Disorder-induced subgap states and Majorana zero-energy edge modes in 2D topological insulator-superconductor hybrid structures

Hoi-Yin Hui, Jay D. Sau, and S. Das Sarma
Department of Physics, Condensed Matter Theory Center and Joint Quantum Institute, University of Maryland, College Park, Maryland 20742-4111, USA
(Dated: January 15, 2014)

Contrary to the widespread belief that Majorana zero-energy modes, existing as bound edge states in 2D topological insulator (TI)-superconductor (SC) hybrid structures, are unaffected by non-magnetic static disorder by virtue of Anderson’s theorem, we show that such a protection against disorder does not exist in realistic multi-channel TI/SC/ferromagnetic insulator (FI) sandwich structures of experimental relevance since the time-reversal symmetry is explicitly broken locally at the SC/FI interface where the end Majorana mode (MM) resides. We find that although the MM itself and the bulk topological superconducting phase inside the TI are indeed universally protected against disorder, disorder-induced subgap states are generically introduced at the TI edge due to the presence of the FI/SC interface as long as multiple edge channels are occupied. We discuss the implications of the finding for the detection and manipulation of the edge MM in realistic TI/SC/FI experimental systems of current interest.

PACS numbers: 74.78.-w, 03.65.Vf, 74.45.+c

Current interest [1–4] in the search for solid-state Majorana modes (MMs) or zero-energy Majorana bound states inside a superconducting gap has increased sharply following a series of experimental reports [5–10] claiming empirical evidence for the existence of MMs in semiconductor (SM) nanowires in proximity to superconductors (SCs) in the presence of an external magnetic field precisely as predicted in earlier theoretical publications [11–13]. Very recent MM-motivated experimental interest [14] has focused also (i.e. in addition to SM nanowires) on 2D (i.e. the so-called quantum spin Hall system) and 3D topological insulators (TIs) in proximity to ordinary s-wave SCs, where, in fact, the first theoretical proposals for the existence of MMs were originally made by Fu and Kane [15, 16], and then followed up by others [17]. One substantive advantage of the TI-based MM proposals over the SM-based proposals is that the explicit presence of time-reversal symmetry (TRS) in the Fu-Kane TI/SC hybrid system makes the bulk topological SC phase in the TI and the associated MMs immune to static non-magnetic elastic disorder arising from impurities and defects invariably present in the environment. By contrast, the SM/SC hybrid topological structures hosting MMs are unprotected from non-magnetic disorder in the environment since time-reversal invariance must be explicitly broken [18] in the SM/SC sandwich structures in order to produce the MM-carrying topological phase. We emphasize, however, that both the TI/SC and the SM/SC systems are protected from elastic disorder residing in the bulk superconductor itself [19] which can have no effect on the proximity-induced topological superconductivity in the TI or the SM material – the protection (or not) for the TI (or SM) system that we are discussing here is specifically from the disorder residing at the interface or inside the bulk TI (or SM) material.

Although there have been many theoretical studies [20] of disorder effects on the MMs and the topological superconductivity in the SM/SC nanowire systems of experimental relevance [5–10], such studies in the TI case are sparse. Some studies in the SM/SC case have even suggested that the experimental observations of the zero bias peak [5–9] in the InSb and InAs SM nanowires in the SM/SC hybrid structures attributed to the existence of zero-energy bound MMs localized at the wire ends as predicted theoretically [11–13] were in fact direct manifestations of anti-localization effects induced by static disorder with nothing whatsoever to do with defect-bound MMs [21]. In contrast, the only analysis of disorder effects in the TI/SC hybrid structures Ref. [22] that we are aware of concluded that TI/SC topological systems are

![Figure 1](image-url)
completely protected from all elastic disorder effects by virtue of Anderson’s theorem [23] in the time-reversal-symmetric superconductivity systems.

In the current work, we consider the realistic experimental TI/SC structures for the existence of bound MMMs at the end where a ferromagnetic insulator (FI) must be deposited in order to localize the MMMs at the system edge (which serves as the defect localizing the MM, and thus achieving non-Abelian braiding statistics). The realistic structure [Fig. 1(a)], first proposed by Fu and Kane in this context [15], thus involves the 2D TI with SC and FI layers deposited on top of it, referred to as the STIM (superconductor-TI-magnet) by Fu and Kane originally. The FI layer in the STIM breaks TRS and, in general, in order to induce spatially separated MMMs for direct characterization of its topological nature, a TRS-breaking mechanism is always necessary in the TI heterostructure [3]. In this scenario, the immunity of the local SC gap in the TI/SC/FI structure against disorder is subtle. Although Anderson’s theorem guarantees that no impurities can degrade the gap deep in the SC region, there is no corresponding argument near the SC/FI interface, since the FI provides a TRS-breaking mechanism. This means that although the MMMs separated by the SC do not hybridize, extra subgap states could potentially appear locally at the SC/FI interface in the TI/SC/FI hybrid structure.

In spite of this argument, we find that disorder in the SC region of the TI edge is unable to generate any extra localized states near the SC/FI interface, and the MM remains the only subgap state there [see Fig. 2(d)], provided that we limit our attention to only single-channel TI edges. Thus, the TI/SC/FI hybrid system is indeed immune to all disorder (in spite of the explicit breaking of TRS at the FI/SC interface) provided there is only a single active edge channel in the system. However, disorder-induced potential fluctuations near the edge of a 2D semiconductor are expected to produce bound states in addition to the 1D edge state. Such extra states or puddles, which have been proposed to explain the temperature dependence of the TI edge conductance [24], can be modeled using a multi-channel TI edge. In fact, in realistic structures, we expect multi-channel edges in the generic 2D TI system.

In this paper we investigate a broad class of TI Hamiltonians by considering a multi-channel TI edge. In addition to disorder, the extra channels could be induced from an intrinsically higher chemical potential near the surface, similar to the 3D TI Bi$_2$Se$_3$ [25, 26]. Also, since the 2D TI is commonly constructed from semiconductors with small band gaps (e.g. HgCdTe or InAs/GaSb [27, 28]), a proximate metallic SC unavoidably induces extra channels due to band-bending [29] [see Fig. 1(b)].

We consider the following Hamiltonian which models a multi-channel TI edge in proximity to SC or FI:

$$H = \int dx \left\{ \psi_{\alpha s}^\dagger(x) \left[ \sum_{\alpha s s' r} (p_\alpha \cdot \sigma_{ss'}) \psi_{\alpha s}^\dagger(x) \right] \left[ -i \partial_x \right] \psi_{\alpha s}(x) + \sum_{\alpha s s'} \Delta_\alpha(x) \left[ \psi_{\alpha s}(x) \psi_{\alpha s'}^\dagger(x) + h.c. \right] \right\}$$

Here, $\psi_{\alpha s}^\dagger(x)$ creates an electron in the $\alpha$th channel with spin $s = \uparrow, \downarrow$ at position $x$. $\sigma = (\sigma^x, \sigma^y, \sigma^z)$ are the three Pauli spin matrices, and $p_\alpha$ is the polarization of the $\alpha$th channel. The $N_{ch} \times N_{ch}$ real symmetric matrix $\mu_{\alpha \beta}(x)$ contains both the chemical potential of each channel (diagonal entries) and inter-channel elastic scattering (off-diagonal entries), while the anti-symmetric matrix $\tilde{\mu}_{\alpha \beta}(x)$ are the coefficients for inter-channel spin-orbit scatterings, which still respect TRS. Static charge impurities and spin-orbit impurities are included through spatial variations in $\mu(x)$ and $\tilde{\mu}(x)$ respectively. The proximate SC and FI induce the local pairing potential $\Delta(x)$ and the Zeeman term $B(x)$ respectively. To model an SC/FI interface, we choose $\Delta(x) = \theta(x)$ and $B(x) = \theta(-x)$, restricting the analysis to an idealized case where there are no spatial fluctuations of $\Delta$ and $B$ on the TI edge [30]. We also assume that there are no inter-channel pairings or Zeeman gaps.

Before presenting the results based on numerical simulations, we first analyze the problem with a scattering matrix approach [20, 31], the details of which are provided in the Supplementary Material. We treat the SC/FI interface as an SC-N-FI problem where the N region has a finite but vanishingly small width. Then, the localized modes at the interface are found by the condition $\det \left( 1 - \tilde{R} \tilde{R} \right) = 0$, where $R(E)$ and $\tilde{R}(E)$ are respectively the reflection matrices at the SC-N and N-FI interfaces. At zero energy, $R$ is constrained by unitarity, time-reversal symmetry, and particle-hole symmetry, while $\tilde{R}$ is constrained by unitarity and $U(1)$ symmetry. From these constraints, we can show that $1 \leq N_0 \leq N_{ch}$, where $N_0$ is the multiplicity of the unity eigenvalue of $\tilde{R} \tilde{R}$, which is equal to the number of localized zero-energy modes. Thus, while Anderson’s theorem does not forbid the presence of subgap states in addition to the MMMs, the scattering matrix argument guarantees that the MM is the only low energy state at the interface.

The exact value of $N_0$ depends on the details of the Hamiltonian, but one can always fine-tune the Hamiltonian near the SC/FI interface by local disorder respecting the symmetries, such that all energies of the localized
modes reach zero. Therefore the constraint derived above implies that the number of localized states is equal to the number of channels in the TI edge, and among them there is always a zero-energy mode, which is the MM. In the following we numerically study how the number of subgap states \( N_0 \) depend on the number of channels and other details of the interface.

We start by considering the simplest case of a TI/SC/FI interface where the TI has a single channel. The local density of states (LDoS) at the SC/FI interface with or without disorder are plotted respectively in Fig. 2(c,d). It shows that the LDoS at the subgap regime \( (E < \Delta) \) is not affected by disorder. Fig. 2(b) shows that the gap deep in the SC region is also completely unaffected. This is consistent with previous results [22].

To understand the interplay of disorder and multiple channels we consider 5-channel and 3-channel TI edge models. The parameters of Eq. (1) are so chosen so that the additional channels cross the Fermi level in the SC region. It is assumed that the FI region is gated so as to place the chemical potential in the gap of the additional bands as well [see Fig. 3(a,b)]. This is necessary to obtain localized MMAs at the interface. The LDoS at the interface for several cases are plotted in Fig. 3(c,d,e,f). For the parameters chosen we typically find that each channel leads to a subgap state. The LDoS in Figs. 3(c,d) show that for the interface parameters chosen, the interface states are close to the edge of the gap away from the interface. On the other hand, introducing disorder leads to the results in Figs. 3(e,f), where one sees interface states that are bound deep inside the bulk gap. The details of the SC/FI interface, such as the chemical potential change at the interface, can lead to scattering in the same way as the disorder potential does. Therefore, depending on the details of the interface it is possible even for an interface without disorder to have subgap states in the middle of the gap. The results plotted in Figs. 3(c,d) correspond to a smooth interface where such scattering is absent and therefore do not contain deep subgap states. In contrast to the interface LDoS, the LDoS in the SC region, away from the SC/FI interface, does not show any subgap states similar to the single channel case as it must obey Anderson’s theorem.

Comparing the results of the disordered SC/FI interface in the single channel case (in Fig. 2) and the multi-channel case (in Fig 3) it is clear that the absence of subgap states in the single channel case is not simply a result of Anderson’s theorem (i.e. TRS). As mentioned in the introduction, since the TRS is locally broken at the interface, it only protects the bulk i.e. the region away from the SC/FI interface, from subgap states. On the other hand, the scattering-matrix interpretation described earlier implies that the presence of a single subgap state in the single channel case is a result of the properties of the scattering matrix near zero energies. Time-reversal invariance does of course play a role in determining the properties of the scattering matrix. The scattering ma-

---

Figure 2. (a) LDoS at the end of a SM nanowire, where the dashed red line shows the result for a clean wire while the solid blue line is for a disordered wire. (b) LDoS deep in the SC region of the TI edge, averaged over many realizations of non-magnetic disorder. (d) LDoS at the SC/FI interface on a TI, without disorder. (c) LDoS at the SC/FI interface on a TI, with a single realization of non-magnetic disorder. (d)

Figure 3. (a,b) Band structures of respectively a 5-channel and 3-channel TI edge. Here the parameters of Eq. (1) is chosen as: \( \mu = \text{diag}(1,-1,0.5,-0.5,0) \) meV, \( \vec{\mathbf{p}}^I = \vec{\mathbf{p}}^F = 0 \), \( \vec{\mathbf{p}}_{\alpha,2,3} = \pm 1 \) meV, \( \mathbf{p}_{\alpha} = (+, -,-,-,+) \) \( \tilde{z} \). The Fermi level in the FI region \( (E_F = 0) \) is shown in dashed line, while that in the SC region \( (E_F = 1.5 \) meV) is shown in dotted line. (c,d) Their corresponding LDoS of at the SC/FI interface, without disorder. (e,f) Their corresponding LDoS at the SC/FI interface, with a single realization of non-magnetic disorder.
The additional subgap states may have consequences for detecting and manipulating MMs at finite temperatures. One of the simplest signatures of an MM is the zero-bias conductance peak that may be measured using zero-bias conductance peak that may be measured using tunnelling-induced broadening of the peak. However since there is a repulsion of these states from zero energy [Fig. 4(a,b)], the likelihood for the states to influence the zero-bias conductance is small unless the temperature is high and the number of channels is large. For utilizing the MM for topological quantum computation, it is known that extra localized states do not interfere with braiding and non-Abelian statistics [32]. However, the additional subgap levels could influence for the read-out schemes that rely on measuring the occupancy of the MMs by tunneling [33]. Direct measurements of the fermion parity [34] associated with pairs of MMs might in-principle be immune to subgap states. In spite of this, the TI/SC structure has an advantage in regards to robustness against disorder, since disorder cannot degrade the SC gap deep in the SC region and hybridize the two MMs at the ends. Also, the disorder-induced subgap states show a repulsion from zero energy, without any complications arising from weak antilocalization [21, 35].

Experimentally, the conductance of a TI edge was measured as $2e^2/h$ [28], which might appear to contradict our assumption of multiple edge channels. However, the simple relation $N_{ch} = \frac{G}{(2e^2/h)}$ between conductance $G$ and the number of channels $N_{ch}$ is valid in the clean limit only. With disorder, the conductance should be analyzed using random matrix theory of the edge transmission matrix [36]. Since the multi-channel TI edge belongs to the symplectic class with odd number of channels, only one channel remains delocalized [37]. A conductance measurement with a length scale greater than the localization length would then produce $G = 2e^2/h$, even if many channels are present. Therefore the conductance measurement alone cannot rule out the possibility of the presence of multiple channels.

To conclude, we have studied the effects of multiple channels and disorder near an SC/FI interface on the edge of a 2D TI, which was proposed to realize MM. We find that a number of localized states, equal to the number of channels, appear at the SC/FI interface. One of these states is the zero-energy MM, while the energies of the other states depend on the specific details of the Hamiltonian. Further, adding disorder in the SC region leads to a distribution of the energies of these extra localized states, potentially reaching the subgap regime. However their effect is less detrimental than those in the SM/SC nanowire structures, in that they are repelled from zero energy and are localized at the boundary.

This work is supported by JQI-NSF-PFC and Microsoft Q.
In this supplementary document we present the details of the derivation of zero-energy modes by scattering matrices.
Consider an SC-N-FI problem, such that the interface between SC and FI is expanded to a finite width. A solution in N is
\[ \psi^\dagger = u_+ \psi^\dagger_+ + u_- \psi^\dagger_- + v_- \psi_- + v_+ \psi_+ , \]  
(1)
in which the multi-component \( \psi_\pm \) are Kramer’s pairs and the subscripts \( \pm \) represents right/left moving modes. The corresponding particle-hole-conjugated and time-reversed solutions are:
\[ C \psi^\dagger C^{-1} = v^*_+ \psi^\dagger_+ + v^*_- \psi^\dagger_- + u^*_+ \psi_- + u^*_- \psi_+ \]  
(2)
\[ T \psi^\dagger T^{-1} = -u^*_- \psi^\dagger_+ + u^*_+ \psi^\dagger_- + v^*_+ \psi_- - v^*_- \psi_+ \]  
(3)
Reflection matrices \( R \) relates the coefficients of \( (\psi^\dagger_+, \psi^\dagger_-) \) to those of \( (\psi^\dagger-, \psi^-) \). At the left (SC/N) interface, the particle-hole symmetry and time-reversal symmetry are both respected, giving three equations for \( R \):
\[ \begin{pmatrix} u_+ \\ v_+ \end{pmatrix} = R \begin{pmatrix} u_- \\ v_- \end{pmatrix} \]  
(4)
\[ \begin{pmatrix} v_+ \\ u_+ \end{pmatrix} = R \begin{pmatrix} v_- \\ u_- \end{pmatrix} \]  
(5)
\[ \begin{pmatrix} -u_- \\ -v_- \end{pmatrix} = R \begin{pmatrix} u_+ \\ v_+ \end{pmatrix} \]  
(6)
from which one can derive
\[ R = \tau_x R^* \tau_x = -R^T . \]  
(7)
Together with the unitarity condition \( R^\dagger R = 1 \), the value of \( \det R \) could be computed. First note that the constraint \( R = \tau_x R^* \tau_x \) requires \( R \) to take the form \( R = \begin{pmatrix} r_{ee} & r_{eh} \\ r_{eh}^* & r_{ee}^* \end{pmatrix} \). The remaining conditions \( R^\dagger R = 1 \) and \( R = -R^T \) implies
\[ r_{eh}^\dagger r_{ee} + r_{ee}^\dagger r_{eh}^* = 0 \Rightarrow r_{ee} r_{eh}^{s-1} = -r_{eh}^{s-1} r_{ee} \]  
(8)
\[ r_{eh}^\dagger r_{eh} + r_{ee}^\dagger r_{ee}^* = 1 \Rightarrow r_{ee} r_{eh}^{s-1} = 1 - r_{eh}^\dagger r_{eh} \]  
(9)
\[ r_{eh} = -r_{eh}^{s-1} \Rightarrow -r_{eh}^{s-1} r_{eh} = 1 \]  
(10)
Now we evaluate \( \det R \):
\[ \det R = (-1)^{N_{ch}} \det \begin{pmatrix} r_{eh} & r_{ee} \\ r_{ee}^* & r_{eh}^* \end{pmatrix} \]  
\[ = (-1)^{N_{ch}} \det (r_{eh} r_{eh}^* - r_{ee} r_{ee}^* r_{eh}^* r_{eh}^{s-1}) \]  
\[ = -1 \]  
(11)
where Eqs. (8, 9, 10) are consecutively used in the last step. \( N_{ch} = \dim R/2 \) is the number of channels in the model, which is restricted to be odd.
At the right (N/FI) interface we have no Andreev reflection or time-reversal symmetry, constraining the form of \( \tilde{R} \) to be \( \tilde{R} = \begin{pmatrix} \tilde{r}_{ee} & 0 \\ 0 & \tilde{r}_{ee}^* \end{pmatrix} \). By the unitarity of \( \tilde{R} \), we have \( \det \tilde{R} = 1 \).

The zero-energy modes are found by solving \( \det (1 - \tilde{R} R) = 0 \), which implies that the multiplicity of the eigenvalue \(-1 \) of \(-\tilde{R} R \) gives the number of zero-energy modes. We now prove that \(-\tilde{R} R \) must have at least one eigenvalue being \(-1 \). To this end we rotate \( R \) to the Majorana basis via \( \Omega = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -i & i \end{pmatrix} \):
\[ \Omega R \Omega^\dagger = 2 \begin{pmatrix} \Re (r_{ee} + r_{eh}) & -\Im (r_{ee} - r_{eh}) \\ \Im (r_{ee} - r_{eh}) & \Re (r_{ee} + r_{eh}) \end{pmatrix} \]  
(12)
This matrix is real and, because it is unitary, it is also orthogonal. Similarly for \( \tilde{R} \) and hence \(-\tilde{R} R \). The eigenvalues of an orthogonal matrix can only be 1, \(-1 \), or pairs of conjugate \( e^{\pm i \phi} \). Since \( \det (-\tilde{R} R) = -1 \), \(-\tilde{R} R \) must have at least one eigenvalue being \(-1 \). On the other hand, because there are no other constraints on the problem, there could be at most \( N_{ch} \) eigenvalues being \(-1 \), and hence, at most \( N_{ch} \) zero modes.