Phonon-induced topological insulation

Kush Saha and Ion Garate
Département de Physique and Regroupement Québécois sur les Matériaux de Pointe, Université de Sherbrooke, Sherbrooke, Québec, Canada J1K 2R1
(Received 17 March 2014; published 8 May 2014)

We develop an approximate theory of phonon-induced topological insulation in Dirac materials. In the weak-coupling regime, long-wavelength phonons may favor topological phases in Dirac insulators with direct and narrow band gaps. This phenomenon originates from electron-phonon matrix elements, which change qualitatively under a band inversion. A similar mechanism applies to weak Coulomb interactions and spin-independent disorder; however, the influence of these on band topology is largely independent of temperature. As applications of the theory, we evaluate the temperature dependence of the critical thickness and the critical stoichiometric ratio for the topological transition in CdTe/HgTe quantum wells and in $BiTl(S_{1-\delta}Se_{\delta})_2$, respectively.

DOI: 10.1103/PhysRevB.89.205103 PACS number(s): 71.10.−w, 65.40.−b, 63.20.kd

I. INTRODUCTION

Nearing a decade of rapid progress [1], the field of topological insulators has become subject of textbooks [2]. Despite its incipient maturity, the subject remains active and broad with ongoing work ranging from exotic field theory [3] to applied physics [4]. Above all, the mechanism of band inversion, responsible for the appearance of topological surface states, continues to captivate the imagination of theorists and experimentalists alike.

At the interface between theory and experiment, there is an intense interest to identify tunable band-gap materials, where band inversion could be done and undone as a function of an experimentally controllable parameter. Pressure [5], electric fields [6], compound stoichiometry [7], lasers [8], and strong random alloying [9] are some of the candidate agents that can switch topological phases on and off. Recently, one of us has argued that temperature is an additional "knob" that may induce a band inversion [10]. This proposal originated from a study of electron-phonon interactions, whose effect in band topology had been previously overlooked. The idea that phonons can change the band topology of an electronic system resonates with an upcoming line of research concerned with the effect of dissipative and thermal baths on topological materials [11].

In Ref. [10], the origin for phonon-induced band inversion remained rather obscure. The main objective of the present work is to provide a simple and intuitive understanding of the mechanism that underlies phonon-induced topological insulation. This same mechanism applies to disorder- and Coulomb-interaction-induced topological insulation as well. We begin (Sec. II) by reviewing some fundamentals of band-gap renormalization in semiconductors. This is a subject that has attracted steady attention in the last four decades [12]; nevertheless, the possibility that phonons could invert the band gap of a Dirac material has not been contemplated. If one considers only the highest valence band and the lowest conduction band of a direct-gap insulator, then simple perturbation theory dictates that intraband (interband) electron-phonon scattering processes reduce (augment) the band gap. The total change in the band gap is given by the sum of the competing interband and intraband parts. In wide-gap semiconductors, interband transitions are suppressed and the band gap decreases with temperature, regardless of the details of electron-phonon matrix elements. However, in narrow-gap semiconductors, the competition between intraband and interband transitions becomes particularly sensitive to the matrix elements. These matrix elements are peculiar in Dirac materials (Sec. III) because the momentum-space texture of the band eigenstates changes topology when the system undergoes a band inversion. Due to this peculiarity, the dominant long-wavelength electron-phonon matrix elements are of intraband (interband) type in trivial (topological) Dirac insulators. Consequently, the magnitude of the band gap of a trivial Dirac insulator decreases with temperature, while it increases in a topological insulator. This idea, which we elaborate in Sec. IV, is the main insight from the present work. In Sec. V, we apply the theory to HgTe/CdTe quantum wells and $BiTl(S_{1-\delta}Se_{\delta})_2$. Because the renormalized Dirac mass and the renormalized band gap differ from one another, the appearance of topological surface states does not occur simultaneously with a band inversion. This may help explain the "topological proximity effect" observed in $BiTl(S_{1-\delta}Se_{\delta})_2$. In Sec. VI, we collect the main ideas and discuss their relevance in materials with complex electronic and phononic structures.

II. BAND-GAP RENORMALIZATION

We begin by reviewing the phonon-induced renormalization of the band gap in semiconductors [12]. The Hamiltonian of a perfectly periodic crystal is

$$\mathcal{H}_0 = \sum_{\mathbf{k}} \sum_{\sigma\sigma'} \sum_{\tau\tau'} \langle \sigma\tau | h_{\mathbf{k}} | \sigma'\tau' \rangle c^{\dagger}_{\mathbf{k}\sigma\tau} c_{\mathbf{k}\sigma'\tau'}, \tag{1}$$

where $\mathbf{k}$ is the crystal momentum, $h_{\mathbf{k}}$ is the Bloch Hamiltonian, and $c_{\mathbf{k}\sigma\tau}$ is an operator that annihilates an electron with momentum $\mathbf{k}$, spin quantum number σ, and orbital quantum number τ. The Bloch Hamiltonian has eigenvalues $E_{\mathbf{k}n}$ and eigenstates

$$|\psi_{\mathbf{k}n}\rangle = \exp(i\mathbf{k}\cdot\mathbf{r})|\mathbf{k}n\rangle/\sqrt{V}, \tag{2}$$

where V is the crystal volume, n is the Bloch band label, and $|\mathbf{k}n\rangle$ is a multicomponent spinor whose spatial dependence has the periodicity of the lattice.

Small deviations of the ions from their equilibrium positions couple to the electron density and result in an electron-phonon interaction $\mathcal{V} \simeq \mathcal{V}^{(1)} + \mathcal{V}^{(2)}$, where

$$\mathcal{V}^{(1)} = \int d\mathbf{r}\, \rho(\mathbf{r}) \sum_j \mathbf{Q}_j \cdot \boldsymbol{\nabla} V_{\mathrm{ei}}\big(\mathbf{r} - \mathbf{R}_j^0\big),$$

$$\mathcal{V}^{(2)} = \frac{1}{2} \int d\mathbf{r}\, \rho(\mathbf{r}) \sum_j (\mathbf{Q}_j \cdot \boldsymbol{\nabla})^2 V_{\mathrm{ei}}\big(\mathbf{r} - \mathbf{R}_j^0\big). \tag{3}$$

Here, R_j^0 is the equilibrium position of the jth ion, $\mathbf{r}$ is the electron coordinate, $\boldsymbol{\nabla} = \partial/\partial\mathbf{r}$, $V_{\mathrm{ei}}(\mathbf{r})$ is the electron-ion potential,

$$\mathbf{Q}_j = i \sum_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{R}_j^{(0)}} \left(\frac{\hbar}{2\rho_A V \omega_{\mathbf{k}}}\right)^{1/2} \mathbf{e}_{\mathbf{k}} \left(a_{\mathbf{k}} + a_{-\mathbf{k}}^\dagger\right) \tag{4}$$

is the ionic displacement from equilibrium (assumed to be small), $a_{\mathbf{k}}$ is an operator that destroys a phonon with momentum $\mathbf{k}$, $\omega_{\mathbf{k}}$ is the phonon frequency, ρ_A is the atomic density, and $\mathbf{e}_{\mathbf{k}}$ is the polarization vector of the phonon mode. In addition, the electron density operator is

$$\rho(\mathbf{r}) = \frac{1}{V} \sum_{\mathbf{q}} e^{-i\mathbf{q}\cdot\mathbf{r}} \rho_{\mathbf{q}} \;\; ; \;\; \rho_{\mathbf{q}} = \sum_{\mathbf{k}\sigma\tau} c_{\mathbf{k}\sigma\tau}^\dagger c_{\mathbf{k}-\mathbf{q}\sigma\tau} \tag{5}$$

in the plane-wave basis. For simplicity, we have considered one atom per lattice site and, for brevity, we have omitted the index that labels different phonon modes. Furthermore, Eq. (3) is local in real space and thus does not capture phonon-induced changes in the electronic hopping amplitude; the effect of these terms will be briefly discussed in Sec. VI.

Equations (3) and (5) together evidence that the electron-phonon interaction conserves spin and orbital quantum numbers in the plane-wave basis. This fact will figure prominently in the mechanism for phonon-induced topological insulation (cf. Sec. IV). At any rate, since the electronic eigenstates do not generally have well-defined spin and orbital quantum numbers, phonons do scatter electrons between different bands. This becomes apparent by rewriting Eq. (5) in the band eigenstate basis,

$$\rho_{\mathbf{q}} = \sum_{\mathbf{k}\mathbf{k}' nn'} \sum_{\mathbf{G}} \delta_{\mathbf{k}'+\mathbf{q}-\mathbf{k},\mathbf{G}} \langle \mathbf{k}n | e^{i\mathbf{G}\cdot\mathbf{r}} | \mathbf{k}'\, n' \rangle c_{\mathbf{k}n}^\dagger c_{\mathbf{k}'n'}, \tag{6}$$

where $c_{\mathbf{k}n}$ annihilates an electron with momentum $\mathbf{k}$ in band n, $\mathbf{G}$ is a reciprocal lattice vector and

$$\langle \mathbf{k}n | e^{i\mathbf{G}\cdot\mathbf{r}} | \mathbf{k}'n' \rangle \equiv \int_{\mathrm{cell}} d\mathbf{r} e^{i\mathbf{G}\cdot\mathbf{r}} u_{\mathbf{k}n}^*(\mathbf{r}) u_{\mathbf{k}'n'}(\mathbf{r}). \tag{7}$$

Here, $u_{\mathbf{k}n}(\mathbf{r}) \equiv \langle \mathbf{r} | \mathbf{k}n \rangle$ and the spatial integration is over the unit cell. If the Bloch eigenstates were plane waves (which would approximately be the case in simple metals), Eq. (7) would be nonzero only for $G = 0$. In Dirac insulators, the Bloch states are not plane waves and the $G \neq 0$ terms (umklapp processes) do not vanish. Yet, hereafter, we neglect umklapp processes on the basis that (i) we consider the coupling of electrons to long-wavelength phonons (deformation potential coupling) and (ii) we model the electronic structure with $\mathbf{k}\cdot\mathbf{p}$ Hamiltonians that are tailored to small momenta in the vicinity of the band-gap minimum. Under this proviso, Eq. (3) can be rewritten as [13,14]

$$\mathcal{V}^{(1)} = \sum_{\mathbf{q}} g_{\mathbf{q}} \rho_{\mathbf{q}} (a_{-\mathbf{q}}^\dagger + a_{\mathbf{q}}),$$

$$\mathcal{V}^{(2)} = \sum_{\mathbf{k},\mathbf{q}} \lambda_{\mathbf{k}\mathbf{q}} \rho_{\mathbf{q}} (a_{\mathbf{k}} + a_{-\mathbf{k}}^\dagger)(a_{-\mathbf{k}+\mathbf{q}} + a_{\mathbf{k}-\mathbf{q}}^\dagger), \tag{8}$$

where the expressions for $g_{\mathbf{q}}$ and $\lambda_{\mathbf{k}\mathbf{q}}$ are listed in Appendix A, and

$$\rho_{\mathbf{q}} \simeq \sum_{\mathbf{k}nn'} \langle \mathbf{k}n | \mathbf{k} - \mathbf{q}\, n' \rangle c_{\mathbf{k}n}^\dagger c_{\mathbf{k}-\mathbf{q}n'}. \tag{9}$$

Since $\langle \mathbf{k}n | \mathbf{k}\, n' \rangle = \delta_{nn'}$, zero-momentum phonons cannot scatter electrons between different bands [15]. In contrast, $q \neq 0$ phonons can induce interband transitions because $\langle \mathbf{k}n | \mathbf{k}'\, n' \rangle \neq \delta_{nn'}$ for $\mathbf{k} \neq \mathbf{k}'$. In the $\mathbf{k}\cdot\mathbf{p}$ Hamiltonians analyzed below, these overlap matrix elements may be evaluated analytically.

Both $\mathcal{V}^{(1)}$ and $\mathcal{V}^{(2)}$ modify the electronic band structure. At zero temperature, the renormalized energy levels are

$$E_{\mathbf{k}n}^* \simeq E_{\mathbf{k}n} + \sum_{n'\mathbf{q}} |g_{\mathbf{q}}|^2 \frac{|\langle \mathbf{k}n | \mathbf{k} - \mathbf{q}n' \rangle|^2}{E_{\mathbf{k}n} - E_{\mathbf{k}-\mathbf{q}n'}}, \tag{10}$$

where we have neglected the frequency of phonon modes in the denominator. This is a good approximation for the purposes of the present work (cf. Sec. IV). In addition, in the derivation of Eq. (10), we have used (cf. Appendix A)

$$\langle \mathbf{k}n; 0 | \mathcal{V}^{(2)} | \mathbf{k}n'; 0 \rangle = 0, \tag{11}$$

where $|\mathbf{k}n; N\rangle \equiv c_{\mathbf{k}n}^\dagger |FS\rangle \otimes |N\rangle$, $|FS\rangle$ is the Fermi sea, and $|N\rangle$ is a Fock state of N phonons. The matrix element in Eq. (11) is known as the Debye-Waller term. When short-wavelength phonons are included, the Debye-Waller term is nonzero and contributes to the renormalization of energy levels at the same order as Eq. (10). In such case, the electron-phonon matrix elements are usually computed from first-principles pseudopotential methods [16], which show that the band-gap renormalization of common semiconductors often contains a significant Debye-Waller component. The possible impact of the Debye-Waller term on band topology will be briefly discussed in Sec. VI.

In Eq. (10), it is instructive to separate the sum over intermediate states onto interband ($n' \neq n$) and intraband ($n' = n$) parts. For a direct-gap semiconductor, a glance at the energy denominators of Eq. (10) and Fig. 1 reveals that intraband transitions decrease the band gap at $k = 0$, while interband transitions increase it. In this paper, "interband transition" refers to a transition that takes place between the valence band and the conduction band edges. Transitions between energy-degenerate bands are counted as "intraband." There are, of course, interband transitions between different (nondegenerate) valence bands as well as between different conduction bands. These interband transitions could *a priori* lead to a decrease of the band gap. However, for Dirac insulators in the vicinity of a topological phase transition, there is often a large energy separation between the lowest conduction band and the rest of conduction bands. Likewise, the highest valence band is typically well separated in energy

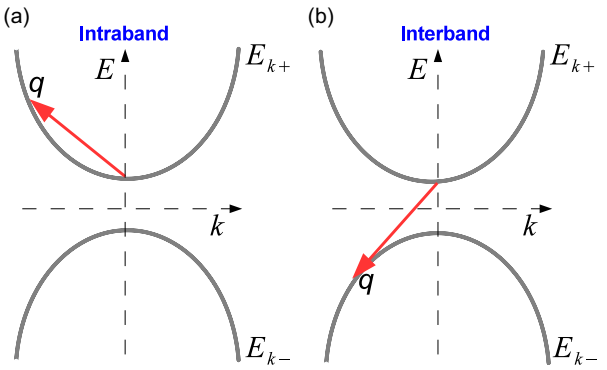

FIG. 1. (Color online) Schematic diagram for (a) intraband and (b) interband electronic transitions induced by phonons in a two-band insulator with a direct gap at the zone center (each band may be degenerate). It is assumed that additional conduction and valence bands are sufficiently far in energy so that they do not contribute substantially to band-gap renormalization. Such band-gap renormalization involves the evaluation of Eq. (10) for $\mathbf{k} = \mathbf{0}$. Accordingly, the relevant electron-phonon scattering matrix elements connect $\mathbf{k} = 0$ with $\mathbf{k} = \mathbf{q}$, where $\mathbf{q}$ is the phonon momentum. Since $E_{\mathbf{0}+} \leqslant E_{\mathbf{k}+}$ and $E_{\mathbf{0}-} \geqslant E_{\mathbf{k}-}$ for any $\mathbf{k}$, intraband transitions lead to a decrease of the band gap ($E_{\mathbf{0}+} - E_{\mathbf{0}-}$). A similar argument reveals that interband transitions increase the band gap.

from the rest of the valence bands. Under these conditions, the leading interband contribution emerges from transitions between the highest valence band and the lowest conduction band, the rest being suppressed by relatively large energy denominators in Eq. (10). In sum, the net change in the band gap depends on the relative strength of the intraband and interband contributions.

For wide-gap semiconductors, the interband contribution is depleted by a large energy denominator in Eq. (10); therefore, electron-phonon interactions decrease the band gap. However, for narrow-gap semiconductors, the energy denominators are no longer large enough to rule out the interband part. Instead, the competition between intraband and interband contributions depends sensitively on the magnitude of the $G = 0$ electron-phonon matrix elements, $g_{\mathbf{k}-\mathbf{k}'}\langle \mathbf{k}n|\mathbf{k}'n'\rangle$. As we explain below (cf. Sec. IV), the key for the phonon-induced topological insulation is that intraband matrix elements dominate on the trivial side of a topological phase transition, whereas interband matrix elements take over on the nontrivial side of said phase transition. In other words, phonons decrease the band gap when the Dirac insulator is trivial, while they increase the band gap when the Dirac insulator is topological. Albeit unusual, this feature seems generic to Dirac insulators with deformation potential coupling to phonons because it arises due to the change in the eigenstates' momentum-space texture across a topological phase transition.

III. DIRAC MASS RENORMALIZATION

The aim of this paper is to extract qualitative insights from low-energy effective models, rather than from more accurate but less transparent first principles calculations. The minimal $\mathbf{k}\cdot\mathbf{p}$ Hamiltonian that captures the low-energy properties of a time-reversal- and inversion-symetric 2D or 3D Dirac insulator with a band-gap minimum at the Brillouin zone center [17] is

$$h_{\mathbf{k}} = d_{0,\mathbf{k}} + \mathbf{d}_{\mathbf{k}}\cdot\boldsymbol{\sigma}\tau^x + M_{\mathbf{k}}\tau^z, \tag{12}$$

where σ^i and τ^i are Pauli matrices in spin and orbital space (respectively), $d_{0,\mathbf{k}} = \gamma k^2 a^2$, $d_{i,\mathbf{k}} = -\alpha k_i a$ ($i \in \{x,y\}$ in 2D and $i \in \{x,y,z\}$ in 3D), $M_{\mathbf{k}} = m + \beta k^2 a^2$, a is the lattice constant that acts as an ultraviolet cutoff, and (γ,β,α,m) are material parameters with units of energy. In particular, m is the Dirac mass of the quasiparticles at $k = 0$. The parameter α determines the velocity of Dirac quasiparticles, whereas γ models the particle-hole asymmetry of the band structure. A lattice regularization of Eq. (12) may be introduced in the usual way, but it does not affect the main results substantially.

The eigenvalues of Eq. (12) are a pair of doubly degenerate conduction and valence bands with energies

$$E_{\mathbf{k}\pm} = d_{0,\mathbf{k}} \pm \epsilon_{\mathbf{k}} \;\; ; \;\; \epsilon_{\mathbf{k}} \equiv \sqrt{\mathbf{d}_{\mathbf{k}}\cdot\mathbf{d}_{\mathbf{k}} + M_{\mathbf{k}}^2}. \tag{13}$$

It is straightforward to obtain the corresponding eigenvectors $|\mathbf{k}n\rangle$ analytically; these will be used to derive the results below. We assume that the band gap at $k = 0$ is parametrically smaller than the gap at any other time-reversal-invariant momentum (TRIM), i.e., we assume $|\beta| \gg |m|$. This scenario comprises most real Dirac insulators. The band topology of the Dirac insulator is then determined by the sign of $m\,\beta$: if positive (negative), the insulator is trivial (topological). From here on, we take $\beta > 0$ without loss of generality, and thus $m > 0$ ($m < 0$) describes a trivial (topological) insulator. Although it plays no role in determining the band topology of noninteracting Dirac insulators, γ can alter the band topology under the presence of electron-phonon interactions.

Phonons affect the band topology of a Dirac insulator by renormalizing the Dirac mass, $m \to m^*$. In particular, if m and m^* have the opposite sign, phonons induce a topological phase transition [10]. For real Dirac insulators, this is susceptible to occur only at the TRIM where the band gap is smallest (at $k = 0$ in our model) because the Dirac masses at all other TRIM are typically large in magnitude compared to the energy-scale of the electron-phonon interaction. The renormalization of the Dirac mass due to electron-phonon interactions can be obtained from the self-energy,

$$\Sigma_{nn'}(\mathbf{k},i\omega) = \sum_{\mathbf{q}n''} g_{\mathbf{q}}^2 \langle \mathbf{k}n|\mathbf{k}-\mathbf{q}\,n''\rangle\langle \mathbf{k}-\mathbf{q}n''|\mathbf{k}n'\rangle \times\left(\frac{1+n_{B\mathbf{q}}-f_{\mathbf{k}-\mathbf{q}n''}}{i\omega-\xi_{\mathbf{k}-\mathbf{q}n''}-\omega_{\mathbf{q}}} + \frac{n_{B\mathbf{q}}+f_{\mathbf{k}-\mathbf{q}n''}}{i\omega-\xi_{\mathbf{k}-\mathbf{q}n''}+\omega_{\mathbf{q}}}\right), \tag{14}$$

where $n,n' \in \{1,2,3,4\}$ are band labels, $\xi_{\mathbf{k}n} = E_{\mathbf{k}n} - \epsilon_F$ with $E_{\mathbf{k}1} = E_{\mathbf{k}2} = E_{\mathbf{k}+}$ and $E_{\mathbf{k}3} = E_{\mathbf{k}4} = E_{\mathbf{k}-}$; ϵ_F is the Fermi energy, $n_{B\mathbf{q}} = [\exp(\omega_{\mathbf{q}}/T) - 1]^{-1}$ is the phonon occupation number, $f_{\mathbf{k}n} = [\exp(\xi_{\mathbf{k}n}/T) + 1]^{-1}$ is the fermion occupation number, and $\omega = (2l+1)\pi T$ ($l \in \mathbb{Z}$) is the fermionic Matsubara frequency.

Hereafter, we concentrate on the $\mathbf{k} = \mathbf{0}$ self-energy,

$$\Sigma_{nn'}(\mathbf{0},i\omega) = \Sigma_0(\mathbf{0},i\omega)\delta_{nn'} + \Sigma_z(\mathbf{0},i\omega)\tau^z_{nn}\delta_{nn'}, \tag{15}$$

where Σ_0 and Σ_z are related to the renormalization of the Fermi energy and the Dirac mass, respectively [cf. Eq. (12)]. The self-energy at $k = 0$ is diagonal and doubly degenerate

due to the combined time-reversal and inversion symmetry of Eq. (12). After simple algebra, we arrive at

$$\Sigma_0(\mathbf{0},i\omega) \simeq \sum_{\mathbf{q}} g^2_{\mathbf{q},\mathrm{eff}}(T)\frac{i\omega+\epsilon_F-d_{0,\mathbf{q}}}{(i\omega+\epsilon_F-d_{0,\mathbf{q}})^2-\epsilon^2_{\mathbf{q}}},$$
$$\Sigma_z(\mathbf{0},i\omega) \simeq \sum_{\mathbf{q}} g^2_{\mathbf{q},\mathrm{eff}}(T)\frac{M_{\mathbf{q}}}{(i\omega+\epsilon_F-d_{0,\mathbf{q}})^2-\epsilon^2_{\mathbf{q}}}, \tag{16}$$

where we have neglected the phonon frequency in the denominators and have defined

$$g^2_{\mathbf{q},\mathrm{eff}}(T) \equiv g^2_{\mathbf{q}}(1+2n_{B\mathbf{q}}). \tag{17}$$

From Eq. (15), the zero-temperature renormalized Dirac mass can be read off as [18]

$$m^* = m + \Sigma_z(\mathbf{0},0). \tag{18}$$

A natural generalization to finite temperature is

$$m^*(T) = m + \mathrm{Re}\,[\Sigma_z(\mathbf{0},i\pi T)], \tag{19}$$

where we have recognized that the lowest Matsubara frequency is $\pm\pi T$. When temperature is low compared to the bandwidth of the electronic bands (which is in fact the case of interest), $\mathrm{Re}\,[\Sigma_z(\mathbf{0},i\pi T)] \simeq \Sigma_z(\mathbf{0},0)$ and the entire temperature dependence of $m^*(T)$ originates from the phonon occupation factor in Eq. (17). The physical consequences of a temperature-dependent Dirac mass will be discussed in Sec. V.

An important aspect of Eq. (16) is that $\Sigma_z(\mathbf{0},0)$ is largely independent of m when $|m|$ is small. Whether phonons favor a trivial or a topological phase is thus independent of whether the bare Dirac insulator is trivial or topological (insofar as the gap is small). The physical reason behind this result will be described in Sec. IV.

Equations (15) and (16) are formally very similar to the ones that appear in the theory of the topological Anderson insulator [9]. The main difference arises in the temperature-dependence of the self-energy, which is negligible for static disorder and significant for phonons. Temperature may be regarded as a knob to effectively tune the strength of electron-phonon coupling [cf. Eq. (17)]. In fact, the experimental fingerprint for the phonon-induced band-gap renormalization in semiconductors is its temperature dependence.

We conclude this section by discussing the relation between the renormalized Dirac mass and the renormalized band gap [18]. For small phonon frequency and at zero temperature, $\Sigma_{nn}(\mathbf{k},\xi_{\mathbf{k}n})$ agrees with Eq. (10). Accordingly, the renormalized band gap at $\mathbf{k}=\mathbf{0}$ is $E^*_g \equiv E^*_c - E^*_v$, where

$$E^*_c \simeq m + \Sigma_0(\mathbf{0},E^*_c-\epsilon_F) + \Sigma_z(\mathbf{0},E^*_c-\epsilon_F),$$
$$E^*_v \simeq -m + \Sigma_0(\mathbf{0},E^*_v-\epsilon_F) - \Sigma_z(\mathbf{0},E^*_v-\epsilon_F). \tag{20}$$

Note that E^*_g depends on temperature and may be either positive (normal band ordering) or negative (inverted band ordering). Due to the frequency dependence of the electron-phonon self-energy, $E^*_g \neq 2m^*$. Conceptually, the difference between the renormalized band gap (which is the bulk gap measured, e.g., in ARPES) and the renormalized Dirac mass (which dictates the existence of topological surface states) implies that the emergence or disappearance of topological surface states does not go hand in hand with the closing of the bulk quasiparticle gap. In Sec. V B, we argue that this difference may help explain recent ARPES measurements [19] that have probed the vicinity of the topological phase transition in $BiTl(S_{1-\delta}\ Se_\delta)_2$.

IV. PHONON-INDUCED TOPOLOGICAL INSULATION

Having reviewed the preliminary concepts, we are ready to discuss when and why phonons favor topological phases in Dirac insulators. The objective of this section is to extract some general principles that govern the fate of band topology in presence of electron-phonon interactions. These principles should be valid beyond the toy model that we use to illustrate them.

Starting from Eq. (12), we denote the self-energy for the positive-energy and negative-energy bands (each of which is doubly degenerate) as Σ_+ and Σ_-, respectively. In the spirit of Sec. II, we separate the intraband and interband contributions:

$$\Sigma_\pm(\mathbf{0},i\omega) = \Sigma^{\mathrm{intra}}_\pm(\mathbf{0},i\omega) + \Sigma^{\mathrm{inter}}_\pm(\mathbf{0},i\omega). \tag{21}$$

From Eq. (14), it follows that

$$\Sigma^{\mathrm{intra}}_\pm(\mathbf{0},i\omega) = \sum_{\mathbf{q}} g^2_{\mathbf{q},\mathrm{eff}}\frac{\left|V^{\mathrm{intra}}_{\mathbf{q}}\right|^2}{i\omega+\epsilon_F-E_{\mathbf{q}\pm}},$$
$$\Sigma^{\mathrm{inter}}_\pm(\mathbf{0},i\omega) = \sum_{\mathbf{q}} g^2_{\mathbf{q},\mathrm{eff}}\frac{\left|V^{\mathrm{inter}}_{\mathbf{q}}\right|^2}{i\omega+\epsilon_F-E_{\mathbf{q}\mp}}, \tag{22}$$

where the intraband and interband matrix elements ($V^{\mathrm{intra}}_{\mathbf{q}}$ and $V^{\mathrm{inter}}_{\mathbf{q}}$) describe the probability amplitude for electronic transitions from $\mathbf{k}=\mathbf{0}$ to $\mathbf{k}=\mathbf{q}$. For intraband transitions, the initial and final scattering states are in bands with the same energy dispersion; for interband transitions, one of them is in the conduction band and the other one is in the valence band. The explicit expressions for the matrix elements are

$$\left|V^{\mathrm{intra}}_{\mathbf{q}}\right|^2 = \frac{1}{2}\left(1+\mathrm{sgn}(m)\frac{M_{\mathbf{q}}}{\epsilon_{\mathbf{q}}}\right),$$
$$\left|V^{\mathrm{inter}}_{\mathbf{q}}\right|^2 = \frac{1}{2}\left(1-\mathrm{sgn}(m)\frac{M_{\mathbf{q}}}{\epsilon_{\mathbf{q}}}\right), \tag{23}$$

where $M_{\mathbf{q}}$ and $\epsilon_{\mathbf{q}}$ were defined below Eq. (12) and in Eq. (13), respectively. Note that $|V^{\mathrm{intra}}_{\mathbf{q}}|^2, |V^{\mathrm{inter}}_{\mathbf{q}}|^2 \in [0,1]$ because they correspond to transition probabilities. In order to interpret Eq. (23), it is useful to recall that τ^z is a good quantum number at $\mathbf{k}=\mathbf{0}$ [cf. Eq. (12)] and that phonons conserve the orbital pseudospin [cf. Eqs. (5) and (8)]. Then, Eq. (23) gives a measure of how parallel the orbital pseudospin at $\mathbf{k}=\mathbf{q}$ is with respect to that at $\mathbf{k}=\mathbf{0}$. This is illustrated in Figs. 2(a) and 2(b). The reason why Eq. (23) depends on the sign of the bare Dirac mass m is because the orbital pseudospin at $\mathbf{k}=\mathbf{0}$ flips direction when m changes sign. This, in turn, is the essence of band inversion.

Combining Eqs. (16), (21), and (22), the renormalized Dirac mass is obtained from

$$\Sigma_z = \mathrm{sgn}(m)(\Sigma_+ - \Sigma_-)/2, \tag{24}$$

where the momentum and frequency arguments are $\mathbf{0}$ and πT, respectively. The sign change in the expression for Σ_z between $m>0$ and $m<0$ is due to the band inversion. When

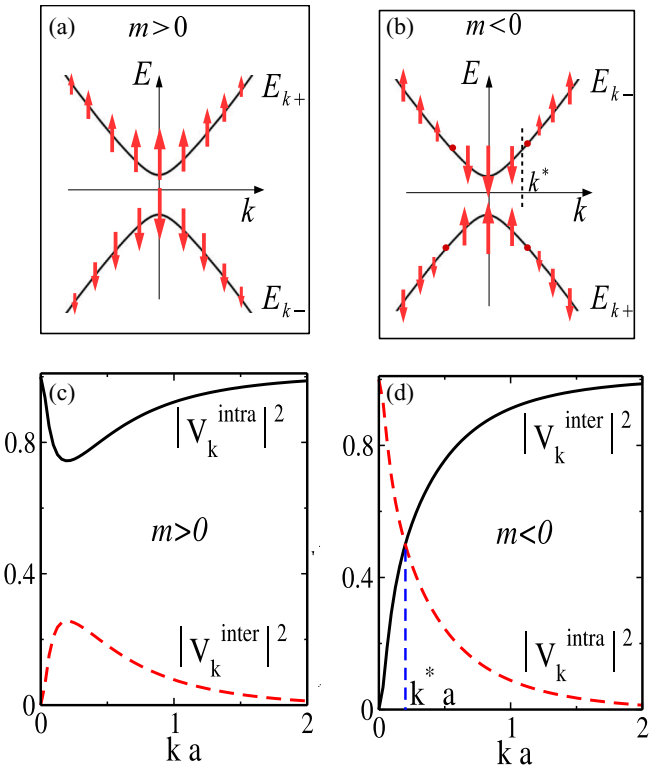

FIG. 2. (Color online) (a) and (b) Expectation value of $\boldsymbol{\tau}$ (arrows) as a function of momentum in (a) a trivial insulator and in (b) a topological insulator. From the analytical expressions for the band eigenstates, we obtain $\langle\boldsymbol{\tau}\rangle_{\mathbf{k}\pm} = \hat{z}M_{\mathbf{k}}/E_{\mathbf{k}\pm}$. Degenerate energy bands give the same contribution to $\langle\tau^z\rangle$; in contrast, $\langle\tau^x\rangle_{\mathbf{k}\pm} = \langle\tau^y\rangle_{\mathbf{k}\pm} = 0$ upon summing over each pair of degenerate bands. (c) and (d) Electronic scattering probability from the zone center to a state with momentum $\mathbf{k}$. The scattering probability is maximized if the spin and pseudospin of the state at momentum $\mathbf{k}$ are parallel to those of the state at the zone center. (c) In a trivial insulator, intraband matrix elements dominate at all momenta transfer because $\langle\tau^z\rangle$ does not change sign with $\mathbf{k}$. (d) In a topological insulator, intraband matrix elements dominate for $k < k^*$, where k^* is defined via $M_{\mathbf{k}^*} = 0$, while interband matrix elements dominate for $k > k^*$.

$m = 0$, the labeling of "intraband" and "interband" becomes ambiguous because the conduction and valence bands touch at $k = 0$. In this case one may take either $\mathrm{sgn}(m) = 1$ or $\mathrm{sgn}(m) = -1$ in Eqs. (23) and (24), and one arrives at the same expression for m^*.

Let us first consider $m > 0$, which corresponds to a trivial bare Dirac insulator. In this case, $M_{\mathbf{q}} > 0$ and thus $|V_{\mathbf{q}}^{\mathrm{intra}}| > |V_{\mathbf{q}}^{\mathrm{inter}}|$ for any $\mathbf{q}$. Namely, phonons tend to scatter the $\mathbf{k} = \mathbf{0}$ electron into another state in the same (or different-but-degenerate) band because the orbital pseudospin is more aligned therein [cf. Fig. 2(a)]. The importance of this point becomes easier to grasp if we consider an undoped ($\epsilon_F = 0$) and particle-hole symmetric ($d_{0,\mathbf{q}} = 0$) insulator at zero temperature. In this case $\Sigma_+^{\mathrm{intra}}(\mathbf{0},0) < 0 < \Sigma_+^{\mathrm{inter}}(\mathbf{0},0)$, with $|\Sigma_+^{\mathrm{intra}}(\mathbf{0},0)| > \Sigma_+^{\mathrm{inter}}(\mathbf{0},0)$. Thus $\Sigma_+(\mathbf{0},0) < 0$ and, by particle-hole symmetry, $\Sigma_-(\mathbf{0},0) = -\Sigma_+(\mathbf{0},0)$. Accordingly, Eqs. (18) and (24) dictate that $m^* < m$. In other words, phonons drive an undoped and particle-hole symmetric trivial Dirac insulator towards the topological phase, solely because the intraband matrix elements prevail over the interband matrix elements. For a sufficiently strong electron-phonon coupling (or sufficiently small m), $m^* < 0$ and phonons induce a topological insulating phase in an otherwise trivial insulator [20]. At the same time, phonons decrease the renormalized gap and may as well invert it, although E_g^* does not change sign at the same time as m^*.

Next, we consider the case $m < 0$. Here, $M_{\mathbf{q}}$ changes from negative to positive as q varies from 0 to π/a. The sign change, which occurs at $q\,a \simeq q^*\,a \equiv \sqrt{|m|/\beta}$, reflects a nontrivial texture of the orbital pseudospin in momentum space [cf. Fig. 2(b)]. Since phonons favor transitions between aligned orbital pseudospins, it follows that $|V_{\mathbf{q}}^{\mathrm{intra}}| > |V_{\mathbf{q}}^{\mathrm{inter}}|$ for $q < q^*$ and $|V_{\mathbf{q}}^{\mathrm{intra}}| < |V_{\mathbf{q}}^{\mathrm{inter}}|$ for $q > q^*$. Let us once again consider an undoped Dirac insulator with particle-hole symmetry. If the main contribution to the self-energy originated from electron-phonon scattering processes with small momentum transfer ($q < q^*$), the intraband contribution would dominate, thereby resulting in $m^* > m$ and $E_g^* > E_g$ (i.e., $|m^*| < |m|$ and $|E_g^*| < |E_g|$). In this scenario, phonons would have driven a topological insulator towards the trivial phase and, in conjunction with the preceding paragraph, we would have concluded that phonons favor the Dirac semimetal phase irrespective of whether the insulator was topological or trivial to begin with. However, in a narrow gap Dirac insulator, $q^*\,a \ll 1$ and typically the most important scattering events are those with $q > q^*$. For one thing, there is more phase space for transitions with higher momentum transfer. Therefore, interband processes make the main contribution to the self-energy and lead to $m^* < m$ ($|m^*| > |m|$), hence stabilizing the topological phase on a system that was already topological to begin with. In terms of the band gap, $E_g^* < E_g$ ($|E_g^*| > |E_g|$). As $|m|$ becomes larger, the phase space for interband transitions shrinks and intraband contributions begin to dominate, thereby restoring the "conventional" behavior of $|m^*| < |m|$ and $|E_g^*| < |E_g|$. Thus $\Sigma_z(\mathbf{0},0)$ changes sign in the topological phase as a function of $|m|$ (cf. Fig. 3).

In sum, electron-phonon interactions of the deformation potential type favor a topological insulator phase in narrow-gap Dirac insulators with particle-hole symmetry. The causes behind this unusual effect are the following: (i) the momentum-space texture of the orbital pseudospin changes across a band inversion, (ii) electron-phonon scattering of the deformation potential type conserves spin and pseudospin degrees of freedom, and (iii) electron-phonon matrix elements with higher momentum transfer are important (in spite of the larger energy denominators associated to them) due to the increased phase-space for scattering [21]. These three principles might help guide the understanding of how phonons influence the band topology in real Dirac materials with complex band structures. Moreover, the ideas developed above apply at finite temperature as well. Given that higher temperature means stronger effective electron-phonon coupling [cf. Eq. (17)], there is the intriguing possibility that heating the system up may drive a trivial insulator into the topological phase [10]. Thus far there are no known materials where helical surface states appear only above certain temperature [22]. In Sec. V we discuss related phonon-effects which may be accessible in some Dirac materials of current experimental interest.

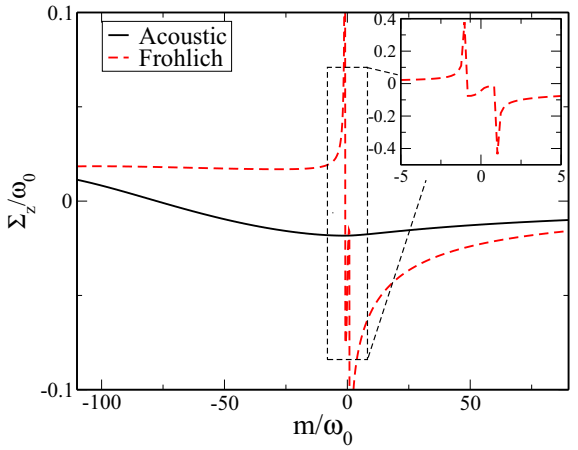

FIG. 3. (Color online) Dependence of the electron-phonon self-energy on the bare Dirac mass for different types of electron-phonon couplings $g_{\mathbf{q}}$. The axes are normalized in units of $\omega_0 = 20$ meV, the optical phonon frequency used for the Fröhlich coupling. The component of the self-energy that is responsible for renormalizing the Dirac mass, Σ_z, changes sign as a function of m when m becomes increasingly negative. The reason for this is explained in the main text. The value of $|m|$ for which the sign change occurs depends on the particular type of electron-phonon coupling $g_{\mathbf{q}}$. If $g_{\mathbf{q}\to\mathbf{0}}$ is finite (or zero), then the sign change occurs when $q^* \sim a^{-1}$. In contrast, if $g_{\mathbf{q}\to\mathbf{0}}$ diverges (which is the case for Fröhlich coupling in absence of screening), then small momenta transitions receive higher weight and Σ_z changes sign at a considerably smaller value of q^* (see inset).

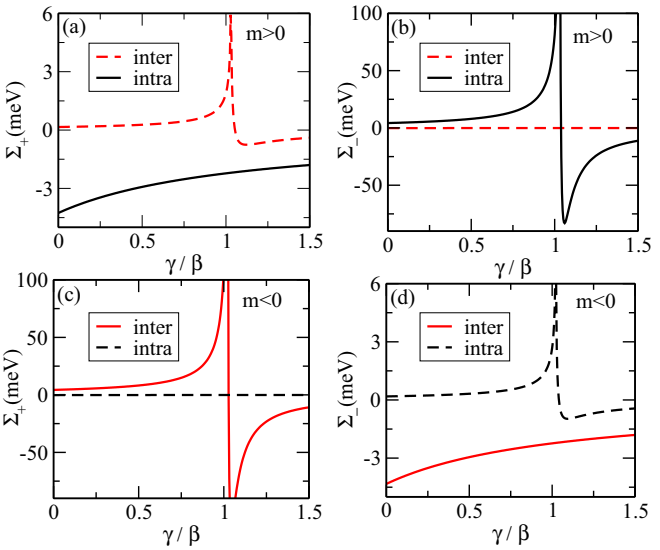

FIG. 4. (Color online) Electron-phonon self-energies for the positive- and negative-energy bands of an undoped Dirac insulator, with deformation potential coupling to acoustic phonons. Intraband and interband contributions are shown separately. The numerical values of the band and phonon parameters are taken from Ref. [10]. (a) and (b) Trivial Dirac insulator with $m = 10$ meV. (c) and (d) Topological insulator with $m = -10$ meV.

The preceding arguments relied on particle-hole symmetry and on the absence of itinerant carriers. Is the mechanism for phonon-induced topological insulation robust under doping ($\epsilon_F \neq 0$) and particle-hole asymmetry ($\gamma \neq 0$)? On one hand, electron-phonon matrix elements are unchanged by $\epsilon_F \neq 0$ and $\gamma \neq 0$. Yet, on the other hand, both γ and ϵ_F change the denominators in Eq. (22). In the remaining part of this section, we argue that the conclusions extracted in the preceding paragraph remain robust under moderate particle-hole asymmetry and doping.

Figure 4 shows the dependence of $\Sigma_\pm(\mathbf{0},0)$ on γ for an undoped insulator. When $m > 0$, the intraband contribution dominates up to a critical value $\gamma_c \simeq \beta$. Likewise, when $m < 0$, the interband contribution dominates up to $\gamma_c \simeq \beta$. Therefore, the status quo derived from $\gamma = 0$ remains qualitatively unchanged until $\gamma \gtrsim \gamma_c$. As shown in Fig. 5(b), the value of γ_c may be shifted away from β via doping. When $\gamma = \gamma_c$, the electron-phonon self-energy at zero frequency develops a singularity due to a elastic transition that connects the Fermi-level band eigenstate at $\mathbf{k} = \mathbf{0}$ with the band edge at the corner of the Brillouin zone [see also Fig. 5(a)]. Close to the singularity, our perturbative result is unreliable. For $\gamma \gtrsim \gamma_c$, the system is first an indirect-gap semiconductor and then a semimetal. In this regime, the sign of the self-energy is reversed and phonons favor a topologically trivial phase. Since β is a significant fraction of the electronic bandwidth, we conclude that the mechanism for phonon-induced topological insulation remains robust for moderate particle-hole asymmetry.

Figure 5(b) illustrates the dependence of Σ_z on ϵ_F and γ. In the lightly doped materials we are interested in, the screening of the electron-phonon matrix elements by the itinerant carriers is weak [23]. Moreover, in this regime the dependence of the self-energy on doping is unremarkable, except close to $\gamma = \gamma_c$. Hence once again the simple mechanism discussed

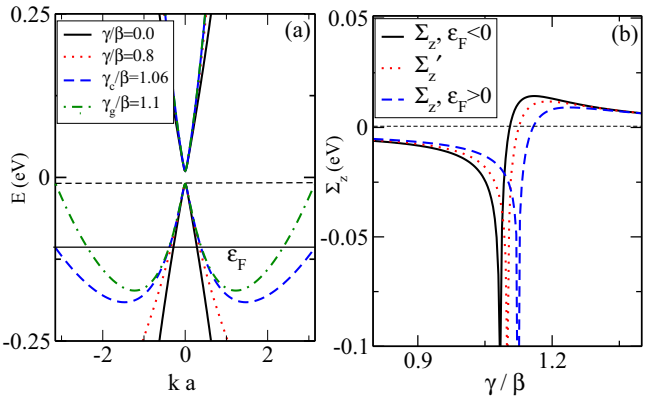

FIG. 5. (Color online) (a) Band structure of a toy Dirac insulator [Eq. (12)] as a function of the particle-hole asymmetry parameter γ. When $\gamma \simeq \beta$, the phonon-induced renormalization of the band gap and the Dirac mass develop singularities. (b) Let us denote the renormalized Dirac mass and band gap as $m^* = m + \Sigma_z$ and $E_g^* = E_g + \Sigma_z'$, respectively, where Σ_z' may be read off Eq. (20). Then, Σ_z diverges and changes sign at $\gamma \equiv \gamma_c$, when the band edge at a corner of the Brillouin zone becomes degenerate with the Fermi energy. Hence, γ_c may be tuned by doping. Likewise, Σ_z' diverges and changes sign at $\gamma \equiv \gamma_g$, when the band edge at the zone center becomes degenerate with the band edge at a corner of the Brillouin zone. Unlike γ_c, γ_g is insensitive to doping. Such different response of γ_c and γ_g to doping will play a role in Sec. V B.

above is applicable for moderately doped Dirac systems. Along the same line, we have confirmed that the neglect of $\omega_{\mathbf{q}}$ on the denominators of Eq. (16) is appropriate except near the singularities of the self-energy (not shown).

V. SOME APPLICATIONS

Thus far we have explained why, in narrow-gap Dirac insulators which are not highly doped or highly particle-hole asymmetric, long-wavelength phonons favor the topological insulating phase. An experimental signature of this phenomenon would be the emergence of helical surface states *above* certain temperature in an insulator that has a topologically trivial ground state. Traces of this hitherto unobserved phenomenon are more likely to be seen in materials with tunable band gaps. HgTe/CdTe quantum wells and $BiTl(S_{1-\delta}\ Se_{\delta})_2$ are examples of such materials in two and three dimensions, respectively. In this section, we discuss signatures of phonon-induced changes in the band topology of these systems.

A. Temperature-dependence of the critical width in HgTe/CdTe quantum wells

Topological insulation in CdTe/HgTe/CdTe quantum wells was predicted [24] by Bernevig, Hughes and Zhang (BHZ) in 2006. The experimental confirmation arrived shortly afterwards [25,26]. In CdTe, as in most tetrahedral semiconductors [27], the p-type valence band edge (Γ_8) lies below the s-type conduction band edge (Γ_6). In this "normal-ordered" electronic structure, the energy gap is $E_g = 1.6$ eV. In contrast, in HgTe, Γ_8 lies above Γ_6 and hence the energy gap at the Γ point is inverted ($E_g = -0.303$ eV). Even though bulk HgTe is semimetallic in absence of strain, it may be coaxed into the insulating phase through quantum confinement in CdTe/HgTe/CdTe quantum wells. The HgTe layer has a small thickness d along the growth direction z and the heterostructure is translationally invariant in the xy plane. Accordingly, $\mathbf{k}_\perp = (k_x, k_y)$ are good quantum numbers. At $k_\perp = 0$, the lowest-energy subbands in the quantum well are denoted as $E1$ and $H1$, their energies being $E_{\rm E1}$ and $E_{\rm H1}$. The dispersions of $E1$ and $H1$ with $k_\perp$ are equivalent to those of Dirac fermions with mass $m = (E_{\rm E1} - E_{\rm H1})/2$. If $d < d_c$, where d_c is some critical thickness, the normal ordering of the CdTe electronic structure prevails. This translates into $m > 0$ and trivial insulation. If $d > d_c$, the ordering between $E1$ and $H1$ subbands is inverted ($m < 0$) and the system becomes a topological insulator. The objective of this subsection is to investigate the effect of electron-phonon interactions on d_c.

The natural starting point for such investigation is the BHZ model [24], which describes the low-energy subbands in the vicinity of $k_\perp = 0$. This model is a special case of Eq. (12), with the Dirac mass given by $m = (E_{\rm E1} - E_{\rm H1})/2$. In presence of phonons, m is renormalized to $m^*(T)$ and the critical thickness for the topological transition or crossover is the one for which $m^*(T) = 0$. Anticipating that $|\gamma| < |\beta|$ for all relevant values of d [cf. Fig. 6(b)], it follows from Sec. IV that $\mathrm{Re}[\Sigma_z(\mathbf{0}, i\pi T)] < 0$ for any temperature T. Namely, within the BHZ model, phonons favor the topological phase and therefore one may expect d_c to *decrease* as the system is heated.

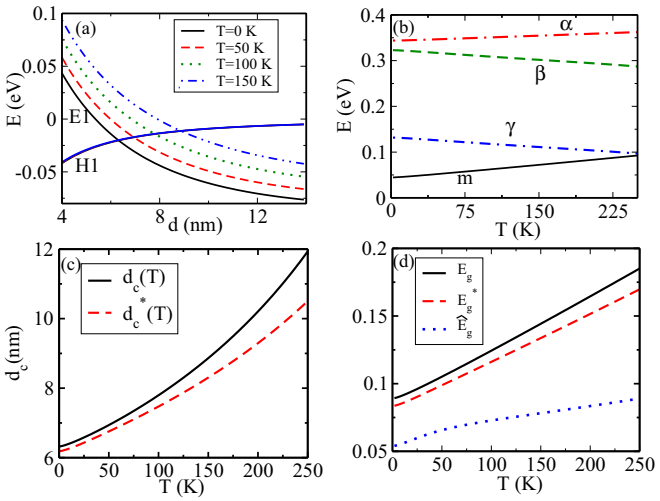

FIG. 6. (Color online) (a) Energies of $E1$ and $H1$ subbands ($k_\perp = 0$) as a function of quantum well width d. The parameters needed for the calculation are taken from Ref. [30]; in addition, we consider $\epsilon_F = 0$. (b) Temperature dependence of the band parameters appearing in the BHZ model [Eq. (12)], for $d = 4$ nm. (c) Plot of the critical quantum well [where $m^*(T)$ changes sign] as a function of temperature. The solid line ignores phonon-induced intersubband transitions, while the dashed curve partly captures them through the evaluation of the electron-phonon self-energy in the BHZ model. (d) Energy-difference between $E1$ and $H1$ subbands, as a function of temperature, for a quantum well of thickness $d = 4$ nm. The solid line is the result without phonon-induced intersubband transitions. The dashed line incorporates the phonon effects in the lowest quantum well states. The doted line ($\hat{E}_g$) takes a slightly larger value of the electron-phonon coupling (by a factor of $\sim$2.5); the change in the slope of the energy gap as a function of temperature becomes noticeable in this case.

This expectation is in stark contrast with the conclusions from a recent theoretical study by Sengupta *et al.* [28], which has claimed that d_c increases with T. These authors considered the effect of electron-phonon interactions solely through the renormalization of the band gaps in *bulk* HgTe and CdTe. Such approach is insufficient because it does not capture the influence of phonon-induced transitions between quantum well states. These transitions, partly included in the electron-phonon self-energy of the BHZ model, are in principle important because they connect states that are close in energy.

We improve on Ref. [28] by taking a two-pronged approach. First, we evaluate the temperature dependence of the band parameters $(m, \alpha, \beta, \gamma)$ appearing in the BHZ model. This T dependence comes through the phonon-induced renormalization of the bulk CdTe and HgTe band gaps. The thickness at which $m(T)$ changes sign will be denoted as $d_c(T)$; this is the quantity that was calculated by Ref. [28] and found to increase with T. Second, we use the temperature-dependent band parameters as input to calculate the electron-phonon self-energy within the BHZ model. In this way, we determine the renormalized Dirac mass via $m^*(T) = m(T) + \mathrm{Re}[\Sigma_z(\mathbf{0}, i\pi T)]$. The actual critical width $d_c^*(T)$ is defined as the thickness for which $m^*(T) = 0$. Since $\mathrm{Re}[\Sigma_z(\mathbf{0}, i\pi T)] < 0$, there is the possibility that $d_c^*(T)$ decreases with T even as $d_c(T)$ increases with T.

In order to determine the temperature dependence of the BHZ parameters, we solve the Schrödinger equation for a $Hg_{0.32}Cd_{0.68}Te/HgTe/Hg_{0.32}Cd_{0.68}Te$ quantum well. This involves diagonalizing a six-band Kane Hamiltonian $h(\mathbf{k}_\perp, i\partial_z)$ [29]. The ensuing procedure is identical to that of Ref. [24], except that we take temperature-dependent band gaps [30] for bulk HgTe and bulk $Hg_{0.32}Cd_{0.68}Te$. The solution at $k_\perp = 0$ reveals discrete quantum well states, from which $E1\pm$ and $H1\pm$ have the lowest energies (here $\pm$ labels Kramers partners that are degenerate at $k_\perp = 0$ due to time-reversal symmetry). The BHZ model follows from applying $\mathbf{k}\cdot\mathbf{p}$ perturbation theory around the $k_\perp = 0$ solution in the Hilbert space spanned by $\{|E1+\rangle, |H1+\rangle, |E1-\rangle, |H1-\rangle\}$. By considering only the lowest electronlike and holelike subbands, the BHZ model is two-dimensional as far as electrons are concerned. However, each of the four states forming the low-energy subspace has an associated spinor, whose six components vary with z. We denote these spinors as $\chi_{\sigma\tau}(z)$, where $\sigma = \pm$ and $\tau = E1, H1$.

Figure 6(a) displays the calculated dependence of E_{E1} and E_{H1} on d and T. It is apparent that $E_{E1} - E_{H1}$ increases with temperature when $d < d_c$, while $|E_{E1} - E_{H1}|$ decreases with temperature when $d > d_c$. Hence, the effect of phonons on the *bulk* states of CdTe and HgTe favors the normal (i.e., topologically trivial) ordering between $E1$ and $H1$. Accordingly, Fig. 6 c shows that $d_c(T)$ increases with temperature, in agreement with the result of Sengupta *et al.* [28]. At any rate, $d_c(T)$ is *not* the actual critical thickness because we have yet to consider the effect of phonons in the low-energy subspace spanned by $\{|E1\pm\rangle, |H1\pm\rangle\}$.

In order to obtain the actual critical thickness d_c^*, we evaluate $\Sigma_z(\mathbf{0}, i\pi T)$ within the BHZ model, using the T- and d-dependent band parameters derived above. Owing to a lack of translational invariance along the growth direction, the expression for Σ_z differs by a form factor [31] from that of Eq. (24). This form factor can be derived by recasting the electron density in Eq. (5) as

$$\rho(\mathbf{r}) = \frac{1}{A}\sum_{\mathbf{k}_\perp\mathbf{q}_\perp}\sum_{\tau\tau'\sigma\sigma'} e^{-i\mathbf{q}_\perp\cdot\mathbf{r}_\perp}(\chi^*_{\sigma\tau}|\chi_{\sigma'\tau'})c^\dagger_{\mathbf{k}_\perp\sigma\tau}c_{\mathbf{k}_\perp-\mathbf{q}_\perp\sigma'\tau'}, \quad (25)$$

where A is the area of the sample in the xy plane, $\mathbf{k}_\perp = (k_x, k_y)$, $\mathbf{q}_\perp = (q_x, q_y)$, $\mathbf{r} = (\mathbf{r}_\perp, z)$ and $(\chi^*_{\sigma\tau}|\chi_{\sigma'\tau'})$ is the z-dependent scalar product between $\chi_{\sigma\tau}(z)$ and $\chi^*_{\sigma'\tau'}(z)$. A direct calculation shows that $(\chi^*_{\sigma\tau}|\chi_{\sigma'\tau'}) = (\chi^*_{\sigma\tau}|\chi_{\sigma\tau})\delta_{\tau\tau'}\delta_{\sigma\sigma'}$. Then, the combination of the first line of Eq. (8) with Eq. (25) yields

$$\Sigma_z(\mathbf{0}, i\omega) \simeq \sum_{\mathbf{q}} F_{q_z}\frac{g^2_{\mathbf{q},\mathrm{eff}}\, M_{\mathbf{q}_\perp}}{(\epsilon_F + i\omega - d_{0,\mathbf{q}_\perp})^2 - \epsilon^2_{\mathbf{q}_\perp}}, \quad (26)$$

where $\mathbf{q} = (\mathbf{q}_\perp, q_z)$. The form factor is given by

$$F_{q_z} = \int_{-\infty}^{\infty} dz dz' |\chi(z)|^2 |\chi(z')|^2 e^{-iq_z(z-z')}, \quad (27)$$

where $|\chi(z)|^2 \equiv \sum_{\tau\sigma}(\chi^*_{\sigma\tau}|\chi_{\sigma\tau})$. In the numerical evaluation of Eq. (26), we sum over the contributions from three types of electron-phonon interactions: deformation potential coupling to acoustic phonons, deformation potential coupling to optical phonons, and polar optical (Fröhlich) coupling. Their respective numerical values are listed in Appendix B.

The final outcome of our calculation is collected in Figs. 6(c) and 6(d), which display d_c^* and the renormalized band gap as a function of T. Despite $\mathrm{Re}[\Sigma_z(\mathbf{0}, i\pi T)] < 0$, we find that d_c^* increases with temperature because $m(T)$ increases rather rapidly with T [cf. Fig. 6(d)]. That is, for a given quantum well that is topological insulating at $T = 0$, increasing temperature produces a crossover into the trivial phase. Hence the conclusion of Ref. [28] is qualitatively correct, although it overestimates the increase of the critical thickness as a function of temperature. In Fig. 6(d), we plot the temperature dependence of the band gap, $E^*_{E1} - E^*_{H1}$, which is experimentally measurable [30]. In the absence of phonon-induced intersubband transitions, the gap increases linearly with temperature starting at low temperature. However, the inclusion of phonon-induced intersubband transitions and their thermal activation results in a kink in the temperature dependence of the gap. The experimental observation of this kink would be an indirect indication of the tendency of phonons to favor a topological phase within the BHZ model.

In sum, the net outcome of electron-phonon interactions in HgTe/CdTe quantum wells is to drive the system closer to the trivial insulating phase. Had we ignored the temperature dependence of the band parameters of the BHZ model, we would have wrongly concluded that phonons favor the topological phase. This is a potentially important lesson that might also impact the theory of topological Anderson insulators [9], where the effect of disorder on the bulk states of CdTe and HgTe has been overlooked.

We close this section with a digression on graphene, which is another canonical two dimensional Dirac insulator. In graphene, the topological invariant is encoded in the relative sign between the masses of the two Dirac fermions located in the first Brillouin zone. In inversion and time-reversal symmetric systems, the magnitudes of the two Dirac masses are the same. Since electron-phonon interactions do not break any symmetries, they cannot change the sign of one Dirac mass without simultaneously changing the other. Hence, phonons cannot change the band topology of inversion-symmetric graphene. In contrast, phonons can alter the band topology of graphenelike systems without inversion symmetry [33]. There is yet another difference between the BHZ model for HgTe/CdTe and the Kane-Mele model [34] for graphene. In the BHZ model, electron-phonon interactions open a band gap even when $m = 0$ because $M_{\mathbf{k}} \simeq m + \beta k^2 \neq 0$ for $k \neq 0$. In graphene, phonons change the band gap only if the bare gap is nonzero to begin with [35], because $\beta = 0$ [36]. Incidentally, these differences are also the reason why strong disorder drives CdTe/HgTe quantum wells into a topological insulating phase [9], while it drives graphene into a metallic phase [37].

B. Fingerprints of phonons at the topological phase transition of BiTl $(S_{1-\delta}Se_\delta)_2$

In 2011, ARPES experiments [7] reported evidence for a topological phase transition in BiTl $(S_{1-\delta}Se_\delta)_2$ as a function of the stoichiometric ratio δ. The material exhibited helical surface states when $\delta > \delta_c \simeq 0.5$, with the bulk energy gap closing and reopening as δ was varied. Due to the finite resolution of ARPES and because of limitations in fine-tuning δ, the value of δ_c could not be measured accurately. However,

the authors reinforced the case for $\delta_c \simeq 0.5$ via first-principles electronic structure calculations. Soon afterwards, the same group completed a more thorough study [19] of δ_c and announced an unexpected finding: in-gap states emerged at $\delta \simeq 0.4$–0.5, *prior* to the bulk gap closing. Those in-gap states showed no dispersion along the direction normal to the surface and displayed the spin helicity characteristic of topological surface states. The authors speculated on a "topological proximity effect" as a possible origin of the phenomenon. In this section, we argue that it may instead be a fingerprint of electron-phonon interactions.

From the arguments of Sec. IV, we infer that δ_c (defined as the stoichiometric ratio for which $m^* = 0$) must depend on the strength of electron-phonon interactions. Moreover, first-principles electronic structure calculations [32] show that the conduction band at the X point of the bulk Brillouin zone is nearly degenerate with the conduction band at the Γ point, which in turn suggests that the phonon-induced renormalization of the band gap can be significant.

Empirically, there are two ways to verify that δ_c depends on electron-phonon interactions. On one hand, the measured value of δ_c (which inevitably incorporates phonon effects) should be different from the value predicted by existing *ab initio* calculations [32] (which ignore phonons). Admittedly, electron-phonon interactions are not the only agents that can shift the value of δ_c with respect to the noninteracting case: Coulomb interactions and short-ranged nonmagnetic disorder may have an impact as well. On the other hand, if phonons are at play, the measured δ_c should be strongly temperature-dependent on the scale of the Debye temperature. The observation of such temperature-dependence would in fact be a true smoking gun for phonon-induced effects in the band topology of BiTl $(S_{1-\delta}Se_\delta)_2$, because neither static disorder nor Coulomb-like electron-electron interactions should produce a significantly temperature-dependent effect. Arguably, it is not easy to measure $\delta_c(T)$ because the thermal smearing of the quasiparticle bands prevents locating the exact point where surface states emerge. Nevertheless, it should be relatively easy to measure the temperature-dependence of the bulk band gap when δ is sufficiently far from δ_c. A phonon-induced reduction of δ_c would manifest itself through $d|E_g^*|/dT < 0$ (if $\delta \ll \delta_c$) and $d|E_g^*|/dT > 0$ (if $\delta \gg \delta_c$). Instead, if phonons increased δ_c with respect to the noninteracting case, the observed temperature-dependence of the band gap would be of opposite sign.

We model the low-energy electronic structure of BiTl $(S_{1-\delta}Se_\delta)_2$ qualitatively by a lattice version of Eq. (12). Upon selecting a value of $|\gamma| \lesssim |\beta|$, we can partially mimic the realistic scenario where the conduction band at the Brillouin zone edge is nearly degenerate with that of the zone center. The corresponding band structure is shown in Fig. 7. Therein, we have identified three different regions. When the Fermi energy is in region I, $|\gamma_c| < |\gamma| < |\gamma_g|$ and thus $\Sigma_z > 0 > \Sigma_z'$. For the definitions of γ_c, γ_g and Σ_z', see the caption of Fig. 5. In this case, phonons favor a trivial insulating phase. In contrast, in both regions II and III phonons favor a topological insulating phase. When the Fermi energy is in region II, $|\gamma| < \gamma_c| < |\gamma_g|$ and $0 > \Sigma_z' > \Sigma_z$. Since $|\Sigma_z| > |\Sigma_z'|$, the Dirac mass renormalizes more strongly than the bulk energy gap. When the Fermi energy is in region III, $|\gamma| < |\gamma_g| < |\gamma_c|$ and $0 > \Sigma_z > \Sigma_z'$. Here, the Dirac mass renormalizes less strongly than the bulk energy gap. The consequences of this will be discussed below.

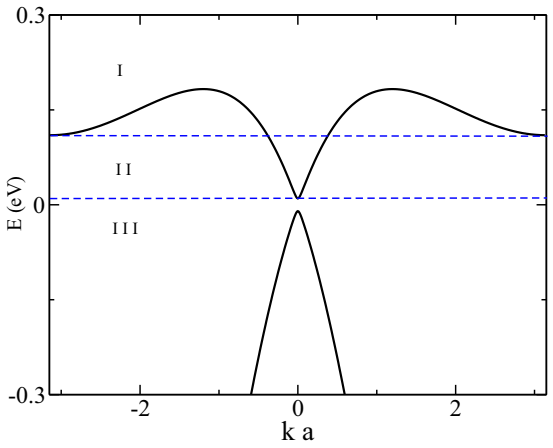

FIG. 7. (Color online) Qualitative modeling of the low-energy bands [32] of BiTl $(S_{1-\delta}Se_\delta)_2$, using the toy model of Eq. (12). The dashed lines separate three distinct regions of doping, which lead to three different types of phonon-induced effects. If the Fermi energy is in region I, phonons favor a topologically trivial phase. If the Fermi energy is in regions II or III, phonons favor a topological phase. The "topological proximity effect" of Ref. [19] takes place when the Fermi energy is in region II. In contrast, a "topological antiproximity effect" is set to occur when the Fermi energy is in region III.

Figure 8(c) illustrates the energies of the conduction and valence band edges [cf. Eq. (20)] as a function of δ, both in presence and absence of electron-phonon interactions. Clearly, phonons favor the topological insulating phase. Along the same line, Fig. 8(d) shows the reduction of $\delta_c(T)$ as a function of temperature and predicts the emergence of helical surface states beyond a crossover temperature, when $\delta < 0.5$.

Throughout these plots, we have extrapolated the experimental data of Ref. [7] into a linear relation between the bare Dirac mass m (which is half the noninteracting band gap) and δ. More complicated $m(\delta)$ functions would not change our conclusions qualitatively. In addition, we have neglected the thermal expansion of the lattice. Thermal expansion renders all the band parameters of Eq. (12) temperature dependent even before the inclusion of the electron-phonon self-energy. In the previous section, we have learned that such "extraneous" temperature-dependence can potentially revert the trend that one would have anticipated solely from the calculation of the self-energy. We justify our approach on the basis that, in most semiconductors, the contribution of thermal expansion to band-gap renormalization is small enough that its neglect poses no risk for qualitative error [38].

We now propose a possible explanation for the "topological proximity effect" reported in Ref. [19]. Topological surface states take place when $m^* < 0$, while the bulk gap measured in ARPES corresponds to $|E_g^*|$. As mentioned in Sec. III, $m^* \neq E_g^*/2$ in presence of interactions. In Figs. 8(a) and 8(b), we plot m^* and E_g^* as a function of δ for different electron densities. Since S and Se belong to the same group in the periodic table, changing δ does not change the carrier density. We find that, when the system is neutral or hole-doped [cf.

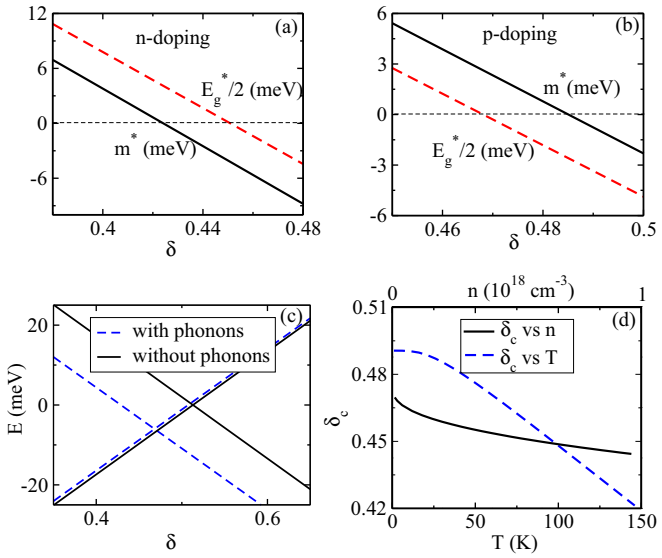

FIG. 8. (Color online) (a) and (b) Zero-temperature renormalized Dirac mass m^* and half-band-gap $E_g^*/2$ of BiTl $(S_{1-\delta}Se_\delta)_2$ as a function of δ, when the Fermi energy is (a) in region II and (b) in region III. The electron and hole densities in (a) and (b) are $n = 10^{19}\text{cm}^{-3}$ and $p = 10^{19}\text{cm}^{-3}$, respectively. In (a), the emergence of topological surface states precedes the band inversion ("topological proximity effect"). In (b), the emergence of topological surface states succeeds the band inversion ("topological antiproximity effect"). (c) Influence of electron-phonon interactions on the renormalized conduction and valence band edges. The solid (dashed) lines correspond to the absence (presence) of electron-phonon interactions. Phonons decrease the magnitude of the band gap in the trivial phase, while they enhance it in the topological phase. The mechanism behind this effect has been explained in Sec. IV. (d) Temperature and density dependence of the critical stoichiometric ratio. Throughout this figure we have taken $\gamma = -0.22$ eV; all other band and phonon parameters are the same as in Ref. [10].

Fig. 8(b)], i.e., when the Fermi energy is located in region III, there exists an interval of δ for which $E_g^* < 0$ and $m^* > 0$. This is a consequence of $0 < \Sigma_z < \Sigma_z'$ and it indicates an onset of topological surface states after (and not simultaneously with) the occurrence of a band inversion. In contrast, in a weakly electron-doped system [such that the Fermi energy is in region II, cf. Fig. 8(a)], we find $m^* < 0$ and $E_g^* > 0$ for $\delta \in (\delta_c, \delta_c + \Delta\delta)$. Assuming a reasonable strength of the electron-phonon coupling [10], $\Delta\delta \simeq 0.2$, the band gap at $\delta = \delta_c$ is $E_g^* \simeq 10$ meV and the band broadening due to electron-phonon interactions remains small. In such scenario, the onset of topological surface states *precedes* the occurrence of a band inversion. These results appear to be consistent with Ref. [19], whose samples are electron-doped. Coulomb interactions are unlikely to be responsible for this effect because their associated self-energy has a weak frequency dependence. On the other hand, although the disorder self-energy is frequency dependent, its temperature dependence is negligible. Hence the observation of a temperature and density dependent $\Delta\delta$ would confirm the key role of phonons.

Finally, we comment on the imaginary part of the self-energy, which gives the broadening of the renormalized quasiparticle bands. In order to observe phonon-induced topological surface states, it is essential that the phonon-induced broadening of the band gap at $k \simeq 0$ be small. It is easy to show that, so long as the width of region II in Fig. 7 is larger than the characteristic phonon energy scale, the phonon-induced band broadening of the $k = 0$ band gap is negligible. This is the case for the parameter values taken in Fig. 8.

VI. SUMMARY AND DISCUSSION

Electron-phonon interactions can induce topological insulation in a narrow-gap Dirac material with an intrinsically trivial electronic structure. The essential ingredients behind this phenomenon are the following: (i) a direct and small band gap, (ii) the change in the momentum-space texture of the band eigenstates from the trivial to the topological phase, (iii) the conservation of spin and orbital degrees of freedom in electron-phonon scattering processes involving long-wavelength phonons, and (iv) the importance of electron-phonon matrix elements with high momentum transfer (due to the increased scattering phase-space associated to them.)

Together, these ingredients produce a peculiar outcome; while the leading electron-phonon matrix elements in a trivial Dirac insulator are those in which a phonon scatters an electron within the same band, the dominant matrix elements in a narrow-gap topological insulator are those in which a phonon scatters an electron between the conduction and the valence band. A direct consequence of this peculiarity in the electron-phonon matrix elements is that the band gap of a trivial (topological) Dirac insulator decreases (increases) as temperature is raised. This, in turn, anticipates the emergence of helical surface states beyond a crossover temperature in a Dirac insulator with a topologically trivial ground state.

The above mechanism is not exclusive to phonons and can be transferred to spin-independent disorder and Coulomb interactions. The large dielectric constant of common Dirac insulators implies that the effect of Coulomb interactions in band topology will often be small in comparison to that of electron-phonon interactions. As for disorder, the signatures of topological Anderson insulation must be accessed by transport experiments because it is difficult to measure the band gap as a function of random impurity concentration. Such transport experiments are rather contrived due to the unintended bulk doping that is prevalent in many Dirac materials. In contrast, measuring the temperature dependence of a bulk band gap is routine. In sum, the main aspect that sets phonons apart from other agents is that their influence on band topology is significantly temperature-dependent.

Signatures of phonon-induced topological insulation have not yet been confirmed in experiment. What are the materials to look for and what should be measured? In principle, any Dirac insulator with a direct and small band gap can display the phenomenon. A small gap is necessary because otherwise intraband electron-phonon scattering processes dominate and lead to a decrease of the band gap regardless of the band topology. Since the typical electron-phonon self-energies are ~10–100 meV, the most spectacular phonon effects occur in the vicinity of a topological phase transition. As of this writing, there is an increased number of Dirac materials where experimentalists are able to apply pressure or change the

stoichiometry continuously in order to tune the band gap from the trivial to the topological phase, and vice versa. Should the band gap decrease (increase) with raising temperature in the trivial (topological) side of the transition, this would confirm that phonons favor a topological insulating phase. Recent experiments in $BiTl(S_{1-\delta}Se_{\delta})$ have reported that topological surface states emerge at $\delta = \delta_c - \Delta\delta < \delta_c$, whereas the bulk gap closes at $\delta = \delta_c$. Does $\Delta\delta$ depend on carrier concentration and on temperature? As explained in Sec. V B, an affirmative response would suggest that phonons are behind the observed effect.

We conclude by assessing the limitations of our theory, which relies on three assumptions: (i) the lowest conduction band and the highest valence band are well separated in energy from the rest of the bands, (ii) short wavelength phonons can be neglected, and (iii) the leading phonon effects originate from the coupling between the density of electrons and the lattice deformation.

The first assumption, which enables the use of low-energy effective models to arrive at a simple picture of phonon-induced topological insulation, is often justified in the vicinity of a topological phase transition. When phonon-induced transitions to high-energy bands become significant, we anticipate that a formally similar low-energy effective model will still be applicable, albeit with renormalized band parameters. Should this renormalization be strong enough, the influence of phonons on band topology might be reversed. For instance, in CdTe/HgTe quantum wells (cf. Sec. V A), the net effect after including electron-phonon interactions in high energy bands is that phonons favor the *trivial* phase.

The neglect of short wavelength phonons and of umklapp processes (and hence of the Debye-Waller term) has been done on the basis of simplicity and is nearly universal in textbooks. Yet, in real materials, the Debye-Waller contribution to band-gap renormalization may be of the same order as the self-energy contribution discussed in this work. Recognizing that Debye-Waller processes involve vertical (zero-momentum-transfer) interband transitions [16], which should be insensitive to the occurrence of a band inversion, we speculate that the Debye-Waller term will renormalize the band gap but not the Dirac mass (thereby not affecting the band topology).

Finally, the third assumption above ignores phonon-induced processes that alter electronic hopping amplitudes. Such terms have been discussed in graphene [39] and can also exist in Dirac insulators. They can be incorporated into our theory in an *ad hoc* fashion by promoting the electron-phonon coupling from an identity matrix in spin and orbital space to a matrix with off-diagonal elements. The off-diagonal elements might lead to phonon-induced *trivial* insulation. For instance, if a phonon mode exists which flips the orbital pseudospin, then this mode will directly oppose phonon-induced topological insulation. However, it is likely that the off-diagonal matrix elements of the electron-phonon coupling in the spin and orbital space are often small compared to the diagonal elements, much like in graphene.

In the future, it would be desirable to resort to first-principles calculations that relax our assumptions. Some initial efforts along this direction are underway [40]. It is our hope that the basic insights for phonon-induced topological insulation, unearthed here in the context of a toy model, will remain relevant after considering the full complexity of the electronic and phononic band structures.

ACKNOWLEDGMENTS

We acknowledge useful discussions or correspondence with S. Adam, G. Antonius, A. Bansil, M. Côté, D. Hsieh, Z. Li, H. Lin, and R. Nourafkan. The numerical calculations were performed on computers provided by Calcul Québec and Compute Canada. This research has been financially supported by Université de Sherbrooke, RQMP, and Canada's NSERC.

APPENDIX A: DEBYE-WALLER CONTRIBUTION

Let us expand the electron-ion potential with respect to small displacements of the ions from their equilibrium positions:

$$V_{ei}\left(\mathbf{r}-\mathbf{R}_j^{(0)}-\mathbf{Q}_j\right) = V_{ei}\left(\mathbf{r}-\mathbf{R}_j^{(0)}\right) - \mathbf{Q}_j\cdot\nabla V_{ei}\left(\mathbf{r}-\mathbf{R}_j^{(0)}\right) + \tfrac{1}{2}(\mathbf{Q}_j\cdot\nabla)^2 V_{ei}\left(\mathbf{r}-\mathbf{R}_j^{(0)}\right)+\cdots. \tag{A1}$$

In textbooks [13], only the first-order term in the expansion of Eq. (A1) is kept, which leads to $\mathcal{V}^{(1)}$ [cf. Eq. (8)] in the limit of long-wavelength phonons, with

$$g_{\mathbf{q}} = \sqrt{\frac{\hbar}{2\rho_A V\omega_{\mathbf{q}}}}\,\mathbf{q}\cdot\mathbf{e}_{\mathbf{q}} V_{ei}(\mathbf{q}). \tag{A2}$$

Thereafter, the effect of electron-phonon interactions on physical observables is computed through second order perturbation theory in $\mathcal{V}^{(1)}$. The term of order Q^2 in Eq. (A1) is ignored in textbooks. However, if treated in first-order perturbation theory, it contributes at the same order as the perturbation kept in textbooks. This contribution receives the name of "Debye-Waller term." In this appendix, we show that the Debye-Waller term vanishes in the limit of long-wavelength phonons.

Following the same steps [13] as in the derivation of $\mathcal{V}^{(1)}$, the electron-phonon interaction emerging from the $O(Q^2)$ term in Eq. (A1) can be written as

$$\mathcal{V}^{(2)} = \frac{1}{2}\sum_{\mathbf{k},\mathbf{q},\mathbf{G}} \rho_{\mathbf{q}+\mathbf{G}} V_{ei}(\mathbf{q}+\mathbf{G})[(\mathbf{q}+\mathbf{G})\cdot\mathbf{e}_{\mathbf{k}}][(\mathbf{q}+\mathbf{G})\cdot\mathbf{e}_{\mathbf{k}-\mathbf{q}}] \times\sqrt{\frac{\hbar}{2\rho_A V\omega_{\mathbf{k}}}}\sqrt{\frac{\hbar}{2\rho_A V\omega_{-\mathbf{k}+\mathbf{q}}}}(a_{\mathbf{k}}+a^{\dagger}_{-\mathbf{k}})(a_{-\mathbf{k}+\mathbf{q}}+a^{\dagger}_{\mathbf{k}-\mathbf{q}}), \tag{A3}$$

where $\mathbf{G}$ is a reciprocal lattice vector. If the short-wavelength phonons are neglected (which implies keeping only the $G = 0$ term [13] in the sum of Eq. (A3)), then Eq. (A3) reduces to the second line of Eq. (8), with

$$\lambda_{\mathbf{k}\mathbf{q}} = \sqrt{\frac{\hbar}{2\rho_A V\omega_{\mathbf{k}}}}\sqrt{\frac{\hbar}{2\rho_A V\omega_{-\mathbf{k}+\mathbf{q}}}}\,V_{ei}(\mathbf{q})(\mathbf{q}\cdot\mathbf{e}_{\mathbf{k}})\,(\mathbf{q}\cdot\mathbf{e}_{\mathbf{k}-\mathbf{q}}). \tag{A4}$$

The change in the electronic structure due to $\mathcal{V}^{(2)}$ involves the following expectation value of the phonon operators:

$$\langle (a_{\mathbf{k}} + a^{\dagger}_{-\mathbf{k}})(a_{-\mathbf{k}+\mathbf{q}} + a^{\dagger}_{\mathbf{k}-\mathbf{q}}) \rangle = \langle a_{\mathbf{k}} a^{\dagger}_{\mathbf{k}-\mathbf{q}} + a^{\dagger}_{-\mathbf{k}} a_{-\mathbf{k}+\mathbf{q}} \rangle \propto \delta_{\mathbf{q},\mathbf{0}}.$$

Since $\lambda_{\mathbf{kq}} = 0$ for $q = 0$, it follows that $\mathcal{V}^{(2)}$ makes a vanishing contribution to the renormalized electronic energy levels. This conclusion does not hold when $G \neq 0$ terms are kept in Eq. (A3).

For completeness, we show an alternative derivation for the vanishing of the Debye-Waller term. Following the work of Allen and Cardona in Ref. [16], the Debye-Waller term involves a vertical interband matrix element of the type

$$\begin{aligned} &\left\langle \psi_{\mathbf{k}n} \left| \frac{\partial V_{ei}(\mathbf{r}-\mathbf{R})}{\partial \mathbf{R}} \right| \psi_{\mathbf{k}n'} \right\rangle \\ &\quad = i \sum_{\mathbf{k}'} \mathbf{k}' V_{ei}(\mathbf{k}') e^{-i\mathbf{k}'\cdot\mathbf{R}} \int_{\text{all}} d\mathbf{r}\, \psi^{*}_{\mathbf{k}n}(\mathbf{r}) \psi_{\mathbf{k}n'}(\mathbf{r}) e^{i\mathbf{k}'\cdot\mathbf{r}} \\ &\quad = i \sum_{\mathbf{G}} \mathbf{G}\, V_{ei}(\mathbf{G}) \int_{\text{cell}} d\mathbf{r}\, e^{i\mathbf{G}\cdot\mathbf{r}} u^{*}_{\mathbf{k}n}(\mathbf{r}) u_{\mathbf{k}n'}(\mathbf{r}), \end{aligned}$$

where $\mathbf{R}$ is the ion coordinate and the spatial integrals in first and second equalities are over the entire crystal and over the unit cell, respectively. In addition, we have used the Bloch's theorem: $\psi_{\mathbf{k}n}(\mathbf{r}) = \exp(i\mathbf{k}\cdot\mathbf{r}) u_{\mathbf{k}n}(\mathbf{r})$. If the eigenstates are plane waves, the spatial integration selects $G = 0$ and the matrix element vanishes. Even for non-plane-wave eigenstates, the $G = 0$ term is clearly zero. Neglecting $G \neq 0$ terms (and hence ignoring the Debye-Waller term) is justified if the electron-ion potential changes slowly in space (which is the case in the deformation potential approximation) and/or if the Bloch eigenstates are approximately plane waves (which is an acceptable approximation for some simple metals, though not for most semiconductors).

APPENDIX B: ELECTRON-PHONON COUPLING PARAMETERS FOR HGTE/CDTE QUANTUM WELLS

For reference, in this appendix, we list the various types of $g_{\mathbf{q}}$ that were used in Sec. V A. Since we are interested in long-wavelength phonons and since the quantum well states have no degeneracies beyond the standard spin degeneracy, we consider the coupling to longitudinal modes only. Because crystals with zinc-blend structure contain two atoms per unit cell, we consider one longitudinal acoustic phonon and one longitudinal optical phonon. The optical phonon can couple to electrons through deformation potential and through Fröhlich-type interaction. The numerical parameters are quoted from Ref. [41].

In the case of deformation potential coupling to longitudinal acoustic phonons,

$$g_{\mathbf{q}} = \sqrt{\frac{\hbar C_{\text{ac}}^2 q^2}{2\rho_A V \omega_{\mathbf{q}}}}, \tag{B1}$$

where $\rho_A = 8100\ \text{kg/m}^3$ is the atomic mass density, $C_{\text{ac}} = 5$ eV is the acoustic deformation potential coupling constant, $\omega_{\mathbf{q}} = c_s q$ is the phonon frequency, and $c_s = 2100$ m/s is the sound velocity.

In the case of deformation potential coupling to nonpolar optical phonons,

$$g_{\mathbf{q}} = \sqrt{\frac{\hbar C_{\text{op}}^2 / a^2}{2\rho_A V \omega_0}}, \tag{B2}$$

where $C_{\text{op}} = 20$ eV is the optical deformation potential, $\omega_0 = 17$ meV is the optical phonon frequency, and $a \simeq 0.646$ nm is the lattice constant for HgTe. In the case of the Fröhlich coupling to polar optical phonons,

$$g_{\mathbf{q}} = \sqrt{\frac{e^2 \hbar \omega_0}{V q^2} \left(\frac{1}{\epsilon_\infty} - \frac{1}{\epsilon_0} \right)}, \tag{B3}$$

where $\epsilon_\infty = 14$ and $\epsilon_0 = 20$ are high-frequency and static dielectric constants (in units of vacuum permittivity).

[1] L. Fu, C. L. Kane, and E. J. Mele, Phys. Rev. Lett. **98**, 106803 (2007); R. Roy, Phys. Rev. B **79**, 195322 (2009); J. Moore, Nature (London) **464**, 194 (2010); M. Z. Hasan and C. L. Kane, Rev. Mod. Phys. **82**, 3045 (2010); X.-L. Qi and S.-C. Zhang, *ibid.* **83**, 1057 (2011).

[2] S.-Q. Sheng, *Dirac Equation in Condensed Matter* (Springer, Berlin, 2012); B. A. Bernevig and T. L. Hughes, *Topological Insulators and Topological Superconductors* (Princeton University Press, Princeton, 2013); *Topological Insulators*, edited by M. Franz and L. Molenkamp, Contemporary Concepts of Condensed Matter Science Vol. 6 (Elsevier, Amsterdam, 2013).

[3] J. Maciejko, X.-L. Qi, A. Karch and S.-C. Zhang, Phys. Rev. Lett. **105**, 246809 (2010); M. A. Metlitski, C. L. Kane, and M. P. A. Fisher, Phys. Rev. B **88**, 035131 (2013).

[4] D. Pesin and A. H. MacDonald, Nat. Mater. **11**, 409 (2012); T. Yokoyama and S. Murakami, Physica E **55**, 1 (2014).

[5] M. Bahramy, B.-J. Yang, R. Arita, and N. Nagaosa, Nat. Commun. **3**, 679 (2012).

[6] M. Kim, C. H. Kim, H.-S. Kim, and J. Ihm, Proc. Natl. Acad. Sci. USA **109**, 671 (2012).

[7] M. Z. Hasan, D. Hsieh, Y. Xia, L. A. Wray, S.-Y. Xu, and C. L. Kane, arXiv:1105.0396; S.-Y. Xu, Y. Xia, L. A. Wray, S. Jia, F. Meier, J. H. Dil, J. Osterwalder, B. Slomski, A. Bansil, H. Lin, R. J. Cava, and M. Z. Hasan, Science **332**, 560 (2011).

[8] S. Diehl, E. Rico, M. A. Baranov, and P. Zoller, Nat. Phys. **7**, 971 (2011); N. Goldman, I. Satija, P. Nikolic, A. Bermudez, M. A. Martin-Delgado, M. Lewenstein, and I. B. Spielman, Phys. Rev. Lett. **105**, 255302 (2010); N. H. Lindner, G. Refael, and V. Galitski, Nat. Phys. **7**, 490 (2011).

[9] J. Li, R.-L. Chu, J. K. Jain, and S.-Q. Shen, Phys. Rev. Lett. **102**, 136806 (2009); C. W. Groth, M. Wimmer, A. R. Akhmerov, J. Tworzydlo, and C. W. J. Beenakker, *ibid.* **103**, 196805 (2009); H.-M. Guo, G. Rosenberg, G. Refael, and M. Franz, *ibid.* **105**, 216601 (2010).

[10] I. Garate, Phys. Rev. Lett. **110**, 046402 (2013).

[11] O. Viyuela, A. Rivas, and M. A. Martin-Delgado, Phys. Rev. B **86**, 155140 (2012); A. Rivas, O. Viyuela, and M. A. Martin-Delgado, *ibid.* **88**, 155141 (2013); O. Viyuela, A. Rivas, and M. A. Martin-Delgado, Phys. Rev. Lett. **112**, 130401 (2014); C.-E. Bardyn, M. A. Baranov, E. Rico, A. Imamoglu, P. Zoller, and S. Diehl, *ibid.* **109**, 130402 (2012); Z. Li and J. P. Carbotte, Phys. Rev. B **88**, 195133 (2013).

[12] For a review, see, e.g., M. Cardona, Solid State Comm. **133**, 3 (2005).

[13] G. D. Mahan, *Many-Particle Physics*, 3rd ed. (Kluwer Academic, New York, 2000).

[14] J. Callaway, *Quantum Theory of the Solid State*, 2nd ed. (Academic Press, San Diego, 1991).

[15] The statement that phonons cannot induce vertical interband transitions follows from our approximation of neglecting umklapp processes.

[16] P. B. Allen and V. Heine, J. Phys. C: Solid State Phys. **9**, 2305 (1976); P. B. Allen and M. Cardona, Phys. Rev. B **23**, 1495 (1981); **27**, 4760 (1983); S. Gopalan, P. Lautenschlager, and M. Cardona, *ibid.* **35**, 5577 (1987); S. Zollner, M. Cardona, and S. Gopalan, *ibid.* **45**, 3376 (1992); S. Krishnamurthy, A.-B. Chen, A. Sher, and M. Van Schilfgaarde, J. Electronic Mater. **24**, 1121 (1995); D. Olguín, M. Cardona, and A. Cantarero, Solid State Commun. **122**, 575 (2002); F. Giustino, S. G. Louie, and M. L. Cohen, Phys. Rev. Lett. **105**, 265501 (2010); S. Poncé, G. Antonius, P. Boulanger, E. Cannuccia, A. Marini, M. Côté, and X. Gonze, Comput. Mater. Sci. **83**, 341 (2014).

[17] G. Rosenberg and M. Franz, Phys. Rev. B **82**, 035105 (2010).

[18] Z. Wang, X.-L. Qi, and S.-C. Zhang, Phys. Rev. B **85**, 165126 (2012); Z. Wang and S.-C. Zhang, Phys. Rev. X **2**, 031008 (2012); Z. Wang and B. Yan, J. Phys. Condens. Matter **25**, 155601 (2013).

[19] S.-Y. Xu, M. Neupane, Chang Liu, S. Jia, L. A. Wray, G. Landolt, B. Slomski, J. H. Dil, N. Alidoust, S. Basak, H. Lin, J. Osterwalder, A. Bansil, R. J. Cava, and M. Z. Hasan, arXiv: 1201:3124.

[20] A calculation based on the self-consistent Born approximation confirms this conclusion. At any rate, the underlying assumption is that there are no lattice instabilities as the renormalized electronic gap closes; in two and three dimensions, such assumption is justifiable.

[21] By "higher-momentum transfer," we mean $q \gg \sqrt{|m|/\beta}$. For Dirac insulators near a topological phase transition, this condition is satisfied even when momentum transfer is small compared to a reciprocal lattice vector. Accordingly, our approximate long-wavelength expressions for the electron-phonon interaction [Eq. (8)] remain applicable.

[22] For example, in Pb_{1-x} Sn_x Se, the band inversion leading from a trivial to a topological phase occurs as temperature is decreased. The source of this phenomenon is believed to be the change of the lattice constant with temperature. See, e.g., P. Dziawa, B. J. Kowalski, K. Dybko, R. Buczko, A. Szczerbakow, M. Szot, E. Łusakowska, T. Balasubramanian, B. M. Wojek, M. H. Berntsen, O. Tjernberg, and T. Story, Nat. Mater. **11**, 1023 (2012).

[23] In presence of screening, $g_{\mathbf{q}} \to g_{\mathbf{q}}/\epsilon_{\mathbf{q}}$, where $\epsilon_{\mathbf{q}} \simeq 1 - (V_{\mathbf{q}}/\epsilon_\infty)\Pi_{\mathbf{q}}$, $V_{\mathbf{q}}$ is the bare Coulomb interaction, ϵ_∞ is the dielectric constant originating from high-energy electronic excitations (not included in our four-band model) and $\Pi_{\mathbf{q}}$ is the static density-density response function in the low-energy subspace (see, e.g., Chap. 7 of Ref. [13]). The intraband part of $\Pi_{\mathbf{q}}$ is suppressed when the density of carriers is small. The interband part of $\Pi_{\mathbf{q}}$ is nonzero even in absence of itinerant carriers; however, $\epsilon \simeq 1$ is still a reasonable approximation for all but the smallest q because $\epsilon_\infty \gg 1$ in narrow-gap systems. The enhanced screening at $q = 0$ removes the long-wavelength singularity of the Fröhlich coupling but makes no significant impact on the deformation potential coupling to acoustic and optical phonons.

[24] B. Andrei Bernevig, T. L. Hughes, S.-C. Zhang, Science **314**, 1757 (2006).

[25] M. König, H. Buhmann, L. W. Molenkamp, T. L. Hughes, C.-X. Liu, X. L. Qi, and S. C. Zhang, J. Phys. Soc. Jpn. **77**, 031007 (2008).

[26] A. Roth, C. Brüne, H. Buhmann, L. W. Molenkamp, J. Maciejko, X.-L. Qi, and S.-C. Zhang, Science **325**, 294 (2009).

[27] P. Yu and M. Cardona, *Fundamentals of Semiconductors*, 4th ed. (Springer-Verlag, Berlin, 2010).

[28] P. Sengupta, T. Kubis, Y. Tan, M. Povolotskyi, and G. Klimeck, J. Appl. Phys. **114**, 043702 (2013).

[29] E. O. Kane, J. Phys. Chem. Solids **1**, 249 (1957).

[30] C. R. Becker, V. Latussek, A. Pfeuffer-Jeschke, G. Landwehr, and L. W. Molenkamp, Phys. Rev. B **62**, 10353 (2000).

[31] S. Das Sarma and A. Madhukar, Phys. Rev. B **22**, 2823 (1980); S. Das Sarma and B. A. Mason, Ann. Phys. **163**, 78 (1985).

[32] See Supplementary Material of Ref. [7].

[33] See the paper of Li and Carbotte in Ref. [11].

[34] C. L. Kane and E. J. Mele, Phys. Rev. Lett. **95**, 226801 (2005).

[35] H. Ochoa, A. H. Castro Neto, V. I. Falko, and F. Guinea, Phys. Rev. B **86**, 245411 (2012).

[36] Although generally nonzero, β is almost always ignored in the effective low-energy model for graphene, see, e.g., H. Min, J. E. Hill, N. A. Sinitsyn, B. R. Sahu, L. Kleinman, and A. H. MacDonald, Phys. Rev. B **74**, 165310 (2006). Even if one were to include it, β would vanish whenever m did (i.e., in absence of intrinsic spin-orbit interaction). This is an important difference with respect to HgTe/CdTe, where $\beta \neq 0$ even when $m = 0$.

[37] J. S. Meyer and G. Refael, Phys. Rev. B **87**, 104202 (2013).

[38] See, e.g., The Landolt-Börnstein Database at http://www.springermaterials.com/docs/bookshelf.html. For an exception in which the thermal expansion is believed to play a dominant role, see, e.g., Ref. [22].

[39] For a review see, e.g., T. Ando, in *Graphene and its Fascinating Attributes* (World Scientific, Singapore, 2011).

[40] J. Kim and S.-H. Jhi, arXiv:1308.0434.

[41] E. Lhuillier, S. Keuleyan, and P. Guyot-Sionnest, Nanotechnol. **23**, 175705 (2012); A. Ingale, M. L. Bansal, and A. P. Roy, Phys. Rev. B **40**, 12353 (1989); M. Rösch, R. Atzmüller, G. Schaack, and C. R. Becker, *ibid.* **49**, 13460 (1994).