculate the phonon relaxation times due to phonon–phonon scatterings, via Fermi's golden rule. For heavily doped samples, the electron scattering of phonons becomes important and will further reduce the phonon relaxation time especially for long-wavelength phonons (37). The total phonon relaxation time $\tau_{\mathbf{q}\lambda}$ combines both phonon–phonon scattering and phonon–electron scattering (for phonon-impurity scattering see *SI Appendix, Supplementary Note 3*), according to Matthiessen's rule.

Figure of Merit for "Preferable" Phonon Modes. Each phonon mode labeled by wave vector $\mathbf{q}$ and branch number λ makes a contribution to phonon drag given by Eq. **2** and to the thermal conductivity given by $\kappa(\mathbf{q},\lambda) = (1/3\Omega N_{\mathbf{q}})v^2_{\mathbf{q}\lambda}\tau_{\mathbf{q}\lambda}\hbar\omega_{\mathbf{q}\lambda}(\partial n_{\mathbf{q}\lambda}/\partial T)$. The mode-specific figure of merit $\zeta_{\mathbf{q}\lambda}$ is defined as the ratio between them:

$$\zeta_{\mathbf{q}\lambda} = \frac{S_{ph}(\mathbf{q},\lambda)}{\kappa(\mathbf{q},\lambda)} = \frac{2e}{\sigma N_k k_B T^2} \frac{\mathbf{v}_{\mathbf{q}\lambda} n^0_{\mathbf{q}\lambda} \cdot \sum_{\mathbf{k}\alpha,\mathbf{k}'\beta} \left(\tau_{\mathbf{k}\alpha}\mathbf{v}_{\mathbf{k}\alpha} - \tau_{\mathbf{k}'\beta}\mathbf{v}_{\mathbf{k}'\beta}\right) \cdot f^0_{\mathbf{k}\alpha}\left(1 - f^0_{\mathbf{k}'\beta}\right) \frac{2\Pi}{\hbar}}{\mathbf{v}^2_{\mathbf{q}\lambda}\left(\partial n_{\mathbf{q}\lambda}/\partial T\right)}. \quad [3]$$

An upper bound for the thermal conductivity can be described as $\sum_{(\mathbf{q},\lambda)\in C} \kappa(\mathbf{q},\lambda) \leq \kappa_{max}$, where set C denotes the phonon modes that are selected. Given this constraint, the largest possible phonon drag contribution that one can achieve is obtained by selecting those modes that have figures of merit as large as possible, noting that

$$\max\left(\sum_{(\mathbf{q},\lambda)\in C} S_{ph}(\mathbf{q},\lambda)\right) = \max\left(\sum_{(\mathbf{q},\lambda)\in C} \zeta_{\mathbf{q}\lambda} \cdot \kappa(\mathbf{q},\lambda)\right). \quad [4]$$

The figure of merit defined above distinguishes the "preferable" phonon modes that are more significant in phonon drag from those that are less important, and serves as the criterion to select phonons if one seeks to maximize the phonon drag contribution to the Seebeck coefficient.

ACKNOWLEDGMENTS. We thank Jonathan Mendoza for helpful discussions. This article is based upon work supported partially by S^3TEC, an Energy Frontier Research Center funded by the US Department of Energy, Office of Basic Energy Sciences, under Award DE- SC0001299/DE-FG02-09ER46577 (for understanding the coupled electron-phonon transport), and partially by the Air Force Office of Scientific Research Multidisciplinary Research Program of the University Research Initiative (AFOSR MURI FA9550-10-1-0533) via Ohio State University (for examining the potential of utilizing the phonon drag effect for cryogenic cooling).

1. Ziman JM (1960) *Electrons and Phonons: The Theory of Transport Phenomena in Solids* (Clarendon, Oxford).
2. Bloch F (1929) Bemerkung zur Elektronentheorie des Ferromagnetismus und der elektrischen Leitfähigkeit. *Z Phys* 57(7–8):545–555.
3. Lundstrom M (2009) *Fundamentals of Carrier Transport* (Cambridge Univ Press, Cambridge, UK).
4. Gurevich L (1945) Thermoelectric properties of conductors. *J Phys (Moscow)* 9:477–480.
5. Frederikse HPR (1953) Thermoelectric power of germanium below room temperature. *Phys Rev* 92(2):248–252.
6. Geballe TH, Hull GW (1954) Seebeck effect in germanium. *Phys Rev* 94(5):1134–1140.
7. Herring C (1954) Theory of the thermoelectric power of semiconductors. *Phys Rev* 96(5):1163–1187.
8. Geballe TH, Hull GW (1955) Seebeck effect in silicon. *Phys Rev* 98(4):940–947.
9. Puri SM (1965) Phonon drag and phonon interactions in n-InSb. *Phys Rev* 139(3A):A995–A1009.
10. Boxus J, Issi J-P (1977) Giant negative phonon drag thermopower in pure bismuth. *J Phys C Solid State Phys* 10(15):L397.
11. Cohn JL, Wolf SA, Selvamanickam V, Salama K (1991) Thermoelectric power of YBa2Cu3O7- δ: Phonon drag and multiband conduction. *Phys Rev Lett* 66(8):1098–1101.
12. Ohta H, et al. (2007) Giant thermoelectric Seebeck coefficient of a two-dimensional electron gas in SrTiO3. *Nat Mater* 6(2):129–134.
13. Hor YS, et al. (2009) p-type Bi2Se3 for topological insulator and low-temperature thermoelectric applications. *Phys Rev B* 79(19):195208.
14. Bentien A, Johnsen S, Madsen GKH, Iversen BB, Steglich F (2007) Colossal Seebeck coefficient in strongly correlated semiconductor FeSb2. *EPL Europhys Lett* 80(1):17008.
15. Pokharel M, et al. (2013) Phonon drag effect in nanocomposite FeSb2. *MRS Commun* 3(01):31–36.
16. Battiato M, Tomczak JM, Zhong Z, Held K (2015) Unified picture for the colossal thermopower compound $FeSb_2$. *Phys Rev Lett* 114(23):236603.
17. Mahan G, Sales B, Sharp J (2008) Thermoelectric materials: New approaches to an old problem. *Phys Today* 50(3):42–47.
18. Goldsmid HJ (2009) *Introduction to Thermoelectricity* (Springer, New York).
19. Boukai AI, et al. (2008) Silicon nanowires as efficient thermoelectric materials. *Nature* 451(7175):168–171.
20. Weber L, Gmelin E (1991) Transport properties of silicon. *Appl Phys, A Mater Sci Process* 53(2):136–140.
21. Wang G, Endicott L, Chi H, Lošt'ák P, Uher C (2013) Tuning the temperature domain of phonon drag in thin films by the choice of substrate. *Phys Rev Lett* 111(4):046803.
22. Poudel B, et al. (2008) High-thermoelectric performance of nanostructured bismuth antimony telluride bulk alloys. *Science* 320(5876):634–638.
23. Biswas K, et al. (2012) High-performance bulk thermoelectrics with all-scale hierarchical architectures. *Nature* 489(7416):414–418.
24. Tian Z, Lee S, Chen G (2013) Heat transfer in thermoelectric materials and devices. *J Heat Transfer* 135(6):061605.
25. Heikes RR, Ure RW (1961) *Thermoelectricity: Science and Engineering* (Interscience, London).
26. Bailyn M (1958) Transport in metals: Effect of the nonequilibrium phonons. *Phys Rev* 112(5):1587–1598.
27. Brinson ME, Dunstant W (1970) Thermal conductivity and thermoelectric power of heavily doped n-type silicon. *J Phys C Solid State Phys* 3(3):483.
28. Behnen E (1990) Quantitative examination of the thermoelectric power of n-type Si in the phonon drag regime. *J Appl Phys* 67(1):287–292.
29. Cantrell DG, Butcher PN (1987) A calculation of the phonon-drag contribution to the thermopower of quasi-2D electrons coupled to 3D phonons. I. General theory. *J Phys C Solid State Phys* 20(13):1985.
30. Cantrell DG, Butcher PN (1987) A calculation of the phonon-drag contribution to the thermopower of quasi-2D electrons coupled to 3D phonons. II. Applications. *J Phys C Solid State Phys* 20(13):1993.
31. Mahan GD, Lindsay L, Broido DA (2014) The Seebeck coefficient and phonon drag in silicon. *J Appl Phys* 116(24):245102.
32. Esfarjani K, Chen G, Stokes H (2011) Heat transport in silicon from first-principles calculations. *Phys Rev B* 84(8):085204.
33. Ward A, Broido DA (2010) Intrinsic phonon relaxation times from first-principles studies of the thermal conductivities of Si and Ge. *Phys Rev B* 81(8):085205.
34. Giustino F, Cohen M, Louie S (2007) Electron-phonon interaction using Wannier functions. *Phys Rev B* 76(16):165108.
35. Marzari N, Mostofi AA, Yates JR, Souza I, Vanderbilt D (2012) Maximally localized Wannier functions: Theory and applications. *Rev Mod Phys* 84(4):1419–1475.
36. Ziman JM (1956) XVII. The effect of free electrons on lattice conduction. *Philos Mag* 1(2):191–198.
37. Liao B, et al. (2015) Significant reduction of lattice thermal conductivity by the electron-phonon interaction in silicon with high carrier concentrations: A first-principles study. *Phys Rev Lett* 114(11):115901.
38. Smith MJ, Butcher PN (1990) Simple models of phonon-drag in 3D and quasi-2D. *J Phys Condens Matter* 2(10):2375.
39. Van Herwaarden AW, Sarro PM (1986) Thermal sensors based on the seebeck effect. *Sens Actuators* 10(3–4):321–346.
40. Trzcinski R, Gmelin E, Queisser HJ (1986) Quenched phonon -drag in silicon microcontacts. *Phys Rev Lett* 56(10):1086–1089.
41. Pei Y, et al. (2011) Convergence of electronic bands for high performance bulk thermoelectrics. *Nature* 473(7345):66–69.
42. Klemens PG (1955) The scattering of low-frequency lattice waves by static imperfections. *Proc Phys Soc A* 68(12):1113.
43. MacDonald DKC (1962) *Thermoelectricity: An Introduction to the Principles* (Wiley, New York).
44. Klemens PG (1996) Theory of phonon drag thermopower. *Proceedings of the Fifteenth International Conference on Thermoelectrics*, ed Caillat T (IEEE, New York), pp 206–208.
45. Hochbaum AI, et al. (2008) Enhanced thermoelectric performance of rough silicon nanowires. *Nature* 451(7175):163–167.
46. Dismukes JP, Ekstrom L, Steigmeier EF, Kudman I, Beers DS (1964) Thermal and electrical properties of heavily doped Ge-Si alloys up to 1300°K. *J Appl Phys* 35(10):2899–2907.
47. Slack GA, Hussain MA (1991) The maximum possible conversion efficiency of silicon-germanium thermoelectric generators. *J Appl Phys* 70(5):2694–2718.
48. Giannozzi P, et al. (2009) QUANTUM ESPRESSO: A modular and open-source software project for quantum simulations of materials. *J Phys Condens Matter* 21(39):395502.
49. Perdew JP, Zunger A (1981) Self-interaction correction to density-functional approximations for many-electron systems. *Phys Rev B* 23(10):5048–5079.
50. Noffsinger J, et al. (2010) EPW: A program for calculating the electron–phonon coupling using maximally localized Wannier functions. *Comput Phys Commun* 181(12):2140–2148.
51. Esfarjani K, Stokes HT (2008) Method to extract anharmonic force constants from first principles calculations. *Phys Rev B* 77(14):144112.
52. Tošić TI, Tjapkin DA, Jevtić MM (1981) Mobility of majority carriers in doped noncompensated silicon. *Solid-State Electron* 24(6):577–582.