Quantitative theory of backscattering in topological HgTe and (Hg,Mn)Te quantum wells: acceptor states, Kondo effect, precessional dephasing, and bound magnetic polaron

Tomasz Dietl1,2,∗

1 International Research Centre MagTop, Institute of Physics, Polish Academy of Sciences, Aleja Lotnikow 32/46, PL-02668 Warsaw, Poland
2 WPI Advanced Institute for Materials Research, Tohoku University, 2-1-1 Katahira, Aoba-ku, Sendai 980-8577, Japan

We present the theory and numerical evaluations of the backscattering rate determined by acceptor holes or Mn spins in HgTe and (Hg,Mn)Te quantum wells in the quantum spin Hall regime. The role of anisotropic s–p and sp–d exchange interactions, Kondo coupling, Luttinger liquid effects, precessional dephasing, and bound magnetic polarons is quantified. The determined magnitude and temperature dependence of conductance are in accord with experimental results for HgTe and (Hg,Mn)Te quantum wells.

I. INTRODUCTION

The experimental discovery of the quantum spin Hall effect (QSHE) in HgTe quantum wells (QWs) [1] provided experimental support for the seminal theoretical predictions about the topological protection [2] and opened prospects for electron transport without scattering backward, the capability of interest for energy-efficient and decoherence-free classical and quantum devices. However, in studied two-dimensional (2D) topological systems, such as aforementioned HgTe QWs [1–4] and 1T’-WTe2 2D monolayers [5–6], the two-terminal conductance approaches the quantized value $2e^2/h$ only in mesoscopic samples, shorter than 10 and 0.1 μm, respectively, without much improvement on cooling below 50 mK.

Understandably, the origin of processes accounting for the unanticipated short topological protection length has attracted a considerable attention. In 2D topological materials, electrons reside in counter-propagating Kramers degenerate helical states adjacent to sample edges. Furthermore, the direction of electron momentum determines the spin orientation. Therefore, mechanisms leading to electron backscattering must contain ingredients that break both time-reversal and spin-rotation symmetry. As reviewed recently [7, 8], previous theories considering backscattering in the linear response regime can be costed into three categories. The first of them contains approaches associating those ingredients to the presence of electron-electron interactions within the helical channels [9–14]. The second class of models considers the effects of external spins coming from extrinsic defects [15–17], magnetic impurities [8, 18–22] or nuclear spins [23–24]. Finally, the influence of random magnetic fluxes has been examined [25]. It can be argued that the abundance of theoretical proposals reflects difficulties in assessing magnitudes of material parameters entering into particular models, precluding a conclusive comparison of the theory to experimental results.

This paper, supporting and extending the companion report [26], presents theory and provides numerical evaluations of conductance in the regime of the quantum spin Hall effect (QSHE) in HgTe [1] and (Hg,Mn)Te quantum wells [27], considering holes localized on acceptors as a source of backscattering. We have quantitatively determined acceptor energies, the Coulomb gap, exchange coupling between edge electrons and acceptor holes, Kondo temperatures for both acceptor holes and Mn spins, spin-flip and backscattering rates in strong and weak coupling regimes, and effects of bound magnetic polarons. In addition to the Kondo effect discussed already in the context of backscattering [16, 23], we incorporate into our approach spin-nonconserving anisotropic components of the exchange coupling, allowing for backscattering even in the spin-momentum locking case [8, 16, 21, 23].

Our theoretical and numerical results substantiate the main conclusion of the companion paper [26] that the value of the exchange interaction between edge electrons and acceptor holes is large enough to drive the system to the strong coupling regime of the Kondo effect, where the spin-flip scattering rate attains the unitary limit. Together with the demonstrated sizable exchange anisotropy, the unitary scattering rate explains the observed topological protection length in HgTe QWs in the Ohmic conductivity regime [1, 28, 29] and the corresponding temperature dependence of conductance, if the influence of Luttinger liquid effects upon the anisotropic exchange are taken into account [30]. A key aspect of our approach is the observation that the acceptor concentration is known experimentally, as it determines hole density in undoped QWs and the gate voltage range separating the electron and hole transport in a given QW. In this way our theory does not involve any fitting parameters. Accordingly, the quantitative agreement with experimental data appears meaningful.

In contrast to occupied acceptor dopants, our results indicate that no Kondo effect is expected for impurities with open d or f shells in dilute magnetic semiconductors (DMSs), at least for typical values of antiferromagnetic sp–df exchange integrals. Furthermore, for magnetic ions with orbital momentum $L = 0$, such as Mn$^{2+}$, anisotropic components of the exchange coupling to edge states vanish, suggesting a minor role of Mn spins in backscattering. Nevertheless, we suggest that backscattering by a dense bath of interacting Mn spins can operate, as for the motionally-narrowed precessional spin-dephasing process, the constraint of local spin momentum conservation is relaxed. At the same time, we show that the formation of bound magnetic polarons by acceptor holes reduces the role of the Kondo effect, the observation elucidating the origin of the recovery of quantized conductance at low temperatures in (Hg,Mn)Te QWs [31].
II. THEORETICAL RESULTS

A. Quantum-well band states

A starting point of our theory is the eight-band Kohn–Luttinger Hamiltonian with boundary conditions previously used to obtain the subband structure in HgTe QWs \[32\] and the four-band model developed for determining, in the axial approximation, acceptor levels in GaAs QWs \[33\]. Within that approximation, discussed in Sec. II A, and in the absence of the impurity potential, the eight relevant electron wave functions \( \psi_i \) can be taken in the form,

\[
\psi_i^{(m)}(\rho, \varphi, z) = h_i^{(m)}(\rho, \varphi) f_i(z) u_i, \tag{1}
\]

where \( m_i \) is the integer orbital quantum number corresponding to the \( z \)-component of the orbital momentum \( L_z \) of the envelope function \( h_i; \rho, \varphi, \) and \( z \) denote the electron position in the cylindrical coordinates with \( z \) along the growth direction; \( u_i \equiv u_{j,z} \) is the set of the Kohn–Luttinger amplitudes for particular angular momenta \( j \) and its \( z \)-component \( j_z: S_1/2, 1/2, -1/2, 3/2, 1/2, -3/2, 2, -2, 1/2, \) and \( p_{1/2, -1/2} \) \[32\], whereas the envelope functions \( h_i^{(m)}(\rho, \phi) \) and \( f_i(z) \) are given by,

\[
h_i^{(m)}(\rho, \varphi) = N_{f}^{(m)} \exp(im_i \varphi) J_{m_i}(k \rho), \tag{2}
\]

where \( N_0 \) is a normalization factor. The Hamiltonian in question can be diagonalized by taking \( m_i = m \) for \( i = 1, 4, \) and \( 7; m_i = m + 1 \) for \( i = 2, 5, \) and \( 8; m_i = m - 1 \) for \( i = 3, \) and \( m_i = m + 2 \) for \( i = 6, \) where \( m \) is an integer and \( J_{m_i}(k \rho) \) is the Bessel function; \( k \) is a module of the in-plane wavevector, and

\[
f_i(z) = (1/\sqrt{L_z}) \sum_{n=-n_{\text{max}}}^{n_{\text{max}}} c_{n}^{(i)} \exp(2\pi i n z/L_z), \tag{3}
\]

where \( L_z \) is the total thickness of the structure, including the two CdTe:HgTe barriers and HgTe QW, assumed here as \( L_z = 60 \text{ nm} + d_{QW}; \) the expansion coefficients \( c_{n}^{(i)} \) are to be determined by a diagonalization procedure; \( n_{\text{max}} = 50 \) ensures an appropriate numerical convergence, as discussed in Sec. II D. As seen by inspection, each of the eight components \( \psi_i^{(m)}(\rho, \varphi, z) \) and, hence, the total wave function corresponds to the same value of the total angular momentum \( F_z = l_z + j_z = m + 1/2, \) meaning that the operator \( F_z \) commutes with the eight-bands’ axial Hamiltonian.

We adopt the identical values of low-temperature \( \vec{k} \cdot \vec{p} \) parameters, as in Ref. \[32\], except for the Luttinger \( \kappa \) that we assume to be the same in the barriers and QW. In this way, no Rashba-like splitting occurs in symmetric QWs. Computations have been performed assuming the absence of strain. Its role is discussed in Sec. II C.

Figures 1(a,b) in the companion paper \[26\] present QW subband dispersions \( E(k) \) for \( d_{QW} = 6 \text{ and } 8 \text{ nm}. \) The magnitudes of \( f_i(z) \) for \( d_{QW} = 8 \text{ nm} \) and \( k = 0 \) are shown in Fig. 1. A character of their inversion symmetry \( (z \rightarrow -z) \) plays an essential role in the anisotropy of the electron-hole exchange interaction, as discussed in Sec. II D.

B. Axial approximation

Figure 2 shows subband dispersions \( E(\vec{k}) \) computed without and with the axial approximation for HgTe QW thicknesses 6 and 8 nm. A slight overestimation of the indirect gap by the axial approximation is visible for \( d_{QW} = 8 \text{ nm}. \) Since, however, the energy differences between \( \vec{k} || (10) \) and \( \vec{k} || (11) \) near valence band top are significantly smaller than the acceptor binding energies in Fig. 1(c) of the companion paper \[26\], the axial approximation holds.

C. Strain effects

We use a conventional sign definition of biaxial epitaxial strain \[34,36\], i.e., a positive value of \( \epsilon_{xx} = \epsilon_{yy} \) corresponds to QW under the tensile strain. In particular, in the case of HgTe QW, \( \epsilon_{xx} = 0.31\% \) corresponds to strain for a CdTe substrate, whereas a Cd\(_{x-z}\)Zn\(_{y}\)Te substrate generates a compressive strain \[37\], \( \epsilon_{xx} < 0 \). The evolution of the QW band structure with biaxial strain is shown in Fig. 3 for HgTe QW of the thickness \( d_{QW} = 6 \text{ and } 8 \text{ nm} \) and \( \epsilon_{xx} = 0.31\%, 0, \) and \(-0.31\%. \) The strain, and also temperature, shift the QW thickness corresponding to the topological phase transition \( (E_g = 0) \), but the overall band picture remains intact, in agreement with previous conclusions \[32,37,38\].

![Figure 1](image-url)
FIG. 2. Verification of axial approximation. Subband dispersions for a HgTe QW of the thickness $d_{QW} = 6$ nm (a) and 8 nm (b) computed for $\vec{k} \| \langle 11 \rangle$ (dashed lines), $\vec{k} \| \langle 10 \rangle$ (dotted lines) and within the axial approximation (solid lines). Character of particular subbands at $k = 0$ is also marked.

FIG. 3. Effects of biaxial strain on band dispersion in HgTe QWs. (a) $\epsilon_{xx} = 0.31\%$ (CdTe substrate); (b) $\epsilon_{xx} = 0$; (c) $\epsilon_{xx} = -0.31\%$. Colors describe the participation of the $p_{3/2,\pm 3/2}$ Kohn-Luttinger amplitude in the electronic wave function; HgTe QW thickness $d_{QW} = 6$ nm.

D. Determination of acceptor level energies

To determine energies of levels brought about by charge dopants, we supplement the Hamiltonian by the Coulomb potential,

$$V(r) = -Ze^2/4\pi\epsilon_0er,$$

and by the potential of image charges in the barriers [33], for which the dielectric constant is $\epsilon_b = (1 - x_b)\epsilon + x_b\epsilon_e$, where in our case $\epsilon = 19.5$, $\epsilon_e = 10.5$, and $x_b = 0.7$. We neglect central cell corrections and the image charge in the gate metal, which is typically more than 100 nm apart.

Furthermore, we replace the Bessel function $J_{|n|}(k\rho)$ in Eq. 2 by

$$J_{|n|}(k\rho) \rightarrow \rho^{|n|} \sum_{l=1}^{l_{max}} d_l^{(i)} \exp(-\rho/a_l),$$

where the coefficients $d_l^{(i)}$ are to be determined by the diagonalization procedure for a given set of $a_l$ values. We take $a_l$ as a geometrical series with a common ratio of $r = 1.5$ and, for a typical number of exponential functions $l_{max} = 15$, the starting value of $a_1 = 0.5$ nm. Since, the exponential functions with real exponents are not orthogonal, a generalized eigenvalue solver has been employed to obtain electronic energies. At the same time, the participation number serves to evaluate an effective in-plane localization radius $a^*$.

The presence of the impurity potential breaks the degeneracy of the states with respect to the quantum number $m$. However, due to time–reversal symmetry, the impurity levels remain at least doubly degenerate. The use of exponential functions (Eq. 5) is suitable for determining the localized levels but not for oscillating extended states. Accordingly, the values of band energies do not converge with increasing $l_{max}$.

In our case, identifying acceptor levels that overlap with the continuum of band states appears difficult [39, 40]. Figure 4 depicts level energies determined for $m = 1$ in a HgTe QW with $d_{QW} = 6$ nm. As shown, the magnitude of $a^*$, and the evolution of the level energies with $Z$ [Fig. (a)], $l_{max}$ [Fig. (b)], and with the acceptor location of the QW center [Fig. (c)] unambiguously tell the band and resonant impurity states. According to the result presented in the companion paper [26], the ground state corresponds to $m = 0$ or $m = -1$, and is predominately built of the $p_{3/2,\pm 1/2}$ and...
Kohn-Luttinger amplitudes. Its energy is referred to as $E_{1/2}$.

E. Coulomb gap

The Efros-Shklovskii Coulomb gap [41] of the width $E_C$ reaching 10 meV was observed in a single-layer of 1T$^-\text{-WTe}_2$ with different coverage of surface by potassium [42]. Furthermore, numerical simulations presented the presence of the Coulomb gap for resonant donor states in HgSe:Fe [43].

In terms of 2D DOS of the acceptor band taken as $\nu_2 = N_a / E_b$, where $E_b$ is the acceptor bandwidth [41],

$$E_C = 2\nu_2 e^4 / (4\pi\epsilon_0 e)^2. \quad (6)$$

In the case of QWs, a non-zero $E_b$ originates from a dependence of the binding energy on the distance between the impurity and QW center. Figure 1(c) in the companion paper [26] shows this effect for acceptors in HgTe QWs, and demonstrates that the bandwidth extends down to a side maximum of the valence band characterized by a heavy mass. For a HgTe QW with $d_{QW} = 8$ nm, where $E_b \approx 20$ meV and the QW dielectric constant $\epsilon = 19.5$, we obtain the value of $E_C = 0.51$ meV for $N_a = 10^{11}$ cm$^{-2}$. The magnitude of $E_C$ sets the temperature scale above which conductance quantization deteriorates and hole mobility $\mu_h$ decreases for QWs at the topological phase transition. Thus, the Coulomb gap model explains the stability of the QSHE up to 100 K in WTe$_2$ monolayers [6], where – as mentioned above – $E_C$ attains 10 meV [42].

F. Electron-hole spin exchange

Taking into account previous insight that flip-flop transitions conserving total spin of edge electrons and, thus, due to spin-momentum locking edge current [19], we are interested in determining the degree of spatial symmetry breaking by the QW edge for the exchange interaction between topological edge electrons and acceptor holes in HgTe QWs. One of possible mechanisms could be kinetic exchange discussed in semiconductors in the context of $p$-$d$ coupling [44] and the Kondo effect in quantum dots [45]. However, we demonstrate here that the Bir–Pikus theory, originally developed for excitons [44], and later extended to the case of exchange coupling between band electrons and acceptor holes in bulk semiconductors [46], satisfactorily explains the experimental results for HgTe QWs.

The wave function of helical states at the edge along the $x$-direction assumes the form,

$$\psi_x^{(\pm)}(\vec{r}) = \exp(ikx) \sqrt{\Omega} [f_1(\vec{x}) u_1 + f_3(\vec{x}) u_3]; \quad (7)$$

$$\psi_x^{(-)}(\vec{r}) = T \psi_x^{(+)}(\vec{r}), \quad (8)$$

where $L_x$ is the edge length; $f_i$ are electron envelope functions that depend on $\vec{x} = (y, z)$; $u_i$ are the relevant Luttinger-Kohn amplitudes, and $T = -i\sigma_y K$ is a time reversal operator.

Similarly, for the acceptor ground state,

$$\psi_h^{(\pm)}(\vec{r}) = C [F_{h1}(\vec{r}) u_1 + F_{h3}(\vec{r}) u_3]; \quad (9)$$

$$\psi_h^{(-)}(\vec{r}) = T \psi_h^{(+)}(\vec{r}), \quad (10)$$

where $C = \sigma_y K$ is a charge conjugation operator transforming the acceptor wave function from the electron representation, employed in previous sections for the determination of electronic states, to the hole representation relevant here. In contrast, in the case of edge states we consider electrons and, thus, the electron representation for the Fermi level both above and below the Dirac point.

With these wave functions we determine a long-range contribution to the electron-hole exchange [44, 46]. This interaction is nonlocal and represented, in the case under consideration, by the matrix elements,

$$\mathcal{H}_{eh}^{\sigma_\sigma, \mu_\mu} = \frac{\hbar^2}{m_0^2 \Omega} \sum_{\alpha, \beta} \int d\vec{r}_1 d\vec{r}_2 \partial_{\alpha(\beta), \alpha(\beta)}^2 V(r_{12}), \quad (11)$$

where $\sigma$ and $\mu = \pm$; $\alpha, \beta = x, y, z$; $\vec{r}_{12} = \vec{r}_1 - \vec{r}_2$. Furthermore, in accord with the effective mass theory, internal matrix elements containing momentum operators $p_\alpha$ and $p_\beta$ are over the elementary cell volume $\Omega$ and involve the amplitudes $u_i$, whereas the envelope functions are assumed constant within this volume. As seen, non-vanishing matrix elements correspond to coupling of $s$-type and $p$-type edge states with $p$- and $s$-type acceptor states, respectively. Neglecting electrostatic image charges, the Coulomb energy has a standard form,

$$V(r_{12}) = \frac{e^2}{4\pi \epsilon_0 c r_{12}}. \quad (12)$$

It is convenient to present a spatial function in the Coulomb term as a sum of the local monopole and non-local dipole components,

$$\partial_{\alpha(\beta), \alpha(\beta)}^2 V(r_{12}) = - \frac{4\pi}{3} \delta_{\alpha\beta} \delta(\vec{r}_{12})$$

$$- \frac{(r_{12})^3}{r_{12}^5} \delta_{\alpha\beta} - 3 \pi \frac{r_{12}^3}{r_{12}^5} \delta_{\alpha\beta}. \quad (13)$$

In the case of electrons at the bottom of the conduction band and holes localized on acceptor impurities in bulk zinc-blende semiconductors, the exchange coupling has a scalar (Heisenberg) form, $H_{eh} = -J \vec{s} \cdot \vec{j}$, where $j = 3/2$ and $J < 0$ [46]. Since the QW and the edge break rotational symmetry, we expect, in the presence of intratomic spin-orbit coupling, a nonscalar form of the exchange interaction,

$$H_{eh} = - \sum_{\alpha, \beta} s_{\alpha} \mathcal{J}^{(\alpha, \beta)} J_{\beta}. \quad (14)$$

where $\mathcal{J}^{(\alpha, \beta)}$ is a real tensor and, in a standard notation, $\mathcal{J}_{\alpha} = \mathcal{J}^{(\alpha, \alpha)}$, and if $\alpha \neq \beta$, $\mathcal{J}_{\beta}^{(\alpha)} = (\mathcal{J}^{(\beta, \alpha)} + \mathcal{J}^{(\beta, \alpha)})/2$; $D_{\alpha} = \sum_{\beta, \gamma} \epsilon_{\alpha\beta\gamma} \mathcal{J}^{(\beta, \gamma)}/2$, where $D_{\alpha}$ are vector components.
of the Dzyaloshinskii-Moriya (DM) contribution and $\epsilon_{\alpha\beta\gamma}$ is the antisymmetric Levi-Civita tensor.

By comparing matrix elements of Hamiltonians given in Eqs. (11) and (14) considering both monopole and dipole contributions of the long-range electron-hole exchange interaction (Eq. (13), and taking into account that $\mathcal{H}_{eh} = -\mathcal{H}_{ee}$, we arrive to final forms of non-zero exchange tensor components and a spin-independent part of the Fock energy $\Delta_{eh},$

\[ \mathcal{J}_{x,y} = A \left( \frac{16\pi}{9} M_{21} + \frac{4}{3} M_{31} - 3M_{52} \pm M_{51} \right); \]  
\[ \mathcal{J}_{z} = A \left[ \frac{8\pi}{9} (M_{11} - 3M_{12}) + \frac{5}{3} M_{41} + M_{42} + 2\sqrt{3i}M_{63} \right]; \]  
\[ \mathcal{J}_{yz}^{(s)} = iA \left( \frac{8\sqrt{3\pi}}{9} M_{43} - \frac{\sqrt{3}}{3} M_{43} - \frac{\sqrt{3i}}{3} M_{53} \right); \]  
\[ D_x = -iA \left( \frac{8\sqrt{3\pi}}{9} M_{13} + \frac{\sqrt{3}}{3} M_{43} + \frac{\sqrt{3i}}{3} M_{53} \right); \]  
\[ \Delta_{eh} = A \left[ \frac{2\pi}{3} (-M_{11} + M_{12}) - \frac{1}{4} M_{41} - \frac{1}{4} M_{42} - \frac{\sqrt{3i}}{2} M_{63} \right]. \]  

In the above formulae the prefactor $A$ contains information about the strength of the Coulomb interaction and $\mathbf{k} \cdot \mathbf{p}$ coupling.

\[ A = \frac{e^2}{4\pi\epsilon_0\epsilon} \frac{P^2}{E_{eh}^2 L_x}. \]  

where $P$ is Kane’s $s$-$p$ momentum matrix element and $E_{eh}$ is the energy distance between hole and electron states in question.

Particular matrix elements $M_{ij}$ are given by

\[ M_{ij} = \int d\vec{r}_1 d\vec{r}_2 g_i(\vec{r}_1, \vec{r}_2)O_j(\vec{r}_1, \vec{r}_2), \]  

where

\[ g_1(\vec{r}_1, \vec{r}_2) = \delta(\vec{r}_{12}); \]  
\[ g_2(\vec{r}_1, \vec{r}_2) = \cos[k(x_1 + x_2)]\delta(\vec{r}_{12}); \]  
\[ g_3(\vec{r}_1, \vec{r}_2) = \cos[k(x_1 + x_2)](r_{12}^2 - 3z_{12}^2)/r_{12}^5; \]  
\[ g_4(\vec{r}_1, \vec{r}_2) = \cos[k(x_1 - x_2)](r_{12}^2 - 3z_{12}^2)/r_{12}^5; \]  
\[ g_5(\vec{r}_1, \vec{r}_2) = \cos[k(x_1 + x_2)](x_{12}^2 - y_{12}^2)/r_{12}^5; \]  
\[ g_6(\vec{r}_1, \vec{r}_2) = \sin[k(x_1 - x_2)]x_{12}z_{12}/r_{12}^5. \]  

Overlap functions between edge-electron and acceptor-hole envelopes read

\[ O_1(\vec{r}_1, \vec{r}_2) = [f_4(\vec{x}_1)F_{1h}(\vec{r}_1) - f_1(\vec{x}_1)F_{4h}(\vec{r}_1)] \cdot [f_4(\vec{x}_2)F_{1h}(\vec{r}_2) - f_1(\vec{x}_2)F_{4h}(\vec{r}_2)]; \]  
\[ O_2(\vec{r}_1, \vec{r}_2) = f_3(\vec{x}_1)F_{1h}(\vec{r}_1)f_3(\vec{x}_2)F_{1h}(\vec{r}_2); \]  
\[ O_3(\vec{r}_1, \vec{r}_2) = f_3(\vec{x}_1)F_{1h}(\vec{r}_1) \cdot [f_4(\vec{x}_2)F_{1h}(\vec{r}_2) - f_1(\vec{x}_2)F_{4h}(\vec{r}_2)], \]  

where for the adopted phase convention, in accord with the results presented in Fig. 1, $f_1(\vec{x})$, $f_3(\vec{x})$, and $F_{1h}(\vec{r})$ are real, $f_3(\vec{x})$ is negative, whereas $f_4(\vec{x})$ and $F_{4h}(\vec{r})$ are imaginary and change sign as a function of $z$. 

Inspection of the above equations shows that exchange integrals $J_{\alpha}$ are negative, implying antiferromagnetic coupling between edge electrons and localized holes, which allows for the Kondo coupling. At the same time, the presence of terms breaking the axial symmetry leads to spin non-conserving transitions ($r \neq 0$) and, hence, to net backscattering of edge electrons. The off-diagonal exchange tensor components $\mathcal{J}_{yz}^{(s)}$ and $D_x$ are non-zero for both monopole and dipole coupling but only if the inversion symmetry is broken, i.e., the acceptor resides away of the QW center, so that the hole envelope functions $F_{1h}$ and $F_{4h}(\vec{r})$ cease to be symmetric and antisymmetric with respect to $z = 0$, respectively. The presence of such terms was noted for dipole interactions of edge electrons with nuclear spins [23] and for the Heisenberg interaction with magnetic impurities [21]. However, in our case, even if inversion symmetry is maintained and, moreover, even if the $s$-type component in the hole wave function is negligible ($F_{1h} = 0$), spin non-conserving transitions are still allowed by the edge-induced breaking of the axial symmetry, leading to $\mathcal{J}_x \neq \mathcal{J}_y$ in the dipole contribution.

For numerical evaluations, guided by theoretical results obtained for the edge states [23, 47, 48] and our data presented in Secs. II A and II D, we assume the electron and hole envelope functions in an approximate form,
Here $b$ describes the penetration length of the edge electron wave function into the QW; $\Theta(x)$ is the Heaviside step function, $y_m$ and $z_m$ are distances of a maximum of the hole wave function from the sample edge located at $y = 0$ and the QW center residing at $z = 0$, respectively, and $N_o$ is an in-plane normalization factor. We note that one expects $|z_m| < d_{QW}/2$ even for acceptors localized outside the QW, $|z_0| > d_{QW}/2$.

Making use of previous results [23] as well as of our data presented in Secs. II A and II D we take $\varepsilon_s = 0.26$; $\varepsilon_p = 0.24$; $\varepsilon_H = -0.93$; $\varepsilon_s = 0.31$; and $\varepsilon_B = 0.92$; as the factors determining the participation of particular orbital components in the total electron and hole wave functions. Furthermore, the computations have been performed for $d_{QW} = 8$ nm and $a^* = 5$ nm. A Monte Carlo method has been used to evaluate six dimensional integrals. Such a method minimizes systematic errors but necessarily leads to a statistical scatter of the results. The chosen number of evaluation points insures that the accuracy of the obtained data is better than one sigma.

Figure 5 presents exchange tensor components as a function of $y_m$ for two values of $k = 0.05$ and 0.1 nm$^{-1}$ and the corresponding magnitudes of $b = 7$ and 5 nm, respectively [23, 47, 48]. As could be expected for the exchange interaction, $\mathcal{J}_a/A$ exponentially decay with $y_m$ but, not surprisingly, this decay is weaker for $r_{xy}$, as shown in Fig. 6. The magnitude of $r_{xy}$ is independent of $A$ and unaffected by a shift of the acceptor away from the QW center. However, for $z_m \neq 0$, other axial symmetry breaking terms appear, $\mathcal{J}_y$ and $\mathcal{J}_z$, where $\mathcal{J}_{yz} \ll \mathcal{J}_{xx}$, and, for $z_m = d_{QW}/4$, $r_{Dx} = [(3\mathcal{J}_x + \mathcal{J}_y + \mathcal{J}_z)/\mathcal{J}_{yz}]^2 \gg r_{xy}$, as depicted in Fig. 6. In the subsequent two sections, the obtained values of $\mathcal{J}_a/A$ and $r_i$ serve to estimate the magnitude of Kondo temperature and of the backscattering rate.

Finally, we comment on the spin-independent part of the Fock energy $\Delta_{eh}$, which together with Hartree terms originating from acceptor, edge, and gate charges, contribute to a self-consistent potential in the edge region, whose determination is beyond scope of the present work. Nevertheless, to have an idea about the energy scale involved, we evaluate a contribution to one-electron energy resulting from the Fock term in a self-consistent way making use of Eqs. [19] and [20] and summing up over all holes with the areal density $N_h$,

\[
E_{ch}^{(F)} = \left[\frac{\varepsilon^2 N_h P^2}{4\pi \epsilon_0 \epsilon} \int dy |f(y)|^2 \Delta_{eh}(y)/A \right]^{1/3},
\]

where the dependence $\Delta_{eh}(y)/A$ is displayed in Fig. 5 for two values of $k$ and $b$. Assuming $f(y) = \exp(-b/y)/\sqrt{b/2}$; $N_h = 0.5 \cdot 10^{11}$ cm$^{-2}$; $P = 8.28 \cdot 10^{-8}$ eV cm, and $\epsilon = 19.5$ we obtain $E_{ch}^{(F)} = 4.8$ and 4.4 meV for $b = 5$ and 7 nm, respectively.

G. Kondo temperatures for acceptor holes and Mn spins

1. Acceptor holes

A Kondo collective state results from antiferromagnetic exchange coupling of Fermi liquid with a single spin localized at $\vec{R}_e$. An antiferromagnetic electron-hole interaction leads also to the Kondo effect, as in the electron picture it corresponds to antiferromagnetic coupling of a surplus localized electron with the electron Fermi liquid. To estimate an order of magnitude of Kondo temperatures $T_K$, we use a time-honored expression [49] costed to the form,

\[
k_BT_K \approx W \exp[-1/(L^d \nu_d |\mathcal{J}_d|)],
\]

where $W$ is a energy width of carrier-containing states; $\nu_d$ is the carrier DOS per spin, and $\mathcal{J}_d$ represents the exchange energy of the antiferromagnetic interaction between spins of one carrier and a single localized paramagnetic center. Both $\nu_d$ and $\mathcal{J}_d$ correspond to values at the Fermi level, and depend on the dimensionality $d$ of the Fermi liquid residing in a structure of the volume $L^d$.

For the case under consideration, $\nu_1 = 1/2\pi h v_F$, where $v_F = 4 \cdot 10^5$ m/s [23, 47, 48], and $\mathcal{J}_1 = (\mathcal{J}_x + \mathcal{J}_y)/2$ at given $y_m$, $k$, $b$, and the energy of the acceptor hole with respect to the Fermi level, $E_{ch}$. Furthermore, together with $P$ and $\epsilon$ values quoted in previous sections, we take $W$ as a half of the bandgap, i.e., 15 meV for $T_K$ estimations.

Several important conclusions emerge from $T_K$ magnitudes presented in Fig. 7 as a function of $y_m$. First, for acceptor holes localized close to the edge, $y_m < b$, $T_K$ reaches 100 K, indicating that the Kondo physics is relevant for the system of edge electrons and acceptor holes, making competing effects, such as exchange interactions between hole spins, irrelevant. Second, due to the exponential dependencies of $T_K$ on $\mathcal{J}_a$ and $\mathcal{J}_0$ on $y_m$ there is a sharp cut-off $y_c$, beyond which strong coupling of holes and electrons vanishes. The magnitude of $y_c$ depends on $b$ and $E_{ch}$, that is, also on the position of the Fermi level in the acceptor band. Third, a broad distribution of $T_K$ values means that at any experimentally relevant temperature, coupling of edge electrons and holes corresponds to a superposition of weak and strong coupling limits, including the crossover between them occurring at $T \simeq T_K$. 

\[
f_1(y, z) = 2ie_s \exp(-y/b)\Theta(y) \cos(\pi z/d_{QW})\Theta(d_{QW}/2 - |z|)/(bd_{QW})^{1/2};
\]

\[
f_2(y, z) = 2ie_p \exp(-y/b)\Theta(y) \sin(2\pi z/d_{QW})\Theta(d_{QW}/2 - |z|)/(bd_{QW})^{1/2};
\]

\[
f_3(y, z) = 2i\mu \exp(-y/b)\Theta(y) \cos(\pi z/d_{QW})\Theta(d_{QW}/2 - |z|)/(bd_{QW})^{1/2};
\]

\[
F_{1h}(\vec{r}) = \hbar \exp[-(a^2 + (y - y_m)^2)/a^*] \Theta(y) \cos[\pi(\epsilon z - z_m)/d_{QW}]\Theta(d_{QW}/2 - \epsilon z)/N_o;
\]

\[
F_{4h}(\vec{r}) = i\hbar \exp[-(a^2 + (y - y_m)^2)/a^*] \Theta(y) \sin[2\pi(\epsilon z - z_m)/d_{QW}]\Theta(d_{QW}/2 - \epsilon z)/N_o.
\]
term ∆ topological HgTe QW, plotted as a function of the distance $y_m$ the Dzyaloshinskii-Moriya energy $D_x$ for a transition metal impurity, such as Mn, residing in a HgTe bulk sample or in a topological edge channel. A $p$-type part of the Bloch function is relevant, as the $p$–$d$ exchange integral $\alpha$ is antiferromagnetic, whereas the $s$–$d$ integral $\beta$ is ferromagnetic [50]. Accordingly, $J_\alpha$ appearing in Eq. [37] for $T_K$, can be written in a form,

$$J_\alpha = \beta [F_{3/2}(\tilde{R}_i)]^2 + |F_{1/2}(\tilde{R}_i)|^2 / 3,$$

(38)

where $F_{jz}$ are envelope functions accompanying the Kohn-Luttinger amplitudes $p_{3/2,jz}$. We see that an upper limit for the bulk case is $J_\alpha = \beta / V$, where $V$ is the sample volume. Similarly, for a Mn ion localized at a topological channel, where $\tilde{R}_i = (x, 0, 0)$ and, thus, $J_\alpha = 4\beta / (L_x b d_{QW})$.

Now, knowing that $N_0\beta = -0.7$ eV in HgTe [51], where $N_0 = 1.49 \cdot 10^{22} \text{cm}^{-3}$ is the cation concentration, we are in position to evaluate a lower limit of the exponent $w$ in Eq. [37] for $T_K$. For the bulk case, where $\nu_i = k_F m^* / 2\pi^2 \hbar^2$ and $m^* = 0.45 m_0$ [52], $w = -1.1 \cdot 10^3$ for the hole concentration $p = 10^{16} \text{cm}^{-3}$ and $w = -49$ for $p = 10^{20} \text{cm}^{-3}$, implying $T_K \sim 10^{-26}$ K. For topological edge channels in HgTe QWs, taking $b = 5$ nm, $d_{QW} = 8$ nm, and $\nu_F = 4 \cdot 10^3 \text{m/s}$, we obtain $w = -350$.

Hence, it appears that for standard values of the $sp$-$df$ exchange integrals, no presence of the Kondo effect is expected for spins tightly localized on $d$ or $f$ shells in semiconductors with magnetic ions. The same conclusion holds for nuclear spins for which the hyperfine coupling constant is at least four orders of magnitude smaller compared to $N_0\beta$.
FIG. 7. Calculated magnitudes (points) of Kondo temperatures, obtained for $J_{s}$ and $J_{F}$ values display in Fig. 5 for edge electrons in HgTe QWs containing holes bound to acceptors as a function of the distance $y_{m}$ between the acceptor wave function maximum and the edge calculated for various values of the electron-hole energy interval $E_{e h}$; full points: electron wavevector $k = 0.05$ nm$^{-1}$ and the penetration of helical state into the QW, $b = 5$ nm; open points: $k = 0.1$ nm$^{-1}$ $b = 5$ nm. Lines connect calculated points.

III. COMPARISON TO EXPERIMENTAL RESULTS

A. Topological protection length: acceptor holes and Mn spins

1. Acceptor holes

Our approach to charge transport by topological edge channels in quantum spin Hall (QSH) materials is built on several pillars put previously forward by others, discussed in the companion paper [26], or elaborated in previous sections of the present paper. First, charge dopants determine pertinent properties of 2D topological systems, including the dependence of carrier density on the gate voltage $V_{g}$ and the magnitudes of electron and hole mobilities. Second, owing to a dependence of the dopant binding energy on the position with respect to the QW center, the acceptors form a band extending over the whole bandgap. The associated Coulomb gap in the acceptor hole spectrum controls a contribution of QW states to charge transport in the QSH regime [26]. Third, due to a close energetic proximity, exchange interactions between edge channel electrons and acceptor holes are strong enough to bring the system to the Kondo regime (Sec. II C). As known, spin-dephasing rate reaches a unitary limit in that regime [15, 53, 55]. Fourth, since the magnitude of Kondo temperature $T_{K}$ exponentially depends on the distance of the acceptor hole to the edge and on the energy interval to the Fermi level, $T_{K}$ values show a broad distribution covering the whole experimentally relevant temperature range (Sec. II C). This effect determines also a spatial region from which acceptor holes contribute to spin-dephasing. Fifth, it has been emphasized by many authors [16, 19, 21, 23, 54] that because of spin-momentum locking, spin-conserving transitions imply momentum conservation, i.e., no net backscattering between helical channels in 2D topological insulators. However, spin-orbit interactions can result in spin-nonconserving processes described by non-scalar terms in the exchange coupling between edge electrons and paramagnetic centers [16, 19, 21, 23, 54]. The magnitude of such terms has been determined in Sec. II F for the case of QW holes bound to acceptor impurities.

We are interested in the topological protection length $L_{p} = \nu_{F}/\gamma_{h}$, where $\gamma_{h}$ is a backscattering rate for helical states. To see its relation to the two terminal conductance $G = 1/R$, we note that $R$ is non-zero due to the quantum contact resistance $R_{q} = e^{2}/2\nu_{F}$ and the backscattering term which, according to the Suhl resummation result, is given by $R_{b} = L^{2-\nu_{F}}(2\gamma_{h}v^{2}_{F}/d)^{-1}$. Hence, in our case, where $d = 1$, $G = (2e^{2}/h)(1 + L_{x}/L_{p})^{-1}$.

To evaluate $\gamma_{h}$ we recall a form of the carrier spin-dephasing rate in Kondo systems,

$$\gamma_{s} = C n_{S} f(T/T_{K})/2\pi\hbar\nu_{d},$$

where $C$ is a numerical coefficient close to one; $n_{S}$ in the concentration of magnetic impurities and $f(x)$ is a function obtained by the numerical renormalization group approach [53, 55] that is more accurate than the original Nagaoka-Suhl resummation result, $f(x) = a/(a + \ln^{2}(T/T_{K}))$, where $a = 3\pi^{2}/4$. In either case, $f(0) = 0$; $f(1) \simeq 1$, and $f(10) \simeq 0.6$.

It has been previously noted that compared to $\gamma_{s}$, the backscattering rate for helical channels is reduced by the exchange anisotropy ratio $r$, $\gamma_{h} = r\gamma_{s}$ [19, 21, 23]. Thus, inverse $L_{p}$ can be written in an appealingly simple form [26],

$$L_{p}^{-1}(T) = \sum_{i} r^{(i)} f(T/T_{K}^{(i)})/L_{x},$$

where the summation is over all QW holes bound to acceptors for a given $V_{g}$. As shown in Fig. 6, the $r$ value is not universal, but varies with the hole position in respect to the edge and QW center, $y_{m}$ and $z_{m}$, respectively, and to a lesser degree with $k$ and $b$. Similarly, according to results presented in Fig. 7, the $T_{K}$ magnitude strongly depends on $y_{m}$ and also on $E_{eh}$ that is controlled by the Fermi level position and $z_{m}$, the distance of the acceptor to the QW center, as shown in Fig. 1 of Ref. 26. However, just to see whether we are on the right track, we take the areal hole density as $N_{h} = 0.5 \cdot 10^{11}$ cm$^{-2}$, an average value of $r^{(i)}$ as $r_{Dx}(z_{m} = 2 \text{ nm}) = 0.13$ (see, Fig. 6), the cut-off length beyond which strong coupling of holes and electrons tends to vanish $y_{c} = 2b = 10 \text{ nm}$ (see, Fig. 7), and an average value of $f(T/T_{K}) = 0.4$. These numbers lead to the linear density $n_{h} = 5/\mu$ of holes participating in backscattering and $L_{p} = 4 \mu$ m, the order of magnitude consistent with experimental findings [11, 29, 56]. A small number of scattering centers has several important consequences, as discussed below.
Figure 8 shows the temperature dependence of two-terminal conductance $G$, as defined in Eq. 39, for a 10 µm long device of HgTe QW. The values of $G$ have been obtained by integrating $L_p^{-1}$ given in Eq. 41 in the $xy$ plane with the areal density of acceptor holes $N_{ah} = 0.5 \cdot 10^{11}$ cm$^{-2}$. The latter changes with the gate voltage and its overall magnitude vary from sample to sample. However, typically, the gate voltage width corresponding to the gap region, $\Delta V_g$, implies the total acceptor density of the order of $10^{11}$ cm$^{-2}$. The theoretical results are presented for three values of $E_{eh}$ and $L_x$. As $r(y)$, we have used the values fitted to $r_{Dy}(z_m = 2)$ nm in Fig. 6. Similarly, $T_K(y)$ is determined from fitted values of $J_{xy}$ and $f(y)$ shown by dashed lines in Fig. 5. Finally, the function $f(T/T_k)$ has been obtained from the data in Fig. 3 for $d = 1$ of Ref [53] and from the proposed there high-temperature extrapolation.

A number of worthwhile conclusions can be drawn from data in Fig. 8. In particular, lowering of temperature eliminates from backscattering spin centers with $T_K > T$, so that $L_p$ and $G$ are seen to steadily increase on cooling. However, the predicted recovery of $G = 2e^2/h$ at $T \to 0$ [15, 30] is not found, as at any temperature there are centers far away from the channel for which $T_K \geq T$. Nevertheless our results are not in accord with experimental findings in a sense that no systematic increase of $G$ on decreasing $T$ has been observed. This disagreement can point to the onset of localization, predicted for disordered Kondo systems [16] or to the presence of Tomonaga-Luttinger effects [15, 30]. In has been shown [30] that renormalization group equations combining Kondo and carrier correlation phenomena imply a growth of the exchange anisotropy ratio $r$ with lowering temperature down to $T^*$ according to,

$$r(T) = r(W)(W/k_BT)^2(1 - K),$$

with

$$k_BT^* = W \left[ \frac{L_x\nu_e|\mathcal{J}_r|}{\sqrt{2}(1 - K)} \right]^{1/(1 - K)},$$

where the Luttinger parameter $K < 1$ for interacting 1D systems. As shown in Fig. 8, incorporation of $r(T)$ into our theory makes $G$ virtually independent of $T$ for a rather moderate interaction magnitude $K = 0.88$. It should be, however, noted that at any temperature only a part of acceptor holes resides sufficiently far from the edge to satisfy the condition $T > T^*$.

Up to now, we have discussed configuration-averaged behavior of an acceptors’ containing system. However, a small number of acceptor holes involved in micron-size samples means that $G(V_g)$ will show strong mesoscopic-like conductance fluctuations, as observed [11, 31, 57, 58]. Furthermore, if at given $V_g$, acceptor holes close to the edge dominate for which $T_K > T$, $f(T/T_k)$ will increase and, hence, $G(T)$ decrease with temperature. By contrast, for distance acceptors $T_K < T$, so that a falling down of $f(T/T_k)$ with $T$ works together with the Tomonaga-Luttinger effects to result in an increase of $G(T)$ on heating. Such changes in sign of $dG(T)/dT$ with $V_g$ have been observed experimentally, though regions with $dG(T)/dT > 0$ appear to prevail [58]. According to Eq. 41 and many experimental observations, conductance quantization can be improved in short samples, $L_x < L_p$. However, a small number of relevant acceptors results in substantial conductance fluctuations. Charge traps biding two electrons in a singlet state might reduce backscattering and ensure pinning of the Fermi level in the QW band gap.

2. Backscattering by Mn spins and precessional spin dephasing

As there are typically $10^4$ more Mn spins with $S = 5/2$ in Hg$_{0.99}$Mn$_{0.01}$Te compared with the number of acceptor holes, a question arises about the role of Mn-induced backscattering of edge electrons. As the number of Mn ions in the edge region is also much larger than the number of edge electrons, we consider Mn spins in a continuous and classical approximation. Such an approximation is not only employed in the description of static and dynamic spintronics functionalities of ferromagnetic metals but has also been found versatile and quantitatively accurate in the case of bound magnetic polarons and single quantum dot electrons immersed in a nuclear spin bath in semiconductors [59] as well as when considering carrier-mediated ferromagnetism in DMSs [36]. Within this approach, it is convenient to expand local magnetization $\vec{M}(\vec{r})$ into a Fourier series,

$$\vec{M}(\vec{r}) = \sum_{\vec{q}} \vec{\eta}_q \exp(i\vec{q} \cdot \vec{r})/\sqrt{V}.$$
For a uniform distribution of Mn ions over the volume $V$ encompassing the QW and under thermal equilibrium conditions, the fluctuation-dissipation theorem implies,

$$\langle \eta_{\vec{q}}^{(\alpha)}\eta_{\vec{q}}^{(\beta)} \rangle_{\beta} = k_B T \chi_{\alpha,\beta}(\vec{q}) \delta_{\vec{q},-\vec{q}}$$

(45)

where in cubic systems and in the absence of a magnetic field, the magnetic susceptibility tensor $\chi_{\alpha,\beta}(\vec{q}) = \delta_{\alpha,\beta} \chi_{\text{Mn}}(q, T)$. We then apply a standard weak-coupling procedure (Fermi’s golden rule) for spin-flip transitions between 1D helical states, $\psi_{\vec{r}}^{(\pm)}(\vec{r}) \rightarrow \psi_{\vec{r}}^{(\mp)}(\vec{r})$, and adopt the form of $\psi_{\vec{r}}^{(\pm)}(\vec{r})$ given in Eq. 8. The scattering rate becomes,

$$\gamma_{\text{Mn}} = \frac{k_B T \chi_{\text{Mn}}(T)}{2\hbar^2 g^2 \mu_B^2 v_F} \int d\vec{r} [\alpha^2 |f_1(\vec{x})|^4 + \beta^2 (|f_3(\vec{x})|^4 + |f_4(\vec{x})|^4)]$$

(46)

where we have assumed that inverse $k$ and characteristic carrier confinement lengths are much longer than an average distance between magnetic ions, so that the $g$-dependence of magnetic susceptibility can be neglected, and $g = 2.0$ is the Landé factor of Mn spins. For a paramagnetic case and using the form of the envelope functions given in Eq. 35

$$\gamma_{\text{Mn}} = \frac{S(S + 1) x N_0}{4\hbar^2 v_F b_{\text{bd QW}}} [\alpha^2 |e_s|^4 + \beta^2 (|e_H|^4 + |e_p|^4)]$$

(47)

For the $s$-$d$ exchange energy $N_0\alpha = 0.3$ eV [51], $x = 0.01$ and other parameter values quoted in the previous paragraphs, the topological protection length becomes, $L_p = 7.2 \mu$m/$\mu_{\text{Mn}}$, where $r_{\text{Mn}} > 1$ is a relevant exchange anisotropy ratio. In principle, $r_{\text{Mn}} \ll 1$ for magnetic ions with the orbital momentum $L = 0$. There are however mechanisms that can enlarge $r_{\text{Mn}}$. In particular, Mn ions residing out of the QW center locally break the inversion symmetry, which leads to the appearance of anisotropic exchange [21], similarly to the case of spin scattering by acceptor holes (Sec. II F). This mechanism should work at $x \rightarrow 0$.

However, beyond the limit $x \rightarrow 0$, interactions between magnetic ions is relevant. While the exchange mediated by edge electrons is weak [36], antiferromagnetic superexchange is significant. This interaction makes $\gamma_{\text{Mn}}$ to decrease with lowering temperature even stronger than $T_\chi(T)$ implies because spin scattering becomes inelastic. In the spin-glass phase $T \leq T_g(x)$, Mn spins cease to contribute to quantum decoherence and, accordingly, a recovery of universal conductance fluctuations was found in nanostructures of Hg$_{0.93}$Mn$_{0.07}$Te at $T < T_g$ [60]. All that might mean that Mn spins play a minor role in backscattering.

Actually, it was noted that the presence of random Rashba fields effectively enhances the $r$ value [21]. We argue that there are two other effects. First, as discussed in the next section, magnetic polaron formation around acceptor holes considerably weakens the Kondo effect and associated backscattering in Hg$_{1-x}$Mn$_x$Te compared to HgTe. Secondly, we suggest that dephasing of carrier spins by a dense bath of interacting magnetic moments originates in semiconductors from a chain of spin precession events generated by local magnetization vectors rather than from flip-flop processes each involving a single magnetic ion.

In order to evaluate the precession-induced dephasing rate, we follow a time-honored Dykonov-Perel motional-narrowing approach to relaxation rates by spin-orbit fields, $\gamma_{\text{so}} = r_m ((g^* \mu_B B_{so}/2\hbar)^2)$, where, in the case under consideration, the motion time of the carrier wave packet extending over $\ell_p$, $\tau_m = \ell_p/v_F$ with $k^{-1} \gg \ell_p \gg (N_{\text{QW}})^{-1/3}$. This approach assumes that dynamics of Mn spins is slow, $\tau_m(\ell_p) < \tau_c(\ell_p)$, where $\tau_c$ is the Mn correlation time. Qualitatively, the magnitude of local magnetization seen by a moving edge carrier, $\langle M_{\alpha,\beta} \rangle = \delta_{\alpha,\beta} k_B T \chi_{\text{Mn}}(T) \int d\vec{r} |\psi(\vec{r})|^4$.

(48)

For the wavepacket,

$$\psi(\vec{r}) = \frac{2 \exp(ik_x x^2 / 4\ell_p^2 - y/b) \Theta(y)}{(2\pi)^{1/4}(b d_{\text{QW}})^{1/2}} \cdot \cos(px/2d_{\text{QW}} + 1/2)$$

(49)

and noting that spin splitting $\Delta = J_{sp-d} M / g \mu_B$ we arrive to,

$$\gamma_{\text{Mn}} = \frac{3S(S + 1) x N_0 J_{sp-d}}{8\sqrt{2\pi} \hbar^2 v_F b_{\text{bd QW}}}$$

(50)

We see that within a numerical factor of the order of one, the precession approach leads to the same expression, as given in Eqs. [60 and 47]. However, within such a model, similarly to the case of spin-transfer torque and electron precession around nuclear spins in a quantum dot, a change of carriers’ spin momentum associated with backscattering is absorbed by the ensemble of Mn spins, so that $r$ approaches 1. We conclude that backscattering by Mn spins may not be negligible in Hg$_{1-x}$Mn$_x$Te QWs, even in the absence of Rashba fields.

Finally, we return to the role of spin dynamics. The magnitude of $\tau_c$ is controlled by Mn spin diffusion and relaxation, determined by scalar and non-scalar terms in the Mn-Mn interaction Hamiltonian, respectively [61]. We note that the mechanisms accounting for finite $\tau_c$ weakens building up of Mn magnetization by electric current. Furthermore, it was suggested that spin dynamics could promote dephasing by edge carriers localized by Kondo impurities [16].

### B. Magnetic polaron gap and zero-field spin-splitting

We consider again isoelectronic magnetic impurities, such as Mn in II-VI compounds, in the paramagnetic phase. The presence of $sp$-$d$ exchange interactions affect, by the bound magnetic polaron (BMP) effect, donor electrons of acceptor holes even in the absence of macroscopic magnetization. Optical studies provided the evidence for the presence of the acceptor BMP in Hg$_{1-x}$Mn$_x$Te [62, 63]. A contribution of the BMP to thermally activated band conductivity [64] and to the Coulomb gap in the hoping region [65] was found in Cd$_{1-x}$Mn$_x$Te.

According to the analytical solution of the central spin problem, the polaron energy $\epsilon_p$ determines BMP energetics
and thermodynamics in the absence of an external magnetic field [68].

\[ \epsilon_p(T) = J_{sp-d}^2 \frac{\chi_{Mn}(T)}{[32(g\mu_B N_0)^2 \pi \alpha^3_{B}]} \tag{51} \]

where the magnitude of the exchange energy \( J_{sp-d} \) is given here by the values of the \( s-d \) and \( p-d \) exchange integrals \( N_0 \alpha \) and \( N_0 j \beta/3 \), respectively, weighted by the corresponding orbital content of the acceptor wave function; \( \chi_{Mn}(T) \) is the Mn susceptibility in the absence of acceptors; \( \alpha_B \) is the acceptor localization radius determined from the participation number. In the case of the doubly-occupied acceptor with \( J = 1 \), \( \epsilon_p(T) \) would be four times greater.

We identify the polaron gap between occupied and non-occupied acceptor centers, \( E_p \), as the twice Fermi energy shift associated with the polaronic effect. For the doubly-degenerate acceptor state \( E_p \) assumes the form [64].

\[ E_p = \epsilon_p + 2k_B T \ln(1 + \epsilon_p/k_B T). \tag{52} \]

For the \( E_{3/2} \) state, \( J_{sp-d} = -0.7 \text{ eV} \) implying, neglecting antiferromagnetic interactions between Mn spins, \( E_p = 1.7 \text{ meV} \) at \( T = 2 \text{ K} \), \( x = 0.024 \), and \( a_B = 5 \text{ nm} \). Similarly, for the more relevant \( E_{1/2} \) case, \( J_{sp-d} = -0.24 \text{ eV} \), for which \( E_p = 0.31 \text{ meV} \). For the \( x \) value in question there are about 500 Mn ions within the volume visited by the acceptor hole. We conclude that the formation of BMPs in Hg\(_{1-x}\)Mn\(_x\)Te may substantially enhance the magnitude of the gap at the Fermi level in the acceptor band compared to the case of HgTe.

We are also interested in the magnitude of acceptor hole spin-splitting \( \Delta \). If larger than \( k_B T_K \), BMPs reduce spin dephasing in the Kondo regime and, thus, backscattering of edge electrons, which improves the precision of resistance quantization in the quantum spin Hall effect regime. The most probable magnitude of zero-field splitting \( \Delta \) is given by an implicit equation [66].

\[ \Delta^2 - 2\Delta \epsilon_p(T) \tanh(\Delta/2k_B T) - 8\epsilon_p(T)k_B T = 0. \tag{53} \]

For the parameters quoted above we obtain \( \Delta = 2.7 \text{ meV} \) and \( 0.54 \text{ meV} \) at \( 2 \text{ K} \) for the \( E_{3/2} \) and \( E_{1/2} \) level, respectively.

In the particular case of the sample with \( x = 0.012 \) studied in Ref. [31] and for the level \( E_{1/2} \), \( \Delta \geq k_B T \) at \( T \leq 3.5 \text{ K} \). It is, therefore, clear that the presence of Mn, via the formation of BMPs and the associated diminishing of the role played by the Kondo effect, can significantly reduce backscattering of edge electrons, leading to the recovery of quantized resistance at low temperature in magnetically doped QWs, as observed in a Hg\(_{0.988}\)Mn\(_{0.012}\)Te QW at 0.2 K [31].

### IV. CONCLUSIONS

The quantitative results presented here support the view that acceptor states play a crucial role in the physics of quantum spin Hall effect in HgTe quantum wells and related systems. On the one hand, the ionization of acceptors accounts for a non-zero width of the quantized plateaus and, on the other, the strong Kondo coupling of edge electrons and acceptor holes leads to the unitary limit of the spin-flip scattering rate. A non-zero orbital momentum specific to \( p \)-type Kohn-Luttinger amplitudes together with breaking of axial and inversion symmetry by the edge and off center hole location allows for flow of edge electron angular momentum to crystal orbital momentum and, thus, for efficient backscattering, despite spin-momentum locking. According to the present insight, lowering of temperature drives the topological edge electrons from the Fermi, to the Kondo, and finally to the Luttinger liquid. Interestingly, the formation of bound magnetic polarons in magnetically doped samples weakness Kondo scattering, which allows for a recovery of conductance quantization at low temperatures, as observed [31].

As discussed in the companion paper [26], the acceptor band model qualitatively elucidates several surprising properties in the vicinity of the topological phase transition, where \( E_p \simeq 0 \) and, thus, acceptor form resonant states. In particular, gate-induced discharging of the acceptor states explains an unexpectedly slow rise of the itinerant hole concentration with increasing negative gate voltage [27] and unusually wide integer quantum Hall plateaus in the same gate region [11] [27] [68]. Furthermore, the presence of resonant bound magnetic polarons in magnetically doped samples diminishes Kondo scattering of electrons by acceptor holes and makes the Coulomb gap harder. We argue that these effects account for the hole mobility as large as \( 0.9 \cdot 10^6 \text{ cm}^2/\text{Vs} \) at hole density as low as \( p = 3.5 \cdot 10^9 \text{ cm}^{-2} \) in a Hg\(_{0.98}\)Mn\(_{0.02}\)Te QW at 20 mK [27]. In the same way, the model qualitatively explains \( \mu = 20 \cdot 10^6 \text{ cm}^2/\text{Vs} \) at the electron concentration \( n = 2.5 \cdot 10^{13} \text{ cm}^{-3} \) at 2 K in a bulk Hg\(_{0.93}\)Mn\(_{0.06}\)Te under hydrostatic pressure that made possible a fine tuning of the system to the topological phase transition [69].

We have also examined backscattering of electrons in helical states by localized spins in DMSs. Our results indicate that precessional spin dephasing by magnetization spacial fluctuations leads to sizable backscattering even in the spin-momentum locking case, as the electron spin momentum is transferred to the magnetic subsystem rather than to an individual magnetic ion.

### ACKNOWLEDGMENTS

This work was supported by the Foundation for Polish Science through the International Research Agendas program co-financed by the European Union within the Smart Growth Operational Programme.
tum wells using superlattice virtual substrates,” Phys. Rev. Lett. 117, 086403 (2016)

[38] A. M. Kadykov, S. S. Krishtopenko, B. Jouault, W. Desrat, W. Knap, S. Ruffenach, C. Consejo, J. Torres, S. V. Morozov, N. N. Mikhailov, S. A. Dvoretskii, and F. Teppe, “Temperature-induced topological phase transition in HgTe quantum wells,” Phys. Rev. Lett. 120, 086401 (2018)

[39] R. Buczko and F. Bassani, “Shallow acceptor resonant states in Si and Ge,” Phys. Rev. B 45, 5838–5847 (1992)

[40] M. S. Zhuludev, D. V. Kozlov, N. S. Kulikov, A. A. Razova, V. I. Gavrilenko, and S. V. Morozov, “Calculation of wave functions of resonant acceptor states in narrow-gap CdHgTe compounds,” Semiconductors 54, 827–831 (2020)

[41] B. I. Shklovskii and A. L. Efros, Electronic Properties of Doped Semiconductors (Springer, Berlin, 1984) pp. 232–237.

[42] Ye-Heng Song, Zhen-Yu Jia, Dongqun Zhang, Xin-Yang Zhu, Zhi-Qiang Shi, Huaqiuqiang Wang, Li Zhu, Qian-Qian Yuan, Haijun Zhang, Ding-Yu Xing, and Shao-Chun Li, “Observation of Coulomb gap in the quantum spin Hall candida single-layer 1T‘-WTe2,” Nat. Commun. 9, 04071 (2018)

[43] Z. Wilamowski, K. Świątek, T. Dietl, and J. Kossut, “Resonant states in semiconductors: A quantitative study of HgSe:Fe,” Solid State Commun. 74, 833–837 (1990)

[44] P. Kacman, “Spin interactions in diluted magnetic semiconductors,” Phys. Rev. B 40, 16,754–16,760 (1989).

[45] M. Pustilnik and L. Glazman, “Kondo effect in quantum dots,” Rev. Mod. Phys. 74, R513–R537 (2002).

[46] I. Yahniuk, A. Kazakov, B. Jouault, J. Wróbel, G. Cywiński, B. Jouault, S. S. Krishtopenko, S. Kret, and P. Becla, “Low-temperature transport properties of Cd0.91Mn0.09Te:In and evidence for a magnetic hard gap in the density of states,” Phys. Rev. Lett. 69, 1800–1803 (1992)

[47] T. Dietl and J. Spalek, “Effect of fluctuations of magnetization on the bound magnetic polaron: Comparison with experiment,” Phys. Rev. Lett. 48, 355–358 (1982)

[48] I. Yahniuk, A. Kazakov, B. Jouault, S. S. Krishtopenko, S. Kret, G. Grabekci, G. Cywiński, N. N. Mikhailov, S. A. Dvoretskii, J. Przybytek, V. I. Gavrilenko, F. Teppe, T. Dietl, and W. Knap, “HgTe quantum wells for QHE metrology under soft cryomagnetic conditions: permanent magnets and liquid He temperatures,” 10.48550/arXiv.2111.07581

[49] I. Yahniuk, S. S. Krishtopenko, G. Grabekci, B. Jouault, C. Consejo, W. Desrat, M. Majewicz, A. M. Kadykov, E. Spinir, V. I. Gavrilenko, N. N. Mikhailov, S. A. Dvoretskii, D. B. But, F. Teppe, J. Wróbel, G. Cywiński, S. Kret, T. Dietl, and W. Knap, “Magnetotransport in inverted HgTe quantum wells,” npj Quantum Mater. 4, 073903 (2019)

[50] M. Sawicki, T. Dietl, W. Plesiewicz, P. Śniadowski, M. Baj, and L. Dmowski, in Application of High Magnetic Fields in Physics of Semiconductors edited by G. Landwehr (Springer, Berlin, 1983) pp. 382-385.