Anderson impurity in the bulk of topological insulators

Igor Kuzmenko,1 Yshai Avishai,1,2 and Tai Kai Ng2

1Department of Physics, Ben-Gurion University of the Negev, Beer-Sheva, Israel
2Department of Physics, Hong Kong University of Science and Technology, Kowloon, Hong Kong

The significance of topological insulators (TIs) as a new state of matter has been stressed in numerous publications [1–7]. So far, the main attention has been focused on the surface states [8–24]. It was pointed out recently that impurity scattering may have nontrivial effects in the bulk of TIs due to its peculiar band structure [25–28]. More concretely, in topological insulators with “inverted-Mexican-hat” band dispersion, a Kondo resonant peak appears simultaneously with an in-gap bound state. The latter generates another spin state thereby screening the Kondo effect. Using a weak-coupling RG scheme, it is shown that the system exhibits complex crossover behavior between different symmetry configurations and may evolve into a self-screened Kondo or an SO(n) (n = 3, 4) low-energy fixed point. Experimental consequences of these scenarios are pointed out.

DOI: 10.1103/PhysRevB.89.035125

PACS number(s): 71.10.Pm, 72.15.Qm, 73.23.—b, 73.43.—f

I. INTRODUCTION

The significance of topological insulators (TIs) as a new state of matter has been stressed in numerous publications [1–7]. So far, the main attention has been focused on the surface states [8–24]. It was pointed out recently that impurity scattering may have nontrivial effects in the bulk of TIs due to its peculiar band structure [25–28]. More concretely, in topological insulators with “inverted-Mexican-hat” band dispersion, a Kondo resonant peak appears simultaneously with an in-gap bound state. The latter generates another spin state thereby screening the Kondo effect. Using a weak-coupling RG scheme, it is shown that the system exhibits complex crossover behavior between different symmetry configurations and may evolve into a self-screened Kondo or an SO(n) (n = 3, 4) low-energy fixed point. Experimental consequences of these scenarios are pointed out.

II. MODEL HAMILTONIAN

Our aim in this section is to derive an effective Hamiltonian that, in addition to the Anderson impurity $d$, contains also the midgap state $f$ as discussed in Ref. [26]. Starting from a Hamiltonian describing an Anderson impurity $d$ in the bulk of a TI, a few manipulations are required to transform it into its workable form, Eq. (15a) below. The bare Hamiltonian is

$$H = H_0 + H_d + H_f^{(0)} + V.$$

Here the first term, $H_0$, describes electrons in the bulk of the TI,

$$H_0 = \sum_k \Psi_k^\dagger h_0(k)\Psi_k,$$

$$h_0(k) = h_{xy}(k \cdot \alpha) + \beta M_k,$$

where $M_k = m v_F^2 - B k^2$, and $\alpha$, $\beta$ are creation operators for electron of momentum $k$ and spin projection $\sigma$. $h_0(k)$ is written in particle-hole $\alpha$ space with $\alpha = t_x \otimes s$ and $\beta = t_z \otimes s_0$, where $t = (t_x, t_y, t_z)$ or $s = (s_x, s_y, s_z)$ are vectors

states of the composite $(d + f)$ impurity. At the end of Sec. II we arrive at an effective Anderson Hamiltonian that will be analyzed further in the next sections. The local density of states (DOS) is calculated in Sec. III for the TI with the potential scattering term. Renormalization group analysis is carried out in Sec. IV. Depending on various energy domains, the RG analysis ends up with various Kondo Hamiltonians that possesses different (dynamical) symmetries, SU(2), SO(3), or SO(4). Section V is devoted to the calculations of electric resistivity and magnetic susceptibility in the relevant energy domains. One of the central results of our work is the occurrence of temperature-driven crossovers of the KE between different symmetry classes. The results are then summarized in Sec. VI. Some details of the calculation technique are presented in the appendices. These include physical origin of the potential scattering induced by an Anderson impurity (Appendix A), discussion of local density of states (Appendix B), the SO(4) Kondo Hamiltonian (Appendix C), and finally, calculations of electric resistivity (Appendix D) and magnetic susceptibility (Appendix E) for different temperature intervals.

$$H = H_0 + H_d + H_f^{(0)} + V.$$
of the Pauli matrices acting in the space of isospins or spins; 
$s_0$ is the $2 \times 2$ identity matrix. $H_d$ is the Hamiltonian for the
Anderson impurity,
$$H_d = \epsilon_d \sum_{\sigma} n_{d\sigma} + U_d n_{d\uparrow} n_{d\downarrow},$$
where $\epsilon_d$ is the impurity energy level and $U_d$ is the interaction
between electrons on the impurity. $n_{d\sigma} = d_{\sigma}^\dagger d_{\sigma}$; $d_{\uparrow}$ or $d_{\downarrow}$ is the creation or annihilation operator of the electron on the
d level. The hybridization between the Anderson impurity $d$ and the band electrons is described by
$$H_d^{(0)} = V_d \sum_{k,\sigma} (a_{k\sigma}^\dagger - i b_{k\sigma}^\dagger)(a_{k\sigma} + i b_{k\sigma}).$$

The last term appearing in the Hamiltonian (1) is an
effective potential scattering between band electrons induced by
the impurity,
$$V = V_0 \sum_{k,k',\sigma} (a_{k\sigma}^\dagger - i b_{k\sigma}^\dagger)(a_{k'\sigma} + i b_{k'\sigma}).$$

A few words on the origin of the potential scattering term $V$ are
in order. When an Anderson impurity is immersed in a metal
it leads to an $s$-$d$ Hamiltonian and to a potential scattering. In
the standard analysis of the Anderson impurity, the induced
potential scattering term is neglected because it is irrelevant as
far as Kondo physics is concerned. The situation is different
when the impurity is immersed in an insulator. Here, as we
shall see in the following, the potential scattering leads to
a midgap bound state that profoundly affects the underlying
physics. The detailed discussions about the physical origin of
the strength $V_0$ and the derivation of the localized in-gap $f$
level are given in Appendix A.

**Diagonalization of $H_0$.** The eigenstate of $H_0$ alone is given by
$$H_0 = \sum_{\nu k} \nu \gamma_\nu^\dagger \gamma_\nu,$$
where $\nu = \pm 1$ or $c/\nu$ denotes the conduction and valence band.
$$\epsilon_\nu = \sqrt{M_\nu^2 + (\hbar v_F k)^2}, \quad M_\nu = m_\nu^2 - B k^2,$$
is the band dispersion. $\gamma_\nu^\dagger$ and $\gamma_\nu$ are creation and annihilation operators of quasiparticle defined through the transformation
$$\tilde{\Psi}_k = S_k U_k \Psi_k, \quad \tilde{\Psi}_k^\dagger = (\gamma_{\nu k\uparrow}, \gamma_{\nu k\downarrow}, \gamma_{\nu k\uparrow}, \gamma_{\nu k\downarrow}),$$
$$U_k = \cos \alpha_k / 2 + i \sin \alpha_k / 2 \gamma \otimes (s \cdot e_k),$$
where $e_k = k/k$,
$$\tan \left( \frac{\alpha_k}{2} \right) = \sqrt{\frac{\epsilon_k - M_k}{\epsilon_k + M_k}}.$$

$S_k$ is a unitary matrix which commutes with $h_0(k)$ and $U_k$. $\epsilon_k$
is gapped and the insulator is topological for $Bm > 0$. For $Bm > 1/2$ (assumed hereafter), the band dispersion has an “inverted-Mexican-hat” form (see Fig. 1) with dispersion minimum at a

![FIG. 1. (Color online) Energy dispersion (6b) for $Bm = 20$. The dark (bright) regions denote the energy levels below (above) the Fermi energy $\epsilon_F$. The band-edges occur at momentum $q$ at energies $\pm \epsilon_q$. (Here it is assumed that $\epsilon_q < \epsilon_F < m v_F^2$.) $k_1$ and $k_2$ are two solutions of the equation $\epsilon_\nu = \epsilon_F$.](image-url)

<table>
<thead>
<tr>
<th>$q/\ell_F$</th>
<th>0.0</th>
<th>0.5</th>
<th>1.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\epsilon_q$</td>
<td>-1.5</td>
<td>-1.0</td>
<td>-0.5</td>
</tr>
<tr>
<td>$\epsilon_F$</td>
<td>0.5</td>
<td>1.0</td>
<td>1.5</td>
</tr>
<tr>
<td>$n$</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>$m$</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
</tbody>
</table>

$\epsilon q$ is a unitary matrix,
$$V_k = \cos \alpha_k - i \sin \alpha_k (s \cdot e_k), \quad \gamma^\dagger \gamma = 1.$$ Choosing $S_k$ as
$$S_k = \begin{pmatrix} \gamma^\dagger_k & 0 \\ 0 & i \gamma^\dagger_k \end{pmatrix},$$
we get
$$H_i = V_d \sum_{k,\nu,\sigma} (\gamma_{\nu k\sigma} d_{\sigma} + d_{\sigma}^\dagger \gamma_{\nu k\sigma}).$$

Applying transformations (6c) and (8) to the potential scattering Hamiltonian (5), we get
$$V = V_0 \sum_{\nu k,\nu' k' \sigma} \gamma_{\nu' k' \sigma} \gamma_{\nu k \sigma}.$$ **Diagonalization of $H_0 + V$: Midgap state.** The next step is
crucial, as it demonstrates the formation of a midgap state $f$
due to the potential scattering term $V$. Because the potential
emerges due to the Anderson impurity, we say that the midgap
state is induced by the impurity. In order to find the eigenstates
of $H_0 + V$, we solve the Heisenberg equation of motion
$$i \hbar \dot{\gamma_{\nu k\sigma}} = [H_0 + V, \gamma_{\nu k\sigma}].$$ Taking $\gamma_{\nu k\sigma}(t) = \gamma_{\nu k\sigma} e^{-i \epsilon t / \hbar}$, we get
$$\epsilon \gamma_{\nu k\sigma} = v \epsilon \gamma_{\nu k\sigma} + V_0 \sum_{\nu' k'} \gamma_{\nu' k' \sigma}.$$
Equation (10) has a nontrivial solution when $\epsilon$ satisfies the secular equation,
\[ \sum_{\sigma} \frac{V_0}{\epsilon - \nu_{\epsilon k}} = 1. \tag{11} \]
The solutions for $|\epsilon| \geq \epsilon_q$ describe the band electrons, whereas the solution for $|\epsilon| < \epsilon_q$ corresponds to the localized $f$ level.

It is seen that when $|\epsilon| < \epsilon_q$, the expression on the left-hand side of the secular equation (11) is positive when $\epsilon$ is negative and negative when $\epsilon$ is positive ($V_0$ is assumed to be positive here). When $|\epsilon| \rightarrow \epsilon_q$, the sum diverges as $(\epsilon_q^2 - \epsilon^2)^{-1/2} \text{sign}(\epsilon)$ for energy bands with the “inverted Mexican hat” structure [26]. As a result, the secular equation gives a midgap energy level $\epsilon_f$, which lies within the interval $-\epsilon_q < \epsilon_f < 0$. When $\epsilon_f + \frac{\epsilon_q}{2} \ll \epsilon_q$, we can write
\[ \epsilon_f \approx -\epsilon_q \sqrt{1 - \left( \frac{V_0 q}{2\pi \hbar^2 B} \right)^2}. \tag{12} \]

This procedure leads to a minor modification of the annihilation and creation operators for the band electrons. Strictly speaking they are respectively expressed as a linear combination of $\gamma_{\epsilon_{\alpha k}}$ and $\gamma_{\epsilon_{\alpha k}}^\dagger$ using perturbation theory with $V_0$ as a small parameter. However, since the potential pulls only one level (out of many) from the band into the gap, the other levels are virtually unaffected. In what follows, we assume that the operators for the modified levels inside the band just slightly differ from the $\gamma_{\epsilon_{\alpha k}}$ and $\gamma_{\epsilon_{\alpha k}}^\dagger$. Thus, the main outcome of the potential scattering is the creation of a midgap level $f$.

The annihilation operator $f_\alpha$ for this localized level is given by
\[ f_\alpha = \sum_{\nu k} \frac{A V_0}{\epsilon_f - \nu_{\epsilon k}} \gamma_{\nu_{\epsilon k} \sigma}, \quad A = \left( \sum_{\nu k} \frac{V_0^2}{(\epsilon_f - \nu_{\epsilon k})^2} \right)^{-1}. \tag{13} \]

**Hybridization term $H_{df}$.** The last ingredient in our quest for constructing an effective tunneling Hamiltonian with $d$ and $f$ impurities is to identify a hopping term $H_{df}$ between the Anderson impurity and the midgap state. The hybridization term $H_{t}^{(0)}$ between the bulk TI electrons and the $d$-impurity level leads to an effective tunneling term $H_{df}$ between the $d$ impurity and the $f$ in-gap bound state, and an effective tunneling term $H_{f}$ between the $d$ impurity and the band states. We assume that $H_{f}$ is still given by Eq. (9), whereas $H_{df}$ is
\[ H_{df} = V_{df} \sum_{\sigma} \left( f_{\sigma}^\dagger d_{\sigma} + d_{\sigma}^\dagger f_{\sigma} \right), \tag{14} \]
where $V_{df} \sim V_f$.

**The effective Hamiltonian.** Finally, to arrive at the desired effective tunneling Hamiltonian we collect all pieces into a sum of three parts structured as “band + composite impurity + hybridization” Hamiltonians,
\[ H = H_t + H_C + H_{df}. \tag{15a} \]

Here $H_t = H_0 + V$ is the Hamiltonian of the two band electrons that includes also the potential scattering term, and
\[ H_C = H_d + H_f + H_{df} \tag{15b} \]
is the Hamiltonian of the composite impurity, including the $d$ and $f$ levels. $H_d$ is given by Eq. (3), $H_{df}$ is given by Eq. (14), and
\[ H_f = \epsilon_f \sum_{\sigma} n_{f \sigma} + U_f n_{f 1} n_{f 1}, \tag{15c} \]
where $\epsilon_f$ is the single-electron energy (12) of the localized $f$ level, $U_f$ is the interaction between electrons on the $f$ level, and $n_{f \sigma} = f_{\sigma}^\dagger f_{\sigma}$. The hybridization term $H_{f}$ is given by Eq. (9).

**Energy scales.** A few words about energy scales are in order: Unless otherwise specified, we shall assume $U_d \rightarrow \infty$ and
\[ \epsilon_f - D_0 < \epsilon_d < \epsilon_f < \epsilon_f + U_f \ll \epsilon_f + D_0, \tag{16} \]
where $D_0$ (the initial bandwidth) is the highest energy cutoff, and $\epsilon_f$ is the Fermi energy (see Fig. 1). We use $\epsilon_f = 2 \epsilon_q, \epsilon_f \approx -\epsilon_q, U_f = 5 \epsilon_q$, and $\epsilon_d = -80 \epsilon_q$ in the following calculations.

**The eigenstates of $H_C$.** Equation (15b) are specified by the configuration numbers $(N_d, N_f)$ representing the number of electrons on the levels $d$ and $f$. With energy scales specified by (16) the ground state has $N_d = N_f = 1$ and there are four possible states, a spin-singlet state $|S\rangle$ and three spin-triplet states $|T_{\alpha}\rangle$ ($\alpha = 0, \pm 1$). The singlet energy is modified when $V_{df} \neq 0$ while the triplet energy is unaffected. Explicitly,
\[ |S\rangle = \left\{ \frac{\alpha_S}{\sqrt{2}} (d_{\uparrow}^\dagger f_{\uparrow}^\dagger - d_{\downarrow}^\dagger f_{\downarrow}^\dagger) - \beta_S f_{\uparrow}^\dagger f_{\downarrow}^\dagger \right\} |0\rangle, \tag{17a} \]
\[ |T_1\rangle = d_{\uparrow}^\dagger f_{\uparrow}^\dagger |0\rangle, \quad |T_{-1}\rangle = d_{\downarrow}^\dagger f_{\downarrow}^\dagger |0\rangle, \tag{17b} \]
where
\[ \epsilon_S = \epsilon_d + \epsilon_f - \frac{2V_{df}^2}{\Delta_f}, \quad \epsilon_T = \epsilon_d + \epsilon_f, \]
\[ \alpha_S = \sqrt{1 - \beta_f^2}, \quad \beta_f = \frac{\sqrt{2} V_{df}}{\Delta_f}, \quad \Delta_f = \epsilon_f - \epsilon_d + U_f. \]

In the absence of hybridization of the $d$ electron with the band electrons (i.e., when $V_{df} = 0$), $\epsilon_S < \epsilon_T$, and the spin singlet state has lower energy.

**III. LOCAL DENSITY OF STATES**

Because of the potential scattering, the local density of states (DOS) depends on the energy $\epsilon$ of an electron and on the distance $r$ from the impurity,
\[ \rho(\epsilon, r) = -\frac{1}{\pi} \text{Im} \sum_{\nu k k' \sigma} G_{\nu k, \nu k'(\epsilon)} e^{i(k - k')r}, \tag{18} \]
\[ G_{\nu k, \nu k'}(\epsilon) = \int G_{\nu k, \nu k'}(t) e^{i\omega_{\nu k}t} dt. \tag{19} \]
Here \( \eta \) is a positive infinitesimal parameter, \( \tilde{G}_{\mathbf{k},\mathbf{k}'}(t) \) is a retarded Green’s function,
\[
\tilde{G}_{\mathbf{k},\mathbf{k}'}(t) = -\frac{i}{\hbar} \Theta(t) \langle \gamma_{\mathbf{k}\sigma} \gamma_{\mathbf{k}'\sigma} \rangle,
\]
where \( \langle \cdots \rangle \) denotes the thermal average with respect to the Hamiltonian \( H_0 + V \).

Applying the Heisenberg equation of motion (10), we get the following expression for the local DOS:
\[
\rho(\epsilon, r) = \rho_0(\epsilon) \frac{1 + \pi^2 V_0^2 [\rho_0^2(\epsilon) - \rho_0^2(\epsilon, r)]}{1 + \pi^2 \rho_0^2(\epsilon)},
\]
(21)
where
\[
\rho_0(\epsilon, r) = \Theta(|\epsilon| - \epsilon_q) \frac{\rho_0(|\epsilon|)}{2 \sqrt{\epsilon^2 - \epsilon_q^2}} \left\{ \Theta(\epsilon_0 - |\epsilon|) \frac{\sin(qr g_1(\epsilon))}{qr} + \Theta(\epsilon_0 + |\epsilon|) \frac{\sin(qr g_2(\epsilon))}{qr} \right\},
\]
\[
\rho_0(\epsilon) = \rho_0(\epsilon, 0).
\]
(22)

Detailed derivation of the local DOS is given in Appendix B.

The local DOS (21) is shown in Fig. 2 for different values of \( qr \). It is seen that \( \rho(\epsilon, 0) \) vanishes as \( (\epsilon^2 - \epsilon_q^2)^2 \) when \( |\epsilon| \to \epsilon_q \). When \( qr > 2 \), \( \rho(\epsilon, r) \) approaches \( \rho_0(\epsilon) \), the bare DOS.

**IV. RENORMALIZATION-GROUP (RG) ANALYSIS**

Within RG analysis, high-energy charge fluctuations are integrated out and one reaches a low-energy spin \((s-d)\) Hamiltonian. The latter is written in terms of the spin operators of the band electrons and a collection of vector operators of the composite impurity. These vector operators generate a dynamical symmetry group that characterize the pertinent Kondo physics [34]. If the ground state of \( H_C \) contains a single electron the symmetry group is \( SU(2) \), while if the ground state of \( H_C \) contains two electrons the symmetry group can be \( SO(3) \) [see Eq. (37)] or \( SO(4) \) [see Eq. (32)]. Technically, the high-energy degrees of freedom are successively integrated out such that \( D_0 \to D < D_0 \) is reduced and energies are renormalized. For the Kondo effect in metals, where the DOS is virtually constant, this is a standard procedure. The fact that \( \rho(\epsilon) \) is a constant requires some modifications. We introduce

\[
\bar{\rho}(D) = \frac{1}{2} [\rho(\epsilon_F + D) + \rho(\epsilon_F - D)],
\]
(23)
\[
\mathcal{L}(D_1, D_2) = \int_{D_1}^{D_2} \frac{2 \bar{\rho}(D) dD}{\rho_i D}.
\]
(24)

The RG flow is divided into the following regimes (see Fig. 3):

(i) \( \epsilon_d > \epsilon_F - D > \epsilon_F - D_0 \), in which charge fluctuations exist in both \( d \) and \( f \) states and there is no KE;

(ii) \( \epsilon_F - D > \epsilon_d \) but \( \epsilon_f (\sim -\epsilon_q) > \epsilon_F - D \) or \( \epsilon_F + D > 2 \epsilon_f + U_f \), where charge fluctuations on the \( d \) orbital are quenched but still exist on the \( f \) orbital. The system is in the single impurity \( SU(2) \) Kondo impurity regime;

(iii) \( \epsilon_F - D > \epsilon_f \) and \( \epsilon_F + D < 2 \epsilon_f + U_f \), but \( D > |J_{df}| \) where charge fluctuations in both \( d \) and \( f \) orbitals are quenched, but the singlet-triplet energy splitting can be neglected. The singlet and triplet states can be considered as degenerate and the system demonstrates the \( SO(4) \) Kondo regime.

(iv, v) The system is at the \( SO(3) \) Kondo regime \([if \ J_{df} < 0, \ interval (iv)]\) or the self-screened Kondo regime \([if \ J_{df} > 0, \ interval (v)]\). These regimes exist only if \( U_f > \epsilon_F - 2 \epsilon_f \sim a \ few \ \epsilon_q \).

The RG analysis for the various regimes now follows.

**Regime (i).** Charge fluctuations on the \( d \) orbital are integrated out as in Ref. [29], but here the spin-singlet and spin-triplet energies are, generically, renormalized distinctly. Renormalization of other quantities such as \( V, \epsilon_f, \) and \( V_{df} \) are weak and can be ignored. The scaling procedure of \( \epsilon_{S(T)} \) then yields [31,32]

\[
\frac{d \epsilon_{S(T)}}{d \ln D} = V_{S(T)}^2 \rho(\epsilon_F + D),
\]
(25)

FIG. 2. (Color online) Local DOS (21) for \( \epsilon_0 = 24 \epsilon_q, \ V_0 \rho_0 = 0.06, \) and different values of \( qr \).

FIG. 3. (Color online) Parametric diagram \( j - D \) for \( \epsilon_d = -80 \epsilon_q, \ \epsilon_q = 24 \epsilon_q, \) and \( \epsilon_F = 2 \epsilon_q \). The curves separate the different temperature intervals: the red curve is \( D_i = \epsilon_y - \epsilon_d \) separating the regimes (i) \[mixed valence regime\] and (ii) \[SU(2) Kondo regime\], the purple line is \( D_{ii} \) separating the regimes (ii) \[SU(2) Kondo regime\] and (iii) \[SO(4) Kondo regime\], the blue curve is \( D_{iii} = |J_{df}| \) separating the regimes (iii) \[SO(4) Kondo regime\] and (iv) \[SO(3) Kondo regime\], whereas the green curve is \( D_{iii} = |J_{df}| \) separating the regimes (iv) \[SO(4) Kondo regime\] and (v) \[self-screened Kondo regime\].
where $V_S^2 = a_S^2 V_2^2$ and $V_T^2 = V_2^2$. The difference between $V_S^2$ and $V_T^2$ originates from the appearance of the normalization factor $a_S$ in the singlet state [see (17a)]. Notice that $V_S^2 < V_T^2$ and the triplet energy level renormalizes faster than the singlet level. This opens the possibility that as RG stops, the ground state of the system may become a triplet if the single-triplet level crossing occurs before quenching of charge fluctuations in the $d$ and $f$ levels. Solving Eq. (25), we find that the singlet-triplet energy spacing $J_{df}^{(i)}(D) = \varepsilon_T(D) - \varepsilon_S(D)$ is given by

$$J_{df}^{(i)}(D) \sim \beta_f^2 \Delta_f \left( 1 - \frac{2V_d^2 J_0}{\Delta_f} \frac{\sqrt{T_0} - \sqrt{D}}{\sqrt{T_0}} \right),$$

where $D = D_i \sim \varepsilon_F - \varepsilon_d$ at the end of scaling.

Regime (ii). Charge fluctuations of the $d$ level are quenched at $D_i \sim \varepsilon_F - \varepsilon_d$. A Schrieffer-Wolf transformation for the $d$ level yields an effective Hamiltonian

$$H_T^{(ii)} = H_0 + H_f + H_{df} + H_{K}^{(ii)},$$

where

$$H_{K}^{(ii)} = \frac{1}{2} \sum_{\alpha, \beta = \uparrow, \downarrow} J_{\alpha\beta} (\sigma_\alpha \cdot \sigma_\beta) \cdot \sigma_d,$$

and scaling stops if

$$T^{(ii)}_K > \delta_D = \min(\epsilon_F - \epsilon_f, U_f - \epsilon_F + \epsilon_f),$$

at which the $d$ spin is quenched by the KE. This happens if $j_f < 0$. Assuming that $J_{\alpha\beta}(D_i) \leq \varepsilon_d \ll |\varepsilon_d|$, and $U_f$ is a few $\varepsilon_d$, we find that $J_f < 0$ is of the same order as $\beta_f^2 \Delta_f$ and $J_{df} < 0$ if

$$|V_d|^2 \geq \frac{3}{2} \beta_f^2 \Delta_f.$$

Regimes (iii) and (iv). The scaling stops at regime (ii) with $D_{ii} \sim \varepsilon_F + \varepsilon_q$ if the chemical potential is located inside the band gap which is the usual case for insulators. A more interesting scenario occurs if the chemical potential $\varepsilon_F$ is located above the band gap which may happen if the insulator is doped by impurities. Scaling continues in this case where charge fluctuations in the $f$ level are also quenched. In this case the mixed spin term $S_m$ becomes ineffective and we are left with an effective Hamiltonian $H^{(iv)} = H_0 + H_{K}^{(iv)}$, where

$$H_{K}^{(iv)} = J_K \sigma_f \cdot \sigma_d + J_d \sigma_f \cdot \sigma_d,$$

and the Kondo temperature $T_K^{(iv)}$ is got by solving the equation

$$j_{\gamma\gamma}(D_i) \mathcal{L}(T_K^{(iv)}, D_i) = 1,$$

where

$$\mathcal{L} = \sum_{\mathbf{k}, \mathbf{k}', \sigma} \left( \gamma_{\mathbf{k}, \mathbf{k}' \sigma} \sigma_{\mathbf{k}'} \gamma_{\mathbf{k} \sigma} \right).$$
and

\[ J_K = J_{yy}(D_i) = \frac{J_{yy}(D_i)}{1 - J_{yy}(D_i)\mathcal{L}(D_i, D_i)}. \]

Notice that \( |J_{df}| \) is in general smaller than \( J_K \) [26].

**SO(3) and SO(4) fixed points.** For \( D_i \gg |J_{df}| \), there exists a regime \( D \gg |J_{df}| \) where the system is governed by a critical point between the SO(3) and the quenched Kondo regimes which has SO(4) symmetry [33]. In this regime the behavior of the system is governed by the SO(4) critical point. The system crosses over to low-temperature SO(3) or self-screened Kondo regime at \( D < |J_{df}| \). We first consider the SO(4) regime \( D \gg |J_{df}| \).

**SO(4) Kondo fixed point.** In this case we may set \( J_{df} = 0 \) and apply the Schrieffer-Wolf transformation directly to the spin-singlet and spin-triplet states to get

\[ H_K^{SO(4)} = J_T(S \cdot S_T) + J_{ST}(R \cdot S_T), \]

where \( S \) and \( R \) are the \( S = 1 \) spin and the Runge-Lenz operators, respectively, that are expressible in terms of Hubbard operators for the composed impurity, and satisfy the SO(4) algebra [34]. The exchange constants \( J_T = J_K/2 \) and \( J_{ST} = \alpha_S J_T \) scale as

\[ \frac{d J_T}{d \ln D} = -\left( \frac{j_T^2 + j_{ST}^2}{\rho_0} \right), \quad (33a) \]

\[ \frac{d j_{ST}}{d \ln D} = -2j_T j_{ST} \frac{\tilde{\rho}}{\rho_0}, \quad (33b) \]

The combinations \( j_n = J_T - (-1)^n j_{ST} \) satisfy

\[ j_n(D) = \frac{j_n(D_i)}{1 - j_n(D_i)\mathcal{L}(D, D_i)}, \quad (34) \]

whence

\[ j_T(D) = \frac{1}{2}j_1(D) + j_2(D), \quad (35a) \]

\[ j_{ST}(D) = \frac{1}{2}j_1(D) - j_2(D). \quad (35b) \]

The corresponding Kondo temperature \( T_{KI} \) is determined from the equation

\[ (J_T(\bar{D}_i) + J_{ST}(\bar{D}_i))\mathcal{L}(\bar{D}_i, \bar{D}_i) = 1 \]

provided \( T_{KI} \gg |J_{df}| \). For \( 0 < T_{KI} < |J_{df}| \) the two spins form a spin singlet (self-screened KE) at \( D \sim |J_{df}| \).

**SO(3) Kondo fixed point.** For \( |J_{df}| < 0 \) and \( T_{KI} < |J_{df}| \), renormalization of \( j_{ST} \) stops at \( D = D_{ii} \sim |J_{df}| \). For \( D < D_{ii} \), the Kondo Hamiltonian becomes

\[ H_K^{SO(3)} = J_T S \cdot S_T. \]

The scaling equation for \( J_T = J_T\rho_0 \) and its solution are

\[ \frac{d J_T}{d \ln D} = -\frac{j_T^2}{\rho_0}, \quad (38a) \]

\[ J_T(D) = \frac{J_T(D_{ii})}{1 - J_T(D_{ii})\mathcal{L}(D, D_{ii})}. \quad (38b) \]

Scaling stops at \( T_{KI} \) determined from the equation

\[ J_T(D_{ii})\mathcal{L}(T_{KI}, D_{ii}) = 1. \quad (39) \]

**V. RESISTIVITY AND IMPURITY MAGNETIC SUSCEPTIBILITY**

Having elaborated upon the theory in the weak-coupling regime \( T \gg T_K \) we are now in a position to carry out perturbation calculations of experimental observables. In 3D, the most accessible ones are the impurity resistivity \( R_{\text{imp}}(T) \) and the impurity magnetic susceptibility \( \chi_{\text{imp}}(T) \). We shall be guided by the quest to find out how the special features of the TIs are reflected in these observables. These features are the occurrence of gap and the structure of the DOS especially near the band edges \( \pm \epsilon_g \). In addition, reducing the temperature results in the crossover between different scaling regimes of the couplings. Explicitly, there are three relevant temperature regimes denoted as (ii), (iii), (iv) in order to match the notation of the corresponding scaling regimes discussed previously.

The first regime, denoted as (ii), is defined by \( [D_i > T > D_{ii}] \) as given by Eq. (28) for the scaling interval (ii). Local moment behavior exists only at the level in this regime and therefore there is Kondo scattering with SU(2) symmetry. The second regime, denoted as (iii), is defined by Eq. (33) for scaling interval (iii) \([D_i > T > D_{iii}]\). Here one may neglect the difference in energies between the singlet and triplet states and the system is at the SO(4) Kondo regime. The third regime, denoted as (iv), is defined by Eq. (38a) for scaling interval (iv) \([D_{iii} > T > T_K]\). Here there is Kondo scattering with SO(3) symmetry (when \( J_{df} < 0 \)) or a self-screened KE if \( J_{df} > 0 \). The temperature dependence of the resistivity and magnetic susceptibility in these three different scaling regimes is distinct.

In the calculation of resistivity, we assume \( \epsilon_F > \epsilon_g \) (the TI is doped) and the system has a Fermi surface. The impurity resistivity as calculated in the framework of the “poor-man’s scaling” formalism is given by

\[ R_{\text{imp}} = \frac{N_v R_0}{\mathcal{L}(T_{KI}, T)}. \quad (40) \]

where \( v = (ii),(iii),(iv) \) denotes the pertinent temperature regime as detailed above. The corresponding Kondo temperatures are \( T_{KI}, \) Eq. (30), \( T_{Kii} \equiv T_{KS}, \) Eq. (36), or \( T_{Kii} \equiv T_{KS}, \) Eq. (39). The numerical factors \( N_v \) are \( N_{ii} = N_{iii} = 3/4 \) and

FIG. 5. (Color online) Resistivity as a function of temperature for \( \epsilon_d = -80\epsilon_g, \epsilon_R = 24\epsilon_g, \epsilon_F = 2\epsilon_g \), and different values of \( j: j = 0.1 \) (bottom red curve), \( j = 0.12 \) (middle purple curve), and \( j = 0.14 \) (top blue curve). The dots denote \( T = D_i \) and \( T = D_{iii} \) separating the temperature intervals (ii), (iii), and (iv). Inset: Behavior of \( R_{\text{imp}} \) for the temperature \( T \sim \epsilon_g \).
the framework of the “poor-man’s scaling” is the temperature intervals (ii), (iii), and (iv). The impurity susceptibility calculated in Eq. (30), \(\chi_{\text{imp}}\) [interval (iii)]. These crossovers appear since the Kondo scattering manifests itself also in the magnetic susceptibility (41) as a function of temperature for \(\epsilon_q = -80\epsilon_d\), \(\epsilon_0 = 24\epsilon_d\), \(\epsilon_F = 2\epsilon_d\), and different values of \(j\): \(j = 0.1\) (red curve), \(j = 0.12\) (purple curve), and \(j = 0.14\) (blue curve). The dots denote \(T = D_{ij}\) and \(T = D_{iii}\) separating the temperature intervals (ii), (iii), and (iv).

\[ N_{\text{iv}} = 2 \text{.} \]

\[
R_0 = \frac{3\pi c_{\text{imp}}}{\hbar^2 \rho_0} \frac{1}{v_1^2 + v_2^2}, \quad \nu_i \approx \frac{1}{\hbar} \left( \frac{\partial \epsilon_{\text{imp}}}{\partial k_i} \right),
\]

and \(k_1, k_2\) are two solutions of \(\epsilon_k = \epsilon_F\) (see Fig. 1).

The resistivity as a function of the temperature is shown in Fig. 5 assuming a low-temperature SO(3) fixed point. It is seen that \(R_{\text{imp}}\) has different temperature dependence within the temperature intervals (ii), (iii), and (iv), with crossover observed at \(T = D_{ij}\) and \(T = D_{iii}\) (the points \(D_{ij}\) and \(D_{iii}\) are denoted by dots). In addition, crossovers are observed at \(T = \epsilon_q\) [interval (iii)]. These crossovers appear since the function \(\mathcal{L}(T_{K_{ii}}, T)\) changes its behavior at \(T = \epsilon_F - \epsilon_q = \epsilon_q\) (we take \(\epsilon_F = 2\epsilon_q\) here).

The Kondo scattering manifests itself also in the magnetic susceptibility \([31]\). The impurity susceptibility calculated in the framework of the “poor-man’s scaling” is

\[ \chi_{\text{imp}} = \frac{K_{ii} \epsilon_0 T_{K_{ii}}}{T} \left\{ P_{\nu} - \frac{1}{\mathcal{L}(T_{K_{ii}}, T)} \right\}, \tag{41} \]

where \(\nu = (\text{ii}), (\text{iii}), (\text{iv})\); the Kondo temperatures are \(T_{K_{ii}}\), Eq. (30), \(T_{K_{ii}} \equiv T_{K_{ii}}\), Eq. (36), and \(T_{K_{ii}} \equiv T_{K_{ii}}\), Eq. (39). The numerical factors are \(K_{ii} = K_{iii} = 1/4, K_{ij} = 2/3, P_{\nu} = P_{\nu + 1} = 1, \) and \(P_{iii} = 2\). The constant \(\chi_0\) is

\[ \chi_0 = \frac{4c_{\text{imp}}\hbar^2}{T_{K_{ii}}} \text{.} \]

The impurity magnetic susceptibility as a function of \(T\) is shown in Fig. 6. The different temperature dependencies of \(\chi_{\text{imp}}\) at different temperature regimes (ii), (iii), and (iv) are obvious, with crossovers observed at \(T = D_{ij}\) and \(T = D_{iii}\) (the points \(D_{ij}\) and \(D_{iii}\) are denoted by dots).

VI. CONCLUSIONS

We have analyzed the interplay between the Anderson impurity and its induced in-gap bound state in a model of 2D topological insulator. Using a weak-coupling RG analysis, it is shown that the exchange interaction \(J_{df}\) between the \(d\) spin and the induced in-gap \(f\) spin may be renormalized dynamically to either positive or negative values. The parameters required to observe the above phenomena are not too restrictive (\(|V_d| \gg \epsilon_0 \sqrt{\epsilon_d}, \epsilon_q < \epsilon_F \ll |U_d|\) and are realistic. The system exhibits complex crossover behaviors at different parameter regimes as a result which can be observed in the temperature dependence of the impurity-induced resistance and magnetic susceptibility. The crossover in the temperature dependence of both the resistivity and the impurity magnetic susceptibility at different regimes is a peculiar feature that can serve as an experimental confirmation of the above analysis. For both screened and underscreened Kondo effect in the weak-coupling regime, the effective coupling constant \(J\) renormalizes as \(1/\mathcal{L}(T_{K_{ii}}, T)\) (or, as \(1/\ln(T_{K_{ii}})\), when the DOS is flat; see Ref. \([31]\)). As a result, the impurity resistivity, \(R_{\text{imp}}\), behaves as \(1/\mathcal{L}(T_{K_{ii}}, T)\) [see Eq. (40)], whereas the susceptibility, \(\chi_{\text{imp}}\), is given by Eq. (41).

The physics described above is not limited to TIs but is a general consequence of (doped) insulators (and semiconductors) with a large electronic density of states at the band edge such that in-gap bound states are easily induced by an Anderson impurity. Similar physics may be found in, for example, two-layer graphene systems. Our paper is just a first step towards understanding the rich physics associated with impurities in these systems.

ACKNOWLEDGMENTS

Discussions with C. M. Varma are highly appreciated. We acknowledge support by HKRGC through Grant No. HKUST03/CRF09. The research of I.K. and Y.A. is partially supported by Grant No. 400/12 of the Israeli Science Foundation (ISF).

APPENDIX A: INTERPRETATION OF \(V_0\) IN EQUATION (5)

Electrons in a lattice move in the periodic potential \(w_i(\mathbf{r})\),

\[ w_i(\mathbf{r}) = \sum_{\mathbf{n}} w_{\mathbf{a}}(\mathbf{r} - \mathbf{n}), \tag{A1} \]

where \(w_{\mathbf{a}}(\mathbf{r})\) is the interaction energy of electrons with an atomic lattice and \(\mathbf{n} = n_1 \mathbf{a}_1 + n_2 \mathbf{a}_2 + n_3 \mathbf{a}_3; \mathbf{a}_{1,2,3}\) are the lattice vectors and \(n_{1,2,3}\) are integers. In this case, electrons tunnel from one atom to another and the single-electron atomic levels \(\epsilon_a\) reduce to the energy band shown in Fig. 7.

When the atom of the lattice at the point \(\mathbf{r} = 0\) is replaced by an impurity atom, the potential energy \(w_i(\mathbf{r})\) of interaction of electrons with the impurity differs from \(w_{\mathbf{a}}(\mathbf{r})\). As a result, the potential energy of electrons in the lattice with the impurity,

\[ w(\mathbf{r}) = \sum_{\mathbf{n} \neq 0} w_{\mathbf{a}}(\mathbf{r} - \mathbf{n}) + w_i(\mathbf{r}), \tag{A2} \]

is not periodic anymore (see the purple curve in Fig. 7).

Then the potential scattering can be estimated as

\[ V_0 = \int d^3\mathbf{r} (w_i(\mathbf{r}) - w_{\mathbf{a}}(\mathbf{r})) \sim w_0 a_0^2, \tag{A3} \]

where \(w_i^{(0)}\) or \(w_{\mathbf{a}}^{(0)}\) is the peak of the potential energy of the impurity or the atom. Here we assume that the electric potential is screened at the interatomic distance \(a_0\).
Then the DOS (18) is
\[ \rho(\epsilon, r) = \rho_0(\epsilon) - \text{Im} \frac{\pi V_0 \left[ R^2(\epsilon, r) + \pi^2 \rho_0^2(\epsilon, r) \right]}{1 + \pi V_0 \left[ R(\epsilon, 0) + i \rho_0(\epsilon, 0) \right]}, \]
where \( R(\epsilon, r) \) and \( \rho_0(\epsilon, r) \) are real and imaginary parts of the Green’s function,
\[ \tilde{g}(\epsilon, r) = -\frac{1}{\pi} \sum_{\nu k} g_{\nu k}(\epsilon) e^{i \nu k r}. \]
Explicitly, \( \rho_0(\epsilon, r) \) is given by Eq. (22) and
\[ R(\epsilon, r) = \frac{\Theta(\epsilon_U - |\epsilon|)}{\pi} \sum_{\nu k} \frac{\sin(kr)}{kr(\epsilon - v_{\nu k})}. \]

When \(|\epsilon| > \epsilon_q\), \( R(\epsilon, r) \) vanishes and Eq. (B3) reduces to Eq. (21). When \(|\epsilon| < \epsilon_q\), the bare DOS vanishes, but the DOS (B3) gets a delta peak due to the localized \( f \) level,
\[ \rho(\epsilon, r) = \pi^2 V_0 \delta A^2(\epsilon_f, r) \delta(1 + \pi V_0 R(\epsilon, 0)), \]
where the condition of vanishing of the argument of the delta function gives us the secular equation (11) for \( \epsilon_f \). The amplitude \( R^2(\epsilon_f, r) \) of the delta peak vanishes when \( r \to \infty \), so that \( \epsilon_f \) is a localized state.

APPENDIX C: THE SO(4) KONDO HAMILTONIAN, EQUATION (32)
The SO(4) Hamiltonian is derived in regime (iii) of the RG analysis when \( \epsilon_f \) is located above the band gap. It has the form
\[ H_K^{SO(4)} = J_T (S \cdot S_f) + J_{ST} (R \cdot S_f), \]
where \( S \) and \( R \) are the \((S = 1)\) spin operator and the Runge-Lenz operator, respectively, with
\[ S^+ = \sqrt{2} (X^{10} + X^{01}), \quad S^- = \sqrt{2} (X^{01} + X^{10}), \]
\[ S^z = X^{11} - X^{-1}, \quad R^+ = \sqrt{2} (X^{15} - X^{51}), \quad R^- = \sqrt{2} (X^{51} - X^{15}), \]
\[ R^z = -(X^{05} + X^{50}). \]
Here \( X^{jk} = |\lambda\rangle \langle \lambda'|, \quad |\lambda\rangle \langle \lambda'|) = |S\rangle, |T_m\rangle \) (\( m = 0, \pm 1 \)) are the spin singlet and triplet states. The operators \( S \) and \( R \) are the generators of the group SO(4), as they satisfied the following commutation relations \((i, j, k = x, y, z, \text{summation convention implied})\):
\[ [S_i, S_j] = i \epsilon_{ijk} S_k, \quad [R_i, R_j] = i \epsilon_{ijk} R_k, \quad [R_i, S_j] = i \epsilon_{ijk} R_k. \]

APPENDIX D: RESISTIVITY
The resistivity for the SU(2) symmetry [regime (ii)] calculated within the third order of the perturbation theory is [31]
\[ R^{SU(2)}_{\text{imp}} = \frac{3R_0}{4} \left[ \frac{1}{b} \right] + \frac{3R_0}{2} \frac{1}{b} \mathcal{L}(T, \bar{D}), \]
where
\[ R_0 = \frac{3\pi e^2}{4\hbar^2}, \quad v_i \approx \frac{1}{\hbar} \left( \frac{\partial \epsilon_{ki}}{\partial k_i} \right); \]
\( k_1 \) and \( k_2 \) are two solutions of the equation \( \epsilon_k = \epsilon_f \) (see Fig. 1 of the main text). The function \( \mathcal{L}(T_1, T_2) \) is given by Eq. (7) of the main text.

Applying the condition of invariance of the resistivity under the “poor-man’s scaling,” we get
\[ R^{SU(2)}_{\text{imp}}(T) = \frac{3R_0}{4} \frac{1}{\mathcal{L}^2(T_1, T_2)}. \]

Here the factor \( N_v = 3/4 \) comes from the factor \( S(S + 1) \) which is 3/4 for \( S = 1/2 \).

The resistivity for the SO(4) symmetry [regime (iii)] calculated within the third order of the perturbation theory
The susceptibility for the SU(2) symmetry calculated within the second order of the perturbation theory is \[ \chi^{\text{imp}}_{\text{SU}(2)}(T) = \frac{\chi_0 T_{K_1}}{4T} \left\{ 1 - \frac{1}{\mathcal{L}(T_{K_1}, T)} \right\} \].

(E1) where the factor $K_{ij} = 1/4$ comes from $S(S + 1)/3$ which is $1/4$ for $S = 1/2$. For SO(4) symmetry, the impurity susceptibility calculated to the second order in $J$ is

\[ \chi^{\text{SO}(4)}_{\text{imp}}(T) = \frac{\chi_0 T_{K_1}}{2T} \left\{ 1 - J_T \rho_0 - \left( j^2_T \rho_0^2 + j^2_T \rho_0^2 \right) \mathcal{L}(T, D) \right\} \].

(E2)

The factors $P_{ii} = 2$ and $K_{ii} = 1/4$ have the following origin: there are two spins $S = 1/2$ (so that $P_{ii} = 2$); every spin gives the factor $K_{ii} = S(S + 1)/3 = 1/4$. For SO(3) symmetry, the impurity susceptibility calculated to the second order in $J$ is

\[ \chi^{\text{SO}(3)}_{\text{imp}}(T) = \frac{2 \chi_0 T_{K_1}}{3T} \left\{ 1 - J_T \rho_0 - J_T^2 \rho_0^2 \mathcal{L}(T, D) \right\} \].

(E3)

The factor $K_{ii} = 2/3$ comes from $S(S + 1)/3$ for $S = 1$.