Electrothermally driven current vortices in inhomogeneous bipolar semiconductors

D. Fu,[1,2,3] A. X. Levander,[1,4] R. Zhang,[2,3] J. W. Ager III,[4] and J. Wu[1,4,*]

[1]*Department of Materials Science and Engineering, University of California, Berkeley, CA 94720, USA*

[2]*Jiangsu Provincial Key Laboratory of Advanced Photonic and Electronic Materials, School of Electronic Science and Engineering, Nanjing University, Nanjing 210093, China*

[3]*Nanjing National Laboratory of Microstructures, Nanjing 210093, China*

[4]*Materials Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA*

(Received 9 February 2011; revised manuscript received 23 March 2011; published 8 July 2011)

We report an effect that occurs in semiconductors where internal electrical fields interact with a temperature gradient. Steady current vortices and a magnetic field develop in the system, even without external carrier injection. The effect is electrodynamic, energy dissipative, and fundamentally distinct from any previously described electrothermal effects. In bipolar structures the effective thermopower can be significantly modified by the vortices. Joule heating arising from the vortices reduces the thermal conductivity by an amount comparable to the electronic thermal conductivity.

DOI: 10.1103/PhysRevB.84.045205 PACS number(s): 72.20.Pa, 73.50.Lw

I. INTRODUCTION

Free carriers motion in a semiconductor can be driven by electric fields ($\boldsymbol{E}$), magnetic fields ($\boldsymbol{B}$), and/or temperature gradients (∇T).[1] Two of these fields applied in orthogonal directions generate a third field in the direction perpendicular to both, which lead to four transverse electrothermomagnetic effects, known as the Hall effect ($E_x \times B_y \rightarrow E_z$), the Righi–Leduc effect ($\partial_x \mathrm{T} \times B_y \rightarrow \partial_z \mathrm{T}$), the Nernst effect ($\partial_x T \times B_y \rightarrow E_z$), and the Ettingshausen effect ($E_x \times B_y \rightarrow \partial_z \mathrm{T}$).[2] A common feature of all these four known effects is that $\boldsymbol{B}$ is on the input side to exert the Lorentz force, which induces the transverse ∇T or $\boldsymbol{E}$. The possibility of inducing a transverse $\boldsymbol{B}$, rather than using it as an input, had been rarely investigated. Using a generic hydrodynamic model, Mohseni *et al.*[3] recently showed that electron-flow vortices may arise in inhomogeneous semiconductors driven by a net torque, which is caused by the fact that inside a differential charge flow element, $\boldsymbol{E}$ acts on the center of charge whereas ∇T acts on the center of mass of the free carriers. In this work, by applying a detailed electrodynamic drift-diffusion model to a practical device structure, we show that when a ∇T is applied in perpendicular to an $\boldsymbol{E}$ field, steady current vortices develop in the system even in the absence of external carrier injection (i.e. open circuit condition). The vortices then give rise to a transverse $\boldsymbol{B}$ field ($\partial_x T \times E_y \rightarrow B_z$), defining an effect that is electrodynamic and energy dissipative. This is fundamentally distinct from the previously described four electrothermomagnetic effects which are all electrostatic in the transverse direction in the steady state.

We show that such current vortices significantly modify the effective thermopower (Seebeck coefficient) in bipolar structures, and Joule heating arising from the vortices can reduce the thermal conductivity by an amount comparable to the electronic thermal conductivity. Internal $\boldsymbol{E}$ fields naturally occur in inhomogeneous materials, which are frequently encountered in the search for high thermopower and simultaneous high electrical conductivity in modern thermoelectric devices.[4–6] Moreover, surface or grain-boundary Fermi-level pinning and composition fluctuation[7] can also render the system electronically inhomogeneous. With potentially high thermoelectric performance from nanostructures[8–11] and nanocomposites,[12,13] it is critically important to understand the interaction between ∇T and $\boldsymbol{E}$ fields, and the consequent electrothermal properties in inhomogeneous semiconductor structures.

II. SIMULATION DETAILS

We consider a prototypical case where the $\boldsymbol{E}$ field is internal and produced by doping variation: a bipolar semiconductor structure with an *N*-type and *P*-type region. In isothermal equilibrium, gradient in free charge carrier density drives the diffusion current, redistributes the carriers, and develops the internal built-in $\boldsymbol{E}$ field, until it is balanced by the drift current driven by the $\boldsymbol{E}$ field. However, when a temperature gradient is present, these local currents are not necessarily balanced and cancelled by the ∇T driven current, because in general they can have nonzero components orthogonal to each other. In steady state, the net current in an *N*-type, isotropic but inhomogeneously doped semiconductor is,[14]

$$\begin{aligned}\mathbf{j}_n(\mathbf{r}) &= -\sigma_n(\mathbf{r})\nabla\varphi(\mathbf{r}) - eD_n(\mathbf{r})\nabla n(\mathbf{r}) - \sigma_n(\mathbf{r})S_n(\mathbf{r})\nabla T(\mathbf{r}) \\ &= -\sigma_n(\mathbf{r})[\nabla E_{Fn}(\mathbf{r})/e + S_n(\mathbf{r})\nabla T(\mathbf{r})] \end{aligned} \quad (1)$$

where $E_{Fn}(\mathbf{r})$ is the electron quasi-Fermi level (electrochemical potential) taking into account both the internal $\boldsymbol{E}$ field $\mathbf{E}(\mathbf{r}) = -\nabla\varphi(\mathbf{r})$ and the electron density gradient $\nabla n(\mathbf{r})$; $\sigma_n(\mathbf{r})$ and $S_n(\mathbf{r})$ are the local electrical conductivity and thermopower, respectively. The electrochemical *emf*, defined by the difference in $E_{Fn}(\mathbf{r})$ between the measured points *A* and *B*, is

$$\begin{aligned} emf_{AB} &= \int_A^B \nabla E_{Fn}(\mathbf{r})d\mathbf{l} \\ &= \int_A^B \frac{\mathbf{j}_n(\mathbf{r})d\mathbf{l}}{\sigma_n(\mathbf{r})} + \int_A^B S_n(\mathbf{r})\nabla T(\mathbf{r})d\mathbf{l}. \end{aligned} \quad (2)$$

It should be noted that this emf_{AB} is the total electrically $[\nabla\varphi(\mathbf{r})]$ and chemically $[\nabla n(\mathbf{r})]$ driven electromotive force and is responsible for the Seebeck voltage builtup between points *A* and *B*. The net current is driven by this electrochemical *emf* in conjunction with the thermal *emf*, the latter being

the second term in Eq. (2), $\int_A^B S_n(\mathbf{r})\nabla T(\mathbf{r})d\mathbf{l}$. For simplicity, emf_{AB} in this paper refers to the electrochemical *emf* defined in Eq. (2). In the presence of internal current $\mathbf{j}_n(\mathbf{r})$, the effective thermopower $S_{AB} = emf_{AB}/(T_B - T_A)$ is generally no longer equal to the simple spatial average of local thermopower $S_n(\mathbf{r})$ as given merely by the second term in Eq. (2). For bipolar conduction, Eq. (2) is extended by including both electron and hole conduction, as discussed in Ref. 14

$$emf_{AB} = \int_A^B \frac{[\mathbf{j}_n(\mathbf{r}) + \mathbf{j}_p(\mathbf{r})]d\mathbf{l}}{\sigma_n(\mathbf{r}) + \sigma_p(\mathbf{r})} + \int_A^B \frac{\sigma_n(\mathbf{r})S_n(\mathbf{r}) + \sigma_p(\mathbf{r})S_p(\mathbf{r})}{\sigma_n(\mathbf{r}) + \sigma_p(\mathbf{r})}\nabla T(\mathbf{r})\, d\mathbf{l}. \quad (3)$$

It is shown below that in inhomogeneous structures under thermal bias, the current term is generally nonzero due to the interaction between the built-in $\boldsymbol{E}$ field and ∇T. When the size of the system is comparable to the carrier depletion length, the effective thermopower is significantly different from what is expected in the static bipolar model given by the second term in Eq. (3), necessitating a dynamic electrothermal model to fully describe the system.

A Si *P-N* junction with a transverse thermal bias, is shown in Fig. 1(a). In contrast to previously simulated *P-N* junctions

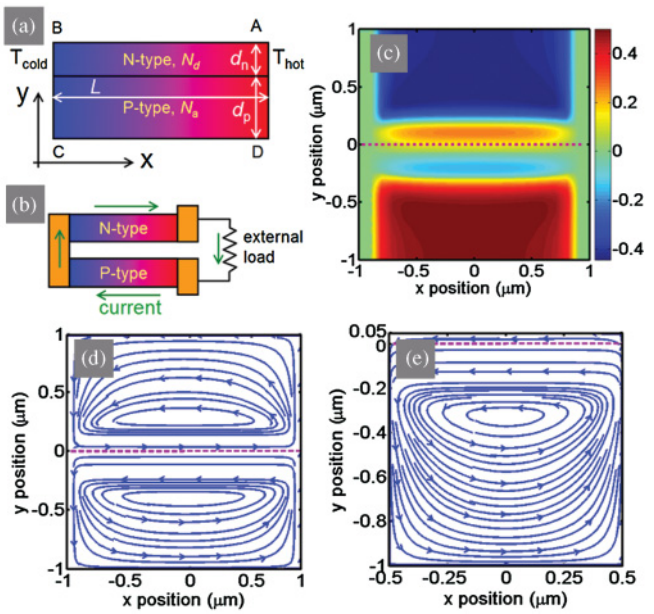

FIG. 1. (Color online) (a) Schematic of the bipolar structure used in the simulation. The physical interface between *N* and *P* layers is at $y = 0$. (b) A Π-shaped thermoelectric module with external load for comparison. When this module is shorted, the current forms an internal loop. The shorted module is similar to the structure in (a). (c) Simulated *x*-direction electric field distribution (in units of V/cm) for a junction biased with $dT/dx = 100$ K/cm and averaged temperature $T_0 = 300$ K. The system is electrically isolated (open circuit). The parameters are: $d_n = 1\ \mu$m, $N_d = 3 \times 10^{16}$ cm^{-3}, $d_p = 1\ \mu$m, $N_a = 6.25 \times 10^{15}$cm^{-3}, and $L = 2\ \mu$m. (d) Electric current density plot for the structure. (e) Electric current density for another junction with $d_n \approx w_n$. Here the parameters are: $d_n = 50$ nm, $N_d = 3 \times 10^{16}$ cm^{-3}, $d_p = 1\ \mu$m, $N_a = 6.25 \times 10^{15}$ cm^{-3}, and $L = 1\ \mu$m. Note the double loops in (d) versus single loop in (e).

with a longitudinal[15] or local transverse thermal bias,[16] here the temperature variation is macroscopic and is not limited to the direction parallel to the built-in $\boldsymbol{E}$ field, so that the thermal driving force [the ∇T term in Eq. (1)] for carrier motion cannot be balanced by electrochemical driving forces [the ∇E_{Fn} term in Eq. (1)]. Similar structures were proposed for thermoelectric power generation,[17] but the formation of internal currents and their effect on the thermopower and thermal conductivity were not explored. We calculated the distribution of electric potential and quasi-Fermi energy by solving the coupled Poisson and electric current continuity equations in 2D. The electric potential $\varphi(\mathbf{r})$ satisfies the Poisson's equation,

$$\varepsilon\nabla^2\varphi(\mathbf{r})/|e| = -\{\pm N_{d,a} - n[\varphi(\mathbf{r})] + p[\varphi(\mathbf{r})]\}, \quad (4)$$

where N_d (positive sign) or N_a (negative sign) is the constant concentration of donors in the N-type region or acceptors in the P-type region, respectively, which are assumed to be fully ionized at the simulated temperature ($T \sim 300$K). The carrier concentrations redistribute in space because the local electric potential modulates the carrier population of the conduction and valence bands, for example,

$$n[\varphi(\mathbf{r})] = \int \frac{\rho_C[E - e\varphi(\mathbf{r})]}{1 + \exp[(E - E_{Fn})/k_BT]}dE, \quad (5)$$

where $\rho_c(E)$ is the density of states for the conduction band. Full Fermi–Dirac carrier statistics are used such that the calculation is valid across all concentrations ranging from nondegenerate to degenerate. The local electron current density is given by Eq. (1), and the local $\sigma_n(\mathbf{r})$ and $S_n(\mathbf{r})$ are calculated from the solution to the Boltzmann transport equation under the relaxation time approximation.[1] The dependence of relaxation time τ_n on electron energy E follows $\tau_n(E) \sim E^\beta$, where the exponent β is taken to be $-1/2$, assuming acoustic phonon dominated scattering mechanism,[18] and the proportional constant is obtained by fitting the calculated conductivity to experimental data. It should be emphasized that we ignored the phonon drag contribution to the Seebeck coefficient because generally it becomes significant only at low temperatures and in lightly doped semiconductors.[19] A similar treatment is adopted for free-hole conduction. The temperature gradient ∇T is set to be constant and along the x direction only, i.e. the thermal transport is assumed to be dominated by the high lattice thermal conduction, such that the effect of free carrier redistribution on ∇T is neglected. The system is electrically isolated from external circuit, so that no electron or hole current is allowed to flow in and out of the boundaries. The continuity equation is separately imposed for electron and hole carrier flow, so as to ensure the continuity of total current density,

$$-\frac{1}{|e|}\nabla\mathbf{j}_n(\mathbf{r}) + R = \frac{1}{|e|}\nabla\mathbf{j}_p(\mathbf{r}) + R = 0, \quad (6)$$

where nonequilibrium electrons and holes recombine nonradiatively at the rate R via midgap traps (the Shockley–Read–Hall mechanism)[20] and the Auger process.[21] Surface recombination rate is neglected because, as shown below, the effect occurs mostly in bulk. A finite difference method with nonuniform meshing is used to numerically solve the problem self-consistently.[22]

III. RESULTS AND DISCUSSION

A. Current vortices

Figure 1(c) shows the simulated x-component electric field in steady state $E_x(\mathbf{r}) = \partial_x\varphi(\mathbf{r})$ for a geometry in which both the N and P layers are relatively thick. It can be seen that at the top and bottom surfaces, a nearly constant $\boldsymbol{E}$ field is established in response to ∇T, with field direction depending on the type of doping. This behavior is identical to that of a homogeneous semiconductor and can be understood because the top and bottom surfaces are far from the space charge region of the P-N junction. However, deep into the structure and near the junction area, regions with electric field pointing in the "wrong" direction are found. This complicated electric field pattern is a direct consequence of the imbalanced electro-chemical-thermal driving forces for the free charge carriers. The total steady-state current $\mathbf{j}(\mathbf{r})$ resulting from this force imbalance is shown in Fig. 1(d). The system exhibits two current vortices distributed in the N- and P-type regions, respectively. Both vortices are counterclockwise and are nearly isolated from each other by the highly resistive space charge region at the P-N junction. It is interesting to note that at the top and bottom surfaces, the current flow direction is opposite to what one would expect for a short-circuited homogeneous semiconductor; namely, free electrons and holes flow against instead of along ∇T. This is due to the fact that the space-charge region deep in the junction area has lower free-carrier concentration than the charge neutral region, and therefore generates stronger *emf* to drive the free carriers to move along $-\nabla T$; but because the neutral region is much more electrically conductive than the junction, the current forms a loop within the P or N layer by driving the carriers against $-\nabla T$ at the top and bottom surface. The development of current vortices in this open-circuited P-N structure is not surprising, considering its similarity to the regular Π-shape thermoelectric module as shown in Fig. 1(b). In the Π-shape thermoelectric module harvesting waste heat into electricity, a P-type semiconductor and an N-type semiconductor are connected in parallel bridging a temperature difference. The P and N arms are electrically shorted at one end (for example, the cold side), while the other end (hot side) output current to the external load. If, however, on the hot side the two arms are also directly electrically shorted to each other, this Π structure would become open to external circuit and be similar to the structure we simulated in Fig. 1(a). It is immediately clear that a current loop would form within this Π structure despite that it is open to external circuit. However, differing from Fig. 1(b), the simulated structure in Fig. 1(a) has a P-N junction interfacing the P and N arms along the entire device length. As discussed above, it is this charge-depleted P-N junction (and its built-in $\boldsymbol{E}$ field) that further redistributes the current flow and causes two vortices as shown in Fig. 1(d), instead of a single vortex as expected from the simple analogy to the Π structure in Fig. 1(b).

B. Effects on thermopower

Despite this unusual current flow, the *emf* on the surfaces is still normal. The majority carrier quasi-Fermi level E_{Fn} for the top surface is equal to that of an isolated, open-circuited homogeneous semiconductor doped at N_d. The effective thermopower $S_{AB} = emf_{AB}/(T_B - T_A)$, calculated using both terms in Eq. (3), is nearly identical to $S_{AB}^{\text{static}} = emf_{AB}^{\text{static}}/(T_B - T_A)$ calculated in the static bipolar model given solely by the second term in Eq. (3),

$$emf_{AB}^{\text{static}} = \int_A^B \frac{\sigma_n(\mathbf{r})S_n(\mathbf{r}) + \sigma_p(\mathbf{r})S_p(\mathbf{r})}{\sigma_n(\mathbf{r}) + \sigma_p(\mathbf{r})}\nabla T(\mathbf{r})d\mathbf{l}. \quad (7)$$

This emf_{AB}^{static} can be easily obtained by solving the isothermal and thus static Poisson equation of the structure for $\sigma(\mathbf{r})$ and $S(\mathbf{r})$ of electrons and holes. The agreement between emf_{AB}^{static} and emf_{AB} is not surprising since the measured points on the surface, A and B, are far from the junction area. The quasi-electric field [not the real electric field $\mathbf{E}(\mathbf{r})$] that causes the deviation of emf_{AB} from emf_{AB}^{static}, which is defined as

$$\mathbf{E}'(\mathbf{r}) \equiv \frac{\mathbf{j}_n(\mathbf{r}) + \mathbf{j}_p(\mathbf{r})}{\sigma_n(\mathbf{r}) + \sigma_p(\mathbf{r})}, \quad (8)$$

is very weak along the integration path from point A to B because the surface is not depleted and thus $\sigma_n(\mathbf{r})$ is large. This agreement, however, vanishes when the thickness of each layer is comparable to its carrier depletion length, because $\mathbf{E}'(\mathbf{r})$ on the surface now becomes considerably stronger. The carrier depletion length on the N side of the P-N junction can be estimated by $w_n = \sqrt{2\varepsilon V_{bi}N_a/[|e|N_d(N_a + N_d)]}$, where $V_{bi} = k_BT\ln(N_aN_d/n_1^2)/|e|$ is the built-in voltage across the junction and n_i is the intrinsic carrier concentration. Figure 1(e) plots $\mathbf{j}(\mathbf{r})$ for such a structure with N-type layer thickness (50 nm) comparable to its depletion width (~75 nm). The effective thermopower calculated is $S_{AB} = -0.55$ compared to $S_{AB}^{\text{static}} = -0.24\text{mV/K}$, a difference of more than a factor of 2. We thereby name the full calculation of S_{AB}, including both terms in Eq. (3) as the dynamic electrothermal model.

Figures 2(a) and 2(b) compare S_{AB} and S_{AB}^{static} as a function of N_d and d_n. When either N_d or d_n is large so that the measured points A and B are far from the depletion region, $S_{AB} \approx S_{AB}^{\text{static}}$. If N_d or d_n decreases so that the surface is partially depleted, S_{AB} deviates significantly from S_{AB}^{static}, invalidating the commonly adopted static bipolar model. However, when N_d or d_n further decreases such that d_n is much smaller than w_n, the N-type doped surface is inverted to unipolar P-type conduction. In this case $\mathbf{E}'(\mathbf{r})$ along the top surface becomes weak again, so that S_{AB} returns to S_{AB}^{static}, which now approaches that of the P layer in the homogeneous limit. Therefore, three regimes exist in S_{AB}: Regime I, $d_n \gg w_n$, $S_{AB} \approx S_{AB}^{\text{static}} = S_{n,\text{bulk}}$; Regime II, $d_n \sim w_n$, S_{AB} is unequal to, and more negative than $S_{AB}^{\text{static}} = (\sigma_nS_{n,\text{bulk}} + \sigma_pS_{p,\text{bulk}})/(\sigma_n + \sigma_p)$ expected from the static bipolar model; and Regime III, $d_n \ll w_n$, $S_{AB} \approx S_{AB}^{\text{static}} = S_{p,\text{bulk}}$. Here $S_{n,\text{bulk}}$ and $S_{p,\text{bulk}}$ are the thermopower calculated for electrons and holes with the surface carrier concentrations but in the homogeneous bulk limit. In Fig. 2 we also show the prediction from a simple bilayer model that is often used to treat multilayer metallic structures, in which the effective thermopower is taken to be the bulk thermopower of each layer weighted by its sheet conduction.[23] It can be seen that the bilayer model gives a poor prediction, especially in Regime II, because it completely neglects charge redistribution across the interface

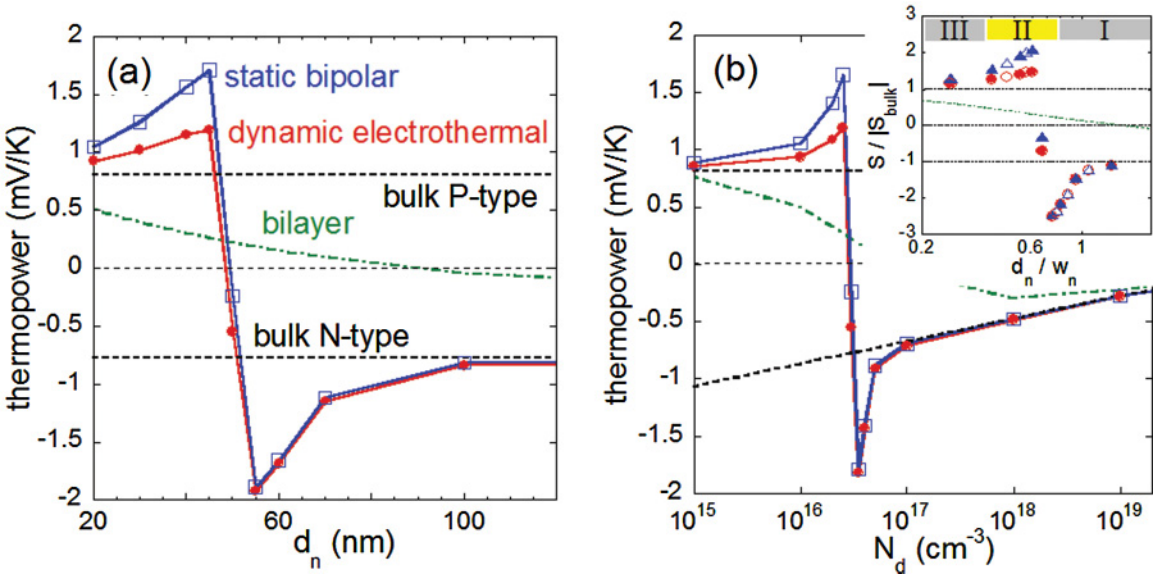

FIG. 2. (Color online) (a) Effective thermopower from the N-side surface as a function of N-Si thickness for the electrothermal, bipolar, and bilayer models. The fixed parameters are: $N_d = 3 \times 10^{16}$ cm^{-3}, $d_p = 1$ μm, $N_a = 6.25 \times 10^{15}$ cm^{-3}, $L = 1$ μm, $dT/dx = 100$ K/cm and $T_{\text{average}} = 300$ K. (b) As a function of N-Si doping concentration N_d. The N-layer thickness is $d_n = 50$ nm, and all other parameters are the same as in (a). The legend is also the same as in (a). Inset is a universal scaling relationship: thermopower normalized by the asymptotic bulk values (N type when negative and P type when positive) as a function of the N-layer thickness normalized by the N-side depletion width. Solid symbols are for varying thickness d_n and empty symbols are for varying doping N_d. Regimes I, II, and III are labeled.

between layers. It is intriguing to notice the critical role of depletion length w_n in gauging the size effect of S_{AB} in both scenarios of varying N_d and d_n. In the inset of Fig. 2(b) the normalized thermopower, $S_{AB}/|S_{\text{bulk}}|$, is plotted as a function of normalized thickness, d_n/w_n. It can be seen that the two sets of curves in Figs. 2(a) and 2(b) collapse onto a universal thickness dependence, regardless of whether d_n/w_n is varied by directly changing d_n or by varying w_n through N_d. It is therefore clear that it is the interaction between the orthogonal built-in $\boldsymbol{E}$ field and ∇T that invalidates the static bipolar model.

C. Conservativeness of electric and temperature fields

We note that in inhomogeneous structures involving bipolar charge conduction on the surface, both emf_{AB} and emf_{AB}^{static} are integration-path dependent.[14] Namely, the integrands in both Eqs. (3) and (7) are nonconservative vector fields, and the loop integration of both along the closed surface loop A-B-C-D-A results in a nonzero value. Their difference $emf_{\text{loop}} \equiv \oint \mathbf{E}'(\mathbf{r})d\mathbf{l}$ along the sample surface and the consequently defined $\Delta S \equiv emf_{\text{loop}}/\Delta T$ is a characteristic thermopower correction to the entire inhomogeneous structure. ΔS can thus be used to represent the effect of inhomogeneity on thermopower of the system.

We also note that the existence of the current vortex indicates a nonzero $\nabla \times \mathbf{j}(\mathbf{r})$ and a nonconservative $\mathbf{j}(\mathbf{r})$ field. However, this does not imply a nonconservative $\mathbf{E}(\mathbf{r}) = -\nabla\varphi(\mathbf{r})$ field or $\nabla T(\mathbf{r})$ field. In fact, both $\nabla\varphi(\mathbf{r})$ and $\nabla T(\mathbf{r})$ are still conservative. For example, from Eq. (1) it can be seen that the Curl of the first term of the current density is

$$\nabla \times [\sigma_n(\mathbf{r})\nabla\varphi(\mathbf{r})] = \nabla\sigma_n(\mathbf{r}) \times \nabla\varphi(\mathbf{r}) + \sigma_n(\mathbf{r})\nabla \times [\nabla\varphi(\mathbf{r})]. \tag{9}$$

Therefore, even though $\nabla \times [\nabla\varphi(\mathbf{r})] = 0$ due to the conservativeness of $\nabla\varphi(\mathbf{r})$, Eq. (9) would still give nonzero $\nabla \times \mathbf{j}(\mathbf{r})$, as long as the vector $\nabla\sigma_n(\mathbf{r})$ is nonzero and has a component perpendicular to $\nabla\varphi(\mathbf{r})$. This could occur for a spatially inhomogeneous system such as the simulated one, where the built-in $\boldsymbol{E}$ field of inhomogeneity redistributes free carriers and causes nonzero $\nabla\sigma_n(\mathbf{r})$. Considering all the current density terms in Eq. (1), $\nabla \times \mathbf{j}(\mathbf{r})$ will include terms of $\nabla\sigma_n(\mathbf{r}) \times \nabla\varphi(\mathbf{r})$ and $\nabla\sigma_n(\mathbf{r}) \times \nabla T(\mathbf{r})$. When $\nabla T(\mathbf{r})$ is zero or nonzero but applied in parallel to $\nabla\varphi(\mathbf{r})$, the system can still find an electrostatic equilibrium where $\nabla\sigma_n(\mathbf{r})$ is adjusted to be always parallel to both $\nabla\varphi(\mathbf{r})$ and $\nabla T(\mathbf{r})$ at all position $\mathbf{r}$, hence $\nabla \times \mathbf{j}(\mathbf{r}) = 0$ and no current vortex is developed. However, if $\nabla T(\mathbf{r})$ is applied not in parallel to $\nabla\varphi(\mathbf{r})$, the developed $\nabla\sigma_n(\mathbf{r})$ cannot be simultaneously parallel to both $\nabla\varphi(\mathbf{r})$ and $\nabla T(\mathbf{r})$, hence the Curl of the total current density given by Eq. (1) must be nonzero, resulting in current vortices. In systems where the conduction is bipolar such as the case shown in Fig. 1(a), this process is further complicated by the existence of two types of carriers, but the current vortices originate from the same mechanism.

D. Effects on thermal conductivity

The current vortices also generate Joule heat in the structure, which effectively reduces the rate of heat transfer along $-\nabla T$. The original electronic contribution to thermal conductivity of the structure $\kappa_{\text{electronic}}$ is calculated using the Wiedemann–Franz law,

$$\kappa_{\text{electronic}} = L_0\overline{\sigma}T_0 = L_0T_0\frac{\int_{-d_p}^{d_n}[\sigma_n(y) + \sigma_p(y)]dy}{d_n + d_p}, \tag{10}$$

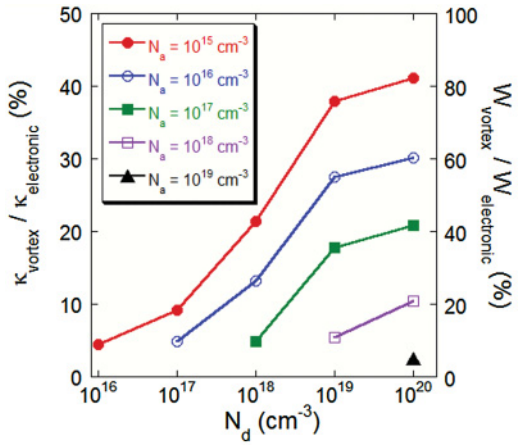

FIG. 3. (Color online) Reduction in thermal conductivity caused by the current vortices normalized by electronic thermal conductivity. The structure is length $L = 2$ cm, d_n and d_p equal to half of the depletion width on the N and P layers, respectively; $dT/dx = 100$ K/cm and $T_0 = 300$ K. W_{vortex} and $W_{\text{electronic}}$ are Joule heat generated by the vortices and heat transported by charge carriers in the structure, respectively.

where L_0 is the Lorenz number[24] and T_0 is averaged temperature (300 K). The electric conductivity $\bar{\sigma}$ is averaged only along the y direction because of the parallel nature of the inhomogeneity in Fig. 1(a). We follow the treatment in Ref. 25 which assumes that half of the internal Joule heat ($W_{\text{vortex}}/2$) flows toward the hot side and the other half toward the cold side; the amount of W_{vortex} out of the originally transferred heat $\kappa_{\text{total}} A|dT/dx|$ is redistributed such that now only $\kappa_{\text{total}} A|dT/dx| - W_{\text{vortex}}/2$ is completely transferred from the hot to the cold side, i.e.,

$$\kappa'_{\text{total}} A|dT/dx| = \kappa_{\text{total}} A|dT/dx| - W_{\text{vortex}}/2. \qquad (11)$$

This defines a Joule-heating induced reduction in thermal conductivity as $\kappa_{\text{vortex}} = W_{\text{vortex}}/(2A|dT/dx|)$. At fixed ΔT, this reduction scales linearly with the length of the structure. In Fig. 3 we plot this reduction κ_{vortex} normalized by the original electronic thermal conductivity $\kappa_{\text{electronic}}$ as a function of doping. It can be seen that the negative contribution κ_{vortex} can become comparable to $\kappa_{\text{electronic}}$ when N_d and N_a are extremely asymmetric. It is noted that for bulk Si, both $\kappa_{\text{electronic}}$ and κ_{vortex} are much lower than the lattice contribution to the thermal conductivity κ_{lattice}, so they are negligible. This justifies our assumption of constant ∇T along the structure for the simulation. However, recently it was shown that κ_{lattice} may be drastically reduced via nanostructuring such as formation of nanowires, nanoporous structures, and nanomeshes.[8,9,26,27] In these efforts, when κ_{lattice} is reduced to the level comparable to or below that of $\kappa_{\text{electronic}}$ in inhomogeneous thermoelectric systems, $\kappa_{\text{electronic}}$ and κ_{vortex} become considerable. Our simulated results in Fig. (3) suggest that in these nanostructures, the current vortices arising from inhomogeneities may start to strongly influence not only the thermopower, but also the electronic thermal transport.

IV. SUMMARY AND EXTENDED DISCUSSION

When the magnetic field arising from these current vortices is considered, this effect has similarities to the Nernst effect, in which a sample subjected to orthogonal ∇T and $\boldsymbol{B}$ field develops a lateral $\boldsymbol{E}$ field in the direction normal to both.[2] However, in contrast to the Nernst effect in open circuit which is static and second order, the new effect is electrodynamic, energy-dissipative, and first order. It should be noted that a possible Hall effect induced in turn by the generated magnetic field itself was neglected in the calculation because it's a second-order effect proportional to $|\mathbf{j}|^2$. The predicted results can also be extended to more complex and unipolar structures. It can be expected that for any open-circuited inhomogeneous structure, internal current vortices always exist as long as ∇T deviates from the direction of inhomogeneity gradient (doping, composition, Fermi-level pinning, or extended defects). However, if the inhomogeneities are deep inside the bulk such that the sample surface is electrostatically screened from their built-in electric field, the thermopower measured from the surface is equal to that of a homogeneous material with the same carrier concentration as on the surface. If the sample size is reduced such that the surface is within the depletion region, the dynamic electrothermal model developed here is needed to understand the effective thermopower. The thermopower in inhomogeneously doped semiconductors is fundamentally determined by electrothermal process in the near surface region. As such, the apparent thermopower depends on the configurational details of inhomogeneity and cannot be predicted from a simple effective medium approximation with knowledge of only the volumetric fraction of each constituent.

ACKNOWLEDGMENTS

This work was supported in part by the National Science Foundation under Grant No. CMMI-1000176 [modeling software]. D. Fu and R. Zhang acknowledge the support by Special Funds for Major State Basic Research Project (Grant No. 2011CB301901) and National Nature Science Foundation of China (Grant No. 60990311). D. Fu also acknowledges the special support of the Graduate Student Research Innovation Project of Jiangsu Province of China (Grant No. CX09B 009Z).

*To whom correspondence should be addressed: wuj@berkeley.edu

[1] J. Cai and G. D. Mahan, Phys. Rev. B **74**, 075201 (2006).
[2] M. R. El-Saden and F. W. Thomas, J. Appl. Phys. **36**, 181 (1965).
[3] K. Mohseni, A. Shakouri, R. J. Ram, and M. C. Abraham, Phys. Fluids **17**, 100602 (2005).
[4] G. D. Mahan and L. M. Woods, Phys. Rev. Lett. **80**, 4016 (1998).
[5] D. J. Bergman and L. G. Fel, J. Appl. Phys. **85**, 8205 (1999).
[6] Z. Bian, H. Wang, Q. Zhou, and A. Shakouri, Phys. Rev. B **75**, 245208 (2007).
[7] S. Bhattacharyya, R. P. Hermann, V. Keppens, T. M. Tritt, and G. J. Snyder, Phys. Rev. B **74**, 134108 (2006).

[8] A. I. Hochbaum, R. Chen, R. D. Delgado, W. Liang, E. C. Garnett, M. Najarian, A. Majumdar, and P. D. Yang, Nature **451**, 163 (2008).
[9] R. Chen, A. I. Hochbaum, P. Murphy, J. Moore, P. Yang, and A. Majumdar, Phys. Rev. Lett. **101**, 105501 (2008).
[10] A. I. Boukai, T. Bunimovich, J. Tahir-Kheli, J. K. Yu, W. A. Goddard, and J. R. Heath, Nature **451**, 168 (2008).
[11] J. K. Yu, S. Mitrovic, D. Tham, J. Varghese, and J. R. Heath, Nature Nanotech. **5**, 718 (2010).
[12] G. Zeng, J. M. O. Zide, W. Kim, J. E. Bowers, A. C. Gossard, Z. Bian, Y. Zhang, A. Shakouri, S. L. Singer, and A. Majumdar, J. Appl. Phys. **101**, 034502 (2007).
[13] D. K. Ko, J. J. Urban, and C. B. Murray, Nano Lett. **10**, 1842 (2010).
[14] Y. G. Gurevich, O. Y. Titov, G. N. Logvinov, and O. I. Lyubimov, Phys. Rev. B **51**, 6999 (1995).
[15] K. P. Pipe, R. J. Ram, and A. Shakouri, Phys. Rev. B **66**, 125316 (2002).
[16] H. K. Lyeo, A. A. Khajetoorians, L. Shi, K. P. Pipe, R. J. Ram, A. Shakouri, and C. K. Shih, Science **303**, 816 (2004).
[17] M. Wagner, G. Span, S. Holzer, and T. Grasser, Semicond. Sci. Technol. **22**, S173 (2007).
[18] D. M. Rowe, *CRC Handbook of Thermoelectrics* (CRC Press, Inc., 1994), p. 27.
[19] L. Weber and E. Gmelin, Appl. Phys. A **53**, 136 (1991).
[20] J. G. Fossum, Solid State Electronics **19**, 269 (1976).
[21] J. Dziewior and W. Schmid, Appl. Phys. Lett. **31**, 346 (1977).
[22] W. L. Engl, H. K. Dirks, and B. Meinerzhagen, Proc. IEEE **71**, 10 (1983).
[23] A. Kyarad and H. Lengfellner, Appl. Phys. Lett. **85**, 5613 (2004).
[24] T. M. Tritt and M. A. Subramanian, MRS Bull. **31**, 188 (2006).
[25] G. J. Snyder, J. P. Fleurial, T. Caillat, R. Yang, and G. Chen, J. Appl. Phys. **92**, 1564 (2002).
[26] H. Lee, D. Vashaee, D. Z. Wang, M. S. Dresselhaus, Z. F. Ren, and G. Chen, J. Appl. Phys. **107**, 094308 (2010).
[27] J. H. Lee, G. A. Galli, and J. C. Grossman, Nano Lett. **8**, 3750 (2008).