Giant Thermoelectric Effect in Graphene-Based Topological Insulators with Heavy Adatoms and Nanopores

Po-Hao Chang,† Mohammad Saeed Bahramy,‡§ Naoto Nagaosa,†§ and Branimir K. Nikolić†§

†Department of Physics and Astronomy, University of Delaware, Newark, Delaware 19716-2570, United States
‡Quantum-Phase Electronics Center and Department of Applied Physics, The University of Tokyo, Tokyo 113-8656, Japan
§RIKEN Center for Emergent Matter Science (CEMS), Wako, Saitama 351-0198, Japan

ABSTRACT: Designing thermoelectric materials with high figure of merit \( ZT = S^2GT / K_{\text{tot}} \) requires fulfilling three often irreconcilable conditions, that is, the high electrical conductance \( G \), small thermal conductance \( K_{\text{tot}} \), and high Seebeck coefficient \( S \). Nanostructuring is one of the promising ways to achieve this goal as it can substantially suppress lattice contribution to \( K_{\text{tot}} \). However, it may also unfavorably influence the electronic transport in an uncontrollable way. Here, we theoretically demonstrate that this issue can be ideally solved by fabricating graphene nanoribbons with heavy adatoms and nanopores. The adatoms locally enhance spin–orbit coupling in graphene thereby converting it into a two-dimensional topological insulator with a band gap in the bulk and robust helical edge states, which carry electrical current and generate a highly optimized power factor \( S^2G \) per helical conducting channel due to narrow boxcar-function-shaped electronic transmission (surpassing even the Mahan-Sofo limit obtained for delta-function-shaped electronic transmission). Concurrently, the array of nanopores impedes the lattice thermal conduction through the bulk. Using quantum transport simulations coupled with first-principles electronic and phononic band structure calculations, the thermoelectric figure of merit is found to reach its maximum \( ZT = 3 \) at low temperatures \( T = 40 \text{ K} \). This paves a way to design high-\( ZT \) materials by exploiting the nontrivial topology of electronic states through nanostructuring.

KEYWORDS: Topological insulators, nanoribbons, graphene, thermoelectricity, phonons, first-principles calculations.
thermoelectrics based on 3D\textsuperscript{7–10} or 2D\textsuperscript{11,12} TIs are mostly qualitative and lack information on their phonons. In addition, many candidate 3D TI materials are unintentionally doped in the bulk, which makes it difficult to reach the topological transport regime where electrons, behaving as massless Dirac fermions, propagate exclusively on their surfaces\textsuperscript{4,10}.

In this Letter, we present a theoretical design of a high-$ZT$ system, using graphene-based 2D TI nanoribbon depicted in Figure 1 for which we accurately obtain all quantities in eq 1 via the nonequilibrium Green function (NEGF) methodology\textsuperscript{13,14} combined with first-principles calculations of both electronic and phononic band structures. Thus far, 2D TIs have been realized experimentally using cumbersome-to-grow

Figure 1. Schematic of 2D TI thermoelectric based on graphene nanoribbons with heavy adatoms and nanopores. (a) The central region, consisting of GNR (of width $W = 6.16$ nm and length $L \approx 1.35$ $\mu$m used in our calculations) with In adatoms and nanopores (253 of them along the chosen length), is attached to hot and cold macroscopic reservoirs. The adatoms shown in the center of randomly chosen ($n_{ad} \approx 19\%$) hexagons locally enhance SOC within those hexagons thereby creating an energy gap $E_G$ and a pair of helical quantum states on each edge channeling flow of electrons of opposite spin in opposite directions. We consider GNRs with (b) perfect edge of zigzag type or (c) disordered edge created by removing one or two edge carbon atoms in each GNR supercell. (d) The total density of states for different versions of ZGNR in (b), where the gap $E_G \approx 17.3$ meV around the Dirac point (at $E = 0$) is filled by contributions from the helical edge states, is insensitive to the randomness of adatom configuration or spatial inhomogeneities.

Figure 2. The electronic transmission function, Seebeck coefficient, and power factor governed by the helical edge states. (a) The local DOS within GNR + heavy-adatoms with a single nanopore in its interior at $E_F = 0.001$ eV. (b) The spatial profile of local currents over the transverse cross section of system in (a). The sum of all bond currents $j_{mn} / V_b$, describing charge flow from site $m$ to site $n$ of the honeycomb lattice at bias voltage $V_b$ gives the electrical conductance $G(E_F) = e^2 T_{el}(E_F)/h$ at zero temperature. Because of helical edge quantum states determining the local DOS and local currents, both quantities have nonzero value only around the edges. (c,d) The zero-bias electronic transmission function $T_{el}(E)$ for setups in Figure 1b,c, respectively. (e,f) The Seebeck coefficient at $T = 40$ K computed by plugging $T_{el}(E)$ from panels c and d into eq 2, respectively. Panels e and f also show the power factor $S^2 G$. The blue and red lines in panels c and f are obtained in the presence of both heavy adatoms and nanopores, while the black dotted line in panels c and d is computed after nanopores in the two setups from Figure 1 are removed.
HgTe/CdTe,15 or somewhat more accessible InAs/GaSb,16 quantum wells. A much simpler system, graphene with randomly distributed heavy adatoms, has been conjectured recently via first-principles studies.17,18 For example, among many possible heavy adatom candidates, In and Ti favor high-symmetry position in the center of the hexagons of honeycomb lattice of carbon atoms, while being nonmagnetic and without inducing Rashba SOC that would compete with the emergence of the 2D TI phase.17 We consider two types of graphene nanoribbons (GNRs) + nanopores as the central region of the two-terminal setup in Figure 1a, where the edge of GNR is either perfectly ordered and chosen to be of zigzag type in Figure 1b or lightly disordered in Figure 1c. We assume that \( n_{ad} \approx 19\% \) of GNR hexagons are randomly selected and covered by heavy adatoms of In, which locally enhance17 tiny (due to lightness of carbon atoms) intrinsic SOC coupling19 already present in graphene. This helps to increase the bulk band gap from \( E_G/k_B \approx 0.28 \) K in pristine graphene19 to \( E_G/k_B \approx 200 \) K for the chosen adatom type and \( n_{ad} \).

Remarkably, despite completely random distribution of heavy adatoms, such disordered system has an extremely stable 2D TI phase (which is actually stabilized by the randomness of adatom distribution21) that does not require disorder averaging12 because each sample has the same bulk energy gap \( E_G \approx 17.3 \) meV. The gap is visible in the total density of states (DOS) in Figure 1d that does not require disorder averaging12 because each sample has the same bulk energy gap \( E_G \approx 17.3 \) meV.

Figure 3. The phononic band structure, transmission function, and thermal conductance. Solid line in both panels a and b plots the phonon band structure for \( 4 \times 4 \) supercell of graphene with one In adatom per supercell, where SOC is not included. Dotted line in panel a plots phonon dispersion for the same supercell of pristine graphene without any adatoms, while dashed line in panel b includes both adatoms and SOC. All curves are obtained from first-principles calculations using VASP.23–25 (c,d) The phononic transmission function \( T_{ph}(\omega) \) for setups in Figure 1b,c, respectively, where solid lines are based on the dashed line dispersion in panel b. Dotted lines in panels c and d plot quantized \( T_{ph}(\omega) \) for an infinite homogeneous ZGNR of the same width and edge shape as in Figure 1 but without any adatoms, nanopores, or edge disorder. (e,f) Solid and dotted lines plot the phononic thermal conductance, computed by plugging the values of \( T_{ph}(\omega) \) from solid and dotted lines in panels c and d, respectively, into eq 4. Additional dashed line in panels e and f plots \( K_B \) for the same setups as in Figure 1b,c, respectively, but where the presence of heavy adatoms and local SOC induced by them is neglected.

The Seebeck coefficient is obtained from the electronic transmission function as14,26

\[
S(E_F) = \frac{L_s(E_F)}{eTL_0(E_F)}
\]

\[
L_s(E_F) = \frac{1}{h} \int_{-\infty}^{\infty} dE T_0(E)(E - E_F)^\frac{1}{2} \frac{dE}{dE} \left( -\frac{d}{dE} \right)
\]

where \( f(E) = \{1 + \exp[(E - E_F)/k_B T]^2 \}^{-1} \) is the Fermi function. The integrals in eq 3 also determine14,26 electronic charge...
$G(E_F) = e^2T_0(E_F)$ and thermal conductance $K_{th}(E_F) = \{T_\alpha(E_F) - [T_\alpha(E_F)]^2/L_\alpha(E_F)/T\}$. The electron-like (for $E > 0$) and hole-like (for $E < 0$) transport gives contributions to $S$ in opposite sign, so that $S \equiv 0$ exactly at the DP (as observed in the experiments on large-area graphene\textsuperscript{27} and 3D TIs\textsuperscript{28}). At low temperatures ($k_B T \ll E_F$), eq 2 can be approximated\textsuperscript{26} by the so-called Mott formula $S(E_F) \approx (\pi^2k_B^2T/3e)[dT_\alpha(E_F)/dE][T_\alpha(E_F)]^{-1}$, which explicitly shows that large values of $S$, plotted in Figure 2c,f, require steep variation of $T_\alpha(E_F)$.

The steep variation from $T_\alpha(E_F) = 2$ to a vanishing value $T_\alpha(E_F) \rightarrow 0$ occurring around $E \approx \pm 4.5$ meV in Figure 2c or $E \approx \pm 3.8$ meV in Figure 2d, whose details are controlled by the nanopores and/or edge disorder, explains our motivation (besides the fact that zigzag edges have already been fabricated experimentally\textsuperscript{29}) to build 2D TI on the top of ZGNRs. Unlike armchair edges, where the width of the edge states \[ \text{i.e., the width of nonzero LDOS around edges illustrated in Figure 2a} \] depends on $E_F$ and can require large nanoribbon widths to ensure desired energy interval with quantized\textsuperscript{12,22} $T_\alpha(E_F) = 2$, for zigzag edges the width of the helical edge states is small and independent of $E_F$ increasing rapidly\textsuperscript{23} only close to $\pm E_F/2$. Thus, when helical edge states start to overlap electrons can be backscattered off nanopores and/or edge disorder, which leads to steep variation of electronic transmission function within a narrower energy interval (shown in Figure 2c,d) than in the case of armchair edges.\textsuperscript{12}

While low $T_\alpha(E_F)$ also helps to increase $S$, such enhancement too far away from the DP is irrelevant because thermoelectric performance depends\textsuperscript{30,31} on the power factor $S^2G$. In fact, $S^2G$ shown in Figure 2e,f has exactly the same shape as the one obtained from the Mahan–Sofo (MS) model\textsuperscript{27} based on $T_\alpha(E_F) = M_0\delta(E_F)$, which ensures $K_{th} \rightarrow 0$ and determines putative\textsuperscript{30,31,33} upper limit for $S^2G$. However, the peak value ($S^2G)_{max} \approx 8k_B T/\hbar$ per spin-polarized conducting channel in Figure 2e,f is larger than the corresponding MS limit\textsuperscript{30} ($S^2G)_{MS} \approx 5.76k_B T/\hbar$. Thus, 2D TI nanoribbons utilize their two spin-polarized edge conducting channels most efficiently for thermoelectricity, which can be traced back to the (approximately) boxcar functional shape\textsuperscript{30,31} of $T_\alpha(E_F)$ in Figure 2c,d. Although only two available conducting channels make up the total power factor of a single graphene-based 2D TI nanoribbon minuscule compared to bulk 3D materials,\textsuperscript{33} quantities shown in Figure 2 do not scale with the nanoribbon width so that one can substantially increase the total $S^2G$ by patterning very large number of very narrow GNRs connected in parallel.

The phononic band structure in Figure 3b for graphene supercell with In adatoms and SOC switched on is computed\textsuperscript{23,25} from first-principles using VASP simulation package.\textsuperscript{23,25} This serves as an input for quantum transport calculation of the phononic transmission function $T_{ph}(\omega)$ plotted in Figure 3c,d, and $K_{ph}$ plotted in Figure 3e,f, which are mutually connected through the Landauer-type formula\textsuperscript{23,24}

$$K_{ph} = \frac{\hbar^2}{2\pi\hbar^2T^2} \int_0^\infty d\omega\omega^2T_{ph}(\omega) \frac{e^{\hbar/\omega T}}{(e^{\hbar/\omega T} - 1)^2}$$

The very high Debye temperature ($\approx 2100$ K) of graphene necessitates quantum transport treatment of phonon propagation captured by eq 4. However, this formula does not take into account the three-phonon scattering processes. Nevertheless, in analogy with carbon nanotubes\textsuperscript{36} we expect that restricted quasi-one-dimensional geometry of GNRs will suppress such processes, so that they become important (for considered GNRs of length $L \approx 1.35 \mu m$) only at temperatures higher than the range shown in Figure 3e,f.

Although the thermal conductivity of freely suspended large-area graphene samples at room temperature is among the highest of any known material\textsuperscript{36} it decreases significantly when this 2D material is in contact with a substrate or confined into nanoribbons. We further reduce the ballistic value of $K_{ph}$ for GNRs [plotted as dotted line in Figure 3e,f] by a factor $\approx 20–90$ (maximized around $T \approx 50$ K) via introduction of nanopores, as demonstrated by comparing the values along the dotted line in Figure 3e,f with the corresponding values along the solid line in Figure 3e,f, respectively. This reduction occurs in sufficiently long GNRs, where we check that decrease of $K_{ph}$ with increasing $L$ and the number of nanopores saturates\textsuperscript{37} around the selected length $L \approx 1.35$ $\mu m$. Although nanopore arrays have been considered theoretically as a way to reduce lattice thermal conductivity of bulk\textsuperscript{9,38} and low-dimensional\textsuperscript{37,39} materials, graphene with its high mechanical stability makes it easy to fabricate nanopore arrays by a variety of recently developed techniques.\textsuperscript{40}

The combination of electronic and phononic transport quantities shown in Figure 2 and 3, respectively, generates maximum $ZT_{max} \approx 3$ for both types of GNRs considered in Figure 1. The difference in $ZT$ for perfect and disordered edge is minuscule since we use only light disorder in Figure 1c with intention to demonstrate (rather than to introduce additional phonon scattering off edge roughness) how GNR-based 2D TIs do not require large nanofabrication efforts\textsuperscript{29} to precisely control the position of edge carbon atoms.\textsuperscript{39,41} The value of $ZT$ can be tuned significantly by changing the charge density (i.e., the corresponding $E_F$) via the gate voltage, where the advantage over other recent proposals\textsuperscript{14,39,41} for thermoelectrics based on topologically trivial GNRs is insensitivity of the position of the peaks of $ZT(E_F)$ to microscopic details of the system. That is, the energy $E_F$ at which $ZT$ reaches maximum in Figure 4 depends only on $T$ and $E_F$ (governed by the adatom coverage and adatom type).

![Figure 4](image-url)
In conclusion, inspired by the recently proposed novel route to create a robust 2D TI phase using the exposed surface of graphene, we have employed quantum transport modeling combined with first-principles electronic and phononic band structure calculations to design in silico a high-ZT thermoelectric where graphene nanoribbons with arbitrary shaped edges and nanopores in their interior are covered with heavy adatoms of In. The adatoms provide sufficiently strong local SOC in some fraction of randomly chosen hexagons, so that such inhomogeneous SOC opens both the bulk band gap $E_G \approx 17.3$ meV (for $n_{ad} \approx 19\%$ of hexagons covered) and generates topologically protected helical edge states. The electronic transmission through helical edge states in the form of the (approximately) boxcar function of width $\leq E_G$ generates power factor $S^2G$ per helical conducting channel, which turns out to be larger than the one obtained from the celebrated Mahan-Sofo model with delta-function-shaped electronic transmission function. This feature combined with 2 orders of magnitude reduction of phononic thermal conductance by the nanopore array leads to thermoelectric figure of merit for this system $ZT \approx 3$ at low temperatures $T \approx 40$ K. Because the existing bulk thermoelectric materials are very inefficient at low operating temperatures $T \approx 10$ K (where they give $\leq 1.5$ $ZT \approx 0.01$), the TI-based high-ZT thermoelectrics at low temperatures designed by our study could be attractive for applications in radioisotope thermoelectric generators on spacecrafts or cooling of electronic satellite components.

While bulk materials are deemed necessary for large-scale power generation, GNRs underlying our proposal are single-atom-thick and with electronic transport properties which do not scale with their width, so that very high packing density of GNRs connected in parallel is possible within a 3D volume. Other choices for heavy adatoms, such as Os, Ir, and Cu–Os or Cu–Ir dimers, are predicted to generate larger gap $E_G \geq 0.2$ eV using smaller adatom coverage $n_{ad} \approx 2\%$ (via different hybridization mechanism between carbon and adatom orbitals than in the case of In or Ti), thereby making it possible to tune the optimal operating temperature while at the same time evading propensity of In atoms to coalesce on graphene because of their weak binding energies and shallow segregation barriers. The operating temperature of 2D TI-based thermoelectrics would also be governed by the requirement on nanoribbon length that has to be smaller than the inelastic scattering length of edge states. Although no symmetry prevents inelastic backscattering off phonons, quantized transmission in Figure 2 is expected to be insensitive to such intrinsic mechanisms to leading order. We note that extrinsic mechanisms, such as inelastic scattering off charge puddles, can cause phase decoherence and additional backscattering, but they can be in principle controlled by proper device fabrication. Interestingly, if the boxcar-shaped transmission function in Figure 2 is preserved at finite bias voltage, the same system could optimize efficiency at finite power output in the nascent field of nonlinear thermoelectrics.

**Methods.** The electronic subsystem of GNR + heavy-adatoms is described by the tight-binding Hamiltonian of Kane–Mele type with a single $p_z$ orbital per site of the honeycomb lattice

$$H = -t \sum_{\langle mn \rangle, \sigma} \epsilon_n \sigma \sigma + \mu_{\text{SO}} \sum_{\langle mn \rangle} \sum_{\sigma} \sum_{\sigma'} \nu_{mn, \sigma} \sigma \sigma' \sigma \sigma'$$

(5)

Here the operator $\hat{c}_m^\dagger (\hat{c}_m)$ creates (annihilates) electron on site $m$ of the lattice in spin state $\sigma$ and $\hat{z}$ is the Pauli matrix. The nearest-neighbor hopping $t = 2.7$ eV in the first term in eq 5 sets the unit of energy scale. The spin-dependent hopping in the second term, where $\nu_{mn} = 1$ for moving counterclockwise around the hexagon and $\nu_{mn} = -1$ otherwise, acts between next-nearest neighbor sites of only those hexagons $P$ of the honeycomb lattice that host In adatoms. The strength of such SOC, which can be viewed as locally enhanced version of the tiny intrinsic SOC in pristine graphene, is parametrized by $\lambda_{am}$. 

This minimal effective model in eq 5 is sufficient to fit, using $\lambda_{am} = 0.0037t$, the low-energy spectrum obtained from first-principles calculations for $4 \times 4$ graphene supercell with two In adatoms using VASP package 23–25 which gives $E_G \approx 11.5$ meV. The electron–core interactions are described by the projector augmented wave (PAW) method46,47 and we use Perdew–Burke–Ernzerhof (PBE)48 parametrization of the generalized gradient approximation (GGA) for the exchange-correlation functional. The cutoff energies for the plane wave basis set used to expand the Kohn–Sham orbitals are 500 eV for all calculations. A 11 $\times$ 11 $\times$ 1 k-point mesh within Monkhorst–Pack scheme is used for the Brillouin zone (BZ) integration. Structural relaxations and total energy calculations are performed ensuring that the Hellmann–Feynman forces acting on ions are less than 0.005 eV/Å.

Starting from the matrix representation $\mathbf{H}$ of the Hamiltonian in eq 5, we compute the electronic retarded GF, $G(E) = \{E - \mathbf{H} - \sum_{n} \delta_{nE} \mathbf{1} - \mathbf{\Sigma}_{LR}(E)\}^{-1}$, where $\mathbf{\Sigma}_{LR}$ are the self-energies introduced by the semi-infinite ideal (without disorder, adatoms, or nanopores) ZGNR leads assumed to be attached to 2D TI nanoribbon in Figure 1bc. The retarded GF and the level broadening matrices $\Gamma_{LR}(E) = i[\mathbf{\Sigma}_{LR}(E) - \mathbf{\Sigma}_{LR}^\dagger(E)]$ allow us to obtain the electronic zero-bias transmission function, $T_{el}(E) = Tr[\Gamma_{LR}(E)G(E)\Gamma_{LR}^\dagger(E)]$, which determines electronic transport quantities through eq 3.

The phononic band structure plotted in Figure 3a,b was computed via first-principles methodology using combined VASP23–25 and Phonopy packages. The details of VASP calculations are the same as delineated above (except that we use $3 \times 3 \times 1$ k-point mesh) but here we start from $4 \times 4$ graphene supercell hosting one In adatom and then enlarged it to $8 \times 8$ supercell in order to capture accurately force constants between a range of neighboring carbon atoms or carbon atoms and In adatoms. Figure 3a,b demonstrate appearance of new low energy bands due to the presence of In adatoms. Although the effect of SOC on phononic band structures can be profound for materials containing heavy elements, especially on surfaces and in thin films (as exemplified by the recent calculations50 on Bi$_2$Te$_3$), the inclusion of SOC in Figure 3b generates only a small difference.

To construct the empirical up to fourth-nearest neighbors force constant (FC) model, we varied and optimized the FCs to fit as closely as possible the phononic dispersions plotted in Figure 3a,b. Using the FC matrix $\mathbf{K}$, the diagonal matrix $\mathbf{M}$ containing atomic masses, and self-energies $\Pi_{LR}$ of the semi-infinite ideal ZGNR leads [obtained using FCs extracted from the dotted line in Figure 3a], we compute the phononic version of the retarded GF, $D(\omega) = \{\omega^2\mathbf{M} - \mathbf{K} - \Pi_{LR}(\omega) - \Pi_{LR}^\dagger(\omega)\}^{-1}$. This, together with the level broadening matrices $\Lambda_{LR}(\omega) = i[\Pi_{LR}(\omega) - \Pi_{LR}^\dagger(\omega)]$, gives the phononic transmission function, $T_{ph}(\omega) = Tr(\Lambda_{LR}(\omega)D(\omega)\Lambda_{LR}^\dagger(\omega))$, which determines $K_{ph}$ through eq 4.
The significant difference between $K_{ph}$ for GNRs with nanopores but neglecting heavy adatoms [dashed line in Figure 3e,f] and $K_{ph}$ when heavy adatoms and the corresponding SOC are included [solid line in Figures 3e,f] confirms the necessity for the procedure delineated above. We note that the values of $K_{ph}$ [solid lines in Figures 3e,f] based on FCs extracted from the phonon band structure of bulk graphene with heavy adatoms in Figure 3b are most likely overestimated; more precise FCs would require a computationally very expensive procedure that considers a large number of atoms confined within the nanoribbon geometry and in the presence of nanoparticles.\textsuperscript{14,39}

The computation of submatrices of $G(E)$ and $D(\omega)$ for ZGNRs of length $L \approx 1.35$ μm and width $W = 6.16$ nm is achieved via standard recursive algorithms.\textsuperscript{31} They make possible to reduce the computational complexity to $O(N_s N_c^5)$, where $N_s$ is the number of supercells along ZGNR length and $N_c$ is the number of atomic orbitals in each supercell, because only specific submatrices of the full $G(E)$ or $D(\omega)$ matrices are required\textsuperscript{31} to obtain $T_0(E)$ and $T_{ph}(\omega)$, respectively.

\section*{AUTHOR INFORMATION}

\textbf{Corresponding Author}

*E-mail: bnikolic@udel.edu.

\textbf{Notes}

The authors declare no competing financial interest.

\section*{ACKNOWLEDGMENTS}

P.-H.C. and B.K.N. were supported by U.S. NSF under Grant ECCS 1202069. M.S.B. and N.N. were supported by Grant-in-Aids for Scientific Research (24224009) from the Ministry of Education, Culture, Sports, Science, and Technology of Japan, Strategic International Cooperative Program (Joint Research Type) from Japan Science and Technology Agency, and also by Funding Program for World-Leading Innovative R&D on Science and Technology (FIRST Program).

\section*{REFERENCES}