Size constraints on a Majorana beam-splitter interferometer: Majorana coupling and surface-bulk scattering

Henrik Schou Røising[*] and Steven H. Simon

Rudolf Peierls Center for Theoretical Physics, Oxford OX1 3NP, United Kingdom

Topological insulator surfaces in proximity to superconductors have been proposed as a way to produce Majorana fermions in condensed matter physics. One of the simplest proposed experiments with such a system is Majorana interferometry. Here we consider two possibly conflicting constraints on the size of such an interferometer. Coupling of a Majorana mode from the edge (the arms) of the interferometer to vortices in the center of the device sets a lower bound on the size of the device. On the other hand, scattering to the usually imperfectly insulating bulk sets an upper bound. From estimates of experimental parameters, we find that typical samples may have no size window in which the Majorana interferometer can operate, implying that a new generation of more highly insulating samples must be explored.

I. INTRODUCTION

There has been an ongoing search for Majorana fermions in condensed matter systems which has been intensified over several years [1]. Vortices in a spinless p-wave superconductor have long been known to bind zero energy Majorana modes [2–5]. With p-wave superconductors being very rare in nature, no experiment has convincingly observed such Majoranas yet [6]. More recently it was predicted that Majorana bound states also exist in vortices in a proximity-induced superconductor on the surface of a topological insulator (TI) [7]. A number of recent experiments on TIs in proximity to superconductors [8–13] and other similar experimental systems [14,15] have increased the interest in this possibility.

The surfaces of TIs support gapless excitations with the dispersion of a Dirac cone [16] (in principle, TIs can have any odd number of Dirac cones in the Brillouin zone but we consider the simplest case of a single cone). The spectrum can be gapped either by applying a magnetic field to give the Dirac fermion either a positive or negative mass, or by placing a superconductor in proximity to the surface. At interfaces between different gapped regions, gapless one-dimensional fermionic channels can develop. For example, an interface between two magnetically gapped regions with opposite mass signs will contain a gapless and a chiral one-dimensional Dirac fermion mode. An interface between a magnetically gapped region and a superconducting region will contain a gapless chiral Majorana mode [7,17,18]. It is the physics of these modes that we are exploring in the current paper.

A very elegant experiment, building an interferometer out of these gapless chiral modes, was proposed by Fu and Kane [18] and simultaneously by Akhmerov *et al.* [17]. The device is depicted schematically in Fig. 1. Incoming particles or holes, biased at a low voltage, flow into a Dirac channel between two oppositely polarized magnetic regions. The Dirac

fermion is split into two Majorana fermions upon hitting the superconductor, one flowing in each direction around the superconducting region (drawn as a disk in Fig. 1). At the other end of the superconducting region the two Majorana modes are recombined into a Dirac mode. The differential conductance of this device was predicted to take the values 0 or $2e^2/h$ depending on whether the number of $\Phi_0 = h/(2e)$ vortices in the superconductor is even or odd, respectively. With the exception of quantum Hall systems, this was the first proposed realistic Majorana interferometry experiment, and it remains a good candidate for the first experiment to successfully establish the existence of chiral Majorana modes (although, we note that promising evidence of chiral Majorana modes was reported very recently [19]).

The most experimentally explored TI materials are the bismuth based compounds [16,20,21], including (among many others) Bi_2Se_3, Bi_2Te_3, and Bi_2Te_2Se. Within this class of materials, several experiments have successfully formed some sort of superconducting interfaces [8–13]. In this paper we mainly have this type of material in mind. However, we note
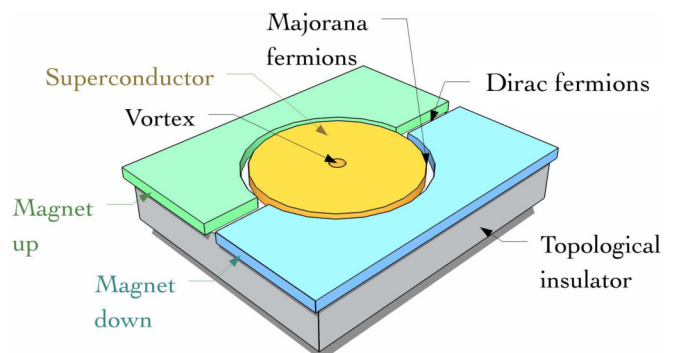

FIG. 1. The Majorana beam-splitter interferometer proposed by Fu and Kane, and Akhmerov *et al.* [17,18]. It consists of a 3D strong TI in proximity to a superconductor and magnets of opposite polarization.
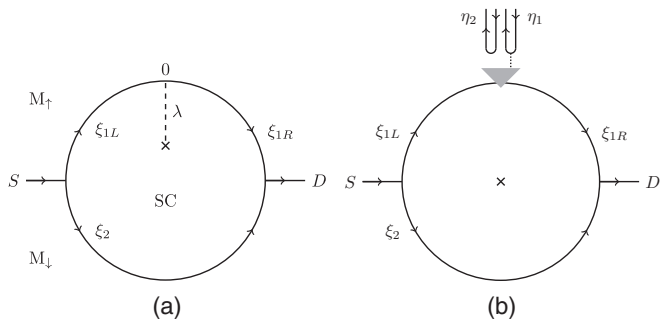
[*]henrik.roising@physics.ox.ac.uk

FIG. 2. Top view of the interferometry device with a superconductor (SC) and magnetic domains (with magnetization $M_{\uparrow/\downarrow}$) that probes charge transport from the source (S) to the drain (D) through a pair of chiral Majorana modes. (a) Single-point coupling of strength λ at coordinate $x = 0$. Here $\xi_{1R} \equiv \xi_1(x = 0^+)$ and $\xi_{1L} \equiv \xi_1(x = 0^-)$. (b) The interferometer coupled to a single-mode conducting lead representing leakage to the bulk of the TI.

that the material SmB_6 [22,23], which is possibly a topological Kondo insulator, will be discussed in the Conclusion.

In this paper we study two effects that restrict the size of the interferometer device, as summarized in Fig. 2. On the one hand, we consider coupling of the Majorana edge mode to the Majorana modes trapped in the core of the vortex in the center of the interferometer. When this coupling is sufficiently strong, i.e., when the interferometer is sufficiently small, the conductance signal will be distorted, rendering the interpretation of the experiment difficult. The coupling is generally strong on the scale of the coherence length $\xi = \hbar v_F / \Delta_0$, where v_F is the Fermi velocity and Δ_0 is the proximity gap. This length scale can be on the order of a micron (we discuss materials parameters in Secs. III A, IV B, and IV C).

Next, we consider surface-bulk scattering due to the unintentionally doped and poorly insulating bulk of TIs [20,24,25]. We find that when current leaks from the Majorana edge channel to the ground, the signal obtained at the end of the interferometer drops exponentially with a length scale set by the surface-to-bulk scattering length from disorder. For typical samples this length scale can be shorter than a micron. Thus we have two potentially conflicting constraints on the size of the proposed interferometer. We will discuss the possible directions forward in the Conclusion.

The outline of this paper is as follows. In Sec. II we review the formalism of Majorana interferometry and the characteristic conductance signal of the experiment. In Sec. III we consider the impact of coupling between the chiral Majorana and a bound Majorana mode in a vortex core. We study how this limits the source-drain voltage and sets a lower bound on the size of the system. In Sec. IV we consider the leakage of current from the device to the TI substrate and we establish an upper limit on the system size due to this leakage. We conclude with numerical estimates and an outlook for Majorana interferometry on the surface of TIs. Appendix A contains a derivation of the average currents. In Appendix B we provide generalizations of the single-point Majorana coupling displayed in Fig. 2(a). Vortex-bound Majorana fermions in a TI/SC hybrid structure and the energy splitting between zero modes is discussed in Appendix C. In Appendix D we present

details of our estimate of the surface-bulk scattering rate based on Fermi's golden rule. Finally, in Appendix E signal loss due to scattering with acoustic phonons is briefly discussed.

II. BACKGROUND ON INTERFEROMETRY WITH MAJORANA FERMIONS

In this section we first review the formalism needed to calculate the conductance and interferometry current [17,18,26,27]. In Fig. 2(a) we show a top view of the interferometer that was described in the Introduction. For the moment we ignore the Majorana coupled to the top arm (marked as an "X") at position 0. Charge transport from the source to the drain is computed by using a transfer matrix which describes transport of particles and holes from the source on the left (L), via the perimeter of the superconductor, to the drain on the right (R), $[\psi_e, \psi_h]_R^T = \mathcal{T}[\psi_e, \psi_h]_L^T$, where the superscript T here means transpose. The matrix $\mathcal{T}$ can be decomposed into three pieces corresponding to the three key steps between the source and the drain:

$$\mathcal{T} = S^\dagger \mathcal{P} S = \begin{pmatrix} \mathcal{T}_{ee} & \mathcal{T}_{eh} \\ \mathcal{T}_{he} & \mathcal{T}_{hh} \end{pmatrix}. \qquad (1)$$

The unitary matrix S relates the Majorana states $[\xi_1, \xi_2]$ running along the upper and lower edge of the superconducting disk to the electron and hole states $[\psi_e, \psi_h]$ that enter via the leads, $[\xi_1, \xi_2]^T = S[\psi_e, \psi_h]^T$. The matrix $\mathcal{P}$ contains plane wave phases that the low energy chiral Majorana modes accumulate as they move along the edge of the superconducting disk. Finally, the matrix $S^\dagger$ reassembles the two Majoranas into outgoing electron and hole states that enter the drain.

The matrices S and $\mathcal{T}$ are functions of energy. Due to the particle-hole symmetry, we must have $S(E) = S^*(-E)\tau_x$, where τ_x is a Pauli matrix in particle-hole space. At $E = 0$ these constraints fix $S(0)$ up to an overall phase that observables do not depend on [17],

$$S(0) = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix} \begin{pmatrix} e^{i\alpha} & 0 \\ 0 & e^{-i\alpha} \end{pmatrix}. \qquad (2)$$

We will apply $S(0)$ with $\alpha = 0$ for convenience. As discussed in Ref. [18] this form is exact when the system has a left-right symmetry. Even in cases where the system breaks this symmetry, corrections are $O(E^2)$ and can thus be ignored at low temperature and low voltage (see Appendix C 1). We study the symmetric situation where the magnetization is $M_0 \equiv M_\uparrow = -M_\downarrow$ throughout the main text. The magnetization enters the model Hamiltonian given in Eq. (C1).

The transfer matrix is used to calculate the average current in the drain as the difference between the electron and hole current, with the source biased at voltage V (see Appendix A for a derivation),

$$I_D = \frac{e}{h} \int_0^\infty dE \, \delta f(E)(|\mathcal{T}_{ee}|^2 - |\mathcal{T}_{eh}|^2). \qquad (3)$$

Here $\delta f = f_e - f_h$ with $f_{e/h}(E) = f(E \mp eV)$, and $f(E) = (1 + e^{\beta E})^{-1}$ is the Fermi-Dirac distribution with $\beta = 1/(k_B T)$ and E measured relative to the Dirac point. The incoming

current is

$$I_S = \frac{e}{h} \int_0^\infty dE \, \delta f(E). \tag{4}$$

By current conservation, a net current of $I_{SC} = I_S - I_D$ is absorbed by the grounded superconductor. As follows from Eqs. (3), (4), and unitarity of $\mathcal{T}$ the differential conductance measured in the grounded superconductor at zero temperature is

$$G_{SC}(V) = \frac{dI_{SC}}{dV}\Big|_{T=0} = \frac{2e^2}{h} |\mathcal{T}_{eh}(eV)|^2. \tag{5}$$

In order to calculate $G_{SC}(V)$ we need only establish the properties of the propagation matrix $\mathcal{P}$. If the arms of the interferometer are of length l_1 and l_2 we have

$$\mathcal{P}(E) = \begin{pmatrix} e^{ik(E)l_1 + 2i\phi(E)} & 0 \\ 0 & e^{ik(E)l_2} \end{pmatrix}. \tag{6}$$

Above the wave vector is of the form $k(E) = E/v_m$ due to the linear dispersion of the Majorana modes [18] where v_m is the Majorana velocity. We have included an additional phase shift ϕ which may come from a number of sources. In Refs. [17,18] the possibility of n vortices being added in the center of the superconducting region was considered. In this case the additional phase is $e^{2i\phi} = (-1)^n$. Inserting $\mathcal{P}$ into Eqs. (1) and (5) at zero temperature yields the conductance

$$G_{SC,0}(V) = \frac{2e^2}{h} \sin^2\left(\frac{n\pi}{2} + \frac{eV\delta L}{2\hbar v_m}\right). \tag{7}$$

Here $\delta L = l_1 - l_2$ is the difference between the lengths of the two arms. At $\delta L = 0$, $G_{SC,0}(V=0) = 2e^2/h$ for n odd, and is zero for n even. This would be a rather clear experimental signature. For the same phase matrix $\mathcal{P}$, the drain current in Eq. (3) evaluates to

$$I_{D,0}(V) = (-1)^n \pi k_B T \frac{e}{h} \frac{\sin\left(\frac{eV\delta L}{\hbar v_m}\right)}{\sinh\left(\frac{\pi \delta L k_B T}{\hbar v_m}\right)}, \tag{8}$$

which holds in the low temperature and low voltage limit (compared to the bulk gap). We emphasise that these results are derived with zero coupling to the central Majorana and to any other degrees of freedom (e.g., phonons or conducting bulk states), hence the subscript 0.

III. EFFECT OF MAJORANA COUPLING

We now consider the effect of coupling the Majoranas trapped in the cores of vortices in the superconductor to the chiral edge states. Each vortex traps a single Majorana mode. If a vortex is close to the edge (roughly within the coherence length) there will be tunneling coupling as shown in Fig. 2(a) where the Majorana zero mode is marked as an "X" and the (tunneling) coupling matrix element is of magnitude λ. The magnitude of the coupling drops exponentially with the distance between the vortex and the edge, see Appendixes C 2 and C 3.

We very generally describe the chiral Majorana on the upper interferometer arm ξ_1 by a Lagrangian density $\mathcal{L}_1 = i\xi_1(\partial_t + v_m\partial_x)\xi_1$, where x is the spatial coordinate along the upper edge. A similar description holds for the state on the lower edge ξ_2. The vortex bound Majorana ξ_0 is described by $\mathcal{L}_0 = i\xi_0\partial_t\xi_0$.
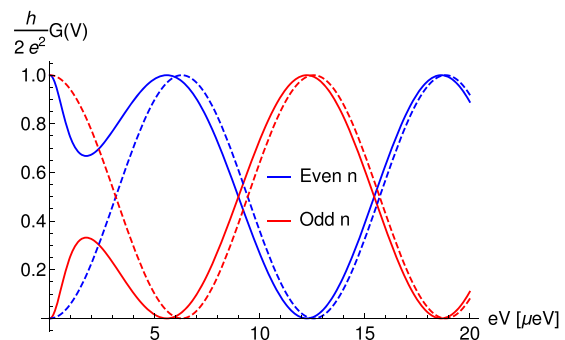

FIG. 3. The differential conductance from Eq. (13) as a function of the voltage. Here $\nu = 1\ \mu$eV and $\hbar v_m/\delta L = 2\ \mu$eV. The dotted lines show the conductance with $\nu = 0$. At low voltage the even-odd effect is reversed.

We add the coupling term between the central bound state and the chiral mode

$$\mathcal{L}_{\text{bulk-edge}} = 2i\lambda\xi_1(x=0)\xi_0. \tag{9}$$

The equations of motion, following from the full Lagrangian $\mathcal{L}_0 + \mathcal{L}_1 + \mathcal{L}_{\text{bulk-edge}}$, are given by [28,29]

$$2\partial_t\xi_0 = \lambda[\xi_{1R} + \xi_{1L}], \tag{10}$$

$$v_m\xi_{1R} = v_m\xi_{1L} + \lambda\xi_0. \tag{11}$$

Here the notation $\xi_{1R} = \xi_1(x=0^+)$ and $\xi_{1L} = \xi_1(x=0^-)$ was introduced. A Fourier transformation yields a phase shift across the coupling point,

$$\xi_{1R}(\omega) = \frac{\omega + i\nu}{\omega - i\nu}\xi_{1L}(\omega) = e^{2i\phi}\xi_{1L}(\omega), \tag{12}$$

with ω the frequency, $\nu \equiv \lambda^2/(2\hbar v_m)$, and $\phi(\omega) = \arctan(\nu/\omega)$. This energy-dependent phase shift is inserted into Eq. (6) and we obtain the zero-temperature result

$$G_{SC}(V) = \frac{2e^2}{h}\sin^2\left[\frac{n\pi}{2} + \frac{eV\delta L}{2\hbar v_m} + \arctan\left(\frac{\nu}{eV}\right)\right]. \tag{13}$$

Observe that the even-odd effect undergoes a crossover when the coupling strength is of the order $\lambda \approx \sqrt{2\hbar v_m eV}$. At zero voltage, or at infinite coupling strength, the even-odd effect is reversed from the value at high voltage or low coupling strength. This crossover is equivalent to shifting n by one in $G_{SC,0}$, causing ξ_1 to acquire a phase shift of π at low energy, and it is assigned the interpretation that a vortex Majorana effectively is absorbed by the edge [28]. Similar results are known from quantum Hall interferometers at filling fraction 5/2 [29–31]. The original conductance is recovered at high voltage (see Fig. 3).

The above result applies when there is no position dependence in the coupling to the edge. If we instead consider a continuous bulk-edge coupling

$$\mathcal{L}_{\text{bulk-edge}} = 2i\int dx\, \lambda(x)\xi_1(x)\xi_0, \tag{14}$$

then the total phase shift is again given by Eq. (12) with

$$\lambda^2 \rightarrow 2\int dx\, \lambda(x)e^{ikx}\int_x dx'\lambda(x')e^{-ikx'} \tag{15}$$

in the numerator and a similar replacement in the denominator, see Appendix B 1. We note that if $x_c \ll 1/k$, where x_c is defined such that $\lambda(|x| > x_c) \approx 0$, then the effective λ^2 above becomes $\langle \lambda(x) \rangle^2$ with $\langle \lambda(x) \rangle = \int dx \lambda(x)$ up to corrections of order $k x_c$.

The scheme above can also be generalized to include more complicated couplings to multiple (vortex) Majorana modes. So long as these modes are coupled to only a single edge, and not to each other, each coupling causes a phase shift of $\arctan[\nu/(eV)]$ in the phase of propagation along the edge, see Appendix B 2. In the case that multiple vortex Majoranas are coupled to each other, an even number of Majoranas will gap out, whereas an odd number will leave a single effective Majorana zero mode.

In the above calculation we assumed Majorana coupling to one edge only. Since coupling varies exponentially with distance to the edge, it is not unreasonable that this will effectively be the case. However, it is also realistic that a vortex will be roughly equal distance from, and hence equally coupled to, both edges. This case is discussed in detail in Appendix B 3. While the general result becomes complicated, at least in the case of equal couplings to both edges and both edges of equal length, the physics of the even-odd crossover found in this section remains unchanged. Finally, coupling between a single vortex Majorana and the edge at finite temperature is discussed in Appendix B 4.

A. Lower bound on size and voltage constraints

The above derived crossover makes the observation of the even-odd conductance effect impossible at low voltage. The proposed experiment is to add a single vortex and observe a change in conductance (say, from zero to $2e^2/h$). However, if the Majorana is then effectively absorbed into the edge, the conductance remains zero even once the vortex is added, destroying the predicted effect. This will occur for interferometers of size comparable to the coherence length, i.e., the length scale of an order-parameter deformation. Assuming that $\Delta_0 = 0.1$ meV is an achievable proximity-induced gap in, e.g., Bi_2Se_3 [10,12], the naive lower size bound for the disk is $\xi = \hbar v_F/\Delta_0 \simeq 4 \, \mu$m [32]. The tunneling coupling is expected to decay like $\lambda \propto \mu \exp(-R/\xi)$ when $R \gg \xi$ [see Eq. (C5)] where μ is the chemical relative to the Dirac point [33]. For disks of size $R \simeq \xi$ the energy splitting is comparable to the energy gap and the notion of stable edge/vortex states breaks down.

We note, however, that the Majorana coupling term vanishes identically at $\mu = 0$ in the topological insulator superconductor hybrid structure [34,35]. This is related to the appearance of an additional symmetry at the Dirac point, bringing the Hamiltonian [Eq. (C1) in the absence of a magnetic field] from symmetry class D to the BDI in the Altland-Zirnbauer classification. If there is disorder inducing local fluctuations in the chemical potential (see Sec. IV C) [36], say on some scale $\delta\mu$, a random coupling term of the type in Eq. (14) will be present and cause energy splitting. Still, the scenario of having the average $\langle \lambda(x) \rangle \sim \langle \mu(x) e^{-r(x)/\xi} \rangle \approx 0$, which would make the parasitic phase shift vanish, is possible but becomes extremely geometry sensitive (e.g., sensitive to the vortex position) if $R \simeq \xi$. Moreover, we should expect $\langle \lambda(x) \rangle$ to

be on the scale of $\delta\mu\sqrt{l/d}$, which in general will not be small. Here d is the length scale associated with the energy fluctuations. Assuming that we cannot control disorder, the only way to assure suppression of unwanted phase shifts and energy splitting is to increase R. Thus, we will use $R = \xi$ as a strict lower size bound on the device.

As far as chiral transport on the interferometer arms is concerned, it seems at this stage that one can operate at high voltage $eV \gtrsim \nu$ to avoid the coupling. Naturally the voltage is constrained from above by the global bulk gap $eV \lesssim \min\{M_0, \Delta_0\}$, to avoid excitation of nontopological states. Thus, if $R \simeq \xi$ the remaining voltage range $\nu \lesssim eV \lesssim \min\{M_0, \Delta_0\}$ might be too limited to have a clear experimental signature. Increasing the disk radius to suppress the coupling (and therefore lower ν) induces the problem of signal leakage to conducting bulk states, which we estimate below.

Finally, we note that the upper voltage bound in practice can be lower. This is due to the Caroli–de Gennes–Matricon excited states of the vortex, characterized by a minigap $\sim \Delta_0^2/\mu$ [37–39], which can be less than 1 mK. The excited vortex bound states can be activated thermally or by tunneling from the edge states, in analogy to the zero mode tunneling in the beginning of this section. By pinning the vortex to a hole in the superconductor, the minigap can be increased to a substantial fraction of Δ_0 [40,41]. Although the details of such a tunneling process is outside the scope of this paper, additional resonances and conductance phase shifts are expected as eV hits the bound state energies.

IV. SURFACE-BULK SCATTERING

Topological qubits are intrinsically protected from decoherence [42]. In protocols based on braiding with topological qubits, leakage of current is harmful since it generically causes entanglement with the environment that potentially corrupt the qubits [7]. Although the experiment we study here does not probe a topological qubit, both types of experiment are sensitive to bulk leakage. Since most TIs are poorly insulating [21], bulk leakage is a relevant problem to consider.

In this section we model leakage of Majorana modes from the surface of the TI to its poorly insulating bulk. This is done by coupling the interferometer arm first to a single metallic lead. Then we take multiple weakly coupled metallic leads to represent a continuously leaking environment. We combine the result of this scattering process with an estimate of the surface-to-bulk scattering rate from disorder in doped TIs. Our results suggest an upper size bound that potentially coincides with the lower bound discussed in Sec. III for many unintentionally doped TIs.

A. Scattering on conducting leads

Let the upper interferometer arm be coupled to a metallic lead as depicted in Fig. 2(b). Referring to the formalism of Ref. [27], the lead fermions are transformed to a Majorana basis $[\eta_1^{(\pm)}, \eta_2^{(\pm)}]^T = S[\psi_e^{(\pm)}, \psi_h^{(\pm)}]^T$ with the superscript indicating incoming $(-)$ or outgoing $(+)$ states. The S matrix here may differ from the one in Eq. (2) at low energy only by having a different phase α, which is irrelevant for observables. The

scattering process in the Majorana basis is denoted by $\mathcal{A}(E)$,

$$\left(\xi_{1R}, \eta_1^{(+)}, \eta_2^{(+)}\right)^T = \mathcal{A}(E)\left(\xi_{1L}, \eta_1^{(-)}, \eta_2^{(-)}\right)^T. \quad (16)$$

By rotating the lead particles and holes into the appropriate Majorana basis, the chiral Majorana mode can be shown to decouple from one of the (artificial) lead Majoranas [27]. In the low energy limit, the scattering matrix is a real rotation matrix $\mathcal{A}(E) \in SO(3)$. If we denote the local reflection amplitude of η_1 (η_2) by r_1 ($r_2 = 1$), this means that the scattering matrix can be parametrized by r_1 at the junction only (the transmission amplitude is $t_1 = \sqrt{1 - r_1^2}$),

$$\mathcal{A}(E) = \begin{pmatrix} r_1 & -t_1 \\ t_1 & r_1 \end{pmatrix} \oplus r_2. \quad (17)$$

Including the state ξ_2 on the lower arm and the plane wave phases acquired across the interferometer, we obtain the matrix $\mathcal{P}$ acting on $[\xi_{1L}, \xi_2, \eta_1^{(-)}, \eta_2^{(-)}]^T$,

$$\mathcal{P} = \begin{pmatrix} r_1 e^{ikl_1 + in\pi} & 0 & -t_1 e^{i\frac{kl_1}{2}} \\ 0 & e^{ikl_2} & 0 \\ t_1 e^{i\frac{kl_1}{2} + in\pi} & 0 & r_1 \end{pmatrix} \oplus 1. \quad (18)$$

Here $r_2 = 1$ was used. The transfer matrix can be computed, and it has the 2×2 subblock structure

$$\mathcal{T} = (S^\dagger \oplus S^\dagger)\mathcal{P}(S \oplus S) = \left(\begin{array}{c|c} \mathcal{T}_{S \to D} & \mathcal{T}_{\ell_{in} \to D} \\ \hline \mathcal{T}_{S \to \ell_{out}} & \mathcal{T}_{\ell_{in} \to \ell_{out}} \end{array}\right). \quad (19)$$

Above, the subscripts indicate the result of transport from/to the source (S), the drain (D), or the incoming/outgoing lead ($\ell_{in/out}$). The blocks in this transfer matrix are used to find the average current contribution as measured in contact β, with

$$|(\mathcal{T}_{\alpha \to \beta})_{ee}|^2 - |(\mathcal{T}_{\alpha \to \beta})_{eh}|^2$$

$$= \begin{cases} r_1(-1)^n \cos\left(\frac{E\delta L}{v_m}\right) & (\alpha, \beta) = (S, D), \\ r_1 & (\alpha, \beta) = (\ell_{in}, \ell_{out}), \\ 0 & (\alpha, \beta) = (S, \ell_{out}), (\ell_{in}, D). \end{cases} \quad (20)$$

Here the first line gives the contribution from combining Majoranas ξ_1 and ξ_2, the second line from combining η_1 and η_2, and the third line from combining ξ_1 and η_1. Combining Majoranas from different sources always give a vanishing contribution to the average current [27].

The drain current is thus $I_D(V) = r_1 I_{D,0}(V)$, where $I_{D,0}$ is defined in Eq. (8); the visibility is coherently reduced by r_1. The current in the outgoing lead is $I_{\ell_{out}} = r_1 e^2 V'/h$ at $T = 0$ when the lead is biased at V'. Subtracting the incoming current $I_{\ell_{in}} = e^2 V'/h$ yields a net current of $I_{\ell_{out}} - I_{\ell_{in}} = (r_1 - 1)e^2 V'/h$ in the conducting lead. When decoupling the lead completely, $r_1 = 1$, no net current goes in the lead.

The matrix $\mathcal{P}$ can be trivially extended to include scattering on many single-mode leads. Repeating the calculation above with two scattering leads of the same reflection amplitude r_1 gives the same reduction of current in both leads, independent of which arms the leads are coupled to. The drain current is reduced by r_1^2. Generalizing these statements by induction [43], with N identical scatterers of reflection amplitude r_1, the total conductance from the collection of leads (representing the bulk of the TI) is $G_{leads} = e^2 N(1 - r_1)/h$. Furthermore,

the N scatterers reduce the visibility of the drain current multiplicatively as $I_D(V) = r_1^N I_{D,0}(V)$.

If we let l_S denote the average length a chiral Majorana travels before it scatters into the bulk, we may by definition express the reflection amplitude as $r_1 = 1 - \frac{1}{N}\frac{l}{l_S}$, which is equivalent to defining the leakage conductance into the collection of leads by $G_{leads} = e^2 l/(h l_S)$. Taking the continuum limit of infinitely many weak scatterers means that $\lim_{N \to \infty} r_1^N = \exp(-l/l_S)$, and the current is exponentially suppressed in the drain,

$$I_D(V) = I_{D,0}(V) e^{-l/l_S}. \quad (21)$$

As for the differential conductance measured in the grounded superconductor, the amplitude of the oscillations are suppressed

$$G_{SC}(V) = G_{SC,0}(V) e^{-l/l_S} + \frac{e^2}{h}(1 - e^{-l/l_S}), \quad (22)$$

and distinguishing even from odd n is rendered difficult as l surpasses l_S. Thus, l_S acts as an upper bound on the length of the interferometer arms.

B. The surface-bulk scattering rate

The remaining step of our argument is to estimate the scattering length l_S. We start by calculating the surface-bulk (SB) scattering length for a TI surface electron ($l_S^{(e)}$) in the absence of any surface superconductor or magnet. Then, we use a similar calculation to estimate the Majorana scattering length ($l_S^{(m)}$). We restrict ourselves to elastic (zero temperature) surface-bulk scattering, where the electron-phonon coupling is irrelevant (see Sec. IV D and Appendix E). Building on the formalism developed in Ref. [44] we consider scattering via screened charge impurities. We note that the consideration in Ref. [44] is restricted to point scatterers where the overall scattering strength is *a priori* unknown, whereas in our approach the effective potential strength is fixed by the screened Coulomb potential and the dopant concentration.

Specifically, we consider a surface state (S) with initial wave vector $\boldsymbol{k} = (\boldsymbol{k}_\parallel, 0)$ that scatters to a lowered conduction band (a charge puddle) n' in the bulk (B) with final wave vector $\boldsymbol{k}' = (\boldsymbol{k}'_\parallel, k'_z)$. The incoming surface state has energy $\epsilon_F = v_F k_\parallel$. For numerical purposes we work with a TI slab of thickness $L = 40$ nm. In Fig. 4(a) we show the form of $|\Psi(z)|^2$ at zero parallel momentum for the lowest three eigenstates (derived in Ref. [45]). One of the TI surface states, penetrating tens of Ås into the bulk, is seen in purple. Assuming randomly distributed screened (dopant) charges with average concentration n_{3D} we get the scattering rate from Fermi's golden rule (cf. Appendix D):

$$\Gamma_{SB}^{imp}(\epsilon_F)$$

$$= n_{3D}\left(\frac{e^2}{2\pi\epsilon_0\epsilon_r k_\parallel^2}\right)^2 \sum_{n' \in B} \frac{k'_\parallel}{|\nabla \xi_{B,n',k'_\parallel}|} \int_0^\infty dk'_z \frac{d\sigma_{long}^{(n')}(\epsilon_F)}{dk'_z}. \quad (23)$$

Here $\xi_{B,n',k'_\parallel}$ is the dispersion of bulk band n' defining the outgoing wave vector $k'_\parallel$ by $\xi_{B,n',k'_\parallel} = \epsilon_F$. The differential
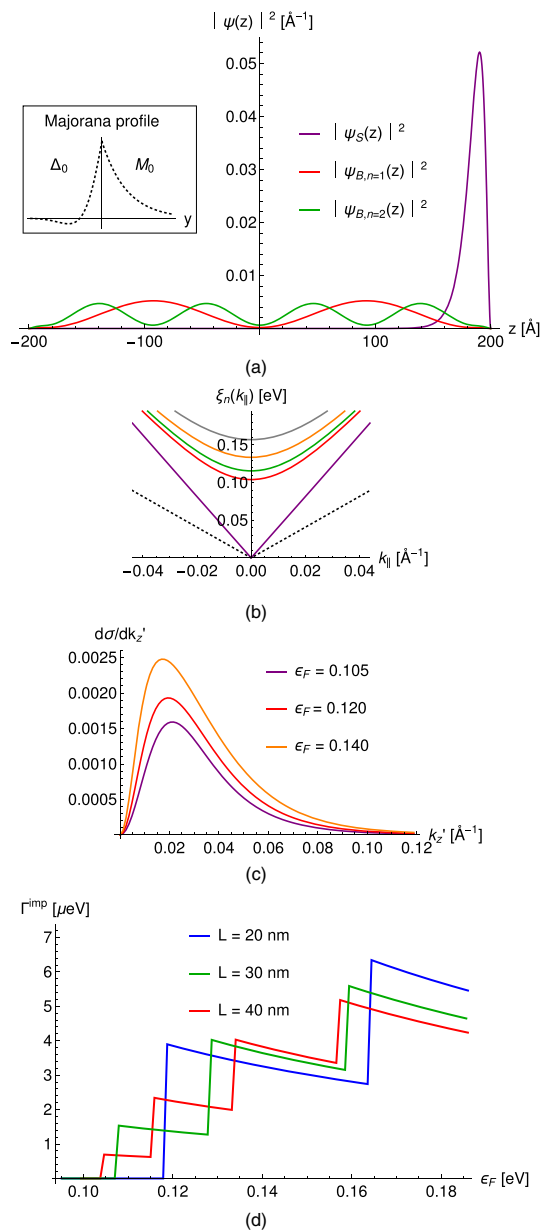
FIG. 4. (a) The three lowest-lying eigenstates in a TI film of thickness $L = 40$ nm. The surface state localized on the TI top is displayed in purple; the bottom surface state is not shown. Inset: A sketch of the typical transverse surface Majorana wave function. Its effect on the scattering rate is discussed in Appendix D 1. The model parameters are adjusted to those of Bi_2Se_3 except for the gap, which is here set to $\Delta_{TI} = 0.1$ eV to model puddles from the conduction band [44,45]. (b) The dispersion of the states displayed in (a) but with two additional bulk states. In black: The typical Majorana surface dispersion. (c) The differential scattering rate $d\sigma_{long}^{(n'=1)}(\epsilon_F)/dk_z'$ defined in Eq. (D5) shown for three values of the Fermi energy (measured in eV). (d) The electronic scattering rate as given in Eq. (23) with $n_{3D} = 10^{19}$ cm^{-3}. Van Hove singularities are seen whenever the Fermi energy hits the bulk bands, but they are softened by the $k_\parallel'$ factor that goes to zero at these points. Here we used $\epsilon_r = 100$ [46].

scattering rate $d\sigma_{long}^{(n')}(\epsilon_F)/dk_z'$ is the screened Coulomb coupling convoluted with the overlap between the bulk and surface states [cf. Eq. (D5)], see Fig. 4(c).

In most topological insulators the Fermi level resides in the bottom of the conduction band or the top of the valence band, and a significant doping of acceptors/donors is needed to reach a bulk insulating state [25,47]. For Bi_2Se_3 films (typically being n doped) we use the dopant concentration $n_{3D} = 10^{19}$ cm^{-3} [25,48,49]. Such doping causes fluctuations in the average (doping) concentration, making the conduction band bend and inducing electron and hole puddles [25,36], i.e., effectively filled pockets from the conduction band at the Fermi level. As a simple model of a typical situation we imagine that chemical doping of the TI bulk lowers the bulk Fermi level to lie in some region $0.10 \lesssim \epsilon_F \lesssim 0.18$ eV relative to the Dirac point. This would naively be bulk insulating since the Dirac point in Bi_2Se_3 is separated from the conduction band by a gap of 0.28 eV [50]. Puddles coming from stretching the bottom of the conduction band are simply modeled by setting the TI gap to $\Delta_{TI} = 0.1$ eV, making elastic scattering to the puddles (i.e., the lowered conduction bands) allowed at the Fermi level, see Fig. 4(b). This captures the generic features in unintentionally doped TIs.

In order to consider scattering from the surface state into the bulk, we need to know how close the surface Fermi energy is to the Dirac point. Note that, as we will see below, the scattering rate increases strongly if we are not very close (compared to Δ_0). Let us therefore assume that the surface Fermi energy is tuned near the Dirac point. The electronic scattering rates as a function of the bulk Fermi energy are shown in Fig. 4(d). We note that the rate stays similar in magnitude for a large TI film thickness range. When the film is made thinner, there are fewer bulk states to scatter into, but this is compensated by an increased surface-bulk overlap. Moreover, the rates in Fig. 4(d) have the same qualitative features as seen for point source scattering [44].

The (elastic) scattering lifetime of the surface electrons is $\tau_{SB}^{(e)} = 1/\Gamma_{SB}^{imp}$, and the scattering length is given by $l_S^{(e)} = \tau_{SB}^{(e)} v_F$. As a typical value of the rates in Fig. 4(d) we use $\Gamma_{SB}^{imp} = 3$ μeV, and we arrive at the unusually long scattering time $\tau_{SB}^{(e)} \approx 0.2$ ns ($l_S^{(e)} \approx 0.1$ mm). This is typically a factor of 10^2–10^3 longer than for bulk scattering lifetimes seen in experiment [49]. Yet, our results could be consistent with what was attributed to be unusually long surface lifetimes as observed after optically exciting bulk states [51,52]. In Ref. [51] one deduced that $\tau_{BS}^{(e)} > 10$ ps after seeing a stable population of the surface state induced by elastic bulk-surface scattering in Bi_2Se_3 (the samples were kept at $T = 70$ K) after subsequent inelastic decays associated with much shorter time scales. It would be desirable to see similar experiments engineered to measure the elastic surface-bulk scattering rate in doped compounds conducted at lower temperatures.

C. Velocity suppression and contradicting requirements

We now use the same expression in Eq. (23) to extract information about the Majorana lifetime and scattering rate. In doing so we ignore the (confined) transverse profile of the Majorana surface state. This is a valid approximation because the reciprocal coherence length is far less than the maximal transverse scattering momentum in the studied energy regime, see Appendix D 1. Moreover, we assume that the

induced superconductivity does not gap the TI bulk such that elastic surface-bulk scattering is not precluded. The Majorana scattering rate will typically be larger or equal to the electronic scattering rate due to a suppression of the Majorana velocity relative to the Fermi velocity.

The Majorana velocity is calculated from the low-energy chiral solution of the Hamiltonian in Eq. (C1): $v_m/v_F = \sqrt{1 - (\mu/M_0)^2}[1 + (\mu/\Delta_0)^2]^{-1}$ [18]. Here μ is the surface chemical potential relative to the Dirac point. From the $k_\parallel \propto 1/v_F$ dependency of $\Gamma_{\rm SB}^{\rm imp}$ in Eq. (23) we deduce that $l_S^{(m)} = \tau_{\rm SB}^{(m)} v_m$ is suppressed approximatively as $l_S^{(m)} \sim (v_m/v_F)^\alpha \, l_S^{(e)}$ with $\alpha \approx 4$ (up to some prefactor of order 1) [53]. Correspondingly, $\tau_{\rm SB}^{(m)} \sim (v_m/v_F)^{\alpha-1} \tau_{\rm SB}^{(e)}$. A nonzero size window for the interferometer size is required by imposing $l_S^{(m)} \gtrsim \xi$, which in turn means that

$$v_m/v_F \gtrsim \left(\xi/l_S^{(e)}\right)^{1/4}. \tag{24}$$

Inserting the electronic rate from the end of the last subsection and $\xi \simeq 4$ μm in this yields $v_m/v_F \gtrsim 0.4$. Hence, we must require a very fine tuning of μ, $|\mu| \lesssim 1.2\Delta_0 \simeq 0.1$ meV given that one can achieve $M_0 \gg \mu$.

When the TI bulk is increasingly doped with screened Coulomb charges to make the Fermi energy approach the Dirac point, fluctuations in the surface electrical potential energy are enlarged. The spatial dependence of the local density of states broadens the Fermi energy into a Gaussian distribution of finite width. The standard deviation of this energy smearing has been estimated in theory [25] and measured in experiment [36] to be $\delta\mu \simeq 10$–20 meV within some 0.1 eV from the average Dirac point ($\delta\mu$ decays inversely proportional to the average chemical potential further away). For Fermi energies close to the average Dirac point the notion of a uniform local density of states breaks down. Tuning of the surface chemical potential is consequently not globally possible within an energy resolution set by the distribution width [49]. Importantly, the smearing greatly exceeds the tuning required above $|\mu| \sim \delta\mu \gg \Delta_0$, even if we relax the assumption of a very small Δ_0 by increasing it one order of magnitude.

Finally, the spatial scale of the chemical potential fluctuations is $n_{\rm 3D}^{-1/3} \approx 5$ nm. Hence, a precise tuning of the chemical potential might be possible within local regions of this size. Still, this would not help in probing the experiment since $n_{\rm 3D}^{-1/3} \ll \xi$. Thus, unless these spatial fluctuations in the local density of states can be brought under control, unintentionally doped TIs are left unsuited for Majorana interferometry.

D. Other limitations

Majorana interactions [35] and coupling between Majorana modes and other degrees of freedom, such as phonons, are other potential sources of decoherence. Local interactions between chiral Majorana fermions are expected to be heavily suppressed at low temperatures and momenta with the leading order term going like $O(k^6)$ [18]. The chiral Majoranas on the interferometer arms can also excite phonons if an electron-phonon coupling is present; see Appendix E for a short discussion. For spatial inversion symmetric materials, e.g., Bi_2Se_3 and Bi_2Te_3, where this coupling is dictated by a deformation potential,

the decay rate of quasiparticles due to scattering on acoustic phonons at $T = 0$ exhibits a $\Gamma^{\rm ph} \sim (eV)^3$ behavior. This is *a posteriori* expected to be small compared to bulk leakage at low voltage. Without spatial inversion symmetry, a piezoelectric interaction can cause the electron-phonon coupling to follow a reciprocal power law in q, making scattering an increasing problem for small momenta.

V. CONCLUSIONS

Two limiting effects in Majorana interferometers are considered. We include a previously neglected coupling between chiral Majoranas and vortex-pinned modes. At low voltage this coupling yields a crossover in the conductance even-odd effect, distorting the interpretation of the experiment. We also find that surface-to-bulk scattering in the TI sets an upper bound on the size of the device. With a proximity gap of $\Delta_0 = 0.1$ meV, the lower bound (the coherence length) is about $\xi = \hbar v_F/\Delta_0 \simeq 4$ μm for Bi_2Se_3. Due to the doping needed to reach a bulk insulating state in many TIs, conduction band puddles lead to large fluctuations in the surface Fermi energy. In turn, this is in conflict with the surface chemical potential fine tuning required to have a nonzero size window for the interferometer. This leaves the possibility of probing the experiment in many poorly bulk insulating bismuth compounds potentially extremely difficult.

The most natural ways to overcome the restrictions considered here are (i) to find superconductors with excellent contact to the TI such that Δ_0 can be made (ideally orders of magnitude) larger, or (ii) to pursue a search for TIs with a highly insulating bulk. Most bismuth based TIs are poorly insulating and are unintentionally doped [20], which results in them being unsuited for Majorana interferometry. However, recently reported mixed bismuth compounds have shown evidence of an appreciably insulating bulk [54], which could potentially open a window of opportunity for this material. We note that the recommendation (i) above is similar to the need for high-quality interfaces in superconductor-semiconductor nanowire devices [55].

Another possible material system to consider is the putative topological Kondo insulator SmB_6, which gives strong sign of surface conduction [23,56]. The bulk resistivity of this material can reach several Ω cm at temperatures below a few Kelvins [23]. Very recently, evidence of the superconducting proximity effect in Nb/SmB_6 bilayers was reported [57], hence providing a possible platform for this experiment. However, the physics of highly interacting topological Kondo insulators is still poorly understood [22], and it is unclear how much of the details of the simple noninteracting TI surface physics will carry through to this more complicated case.

APPENDIX A: DERIVATION OF THE AVERAGE CURRENT

The derivation of Eqs. (3) and (4) goes along the lines of Ref. [58]. The scattering states of the interferometer are (chiral) plane waves in the one-dimensional channel convoluted with a transverse wave function (we let $k > 0$ by definition below) (Fig. 5),

$$\psi_L^e = e^{ik_L x} \varphi_L^e(y), \tag{A1}$$

$$\psi_L^h = e^{-ik_L x} \varphi_L^h(y), \tag{A2}$$

$$\psi_R^e = \mathcal{T}_{ee} e^{ik_R x} \varphi_R^e(y) + \mathcal{T}_{eh} e^{-ik_R x} \varphi_R^h(y), \tag{A3}$$

$$\psi_R^h = \mathcal{T}_{hh} e^{-ik_R x} \varphi_R^h(y) + \mathcal{T}_{he} e^{ik_R x} \varphi_R^e(y). \tag{A4}$$

We have assumed that the contacts $\alpha \in \{L, R\}$ contain only single mode states and that the incident state in the L contact is either an electron or a hole.

One can construct arbitrary states by expanding in the scattering states above. This is incorporated in the (second quantized) field operator

$$\hat{\Psi}_{\alpha,\sigma}(\boldsymbol{r}, t) = \frac{1}{\sqrt{2\pi}} \int \frac{dE_\alpha}{\sqrt{\hbar v_\alpha}} \, \psi_\alpha^\sigma(E_\alpha, \boldsymbol{r}) \hat{a}_{\alpha,\sigma}(E_\alpha) e^{-i\omega_\alpha t}, \tag{A5}$$

where we introduced $\omega_\alpha = E_\alpha/\hbar = v_\alpha k_\alpha$ and the annihilation operator $\hat{a}_{\alpha,\sigma}(E_\alpha)$ of type $\sigma \in \{e, h\}$ satisfying

$$\{\hat{a}_{\alpha,\sigma}^\dagger(E), \hat{a}_{\beta,\sigma'}(E')\} = \delta_{\alpha,\beta} \delta_{\sigma,\sigma'} \delta(E - E'). \tag{A6}$$

The current operator of type σ in contact α is defined by

$$\hat{I}_{\alpha,\sigma} = \frac{\hbar e}{2im} \int d\boldsymbol{r}_{\perp,\alpha} [\hat{\Psi}_{\alpha,\sigma}^\dagger \partial_x \hat{\Psi}_{\alpha,\sigma} - (\partial_x \hat{\Psi}_{\alpha,\sigma}^\dagger) \hat{\Psi}_{\alpha,\sigma}]. \tag{A7}$$

Here m is the effective mass, $mv_\alpha = \hbar k_\alpha$. Assuming further that the contacts act as thermal reservoirs kept at equal temperature, we can average the density operator by

$$\langle \hat{a}_{\alpha,\sigma}^\dagger(E) \hat{a}_{\alpha',\sigma'}(E') \rangle = \delta_{\alpha,\alpha'} \delta_{\sigma,\sigma'} \delta(E - E') f_\sigma(E), \tag{A8}$$

where $f_{e/h}(E) = [\exp(\beta[E \mp \mu]) + 1]^{-1}$ is the Fermi function for particles and holes. Finally, we assume the transverse wave functions to be orthonormal,

$$\int dy \, (\varphi_\alpha^\sigma(y))^* \varphi_{\alpha'}^{\sigma'}(y) = \delta_{\alpha,\alpha'} \delta_{\sigma,\sigma'}. \tag{A9}$$

The source and the drain current are defined by the average particle minus hole current $I_D \equiv \langle \hat{I}_{D,e} - \hat{I}_{D,h} \rangle$ and $I_S \equiv \langle \hat{I}_{S,e} - \hat{I}_{S,h} \rangle$. Calculating these explicitly by using Eqs. (A7), (A8), (A9), and unitarity of $\mathcal{T}$ leads exactly to the expressions in Eqs. (3) and (4).
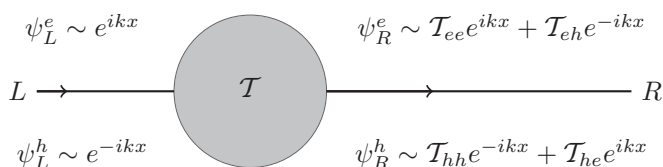

$\psi_L^e \sim e^{ikx}$ $\psi_R^e \sim \mathcal{T}_{ee} e^{ikx} + \mathcal{T}_{eh} e^{-ikx}$

$L \rightarrow$ $\mathcal{T}$ $\rightarrow$ R

$\psi_L^h \sim e^{-ikx}$ $\psi_R^h \sim \mathcal{T}_{hh} e^{-ikx} + \mathcal{T}_{he} e^{ikx}$

FIG. 5. The scattering states at the two contact points.

APPENDIX B: CHARGE TRANSPORT WITH MAJORANA COUPLING DISORDER

1. One Majorana with smeared coupling to the edge

Consider the case where the edge Majorana is coupled continuously to the bound state as in Eq. (14). Using the ansatz $\xi_1(x) = f(x) e^{i(kx - \omega t)}$ with $f(-\infty)$ being the Majorana (fermion) field on the far left and $\omega = k/v_m$ for the linearly dispersing modes. The equations of motion can be combined to give the continuum version of Eq. (12):

$$-iv_m \omega \partial_x f(x) = \lambda(x) e^{-ikx} \int dx' \, \lambda(x') f(x') e^{ikx'}. \tag{B1}$$

Here the plane wave phase contribution that can be neglected in the discrete case where $\lambda(x) = \lambda \delta(x)$ is included. The equation above can be solved as follows. Integrating both sides gives the solution implicitly by

$$f(x) = f(-\infty) - \frac{\zeta}{i\omega v_m} \int^x dx' \, \lambda(x') e^{-ikx'}, \tag{B2}$$

with $\zeta = \int dx \lambda(x) f(x) e^{ikx}$. Inserting this back into (B1) yields

$$\zeta = \frac{f(-\infty) \int dx \, \lambda(x) e^{ikx}}{1 + \frac{1}{i\omega v_m} \int dx \, \lambda(x) e^{ikx} \int^x dx' \lambda(x') e^{-ikx'}}. \tag{B3}$$

Finally, this expression is used in (B2), and we obtain

$$f(+\infty) = \frac{\omega + \frac{i}{\hbar v_m} \int dx \lambda(x) e^{ikx} \int_x dx' \lambda(x') e^{-ikx'}}{\omega - \frac{i}{\hbar v_m} \int dx \lambda(x) e^{ikx} \int^x dx' \lambda(x') e^{-ikx'}} f(-\infty). \tag{B4}$$

Comparing this to Eq. (12) proves the statement in Eq. (15).

2. Multiple Majoranas coupled to each edge

In the main text we obtained a phase contribution of $\phi = \arctan(\frac{v}{eV})$ in the differential conductance when the edge was coupled to one vortex [Eq. (13)]. This can be generalized if the chiral Majoranas have single-point couplings to several vortices. In that case ϕ is replaced by $\sum_i \phi_i - \sum_j \phi_j$, where $\phi_i = \arctan(\frac{v_i}{eV})$ comes from vortex-edge coupling with the upper edge and ϕ_j from coupling with the lower edge. Hence, multiple phase crossovers will occur if several vortices are located close to the edge.

3. One Majorana coupled to both edges

Another generalization is to study point couplings to both the lower and the upper arm, in which case the coupling term in the Lagrangian is $\mathcal{L}_{\text{bulk-edge}} = 2i\lambda_1 \xi_1(x = 0)\xi_0 + 2i\lambda_2 \xi_2(x = 0)\xi_0$. The corresponding equations of motion are

$$2\partial_t \xi_0 = \lambda_1 [\xi_{1R} + \xi_{1L}] + \lambda_2 [\xi_{2R} + \xi_{2L}], \tag{B5}$$

$$v_m \xi_{1R} = v_m \xi_{1L} + \lambda_1 \xi_0, \tag{B6}$$

$$v_m \xi_{2R} = v_m \xi_{2L} + \lambda_2 \xi_0. \tag{B7}$$

Here we again use the notation $\xi_{1R} = \xi_1(x = 0^+)$, $\xi_{2R} = \xi_2(x = 0^+)$, $\xi_{1L} = \xi_1(x = 0^-)$, and $\xi_{2L} = \xi_2(x = 0^-)$. In frequency space we find a relation between ξ_1 and ξ_2 across the

coupling points $[\xi_1,\xi_2]_R^T = U(\omega)[\xi_1,\xi_2]_L^T$, where $U(\omega)$ is the unitary matrix

$$U(\omega) = \frac{1}{W(\omega)}\begin{pmatrix} -\nu_1 + \nu_2 + i\omega & -2\sqrt{\nu_1\nu_2} \\ -2\sqrt{\nu_1\nu_2} & \nu_1 - \nu_2 + i\omega \end{pmatrix}. \quad (B8)$$

Above, $\nu_i \equiv \lambda_i^2/(2\hbar v_m)$ and $W(\omega) = \nu_1 + \nu_2 + i\omega$. Notice how $U(\omega)$ is off-diagonal in the low-energy limit for the configuration $\nu_1 = \nu_2$. The chiral Majoranas on the perimeter therefore switch place by tunneling across the vortex Majorana in this case. The phase matrix is given by

$$\mathcal{P} = \begin{pmatrix} e^{i\frac{kl_1}{2}} & 0 \\ 0 & e^{i\frac{kl_2}{2}} \end{pmatrix} U(\omega) \begin{pmatrix} e^{i\frac{kl_1}{2}+in\pi} & 0 \\ 0 & e^{i\frac{kl_2}{2}} \end{pmatrix}, \quad (B9)$$

which leads to the differential conductance

$$G_{SC}(V) = \frac{2e^2}{h}\frac{1}{(eV)^2 + (\nu_1 + \nu_2)^2}\Bigg\{(\nu_1 - \nu_2)^2$$
$$+ [(eV)^2 - (\nu_1 - \nu_2)^2]\sin^2\left(\frac{n\pi}{2} + \frac{eV\delta L}{2\hbar v_m}\right)$$
$$+ eV(\nu_1 - \nu_2)\sin\left(n\pi + \frac{eV\delta L}{\hbar v_m}\right)$$
$$+ 2\nu_1\nu_2[1 + (-1)^n]\Bigg\}. \quad (B10)$$

This reduces to Eq. (13) when $\nu_2 = 0$. For $\delta L = 0$ the expression above is identical to Eq. (13) with the replacement $\nu \to \nu_1 + \nu_2$.

4. One Majorana coupled to one edge at finite temperature

We return to the case with a one-point Majorana coupling [Fig. 2(a)]. The drain current can generally be reexpressed as a residue sum over poles z_j^+ in the upper complex half-plane. Rewriting Eq. (3),

$$I_D(\nu) = (-1)^n \frac{e}{4} i \sinh(\beta eV) \sum_j \text{Res}\{h(z), z_j^+\}, \quad (B11)$$

where $h(z)$ is the function

$$h(z) = \frac{1}{z^2 + \nu^2}\frac{(z^2 - \nu^2)\cos\left(\frac{\delta Lz}{v_m}\right) - 2z\nu\sin\left(\frac{\delta Lz}{v_m}\right)}{\cosh\left[\frac{\beta}{2}(z - eV)\right]\cosh\left[\frac{\beta}{2}(z + eV)\right]}. \quad (B12)$$

The set of simple poles of $h(z)$ in the upper half-plane is $z_j^+ \in \{i\nu\} \cup \{i\pi(2m - 1)/\beta + eV\}_{m=1}^{\infty}$. For $\nu = 0$, the sum is obtainable in closed form and stated in Eq. (8).

For weak coupling at finite temperature $\nu \ll k_BT, eV$, we expand $h(z)$ in powers of $\beta\nu$, with $\beta^{-1} = k_BT$. For convenience we define the reduced current as $R \equiv I_D(\nu)/I_{D,0}$. To second order in $\beta\nu$ with $\delta L = 0$ we find

$$R = 1 - \frac{\nu\pi}{eV}\tanh(\beta eV/2)$$
$$+ \frac{\beta\nu^2}{\pi eV}\text{Im}\left\{\psi^{(1)}\left(\frac{1}{2} - i\frac{\beta\nu}{2\pi}\right)\right\} + O(\beta\nu)^3. \quad (B13)$$

Above, $\psi^{(1)}(z) = \frac{d^2}{dz^2}\log\Gamma(z)$ is the trigamma function. In the infinite coupling limit the sign of the drain current is flipped, $R \to -1$, which can be seen from Eq. (B12). The reduced current decreases monotonically, a feature of setting
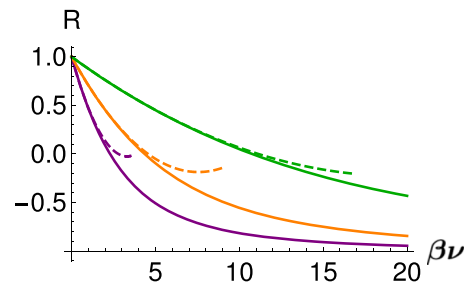

FIG. 6. The reduced drain current R as function of $\beta\nu$ for symmetric arms $\delta L = 0$. The dotted lines represent the weak coupling result from Eq. (B13) and the full lines are numerical results for βeV being 5 (purple), 10 (orange), and 25 (green).

$\delta L = 0$, with coupling strength until the vortex Majorana is fully absorbed by the edge (Fig. 6).

At $T = 0$ the current is found to be (with $\hbar = 1$ here)

$$(-1)^n \frac{2\pi}{e}I_D = \frac{v_m}{\delta L}\sin\left(\frac{\delta LeV}{v_m}\right)$$
$$+ 2\nu\left[\cosh\left(\frac{\delta LeV}{v_m}\right) - \sinh\left(\frac{\delta LeV}{v_m}\right)\right]$$
$$\times \left[\text{Im}\left\{\text{Ci}\left(\frac{\delta L}{v_m}(eV + i\nu)\right)\right\}\right.$$
$$+ \text{Im}\left\{\text{Ci}\left(-i\nu\frac{\delta L}{v_m}\right)\right\}$$
$$\left. - \text{Re}\left\{\text{Si}\left(\frac{\delta L}{v_m}(eV + i\nu)\right)\right\}\right]. \quad (B14)$$

Here Ci and Si are trigonometric integral functions. For symmetric arms the above result simplifies to give the reduced current $R = 1 + 2\frac{\nu}{eV}[\arctan\left(\frac{\nu}{eV}\right) - \frac{\pi}{2}]$. From this, the aforementioned result $\lim_{\nu\to\infty} R = -1$ follows trivially.

APPENDIX C: MAJORANA FERMIONS IN A TI/SC HYBRID STRUCTURE

Superconductivity induced on the surface of a strong TI support chiral Majorana fermions on domain walls and Majorana bound states localized in vortices. The Fu and Kane Hamiltonian of a TI/SC hybrid structure with a single Dirac-like dispersion is given by [7,18]

$$\mathcal{H} = (v_F\boldsymbol{\sigma}\cdot\boldsymbol{p} - \mu)\tau_z + M(\boldsymbol{r})\sigma_z + \Delta(\boldsymbol{r})\tau_+ + \Delta^*(\boldsymbol{r})\tau_-, \quad (C1)$$

where τ_j and σ_j are Pauli matrices acting in particle-hole and spin space, respectively. Moreover, $\tau_\pm = (\tau_x \pm i\tau_y)/2$ and M is the Zeeman energy associated with a magnetic field applied in the z direction. The standard procedure is to introduce $u_\sigma(\boldsymbol{r})$ and $v_\sigma(\boldsymbol{r})$ that define the quasiparticles, conveniently arranged in $\boldsymbol{\Psi} = [u_\uparrow, u_\downarrow, v_\uparrow, -v_\downarrow]^T$ satisfying the BdG equations $\mathcal{H}\boldsymbol{\Psi} = E\boldsymbol{\Psi}$.

1. Corrections to the S matrix

One may solve Eq. (C1) on a magnetic domain wall in the absence of a superconductor. There are then *two* chiral solutions: one in the particle sector $|\tau_z = +1\rangle$, and one in the

hole sector $|\tau_z = -1\rangle$, both localized at the interface [18]. On a magnet-superconductor interface there exists a single chiral Majorana solution with linear dispersion. By the definition of the S matrix,

$$\begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} = \begin{pmatrix} S_{ee} & S_{eh} \\ S_{he} & S_{hh} \end{pmatrix} \begin{pmatrix} \psi_e \\ \psi_h \end{pmatrix}, \quad \text{(C2)}$$

one can estimate the scattering elements at the trijunction as overlaps between the chiral states $S_{ee} = \langle \tau_z = +1|\xi_1\rangle$, $S_{eh} = \langle \tau_z = -1|\xi_1\rangle$, etc. Here ξ_1 (ξ_2) is the chiral Majorana fermion on the upper (lower) interferometer arm. The Majorana wave function decays exponentially with the length scale $v_F/\sqrt{M^2 - \mu^2}$ (v_F/Δ_0) in the magnetic (superconducting) region [inset of Fig. 4(a)]. Interestingly, if the magnetization on each side of the Dirac channel are of equal magnitude (with opposite polarization), compactly expressed in terms of the symmetry $\mathcal{H}(-y) = \mathcal{M}^{-1}\mathcal{H}(y)\mathcal{M}$ where $\mathcal{M} = i\sigma_y$ [18], the zero energy result in Eq. (2) is obtained at all energies. Experimentally, this is likely a weak assumption, and we can therefore safely apply $S(E = 0)$ throughout. If the two magnetic fields in the Dirac region are of different strengths there will be corrections to S that are small when $E \ll \frac{v_m}{v_F} \min\{M_0, \Delta_0\}$. One can heuristically add a correction term to $S(E)$ proportional to E/Λ (where $\Lambda \sim \frac{v_m}{v_F} \min\{M_0, \Delta_0\}$ is some phenomenological scale) and impose unitarity to order $O(E^2)$ and particle-hole symmetry. This leads to order $O(E^2)$ corrections to the drain current.

2. Majorana bound states and energy splitting

If a vortex is present in the system described by Eq. (C1), $\Delta(r) = \Delta(r)e^{i\ell\theta}$, a single Majorana bound state will be localized in the vortex core when the vorticity ℓ is odd [33,34]. Following the procedure in Ref. [59], the zero energy BdG equations for the model in Eq. (C1) with a central vortex of vorticity $\ell = 1$ are expressed in terms of two real and coupled radial equations,

$$\begin{pmatrix} -\mu & \eta\Delta(r) + v_F\left(\partial_r + \frac{1}{r}\right) \\ -\eta\Delta(r) - v_F\partial_r & -\mu \end{pmatrix} \begin{pmatrix} u_\uparrow(r) \\ u_\downarrow(r) \end{pmatrix} = 0. \quad \text{(C3)}$$

Here $v_\sigma(r) = \eta u_\sigma(r)$ with $\eta = \pm 1$ dictating two possible solution channels. Modeling the vortex by a hard step of size $R_1 \approx \xi$, $\Delta(r) = \Delta_0\Theta(r - R_1)$, the resulting zero mode is expressed in terms of Bessel functions,

$$(u_\uparrow(r), u_\downarrow(r))^T$$
$$= \mathcal{N}\begin{pmatrix} J_0(\frac{\mu r}{v_F}) \\ J_1(\frac{\mu r}{v_F}) \end{pmatrix} [\Theta(R_1 - r) + e^{-\frac{\Delta_0}{v_F}(r-R_1)}\Theta(r - R_1)]. \quad \text{(C4)}$$

Above, $\mathcal{N}$ is given by normalization of the two-component spinor. Two vortex bound Majoranas separated by a distance R will generally lift from zero energy by an energy splitting exponentially small in R/ξ when $R \gg \xi$. If μ is tuned close to the Dirac point (typically achieved by bulk doping of screened charges [36]), this splitting has previously been estimated to asymptotically approach $\varepsilon_+ \sim -\mu(R/\xi)^{3/2}\exp(-R/\xi)$ up to some prefactor of order one [34].

Similarly, a superconducting disk of radius R (with a central vortex) deposited on a TI supports a Majorana edge state located exponentially close to the edge. For weak magnetic fields the splitting between the central bound state and the edge state has been estimated to decay like [33]

$$\varepsilon_+ \sim -\mu l_B e^{-\frac{R}{\xi}}/\xi, \quad \text{(C5)}$$

with l_B the magnetic length.

3. Toy model calculation: Energy splitting in a spinless p-wave superconductor

For completeness, and serving as an illuminating example not found elsewhere, we explain how the energy splitting between edge states in a Corbino geometry can be calculated in the spinless $p + ip$ superconductor. See, e.g., Ref. [1] for general aspects of this system and Ref. [60] for a presentation of the similar intervortex splitting. Finally, the problem of the Majorana energy hybridization in superconductor-semiconductor hybrid structures is addressed in Ref. [61].

The spinless p-wave superconductor realizes a topological phase for $\mu > 0$ and the trivial phase for $\mu < 0$. The model has the corresponding zero energy BdG equations [42]

$$\begin{pmatrix} -\frac{1}{2m}\nabla^2 - \mu & \frac{1}{2p_F}\{\Delta(r), \partial_{z^*}\} \\ -\frac{1}{2p_F}\{\Delta^*(r), \partial_z\} & \frac{1}{2m}\nabla^2 + \mu \end{pmatrix} \begin{pmatrix} u(r) \\ v(r) \end{pmatrix} = 0, \quad \text{(C6)}$$

with the operator $\partial_{z^*} = e^{i\theta}(\partial_r + \frac{i}{r}\partial_\theta)$. We let a radial annulus geometry between R_1 and R_2 be superconducting with a vortex located in the center hole $\Delta(r) = \Delta_0 e^{i\theta}\Theta(r - R_1)\Theta(R_2 - r)$, with chemical potential $2mv_F^2\mu > \Delta_0^2$ (causing the wave functions to oscillate) and $\mu < 0$ outside the disk. When $R \equiv R_2 - R_1 \gg \xi$, we treat the two single-edged systems separately and construct the ground state candidates $\phi_\pm = \frac{1}{\sqrt{2}}(\psi_1 \pm i\psi_2)$, where ψ_1 (ψ_2) is the solution localized on the inner (outer) edge [60]. The two solutions are exponentially damped away from their respective edges and they oscillate with frequency $k = \sqrt{2m\mu - (\Delta_0/v_F)^2}$. Moreover, if both R_1, $R_2 \gg 1/k$, the Bessel functions that appear as radial solutions can be expanded asymptotically. The splitting integral can be calculated, and we send $\mu \to -\infty$ outside the annulus in the end. Upon evaluating the splitting integrals and then taking the limit, we obtain

$$\varepsilon_\pm = \langle \phi_\pm | \mathcal{H} | \phi_\pm \rangle \approx \mp \frac{4\Delta_0\mu}{v_F k}\sin(kR)e^{-R/\xi}. \quad \text{(C7)}$$

The splitting is zero for particular values of the domain separation. Whenever $R = \pi n/k$ for $n \in \mathbb{N}$, there are two degenerate ground states. The zero mode condition gives associations of an interference phenomenon, caused by the oscillating edge modes that convolute in a destructive manner for certain values of the separation. The expression in Eq. (C7) agrees with numerical diagonalization, already when R is a small multiple of ξ. The Corbino geometry has been studied for small disk size, in which case the same zero energy criterion is found in the limit $2mv_F^2\mu \gg \Delta_0^2$ by imposing Dirichlet boundary conditions on the wave functions directly [62].

APPENDIX D: SURFACE-BULK SCATTERING WITH SCREENED DISORDER

In Ref. [44] Fermi's golden rule is used to find the scattering rate from the surface (S) to the bulk (B) of a TI in the presence of static and dilute point impurities. Here we take Ref. [44] as a starting point (the reader is referred to this reference for further details) and apply the formalism to the case of long range scatterers. As described in the main text we study scattering from surface initial wave vector $\boldsymbol{k} = (\boldsymbol{k}_\parallel, 0)$ [in practice we let $\boldsymbol{k}_\parallel = (k_x, 0)$] to the final bulk wave vector $\boldsymbol{k}' = (\boldsymbol{k}'_\parallel, k'_z)$. The energy of the incoming surface state determines the Fermi energy, $\xi_{S,\boldsymbol{k}_\parallel} = v_F k_\parallel \equiv \epsilon_F$. Assuming low energy elastic scattering, and making use of the continuum limit $\sum_{\boldsymbol{k}'} \to \frac{V}{(2\pi)^3} \int d^3\boldsymbol{k}'$, we find the scattering rate

$$\Gamma_{\mathrm{SB}}^{\mathrm{imp}}(\epsilon_F) = 2\pi \sum_{\boldsymbol{k}',n'} \left| g_{\boldsymbol{k}-\boldsymbol{k}'}^{\mathrm{imp}} \right|^2 (|F_{S,\boldsymbol{k}_\parallel;B,\boldsymbol{k}',1n'}|^2 + |F_{S,\boldsymbol{k}_\parallel;B,\boldsymbol{k}',2n'}|^2) \delta(\xi_{B,n',k'_\parallel} - \epsilon_F)$$

$$\approx \frac{V}{(2\pi)^2} \sum_{\substack{n': \\ \min\{\xi_{B,n',k'_\parallel}\} < \epsilon_F}} \frac{k'_\parallel}{|\nabla \xi_{B,n',k'_\parallel}|} \int_0^\infty dk'_z \int_0^{2\pi} d\varphi \left| g_{\boldsymbol{k}_\parallel - k'_\parallel, k'_z, \varphi}^{\mathrm{imp}} \right|^2 (|F_{S,\boldsymbol{k}_\parallel;B,\boldsymbol{k}',1n'}|^2 + |F_{S,\boldsymbol{k}_\parallel;B,\boldsymbol{k}',2n'}|^2). \quad \text{(D1)}$$

In going from the first to the second line we integrated over $k'_\parallel$, which in the last line is then defined implicitly by $\xi_{B,n',k'_\parallel} = \epsilon_F$. Above, φ is the polar angle between the incoming and the outgoing momentum projected onto the $k_x k_y$ plane, i.e., $k'_x = k'_\parallel \cos\varphi$ and $k'_y = k'_\parallel \sin\varphi$. The F's are defined as convoluted overlaps between the surface and the bulk states,

$$F_{S,\boldsymbol{k}_\parallel;B,\boldsymbol{k}',1n'} = \langle \Psi_{S,\boldsymbol{k}_\parallel} | e^{ik'_z z} | \Psi_{B,\boldsymbol{k}',1n'} \rangle, \quad \text{(D2)}$$

$$F_{S,\boldsymbol{k}_\parallel;B,\boldsymbol{k}',2n'} = \langle \Psi_{S,\boldsymbol{k}_\parallel} | e^{ik'_z z} | \Psi_{B,\boldsymbol{k}',2n'} \rangle, \quad \text{(D3)}$$

where two bulk bands $1n'$ and $2n'$ are degenerate. The four-component wave functions in these expression are found by solving the BdG equations for the 3D TI exactly at $k_\parallel = 0$ and then applying perturbation theory to leading order in the wave vector (see Appendix C in Ref. [44] where the full expressions are listed in Eqs. (C1)–(C11)). This procedure also yields the dispersion relations to leading order in the parallel wave vector. In Figs. 4(a) and 4(b) the surface and some of the bulk states are visualized in a thin film with model parameters as for Bi$_2$Se$_3$ [45], except for the bulk gap which is set to $\Delta_{\mathrm{TI}} = 0.1$ eV as motivated in the main text.

The coupling $g_{\boldsymbol{k}-\boldsymbol{k}'}^{\mathrm{imp}}$ is here an ensemble averaged Coulomb potential due to screened charge impurities. Assuming randomly distributed impurities with zero mean, the coupling should be proportional to $|g_{\boldsymbol{k}-\boldsymbol{k}'}^{\mathrm{imp}}|^2 \propto n_{\mathrm{3D}}[Ze^2/(\epsilon_0\epsilon_r)]^2(|\boldsymbol{k}-\boldsymbol{k}'|^2 + k_{\mathrm{TF}}^2)^{-2}$ [63], where ϵ_r is the relative permittivity, k_{TF} is the Thomas-Fermi wave vector, and n_{3D} is the average dopant concentration. In cylindrical coordinates the coupling is expressed as

$$\left| g_{\boldsymbol{k}_\parallel - k'_\parallel, k'_z, \varphi}^{\mathrm{imp}} \right|^2 = \frac{n_{\mathrm{3D}}}{V} \left(\frac{Ze^2}{\epsilon_0 \epsilon_r k_\parallel^2} \right)^2 \left[(1 - k'_\parallel/k_\parallel)^2 + (k'_z/k_\parallel)^2 + r_s^2 \right. $$
$$\left. + 4\frac{k'_\parallel}{k_\parallel} \sin^2 \frac{\varphi}{2} \right]^{-2}, \quad \text{(D4)}$$

where $r_s = k_{\mathrm{TF}}/k_\parallel$ was introduced. In the main text we assume that the screened charges have $Z = 1$. With this coupling established we define the differential cross section for the screened Coulomb scattering as

$$\frac{d\sigma_{\mathrm{long}}^{(n')}(\epsilon_F)}{dk'_z}$$
$$\equiv \int_0^{2\pi} d\varphi \frac{|F_{S,\boldsymbol{k}_\parallel;B,\boldsymbol{k}',1n'}|^2 + |F_{S,\boldsymbol{k}_\parallel;B,\boldsymbol{k}',2n'}|^2}{\left[\left(1 - \frac{k'_\parallel}{k_\parallel}\right)^2 + \left(\frac{k'_z}{k_\parallel}\right)^2 + r_s^2 + 4\frac{k'_\parallel}{k_\parallel} \sin^2 \frac{\varphi}{2} \right]^2}. \quad \text{(D5)}$$

This function is shown in Fig. 4(c) for $r_s = 0.035$, which corresponds to $\epsilon_r = 100$ [46]. Note that in the case of point scatterers, the differential cross section is obtained simply by replacing the denominator in Eq. (D5) by 1.

Using Eqs. (D4) and (D5) in (D1) assembles the expression for the scattering rate in the main text, Eq. (23). We note that the rates as shown in Fig. 4(d) are seen to be in good agreement with Ref. [44].

1. Effect of spatially confined transverse wave function

When applying the formula of the electronic scattering rate in Eq. (23) to the case of the surface Majorana we neglect the transverse profile of the surface wave function [inset of Fig. 4(a)]. Here we argue why this is a valid approximation in our parameter regime.

Assume for simplicity that the two surface gaps are of similar magnitude $\Delta_0 \approx M_0 \gg \mu$, so that the transverse wave function is confined over a the scale $\sigma_y \sim \xi = v_F/\Delta_0$. By the uncertainty principle, this confinement leads to an uncertainty in k'_y of $\sigma_{k'_y} \sim \xi^{-1}$. To simulate this uncertainty we draw random additions to $k'_y = k'_\parallel \sin\varphi$ from a normal distribution with the standard deviation above and zero mean for each scattering direction φ. The resulting scattering rates display an average absolute deviation $\langle |\varepsilon| \rangle$ [averaged over the Fermi energy range in Fig. 4(d)] from the sharp k'_y curve of roughly $\langle |\varepsilon| \rangle \approx (\sigma_{k'_y}/\max\{k'_y\})^2$. Since $\max\{k'_y\} \gg \xi^{-1}$ in the energy range and with the proximity gaps we consider, the effect of uncertainty in k'_y, and hence confinement in the y direction, can be ignored.

APPENDIX E: SCATTERING WITH ACOUSTIC PHONONS

Consider the toy model coupling electrons to (acoustic) phonons through a deformation potential,

$$H_{\text{ep}} = \sum_{\boldsymbol{q}_1, \boldsymbol{q}_2, s, s'} M_{\boldsymbol{q}_1, \boldsymbol{q}_2}^{s, s'} c_{\boldsymbol{q}_1 + \boldsymbol{q}_2, s}^{\dagger} c_{\boldsymbol{q}_1, s'} \left(a_{\boldsymbol{q}_2} + a_{-\boldsymbol{q}_2}^{\dagger} \right). \quad \text{(E1)}$$

For small momenta, the electron-phonon coupling goes as $|M(q)| \sim q$ for surface phonons [64,65]. Acoustic phonons have linear dispersion $\omega(q) = c_R |\boldsymbol{q}|$ and are expected to dominate the coupling Hamiltonian at low temperatures [64]. The scattering rate follows once again from Fermi's golden rule $\Gamma_{i \to f}^{\text{ph}} = 2\pi \nu_f(E) |\langle f | H_{\text{ep}} | i \rangle|^2$, where $\nu_f(E)$ is the final density of states. Recall also that states with a linear dispersion in two dimensions have $\nu(E) \propto E$. In a superconducting system, there will be a nonzero amplitude for the creation of a phonon with the cost of annihilating two quasiparticles. For illustrative purposes, we consider quasiparticles excitations of the (s-wave) Bardeen-Cooper-Schrieffer (BCS) ground state $|\Omega\rangle$,

$$|i; \sigma_1, \sigma_2\rangle = \gamma_{\boldsymbol{k}_1, \sigma_1}^{\dagger} \gamma_{\boldsymbol{k}_2, \sigma_2}^{\dagger} |\Omega\rangle, \quad \text{(E2)}$$

$$|f\rangle = a_{\boldsymbol{q}}^{\dagger} |\Omega\rangle, \quad \text{(E3)}$$

$$|\Omega\rangle = \prod_{\boldsymbol{k}} (u_{\boldsymbol{k}} + v_{\boldsymbol{k}} c_{\boldsymbol{k}, \uparrow}^{\dagger} c_{-\boldsymbol{k}, \downarrow}^{\dagger}) |0\rangle. \quad \text{(E4)}$$

Above, the quasiparticle creation operators are $\gamma_{\boldsymbol{k}, +}^{\dagger} = u_{\boldsymbol{k}}^{*} c_{\boldsymbol{k}, \uparrow}^{\dagger} - v_{\boldsymbol{k}}^{*} c_{-\boldsymbol{k}, \downarrow}$ and $\gamma_{\boldsymbol{k}, -}^{\dagger} = u_{\boldsymbol{k}}^{*} c_{-\boldsymbol{k}, \downarrow}^{\dagger} + v_{\boldsymbol{k}}^{*} c_{\boldsymbol{k}, \uparrow}$. The scattering element is found to be

$$
\begin{aligned}
&\langle f | H_{\text{ep}} | i; \sigma_1, \sigma_2 \rangle \\
&= \left(M_{\sigma_1 \boldsymbol{k}_1, -\sigma_1 \boldsymbol{k}_1 - \boldsymbol{k}_2}^{\uparrow, s(\sigma_1)} \delta_{\boldsymbol{q}, \sigma_1 \boldsymbol{k}_1 + \boldsymbol{k}_2} - M_{\sigma_1 \boldsymbol{k}_1, -\sigma_1 \boldsymbol{k}_1 + \boldsymbol{k}_2}^{\downarrow, s(\sigma_1)} \delta_{\boldsymbol{q}, \sigma_1 \boldsymbol{k}_1 - \boldsymbol{k}_2} \right) \\
&\quad \times v_{-\boldsymbol{k}_2}^{*} u_{\boldsymbol{k}_1}^{*} \\
&\quad - \left(M_{\sigma_2 \boldsymbol{k}_2, -\sigma_2 \boldsymbol{k}_2 - \boldsymbol{k}_1}^{\uparrow, s(\sigma_2)} \delta_{\boldsymbol{q}, \sigma_2 \boldsymbol{k}_2 + \boldsymbol{k}_1} - M_{\sigma_2 \boldsymbol{k}_2, -\sigma_2 \boldsymbol{k}_2 + \boldsymbol{k}_1}^{\downarrow, s(\sigma_2)} \delta_{\boldsymbol{q}, \sigma_2 \boldsymbol{k}_2 - \boldsymbol{k}_1} \right) \\
&\quad \times v_{-\boldsymbol{k}_1}^{*} u_{\boldsymbol{k}_2}^{*}. \quad \text{(E5)}
\end{aligned}
$$

Above we introduced the symbol $s(+) = \uparrow$ and $s(-) = \downarrow$. The scattering amplitude depends only on the momentum transfer for small energies, in which case the coupling above vanishes identically. This is presumably because the average charge of the quasiparticles becomes zero in the low energy limit.

A careful analysis for the electron decay rate alone yields a $\Gamma^{\text{ph}} \sim T^3$ law well below the Bloch-Grüneisen temperature at the Fermi surface [64]. Away from the Fermi surface, by biasing the electrons at a small voltage V, the decay rate is finite at $T = 0$ and goes as $\Gamma^{\text{ph}} \sim (eV)^3$. Inserting the exact prefactor (for $Bi_2 Te_3$), by following the steps in Ref. [64], we find a decay rate in the peV range when $V \simeq 1\,\mu\text{V}$. This means that the electron lifetime due to acoustic phonon scattering $\tau_{\text{ph}} = 1/\Gamma^{\text{ph}}$ is in the microsecond range, making this effect negligible at $T = 0$ at low voltage.

[1] J. Alicea, Rep. Prog. Phys. **75**, 076501 (2012).

[2] N. Read and D. Green, Phys. Rev. B **61**, 10267 (2000).

[3] N. B. Kopnin and M. M. Salomaa, Phys. Rev. B **44**, 9667 (1991).

[4] G. E. Volovik, J. Exp. Theor. Phys. **70**, 609 (1999).

[5] V. Gurarie and L. Radzihovsky, Phys. Rev. B **75**, 212509 (2007).

[6] C. Kallin and J. Berlinsky, Rep. Prog. Phys. **79**, 054502 (2016).

[7] L. Fu and C. L. Kane, Phys. Rev. Lett. **100**, 096407 (2008).

[8] F. Qu, F. Yang, J. Shen, Y. Ding, J. Chen, Z. Ji, G. Liu, J. Fan, X. Jing, C. Yang, and L. Lu, Sci. Rep. **2**, 339 (2012).

[9] A. D. K. Flinck, C. Kurter, E. D. Huemiller, Y. S. Hor, and D. J. V. Harlingen, Appl. Phys. Lett. **108**, 203101 (2016).

[10] S. Sasaki, M. Kriener, K. Segawa, K. Yada, Y. Tanaka, M. Sato, and Y. Ando, Phys. Rev. Lett. **107**, 217001 (2011).

[11] J.-P. Xu, C. Liu, M.-X. Wang, J. Ge, Z.-L. Liu, X. Yang, Y. Chen, Y. Liu, Z.-A. Xu, C.-L. Gao, D. Qian, F.-C. Zhang, and J.-F. Jia, Phys. Rev. Lett. **112**, 217001 (2014).

[12] G. Koren, T. Kirzhner, E. Lahoud, K. B. Chashka, and A. Kanigel, Phys. Rev. B **84**, 224521 (2011).

[13] P. Zareapour, A. Hayat, S. Y. F. Zhao, M. Kreshchuk, A. Jain, D. C. Kwok, N. Lee, S. Cheong, Z. Xu, A. Yang, G. D. Gu, S. Jia, R. J. Cava, and K. S. Burch, Nat. Commun. **3**, 1056 EP (2012).

[14] V. Mourik, K. Zuo, S. M. Frolov, S. R. Plissard, E. P. A. M. Bakkers, and L. P. Kouwenhoven, Science **336**, 1003 (2012).

[15] S. Nadj-Perge, I. K. Drozdov, J. Li, H. Chen, S. Jeon, J. Seo, A. H. MacDonald, B. A. Bernevig, and A. Yazdani, Science **346**, 602 (2014).

[16] M. Z. Hasan and C. L. Kane, Rev. Mod. Phys. **82**, 3045 (2010).

[17] A. R. Akhmerov, J. Nilsson, and C. W. J. Beenakker, Phys. Rev. Lett. **102**, 216404 (2009).

[18] L. Fu and C. L. Kane, Phys. Rev. Lett. **102**, 216403 (2009).

[19] Q. L. He, L. Pan, A. L. Stern, E. C. Burks, X. Che, G. Yin, J. Wang, B. Lian, Q. Zhou, E. S. Choi, K. Murata, X. Kou, Z. Chen, T. Nie, Q. Shao, Y. Fan, S.-C. Zhang, K. Liu, J. Xia, and K. L. Wang, Science **357**, 294 (2017).

[20] Y. Ando, J. Phys. Soc. Jpn. **82**, 102001 (2013).

[21] F. Ortmann, S. Roche, S. O. Valenzuela, L. W. Molenkamp, and I. Aguilera, *Topological Insulators: Fundamentals and Perspectives* (Wiley, Weinheim, Germany, 2015).

[22] S. Wolgast, Ç. Kurdak, K. Sun, J. W. Allen, D.-J. Kim, and Z. Fisk, Phys. Rev. B **88**, 180405 (2013).

[23] X. Zhang, N. P. Butch, P. Syers, S. Ziemak, R. L. Greene, and J. Paglione, Phys. Rev. X **3**, 011011 (2013).

[24] J. G. Checkelsky, Y. S. Hor, M.-H. Liu, D.-X. Qu, R. J. Cava, and N. P. Ong, Phys. Rev. Lett. **103**, 246601 (2009).

[25] B. Skinner, T. Chen, and B. I. Shklovskii, J. Exp. Theor. Phys. **117**, 579 (2013).

[26] K. T. Law, P. A. Lee, and T. K. Ng, Phys. Rev. Lett. **103**, 237001 (2009).

[27] J. Li, G. Fleury, and M. Büttiker, Phys. Rev. B **85**, 125440 (2012).

[28] P. Fendley, M. P. Fisher, and C. Nayak, Ann. Phys. **324**, 1547 (2009), July 2009 Special Issue.

[29] W. Bishara and C. Nayak, Phys. Rev. B **80**, 155304 (2009).

[30] B. Rosenow, B. I. Halperin, S. H. Simon, and A. Stern, Phys. Rev. Lett. **100**, 226803 (2008).

[31] B. Rosenow, B. I. Halperin, S. H. Simon, and A. Stern, Phys. Rev. B **80**, 155305 (2009).

[32] H. Zhang, C.-X. Liu, X.-L. Qi, X. Dai, Z. Fang, and S.-C. Zhang, Nat. Phys. **5**, 438 (2009).

[33] R. S. Akzyanov, A. L. Rakhmanov, A. V. Rozhkov, and F. Nori, Phys. Rev. B **92**, 075432 (2015).

[34] M. Cheng, R. M. Lutchyn, V. Galitski, and S. Das Sarma, Phys. Rev. B **82**, 094504 (2010).

[35] C.-K. Chiu, D. I. Pikulin, and M. Franz, Phys. Rev. B **91**, 165402 (2015).

[36] H. Beidenkopf, P. Roushan, J. Seo, L. Gorman, I. Drozdov, Y. S. Hor, R. J. Cava, and A. Yazdani, Nat. Phys. **7**, 939 (2011).

[37] C. Caroli, P. D. Gennes, and J. Matricon, Phys. Lett. **9**, 307 (1964).

[38] G. Möller, N. R. Cooper, and V. Gurarie, Phys. Rev. B **83**, 014513 (2011).

[39] T. Mizushima and K. Machida, Phys. Rev. A **81**, 053605 (2010).

[40] A. L. Rakhmanov, A. V. Rozhkov, and F. Nori, Phys. Rev. B **84**, 075141 (2011).

[41] R. S. Akzyanov, A. V. Rozhkov, A. L. Rakhmanov, and F. Nori, Phys. Rev. B **89**, 085409 (2014).

[42] C. Nayak, S. H. Simon, A. Stern, M. Freedman, and S. Das Sarma, Rev. Mod. Phys. **80**, 1083 (2008).

[43] This can be proven formally; the proof is a straightforward generalization of Eqs. (18) and (19) when identifying the structure of the subblocks in $\mathcal{T}$. Blocks characterizing transport between different leads always give a vanishing contribution to the average current, related to the observation right below Eq. (20). If two leads are not causally connected by the chiral arm, the corresponding block in the $\mathcal{T}$ matrix is zero. Finally, no matter which arms the N leads are coupled to, the drain current is reduced by the factor $\prod_{j=1}^{N} r_j$, where r_j is the local reflection amplitude of lead j.

[44] K. Saha and I. Garate, Phys. Rev. B **90**, 245418 (2014).

[45] W.-Y. Shan, H.-Z. Lu, and S.-Q. Shen, New J. Phys. **12**, 043048 (2010).

[46] D. Culcer, E. H. Hwang, T. D. Stanescu, and S. Das Sarma, Phys. Rev. B **82**, 155457 (2010).

[47] D. Hsieh, Y. Xia, D. Qian, L. Wray, J. H. Dil, F. Meier, J. Osterwalder, L. Patthey, J. G. Checkelsky, N. P. Ong, A. V. Fedorov, H. Lin, A. Bansil, D. Grauer, Y. S. Hor, R. J. Cava, and M. Z. Hasan, Nature (London) **460**, 1101 (2009).

[48] H. Li, T. Zhou, J. He, H.-W. Wang, H. Zhang, H.-C. Liu, Y. Yi, C. Wu, K. T. Law, H. He, and J. Wang, Phys. Rev. B **96**, 075107 (2017).

[49] N. Borgwardt, J. Lux, I. Vergara, Z. Wang, A. A. Taskin, K. Segawa, P. H. M. van Loosdrecht, Y. Ando, A. Rosch, and M. Grüninger, Phys. Rev. B **93**, 245149 (2016).

[50] Y. Xia, D. Qian, D. Hsieh, L. Wray, A. Pal, H. Lin, A. Bansil, D. Grauer, Y. S. Hor, R. J. Cava, and M. Z. Hasan, Nat. Phys. **5**, 389 (2009).

[51] J. A. Sobota, S. Yang, J. G. Analytis, Y. L. Chen, I. R. Fisher, P. S. Kirchmann, and Z.-X. Shen, Phys. Rev. Lett. **108**, 117403 (2012).

[52] M. Neupane, S.-Y. Xu, Y. Ishida, S. Jia, B. M. Fregoso, C. Liu, I. Belopolski, G. Bian, N. Alidoust, T. Durakiewicz, V. Galitski, S. Shin, R. J. Cava, and M. Z. Hasan, Phys. Rev. Lett. **115**, 116801 (2015).

[53] The prefactor of $\Gamma_{\mathrm{SB}}^{\mathrm{imp}}$ in Eq. (23) has a $k_\parallel^{-4}$ dependency. There is, however, also a less significant dependency on $k_\parallel$ in $d\sigma_{\mathrm{long}}^{(n')}(\epsilon_F)/dk_z'$, which can be inferred from Fig. (b) (recall that $\epsilon_F \propto k_\parallel$) to be very close to linear.

[54] S. K. Kushwaha, I. Pletikosić, T. Liang, A. Gyenis, S. H. Lapidus, Y. Tian, H. Zhao, K. S. Burch, J. Lin, W. Wang, H. Ji, A. V. Fedorov, A. Yazdani, N. P. Ong, T. Valla, and R. J. Cava, Nat. Commun. **7**, 11456 (2016).

[55] R. M. Lutchyn, E. P. A. M. Bakkers, L. P. Kouwenhoven, P. Krogstrup, C. M. Marcus, and Y. Oreg, arXiv:1707.04899.

[56] D. J. Kim, S. Thomas, T. Grant, J. Botimer, Z. Fisk, and J. Xia, Sci. Rep. **3**, 3150 EP (2013).

[57] S. Lee, X. Zhang, Y. Liang, S. W. Fackler, J. Yong, X. Wang, J. Paglione, R. L. Greene, and I. Takeuchi, Phys. Rev. X **6**, 031031 (2016).

[58] M. Büttiker, Phys. Rev. B **46**, 12485 (1992).

[59] J. D. Sau, S. Tewari, R. M. Lutchyn, T. D. Stanescu, and S. Das Sarma, Phys. Rev. B **82**, 214509 (2010).

[60] M. Cheng, R. M. Lutchyn, V. Galitski, and S. Das Sarma, Phys. Rev. Lett. **103**, 107001 (2009).

[61] C. Knapp, T. Karzig, R. Lutchyn, and C. Nayak, arXiv:1711.03968 [Phys. Rev. B (to be published)].

[62] B. Rosenstein, I. Shapiro, and B. Y. Shapiro, Europhys. Lett. **102**, 57002 (2013).

[63] J. M. Ziman, *Electrons and Phonons*, 5th ed. (Oxford University Press, Oxford, 1960).

[64] S. Giraud and R. Egger, Phys. Rev. B **83**, 245322 (2011).

[65] V. Parente, A. Tagliacozzo, F. von Oppen, and F. Guinea, Phys. Rev. B **88**, 075432 (2013).