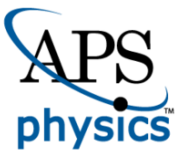

Topological modes in one-dimensional solids and photonic crystals

Timothy J. Atherton, Celia A. M. Butler, Melita C. Taylor, Ian R. Hooper, Alastair P. Hibbins, J. Roy Sambles, and Harsh Mathur

Phys. Rev. B **93**, 125106 — Published 2 March 2016

DOI: 10.1103/PhysRevB.93.125106

Topological Modes in One Dimensional Solids and Photonic Crystals

Timothy J. Atherton
Tufts University, Department of Physics & Astronomy, 574 Boston Avenue, Medford, Massachusetts, USA. 02155

Celia A. M. Butler, Melita C. Taylor, Ian R. Hooper, Alastair P. Hibbins, and J. Roy Sambles
University of Exeter, Electromagnetic and Acoustic Materials Group, Department of Physics and Astronomy, Stocker Road, Exeter, United Kingdom. EX4 4QL

Harsh Mathur
Case Western Reserve University, Department of Physics, 10900 Euclid Avenue, Cleveland, Ohio, USA 44106

It is shown theoretically that a one-dimensional crystal with time-reversal and particle-hole symmetry is characterized by a topological invariant that predicts the existence or otherwise of edge states. This is confirmed experimentally through the construction and simulation of a photonic crystal analogue in the microwave regime. It is shown that the edge mode couples to modes external to the photonic crystal via a Fano resonance.

I. INTRODUCTION

Topological phases have been shown to arise in a number of condensed matter systems: in the quantum Hall effect[1] where electrons are confined to two dimensions and subject to a perpendicular applied magnetic field and in so-called topological insulators[2,3] which are materials that possess conducting metallic surfaces despite being insulators in the bulk. Experimental studies of these states, in materials such as graphene[4] and Bi_2Se_3 have recently been an area of considerable focus both for fundamental reasons, because the topological states in these materials lead to exotic quasi-particles, and also for applications such as quantum computing.

The states arise in these systems as follows: consider a map from the Brillouin zone to a space of nondegenerate Bloch Hamiltonians $H(k)$. If $|k\rangle$ is an eigenstate of $H(k)$, then a vector potential $A(k) = -i \langle k | \partial_k | k \rangle$ may be defined following Berry[5]. The topological index of this map is $Q = \int \nabla \times A(k)\ \mathrm{d}^2k$. Thouless et al. discovered that non-trivial topological invariants can arise in the Brillouin zone with time reversal symmetry broken by the application of a strong magnetic field as in the quantum Hall effect[1]. More recently Balents and Moore[6] applied this paradigm to systems with strong spin-orbit interaction but with time reversal symmetry intact and thereby clarified an earlier proposition by Kane and Mele[7] that a $\mathbb{Z}_2$ invariant of the band structure divides insulators into two classes: an even class corresponding to conventional insulators and an odd topological insulating phase that possesses conductive surface states.

In this paper, we apply this paradigm to a system with different symmetry, a *one*-dimensional crystal with time reversal and charge conjugation symmetry. In the language of random matrix theory[8], our system corresponds to class BDI in contrast to the unitary class A in the work of Thouless et al.[1] and the symplectic class AII in Balents and Moore[6]. By an analogous argument, it is shown that such a structure may possess edge states characterized by a $\mathbb{Z}$ topological invariant. We also develop a coarser topological classification based on a $\mathbb{Z}_2$ invariant. The system is then realized experimentally through a photonic analogue. Photonic topological insulators in the same symmetry class have been studied previously; for a review see[9]. We stress that while edge states such as Tamm states are well known in 1D crystals[10], it is the topologically-protected nature of those presented here that primarily concerns us.

II. MODEL

A. Topological Argument

We begin with the topological argument. Consider a tight binding model on a one dimensional-lattice with nearest neighbor hopping and a constant on-site energy. Such a model is described by the Schrödinger equation,

$$-\tau_n\psi_{n+1} - \tau_{n-1}\psi_{n-1} = E\psi_n, \tag{1}$$

where the ψ_n represent the amplitude of the wavefunction at the n-th lattice site and τ_n is the real hopping coefficient between the n-th site to the $n+1$th site. For arbitrary τ_n, the model (1) possesses a time-reversal symmetry with associated operator $\mathcal{T}\psi_n = \psi_n^*$ as well as a particle-hole symmetry $\mathcal{C}$ represented by the anti-unitary charge conjugation operator $\mathcal{C}\psi_n = (-1)^n\psi_n^*$ where $\mathcal{C}^2 = 1,\ \mathcal{T}^2 = 1$. The particle-hole symmetry restricts the Hamiltonian's spectrum because if $|\psi\rangle$ is an eigenfunction of (1) with energy E, then $\mathcal{C}|\psi\rangle$ is also an eigenfunction with energy $-E$.

A bi-partite lattice is now considered [fig. 1(a)i)] where the bond strengths are alternately τ_1 and τ_2; this is known as the Schrieffer-Su model[11,12]. The Schrödinger equation for this situation has the form,

$$-\tau_1\phi_n^B - \tau_2\phi_{n-1}^B = E\phi_n^A,\ -\tau_1\phi_n^A - \tau_2\phi_{n+1}^A = E\phi_n^B. \tag{2}$$

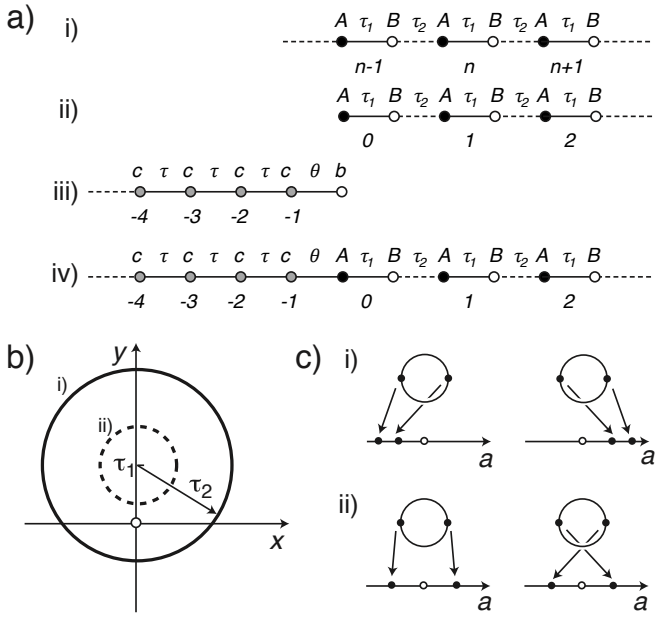

Figure 1. (a) 1D lattice models considered in this paper: i) an infinite bi-partite lattice corresponding to the Schrieffer Su model, a simple one dimensional topological insulator; ii) A semi-infinite Schrieffer-Su model where a bound surface state forms due to the non-trivial topology of the Brillouin zone; iii) A simple model with a Fano resonance. A continuum of states on the external lattice is coupled to a bound state on a single site; iv) In order to observe the bound surface state the semi-infinite topological insulator is connected to an external lattice. The bound state then shows up as a resonant jump in the reflection phase. (b) Maps from the Brillouin zone to the space of 2×2 Hamiltonians (3) fall into 2 classes depending on τ_1 and τ_2: case i) encircles the origin and ii) doesn't. (c) Maps from the Brillouin zone to the space of C-symmetric Hamiltonians are i) trivial if both points map to the same side of the origin $a = 0$ and ii) nontrivial otherwise.

This may be diagonalized by introducing the Bloch ansatz, $\phi_n^A = \alpha e^{ikn}$, $\phi_n^B = \beta e^{ikn}$, yielding,

$$\begin{pmatrix} 0 & z^*(w) \\ z(w) & 0 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = E \begin{pmatrix} \alpha \\ \beta \end{pmatrix}, \tag{3}$$

where the 2×2 matrix on the left of eq. (3) is the Bloch Hamiltonian, $w = e^{ik}$ and $z(w) = -(\tau_1 + \tau_2 w)$. Eq. (3) thus defines a map from the Brillouin zone $-\pi \leq k < \pi$ to the space of Bloch Hamiltonians. Equivalently, this may be viewed as a map $z(w)$ from a unit circle in the w plane (the Brillouin zone) to the complex plane with the origin excluded. The origin is excluded provided, as we assume, the bands are non-degenerate. The π_1 homotopy of the punctured plane is well known to be non-trivial. The topological index Q in this case corresponds to the number of times the loop $z(e^{ik})$ winds about the origin. It is easy to see that for the case $\tau_1 < \tau_2$, $Q = 1$ [fig. 1(b)(i)]; for $\tau_1 > \tau_2$, $Q = 0$ [fig. 1(b)(ii)]. This can also be determined more formally by constructing the eigenspinors of the Bloch Hamiltonian, computing the corresponding Berry connection, and evaluating the Wilson loop $\int dk A(k)$. Higher values of Q can be obtained if we perturb the model by e.g. incorporating longer range hopping that respects $\mathcal{C}$ and $\mathcal{T}$ symmetry.

Now let us reformulate the argument with more generality. We continue to assume that the crystal has a bipartite lattice and that the Hamiltonian commutes with $\mathcal{T}$ and anti-commutes with $\mathcal{C}$. $\mathcal{C}$ and $\mathcal{T}$ transform a plane wave of wave vector k into one of wave-vector $-k$. At the same time the amplitudes (α, β) are transformed to $(\alpha^*, -\beta^*)$ and (α^*, β^*) respectively. The two symmetries relate the Bloch Hamiltonians at wave-vectors k and $-k$ via $H(-k) = \mathcal{T} H(k) \mathcal{T} = -\mathcal{C} H(k) \mathcal{C}$. Together these relations and the requirement of hermiticity constrain the 2×2 Bloch Hamiltonian $H(k)$ to be off-diagonal. At the special points $k = 0$ and $k = \pi$ which are invariant under $k \to -k$ the Bloch Hamiltonian is required to have the form,

$$H = \begin{pmatrix} 0 & a \\ a & 0 \end{pmatrix}. \tag{4}$$

Here a is real and $a \neq 0$ since we are assuming that the bands have no accidental degeneracies. Topologically, this space is a punctured line. Rather than considering the mapping of the entire Brillouin zone to the space of Bloch Hamiltonians as before we may instead consider the mapping from the two special points $k = 0$ and $k = \pi$ to the punctured line. Such maps fall into two classes: a trivial one [fig. 1(c)(i)] where the special points are mapped to the same side of the origin, $a = 0$, and a non-trivial case [fig. 1(c)(ii)] such that the special points are mapped to opposite sides[13]. The edge mode, as we shall show in the following section, occurs in the non-trivial case.

Our argument mirrors that of Balents and Moore[6] who considered maps from an effective Brillouin zone in two dimensions which had the topology of a cylinder to the space of Bloch Hamiltonians. The circular boundaries of the effective Brillouin zone in their argument are analogous to our special points $k = 0$ and π. Apart from the difference in dimensionality another key difference between their work and ours is in the symmetry. As noted above they considered bands with odd time reversal symmetry (the symplectic class AII) whereas we consider even time reversal symmetry accompanied by charge conjugation symmetry (class BDI).

B. Topological Edge State in the Bipartite Lattice

Having made the topological argument, we wish to verify that a winding number of one (or more generally an odd winding number) is associated with a zero energy bound state. First, let us return to the infinite system of fig. 1(a)(i). The model has two bands symmetric about zero energy with a band gap of $2(\tau_2 - \tau_1)$. For later use we describe the mid-gap states that decay as $n \to \infty$. For

these states we adopt the ansatz,

$$\phi_n^A = \alpha(-1)^n \exp(-\kappa n) \quad \text{and} \quad \phi_n^B = \beta(-1)^n \exp(-\kappa n). \tag{5}$$

A simple calculation reveals that this solution must have the amplitude ratio,

$$\frac{\beta}{\alpha} = \pm \frac{(\tau_2 e^{-\kappa} - \tau_2)^{1/2}}{(\tau_2 e^{\kappa} - \tau_1)^{1/2}}, \tag{6}$$

and energy,

$$E = \pm(\tau_2 e^{-\kappa} - \tau_2)^{1/2}(\tau_2 e^{\kappa} - \tau_1)^{1/2}. \tag{7}$$

As eqs (6) and (7) imply the positive energy solutions have a positive amplitude ratio and the negative energy solutions have a negative amplitude ratio. The decay constant κ must lie in the range $0 \leq \kappa \leq \kappa_c$ where $\kappa_c = \log(\tau_2/\tau_1)$. $\kappa = 0$ corresponds to a solution that lies at the band edge; $\kappa = \kappa_c$ to a mid gap solution of zero energy. Needless to say the mid gap states are not permissible states for an infinite crystal since the wavefunction diverges as $n \to -\infty$.

Now consider the semi-infinite crystal briefly discussed in the previous section and depicted in fig. 1(a)(ii). Here, the first bond is τ_1 and the sites are numbered sequentially from zero to infinity. The interior of the crystal is described by the Schrödinger equation,

$$\begin{aligned} E\phi_n^A &= -\tau_1 \phi_n^B - \tau_2 \phi_{n-1}^B \quad \text{for} \quad n \geq 1, \\ E\phi_n^B &= -\tau_1 \phi_n^A - \tau_2 \phi_{n+1}^A \quad \text{for} \quad n \geq 0. \end{aligned} \tag{8}$$

together with the boundary condition,

$$E\phi_0^A = -\tau_1 \phi_0^B. \tag{9}$$

A simple calculation reveals that of all the mid-gap solutions only the zero energy solution also satisfies the boundary condition eq. (9). The zero energy solution corresponds to $\kappa = \kappa_c$. More explicitly this solution has the form,

$$\phi_n^A = \alpha(-1)^n \left(\frac{\tau_1}{\tau_2}\right)^n, \quad \phi_n^B = 0, \tag{10}$$

with the normalization constant,

$$\alpha = \frac{\sqrt{\tau_2^2 - \tau_1^2}}{\tau_2}. \tag{11}$$

The solution (10) is finite as $n \to +\infty$ only if $\tau_1 < \tau_2$, of the previous section, and hence is manifestly localized in character. The effect of curtailing the lattice is therefore as follows: while the infinite Schrieffer-Su crystal has a band gap, the semi-infinite Schrieffer-Su model has a mid-gap bound surface state at zero energy corresponding to the topological argument of the previous section. In the following section we consider one means of observing the surface state: we will show that if the topological surface state is coupled to an external continuum of states it dissolves into the continuum in accordance with Fano's general analysis[14] but leaves behind a scattering resonance which manifests itself as a phase jump of π in the reflection coefficient.

C. Fano Resonance

1. *Single site model of Fano resonance*

We model the external continuum as a semi-infinite lattice along the negative x-axis. The eigenstates of the continuum satisfy the Schrödinger equation,

$$E\psi_n = -\tau\psi_{n-1} + V\psi_n - \tau\psi_{n+1} \quad \text{for} \quad n \leq -2. \tag{12}$$

As shown in the fig.1(a)(iii) the sites of the lattice are numbered consecutively from $-\infty$ to -1 for the last site of the lattice. Ultimately we want to couple this continuum to the topological bound state of the semi-infinite Schrieffer-Su model. As a prelude let us couple it to a bound state consisting of a single site with a hopping element θ. Thus eq. (12) must be supplemented by the boundary conditions,

$$\begin{aligned} E\psi_{-1} &= -\tau\psi_{-2} + V\psi_{-1} - \theta\phi_b \\ E\phi_b &= -\theta\psi_{-1}. \end{aligned} \tag{13}$$

In the limit $\theta = 0$ the bound state is decoupled from the continuum. In this limit the bound state $|b\rangle$ is described by the wave function $\phi_b = 1$ and $\psi_n = 0$ for all n and it has zero energy. The continuum states have wave functions,

$$|k\rangle \quad \to \quad \phi_b = 0 \quad \text{and} \quad \psi_n = \frac{1}{\sqrt{\pi}} \sin kn \quad \text{for} \quad n \leq -1 \tag{14}$$

and energy $E = V - 2\tau\cos k$. The wave-vector k is restricted to lie in the range $0 \leq k \leq \pi$. A non-zero value of the hopping element θ couples the bound state to the continuum and leads to its dissolution. In fact we can map this model exactly onto Fano's model and then apply his general analysis[14]. But it is more instructive to directly solve the model defined by eqns (12) and (13). To this end we make the ansatz that,

$$\psi_n = e^{ikn} + re^{-ikn} \quad \text{for} \quad n \leq -1. \tag{15}$$

Here r the reflection coefficient is expected to be a complex number of magnitude one due to unitarity. We therefore write,

$$r = -\exp(-i2\delta) \tag{16}$$

which established our definition of the reflection phase shift δ. Note that for the decoupled case $\theta \to 0$ the phase shift is $\delta = 0$. In general a simple calculation substituting eq. (16) in eqns (12) and (13) reveals the exact result,

$$\tan\delta = \frac{(\theta^2/\tau)\sin k}{E(1 - \theta^2/2\tau^2) + V(\theta^2/2\tau^2)}. \tag{17}$$

The general expression is not particularly edifying. Recall that the energy is related to k via $E = V - 2\tau\cos k$ with k constrained to lie in the interval $0 \leq k \leq \pi$. Let $E = 0$ for $k = k_c$. Also assuming that the lattice is

weakly coupled to the bound state ($\theta \ll \tau$). Near zero energy the expression for the phase shift simplifies to,

$$\tan\delta = \frac{(\theta^2/\tau)\sin k_c}{E + V(\theta^2/2\tau^2)}. \quad (18)$$

Eq. (18) reveals that the resonance is slightly shifted from zero energy to the energy $-V(\theta^2/2\tau)$ and that the phase jumps by π over an energy width of $(\theta^2/\tau)\sin k_c$.

2. Fano resonance in Schrieffer-Su model

Finally let us couple an external continuum to the semi-infinite Schrieffer-Su model as shown in fig. 1(a)(iv). Deep in the interior the Schrodinger equation has the form,

$$\begin{aligned} E\phi_n^A &= -\tau_1\phi_n^B - \tau_2\phi_{n-1}^B \quad \text{for} \quad n \geq 1, \\ E\phi_n^B &= -\tau_1\phi_n^A - \tau_2\phi_{n+1}^A \quad \text{for} \quad n \geq 0, \\ E\psi_n &= -\tau\psi_{n-1} + V\psi_n - \tau\psi_{n+1} \quad \text{for} \quad n \leq -2. \end{aligned} \quad (19)$$

This equation must be solved subject to the following matching conditions at the interface of the external lattice and the topological insulator,

$$\begin{aligned} E\phi_0^A &= -\tau_1\phi_0^B - \theta\psi_{-1} \\ E\psi_{-1} &= -\tau\psi_{-2} + V\psi_{-1} - \theta\phi_0^A. \end{aligned} \quad (20)$$

We are interested in solutions at mid-gap energies close to zero energy. Since there are no propagating modes in the topological insulator we expect incoming plane waves from the external lattice to be perfectly reflected with the phase of the reflection coefficient encoding important information about the topological insulator as in the simple example above. We introduce the scattering ansatz,

$$\begin{aligned} \psi_n &= e^{ikn} + re^{-ikn} \quad \text{for} \quad n \leq -1, \\ \phi_n^A &= \alpha(-1)^n e^{-\kappa n} \quad \text{for} \quad n \geq 0, \\ \phi_n^B &= \beta(-1)^n e^{-\kappa n} \quad \text{for} \quad n \geq 0. \end{aligned} \quad (21)$$

Here κ and k are related to the energy by eq. (7) and by $E = V - 2\tau\cos k$ respectively. As before we define the scattering phase shift via $r = -\exp(-i2\delta)$. By substituting the ansatz (21) into eq. (20) a short calculation yields the exact result,

$$\tan\delta = \frac{\theta^2}{\tau^2}\sin k / \left[\frac{\tau_1 e^\kappa}{\tau(\tau_2 e^\kappa - \tau_1)}E + \frac{\theta^2}{\tau^2}\cos k\right]. \quad (22)$$

Once again the general expression is not particularly edifying but close to zero energy the phase shift simplifies to,

$$\tan\delta \approx \frac{\theta^2}{\tau^2}\sin k_c / \left[\frac{\tau_1\tau_2}{\tau(\tau_2^2 - \tau_1^2)}E + \frac{\theta^2}{\tau^2}\cos k_c\right]. \quad (23)$$

Here k_c is defined as the value of k that corresponds to zero energy ($V - 2\tau\cos k_c = 0$) and we have assumed that the topological insulator is weakly coupled to the external lattice ($\theta \ll \tau$). The energy difference of the phase shift is now explicit and from eq. (23) we can see that the phase jumps by π as the energy is swept through the resonance and we can read off the energy of the resonance, which is again slightly shifted from zero energy, as well as the energy width of the resonance. We note that this π-phase shift is related to scattering matrix expressions for topological invariants discussed in[15,16]. The phase shift provides one potential means of observing the topological state; in the following section we shall demonstrate another.

D. Observing the Topological States

We now address the issue of how the edge state might be observed in a finite system. The hopping model on a bipartite lattice is actualized in a 1D solid where an electron is subject to a periodic array of potential barriers of alternating height. The reflection and transmission of incident free particle wavefunctions of wavevector p through such a structure is readily determined by matching local solutions to the Schrödinger equation at interfaces using matrix methods[17]. To do so, define,

$$T(\theta) = \begin{pmatrix} e^{i\chi}\cosh\theta & \sinh\theta \\ \sinh\theta & e^{-i\chi}\cosh\theta \end{pmatrix}, \; U = \begin{pmatrix} e^{ipb} & 0 \\ 0 & e^{-ipb} \end{pmatrix}. \quad (24)$$

Here $T(\theta)$ is the transfer matrix for a single symmetric barrier located at the origin and U a translation operator; θ is the opacity of the barrier, b is the lattice spacing and p is the square root of the energy. In the limit of large barriers, $\chi = pa + \xi$ where a is the barrier width and ξ is an overall phase shift.

The band structure of the bipartite lattice, with alternating barriers of opacity θ_1 and θ_2 respectively, is determined by finding the eigenvalues of the transfer matrix corresponding to the unit cell, i.e. $UT(\theta_1)UT(\theta_2)$. For $\theta_2 \neq \theta_1$, the usual band, i.e. values of k such that the eigenvalues are complex, is found to split into two subbands symmetrically placed around the point $p_0 = (\pi/2 - \xi)/(a+b)$ and with edges corresponding to the roots of $\cos(2p(a+b) + 2\xi)\cosh\theta_1\cosh\theta_2 + \sinh\theta_1\sinh\theta_2 = \pm 1$. Between these subbands, including the point p_0, the reflection coefficient $|r|^2 \to 1$; the phase shift δ of a reflected wave, however, experiences a jump of π around p_0 if $\theta_1 < \theta_2$ and 0 otherwise. The jump leads to a Lorentzian feature in the time delay $\frac{\partial\delta}{\partial p}$ of a reflected wave; the width was determined by routine calculations to be,

$$L = 2e^{\theta_2}\cosh\theta_1 \frac{\sinh(\theta_1 + \theta_2) - \cosh\theta_1\cosh\theta_2}{\sinh^2(\theta_2 - \theta_2)}. \quad (25)$$

Hence, the topological mode may be observed experimentally by a technique such as Time Delay Reflection Spectroscopy. As will be seen shortly, for a finite struc-

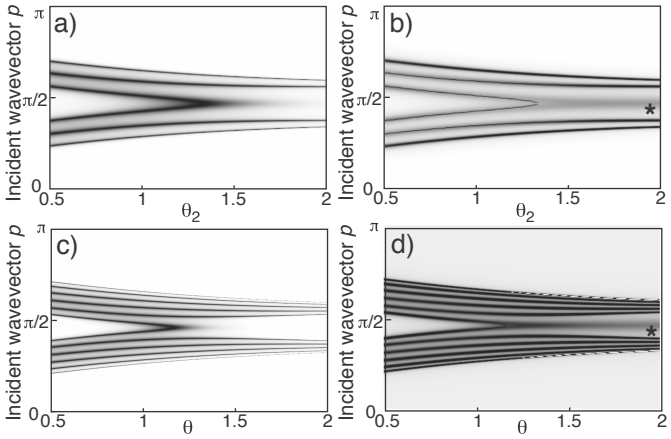

Figure 2. (a) Reflected intensity as a function of wavevector for opacity parameters $\theta_1 = 1$ and varying θ_2. (b) Time delay as a function of wavevector and θ_2; the lorentzian feature indicative of the topological mode for $\theta_2 > \theta_1$ is indicated with a *. Corresponding plots are shown for 6.5 units in (c) and (d).

ture evidence of the topological mode is also available in some circumstances from the reflection profile.

To verify the above arguments, we numerically evaluated the reflection coefficient and time delay for a 3.5 period structure from the transfer matrix as a function of p and θ_2 with fixed $\theta_1 = 1$, $a = 0.1$, $b = 1$. Results are plotted in fig. 2. When $\theta_1 = \theta_2$ the reflection spectrum [fig. 2(a)] consists of 6 minima, which correspond to a band in a structure with an infinite number of unit cells. As θ_2 is increased, the central two minima move closer together and merge around $\theta_2 \approx 1.3$; as $\theta_2 \to \infty$, this mid-band mode vanishes in reflection, contributing only a phase shift to the reflected light as discussed earlier. Conversely, if $\theta_2 < \theta_1$, the central minima become separated and, as predicted by the topological argument above, no mid-band mode exists. Results for 6.5 units are shown in fig. 2(c) and (d); the additional layers generate corresponding modes, but the topological mode remains for $\theta_2 > \theta_1$. If the number of periods is increased further, more modes appear to eventually yield the two sub-bands of the infinite bipartite lattice and the mid-gap mode persists confirming its topological nature.

E. Topological States in Other Lattice Types

1. *Tripartite Lattice*

Having studied the bipartite Schrieffer-Su lattice, we now consider other lattice types to establish whether the topological modes observed occur in other systems with the same symmetries. As a first step, we consider the tripartite ABC lattice [fig. 3(a)] for which the Schrödinger equation has the form,

$$-\tau_3\phi_{n-1}^C - \tau_1\phi_n^B = E\phi_n^A$$

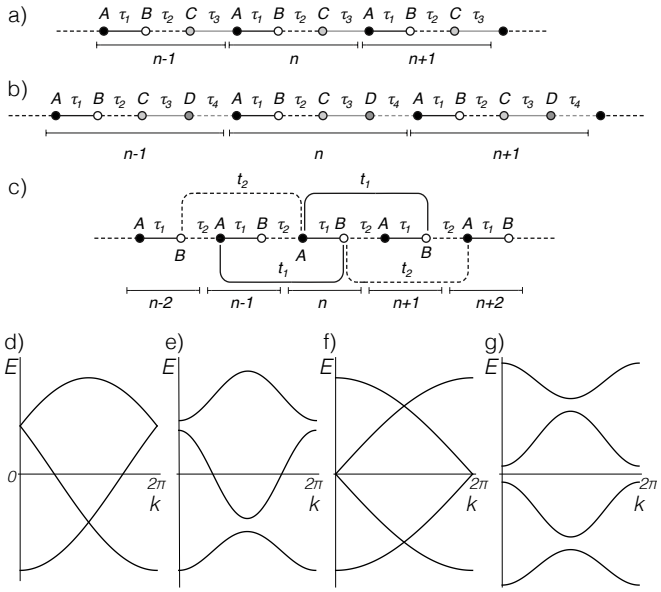

Figure 3. Further lattices considered in this section: (a) a tripartite lattice; (b) a quadripartite lattice and (c) a bipartite lattice with nearest neighbor hops. Band structure of the tripartite lattice (d) with equal hopping strengths $\tau_1 = \tau_2 = \tau_3$ for which the Brillouin zone is folded into three connected parts, and (e) with different hopping coefficients for which the band splits into three subbands. Non-topological edge modes may exist in these gaps, but can be arbitrarily shifted into the continuum by approproriate choice of coefficients. For the quadripartite lattice, the band structure for equal hopping strengths (f) is folded into four components that become distinct sub-bands (g) if the hopping strengths differ. Topological edge states may exist at $E = 0$; non topological edge states may be found in the remaining band gaps.

$$\begin{aligned} -\tau_1\phi_n^A - \tau_2\phi_n^C &= E\phi_n^B \\ -\tau_2\phi_n^B - \tau_3\phi_{n+1}^A &= E\phi_n^C. \end{aligned} \tag{26}$$

As before, a plane wave Bloch ansatz solution,

$$\phi_n^A = \alpha e^{ikn},\ \phi_n^B = \beta e^{ikn},\ \phi_n^C = \gamma e^{ikn}, \tag{27}$$

is inserted into eq. (26) to obtain a set of linear equations,

$$\begin{pmatrix} 0 & -\tau_1 & -\tau_3 e^{-ik} \\ -\tau_1 & 0 & -\tau_2 \\ -\tau_3 e^{ik} & -\tau_2 & 0 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} = E \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix}. \tag{28}$$

Eq. (28) defines a map from the first Brillouin zone $-\pi \leq k < \pi$ to the space of 3×3 Bloch Hamiltonians, just as eq. (3) defined a map from the Brillouin zone to the space of 2×2 Bloch Hamiltonians. The topology of this map will be established momentarily, but first let us examine the allowed solutions.

It is straightforward to show that the tripartite ABC lattice may possess edge states, though it will be shown later that these are *not* topological in nature. To do so, the Schrödinger equation in eq. (26) may be reformulated

as a transfer matrix by routine manipulations,

$$\begin{pmatrix} \phi^A_{n+1} \\ \phi^C_n \end{pmatrix} = \begin{pmatrix} \frac{E\tau_1^2+E\tau_2^2-E^3}{\tau_1\tau_2\tau_3} & \frac{\tau_2^2-E^2}{\tau_1\tau_2} \\ \frac{E^2-\tau_1^2}{\tau_1\tau_2} & \frac{E\tau_3}{\tau_1\tau_2} \end{pmatrix} \begin{pmatrix} \phi^A_n \\ \phi^C_{n-1} \end{pmatrix}. \quad (29)$$

For $(1,0)$ to be an eigenvector of the transfer matrix, $E = \pm\tau_1$ with associated eigenvalue $\pm\tau_2/\tau_3$ corresponding to two states of the form,

$$\phi^A_n = \left(\frac{\tau_2}{\tau_3}\right)^n \phi^A_0,\ E = +\tau_1$$
$$\phi^A_n = (-1)^n \left(\frac{\tau_2}{\tau_3}\right)^n \phi^A_0,\ E = -\tau_1 \quad (30)$$

which are manifestly localized in character and finite as $n \to +\infty$ only if $|\tau_2| < |\tau_3|$.

We now turn to the remaining states. Energy eigenvalues of (28) are found by constructing the characteristic polynomial,

$$E^3 - E(\tau_1^2 + \tau_2^2 + \tau_3^2) + 2\tau_1\tau_2\tau_3 \cos k = 0. \quad (31)$$

For $\tau_1 = \tau_2 = \tau_3 \equiv \tau$, we obtain the solution $E = -2\tau\cos(k/3)$. Distinct solutions span $0 \le k < 6\pi$, which can be folded back into the first Brillouin zone $0 \le k < 2\pi$ yielding three connected components [fig. 3(c)],

$$E_1 = -2\tau\cos(k/3)$$
$$E_2 = -2\tau\cos(k/3 + 2\pi/3)$$
$$E_3 = -2\tau\cos(k/3 + 4\pi/3), \quad (32)$$

Suppose the hopping strengths are now perturbed so that $\tau_1 = \tau + \delta\tau_1$, $\tau_2 = \tau + \delta\tau_2$ and $\tau_3 = \tau + \delta\tau_3$, subject to $\delta\tau_1 + \delta\tau_2 + \delta\tau_3 = 0$. The single band now becomes three distinct sub-bands as $\tau_1 \neq \tau_2 \neq \tau_3$ [fig. 3(d)]. Perturbative analysis of the characteristic polynomial eq. (31) shows that gaps appear at $k = 0$ and $k = \pi$,

$$E = \tau \pm \sqrt{\frac{2}{3}}\sqrt{\delta\tau_1^2 + \delta\tau_2^2 + \delta\tau_3^2},\ k = 0$$
$$E = -\tau \pm \sqrt{\frac{2}{3}}\sqrt{\delta\tau_1^2 + \delta\tau_2^2 + \delta\tau_3^2},\ k = \pi. \quad (33)$$

If, therefore, we set $\delta\tau_1 = -\delta\tau_2 - \delta\tau_3$ and require that $\delta\tau_2 < \delta\tau_3$, bound states will exist at $E = \pm(\tau - \delta\tau_2 - \delta\tau_3)$ between the band gaps at $E = \tau \pm \frac{2}{\sqrt{3}}\sqrt{\delta\tau_2^2 + \delta\tau_3^2 + \delta\tau_2\delta\tau_3}$ and $E = -\tau \pm \frac{2}{\sqrt{3}}\sqrt{\delta\tau_2^2 + \delta\tau_3^2 + \delta\tau_2\delta\tau_3}$. As $\delta\tau_2 \to \delta\tau_3 \equiv \delta$, however, the bound states have energy $E = \pm(\tau - 2\delta)$ while the band edges are at $E = \tau \pm 2\delta$ and $E = -\tau \pm 2\delta$. Hence, the edge states can be merged smoothly into the continuum *without* closing the gap as $\tau_2 \to \tau_3$. In contrast, the edge state obtained in subsection II B for the bipartite lattice exists if $\tau_1 < \tau_2$ in the center of the gap for any finite gap size and *cannot* be merged into the continuum by adjusting the hopping parameters. The ABC lattice is therefore an interesting counterexample because it may possess edge states, but these are not topological in nature.

We now demonstrate that the topology of the map eq. (28) is indeed trivial. Applying the particle-hole symmetry $\mathcal{C}$ to the Bloch ansatz eq. (27) yields,

$$\mathcal{C}\phi^A_n = (-1)^n\alpha^* e^{-ikn},$$
$$\mathcal{C}\phi^B_n = -(-1)^n\beta^* e^{-ikn},$$
$$\mathcal{C}\phi^C_n = (-1)^n\gamma^* e^{-ikn}. \quad (34)$$

Hence, the action of $\mathcal{C}$ on a Bloch state (α,β,γ) is $\mathcal{C}(\alpha,\beta,\gamma) \to (\alpha^*,-\beta^*,\gamma^*)$ and replace the wavevector $k \to -\pi - k$ modulo 2π. The Bloch Hamiltonian of eq. (28) under $\mathcal{C}$ therefore possesses fixed points at $k = \pm\frac{\pi}{2}$. Writing an arbitrary 3×3 complex matrix H and requiring that $\mathcal{C}H\mathcal{C} = -H$, we determine the most general form of the Bloch Hamiltonian at these points,

$$H = \begin{pmatrix} 0 & a & ib \\ a & 0 & c \\ -ib & c & 0 \end{pmatrix}, \quad (35)$$

where a, b and c are real parameters. The eigenvalues of this matrix are $E = 0$ or $E = \pm\sqrt{a^2+b^2+c^2}$ and so it is degenerate if and only if $(a,b,c) = 0$. Topologically, this space is $\mathbf{R}^3$ minus the origin: it is simply connected so possesses trivial π_0 homotopy in agreement with the analysis of the edge states above.

2. *Quadripartite Lattice*

The quadripartite lattice [fig. 3(b)] is interesting because it possesses bound states of *both* topological and non-topological nature, and hence demonstrates that the topological modes of the bipartite lattice are not an isolated special case. We begin with the topological argument: As before, we construct a plane wave ansatz,

$$\phi^A_n = \alpha e^{ikn},\ \phi^B_n = \beta e^{ikn},.$$
$$\phi^C_n = \gamma e^{ikn},\ \phi^D_n = \delta e^{ikn} \quad (36)$$

and determine the action of the charge conjucation operator on a Bloch state,

$$\mathcal{C}\phi^A_n = -\alpha^* e^{-ikn},\ \mathcal{C}\phi^B_n = -\beta^* e^{-ikn},$$
$$\mathcal{C}\phi^C_n = -\gamma^* e^{-ikn},\ \mathcal{C}\phi^D_n = -\delta^* e^{-ikn}. \quad (37)$$

Hence, $\mathcal{C}(\alpha,\beta,\gamma,\delta) \to (\alpha^*,-\beta^*,\gamma^*,-\delta^*)$ and replaces the wavevector $k \to -k$ modulo 2π. As for the bipartite case, the fixed points of the mapping are $k = 0$ and $k = \pi$. The time-reversal operator $\mathcal{T}$, conversely, acts on a Bloch state by conjugating the coefficients $\mathcal{T}(\alpha,\beta,\gamma,\delta) \to (\alpha^*,\beta^*,\gamma^*,\delta^*)$ and reversing the sign of the wavevector $k \to -k$; hence the fixed points of $\mathcal{C}$ coincide with those of $\mathcal{T}$. Writing an arbitrary 4×4 complex matrix H and requiring that $\mathcal{C}H\mathcal{C} = -H$ and $\mathcal{T}H\mathcal{T} = H$, the most general form of the Bloch Hamiltonian at these points is,

$$H = \begin{pmatrix} 0 & a & 0 & b \\ a & 0 & c & 0 \\ 0 & c & 0 & d \\ b & 0 & d & 0 \end{pmatrix} \quad (38)$$

where a, b, c and d are real parameters. The characteristic polynomial of this matrix is,

$$\lambda^4 - \lambda^2 P + Q^2 = 0, \tag{39}$$

where $P = a^2+b^2+c^2+d^2$ and $Q = (ad-bc)$ and hence the eigenvalues are,

$$E = \pm\frac{1}{\sqrt{2}}\sqrt{P \pm \sqrt{P^2-4Q^2}}. \tag{40}$$

Degenerate pairs of eigenvalues occur under several circumstances: (i) If $Q = 0$, i.e. $ad = bc$, a degenerate pair exists with $E = 0$; (ii) if $P = 2Q$, i.e. $a = d$ and $b = -c$, and (iii) if $P = -2Q$, i.e. if $a = -d$ and $b = c$. Condition (i) divides the space (a, b, c, d) into two disconnected regions $ad > bc$ and $ad < bc$. The remaining conditions define 2D slices through the space (a, b, c, d) and hence do not disconnect the space any further. Topologically, the space of $\mathcal{C}$ symmetric Bloch Hamiltonians contains two disconnected pieces and the π^0 homotopy is hence characterized by a $\mathbb{Z}_2$ invariant.

The $\mathbb{Z}_2$ invariant predicts the existence or otherwise of a topological surface state as follows: The Schrödinger equation in Bloch matrix form for arbitrary k is,

$$\begin{pmatrix} 0 & -\tau_1 & 0 & \tau_4 e^{-ik} \\ -\tau_1 & 0 & -\tau_2 & 0 \\ 0 & -\tau_2 & 0 & -\tau_3 \\ -\tau_4 e^{ik} & 0 & -\tau_3 & 0 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \\ \gamma \\ \delta \end{pmatrix} = E \begin{pmatrix} \alpha \\ \beta \\ \gamma \\ \delta \end{pmatrix}. \tag{41}$$

Comparing eq. (41) with (38), $a = -\tau_1$, $b = -\tau_2$, $d = -\tau_3$ and $c = +\tau_4$ at $k = 0$ and $c = -\tau_4$ at $k = \pi$ respectively. If $\tau_1\tau_3 > \tau_2\tau_4$, then $ad > bc$ for both $k = 0$ and $k = \pi$ and no surface state exists. Conversely, if $\tau_1\tau_3 < \tau_2\tau_4$ then $ad < bc$ for $k = 0$ and $ad > bc$ for $k = \pi$, and a surface state will exist.

To confirm the existence of the state, we recast the Schrödinger equation in transfer matrix form,

$$\begin{pmatrix} \phi^A_{n+1} \\ \phi^D_n \end{pmatrix} = \begin{pmatrix} \frac{E^4 - E^2(\tau_1^2+\tau_2^2+\tau_3^2)+\tau_1^2\tau_3^2}{\tau_1\tau_2\tau_3\tau_4} & \frac{E^3 - E\tau_2^2 - E\tau_3^2}{\tau_1\tau_2\tau_3} \\ \frac{E\tau_1^2 + E\tau_2^2 - E^3}{\tau_1\tau_2\tau_3} & \frac{\tau_2^2\tau_4 - E^2\tau_4}{\tau_1\tau_2\tau_3} \end{pmatrix} \begin{pmatrix} \phi^A_n \\ \phi^D_{n-1} \end{pmatrix}. \tag{42}$$

For $(1, 0)$ to be an eigenvector of the transfer matrix, the lower left element must vanish, i.e. $E(E^2 - \tau_2^2 - \tau_3) = 0$. Hence bound states exist with $E = 0$ and $E = \pm\sqrt{\tau_1^2+\tau_2^2}$. Inserting these values back into the transfer matrix, the associated eigenvalue for $E = 0$ is $\frac{\tau_1\tau_3}{\tau_2\tau_4}$ and hence the associated state is only finite as $n \to \infty$ if $\tau_1\tau_3 < \tau_2\tau_4$ in agreement with the topological argument above. The eigenvalue for $E = \pm\sqrt{\tau_1^2+\tau_2^2}$ is $-\frac{\tau_2\tau_3}{\tau_1\tau_4}$ and hence these states exist if $\tau_2\tau_3 < \tau_1\tau_4$. Unlike the $E = 0$ state, the position of these states can be adjusted arbitrarily with respect to the edge of the adjacent sub-bands through appropriate choice of hopping parameters; thus, while these states are edge-localized, they are not topological in origin.

3. *Lattice of Arbitrary Periodicity*

We now consider the existence of edge modes for a lattice of arbitrary periodicity M. The argument follows closely the approach developed in the above sections: For a given periodicity, first identify the relevant points of the Brillouin zone for which the Bloch Hamilton commutes with $\mathcal{C}$ and $\mathcal{T}$ by considering the action of these operators on the Bloch ansatz. Then consider two distinct fixed points in the Brillouin zone and construct the most general Bloch Hamiltonian that possesses the relevant symmetries at these points. By determining the connectivity of this space, we establish the topology of the full map from the Brillouin zone to the space of Bloch Hamiltonians.

The odd periodicity case is rapidly dismissed. As for the tripartite lattice, the fixed points of $\mathcal{C}$ at $k = \pm\frac{\pi}{2}$ and $\mathcal{T}$ at $k = 0, \pi$ do not coincide. Furthermore at $\mathcal{C}$ invariant points the eigenvalues of the Bloch matrix have to come in equal and opposite pairs. For odd M that means one of the eigenvalues has to be zero implying that the positive and negative energy bands are degenerate at these points in k-space. Since there is no band gap there is obviously no question of having mid gap states, topological or otherwise.

For even periodicity $2n$, the fixed points of $\mathcal{C}$ and $\mathcal{T}$ coincide at $k = 0, \pi$ as was shown above for the bipartite and quadripartite lattices. At these fixed points, the Bloch Hamiltonian matrix H_{ij} must have the following structure,

$$H_{ij} = 0, \text{ if } (-1)^i = (-1)^j \tag{43}$$
$$H_{ij} = H_{ji} \tag{44}$$
$$H_{ij} \in \mathbb{R}. \tag{45}$$

Note that here we have in fact generalized the problem and allowed hops over more than one lattice site. There are $2n^2$ non-zero elements and n^2 independent elements due to the symmetry. From these matrices, one can define two $n \times n$ matrices,

$$L_{ij} = H_{2i-1,2j} \tag{46}$$
$$M_{ij} = H_{2i,2j-1} \tag{47}$$

where clearly $L = M^T$. For example for the quadripartite

lattice Hamiltonian in eq. 38 the associated matrices are,

$$L = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \; M = \begin{pmatrix} a & c \\ b & d \end{pmatrix}. \tag{48}$$

We now show that the determinant of the Bloch Hamiltonian H_{ij} at the fixed points can conveniently be written as a perfect square. To do so we use the Leibniz formula for the determinant,

$$\det H = \sum_{\sigma \in S_{2n}} \mathrm{sgn}(\sigma) M_{1,\sigma(1)} M_{2,\sigma(2)} \cdots M_{2n,\sigma(2n)}, \tag{49}$$

where the sum is taken over the permutation group S_{2n}, i.e. all possible permutations of the integers $[1, 2n]$. We note that the factors in each term can be rearranged into odd and even n,

$$\det H = \sum_{\sigma \in S_{2n}} \mathrm{sgn}(\sigma) M_{1,\sigma(1)} M_{3,\sigma(3)} \cdots M_{2n-1,\sigma(2n-1)} \times M_{2,\sigma(2)} M_{4,\sigma(4)} \cdots M_{2n,\sigma(2n)}, \tag{50}$$

and hence every nonzero term in $\det H$ can be written as a product,

$$(-1)^n \prod_{i=1}^{n} H_{2i-1,2\sigma(i)} H_{2\sigma'(i)-1} \mathrm{sgn}(\sigma)\mathrm{sgn}(\sigma') \tag{51}$$

where here $\sigma \in S_n$. The determinant of H can therefore be factorized,

$$\begin{aligned} \det H &= (-1)^n \sum_{\sigma,\sigma' \in S_n} \left[\prod_{i=1}^{n} \mathrm{sgn}(\sigma) H_{2i-1,2\sigma(i)}\right] \left[\prod_{i=1}^{n} \mathrm{sgn}(\sigma') H_{2i,2\sigma'(i)-1}\right] \\ &= (-1)^n \det L \det M \\ &= (-1)^n (\det L)^2 . \end{aligned} \tag{52}$$

which is a perfect square up to a factor of $(-1)^n$. Recollect that L is an arbitrary $n \times n$ real matrix. The space of H is therefore divided into two disconnected pieces defined by $\det L > 0$ and $\det L < 0$; hence the point homotopy of H $\pi_0(H) = \mathbb{Z}_2$ corresponding to the sign of $\det L$. This result is valid for arbitrary n and hence, since the homotopy of H does not depend on n, there exists either zero or one topological modes per edge, regardless of the number of bands, in a 1D hopping model with real coefficients and constant on-site potential. This is another interesting difference between the present scenario and the work of Thouless and of Balents and Moore, whose system supported multiple topological modes.

The results above for the bipartite and quadripartite lattices are now seen as special cases of this result. For the bipartite lattice, the matrix L constructed from eq. (4) is just the 1×1 matrix (a); for this case the topological invariant is $\mathrm{sgn}(a)$ as was argued above. For the quadripartite lattice, L was displayed in eq. (48) and has $\det L = ad - bc$, precisely corresponding to the quantity identified in the previous subsection as the discriminant for the existence of topological states.

It is worth noting that Kitaev in his seminal work on topological edge modes[18] found that the determinant of a class D matrix is also a perfect square of a quantity called the Pfaffian; Kitaev used the sign of the Pfaffian as a topological invariant as we do here. Class D matrices have a charge conjugation symmetry but broken time reversal symmetry. In a suitable basis such a matrix is antisymmetric and Hermitian. The matrices we consider belong to class BDI. They possess both time reversal and charge conjugation symmetry and in a suitable basis have the checkerboard structure we have identified. As a result the determinant of these matrices, following Kitaev's analysis, is also a perfect square.

Our $\mathbb{Z}_2$ topological index is related to the already known $\mathbb{Z}$ topological classification for class BDI insulators[19] as follows. In this work, we considered two invariants for class BDI insulators: The first, introduced in section II A, was the winding number Q of the map from the Brillouin zone to the space of Bloch Hamiltonians. This is a Z invariant. The second, based on comparing the Hamiltonian at two symmetric points in the Brillouin zone, is a $\mathbb{Z}_2$ invariant. It is, essentially, the winding number invariant modulo 2. The Z_2 invariant therefore provides a coarser classification of BDI insulators but it is the relevant classification if one is only interested in the existence of mid-gap topological modes. The relevance of $\mathbb{Z}$ vs. $\mathbb{Z}_2$ invariants for BDI insulators has been discussed before in the literature in context of other physical quantities such as the polarization[20,21].

4. Next-nearest Neighbor Hopping

In order to elucidate the connection between the winding number, the $\mathbb{Z}_2$ invariant, and the number of zero modes, it is instructive to consider a simple generalization of the

model analyzed in section IIA, namely, a bipartite lattice wherein hopping between next nearest neighbor sites of the two sub lattices is permissible [see fig. 3(c)].

Winding number—For an infinite crystal the bands of this model are determined by solving the Schrödinger equation,

$$\begin{aligned} -\tau_1\phi_n^B - \tau_2\phi_{n-1}^B - t_1\phi_{n+1}^B - t_2\phi_{n-2}^B &= E\phi_n^A \\ -\tau_1\phi_n^A - \tau_2\phi_{n+1}^A - t_2\phi_{n+2}^A - t_1\phi_{n-1}^A &= E\phi_n^B. \end{aligned} \quad (53)$$

Note that eq. (53) respects both time reversal and charge conjugation invariance and therefore this model belongs to BDI, the symmetry class of interest. As in section IIA we introduce the Bloch ansatz, $\phi_n^A = \alpha e^{ikn}$ and $\phi_n^B = \beta e^{ikn}$, and obtain,

$$\begin{pmatrix} 0 & z^*(w) \\ z(w) & 0 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = E \begin{pmatrix} \alpha \\ \beta \end{pmatrix}. \quad (54)$$

The 2×2 matrix on the left is the Bloch Hamiltonian, $w = e^{ik}$, and $z(w) = -(\tau_1 + \tau_2 w + t_2 w^2 + t_1 w^{-1})$. As in section IIA we view $z(w)$ as a map from the unit circle in the w-plane to the z-plane with the origin excluded. Formally the number of times the loop in the z-plane winds about the origin is given by,

$$Q = \Im\left(\frac{1}{2\pi}\int_0^{2\pi} dk\, \frac{1}{z}\frac{dz}{dk}\right) = \Im\left(\frac{1}{2\pi}\oint_{\text{unit circle}} dw\, \frac{1}{z}\frac{dz}{dw}\right). \quad (55)$$

Eq. (55) can be justified by using the polar representation $z = \rho e^{i\theta}$ and noting that,

$$\frac{d\theta}{dk} = \Im\left(\frac{1}{z}\frac{dz}{dk}\right). \quad (56)$$

To facilitate explicit computation of Q it is helpful to verify that,

$$\frac{1}{z}\frac{dz}{dw} = -\frac{1}{w} + \frac{1}{w-\alpha} + \frac{1}{w-\beta} + \frac{1}{w-\gamma}. \quad (57)$$

Here α, β and γ are the roots of the polynomial,

$$-wz(w) = t_1 + \tau_1 w + \tau_2 w^2 + t_2 w^3. \quad (58)$$

The winding number then simply counts the number of roots of this polynomial that lie inside the unit circle. Explicitly,

$$Q = -1 + \overline{\Theta}(|\alpha|) + \overline{\Theta}(|\beta|) + \overline{\Theta}(|\gamma|) \quad (59)$$

Here the complementary step function $\overline{\Theta}(x) = 1$ for $|x| < 1$ and $\overline{\Theta}(x) = 0$ for $|x| > 1$. Evidently for this model the winding number can take four possible values $Q = -1, 0, 1$, and 2. We will see below that the winding number equals the difference between the numbers of zero modes that live on the A and B sub lattices. On the other hand one can argue that the $\mathbb{Z}_2$ invariant introduced in section IIA is simply Q (mod 2). Hence if the $\mathbb{Z}_2$ invariant is odd the existence of zero modes is assured. However an even value of the $\mathbb{Z}_2$ invariant is consistent with either the absence of zero modes or of an even difference in the numbers of zero modes on the two sub lattices.

Zero modes—In order to investigate the existence of zero modes we wish to solve eq. (53) at zero energy, namely,

$$\begin{aligned} &\tau_1\phi_n^A + \tau_2\phi_{n+1}^A + t_2\phi_{n+2}^A + t_1\phi_{n-1}^A = 0, \text{ for } n = 1, 2, 3, \ldots \\ &\tau_1\phi_n^B + \tau_2\phi_{n-1}^B + t_1\phi_{n+1}^B + t_2\phi_{n-2}^B = 0, \text{ for } n = 2, 3, 4 \end{aligned} \quad (60)$$

These equations must be supplemented by the boundary conditions,

$$\begin{aligned} \tau_1\phi_0^A + \tau_2\phi_1^A + t_2\phi_2^A &= 0, \\ \tau_1\phi_0^B + t_1\phi_1^B &= 0, \\ \tau_1\phi_1^B + \tau_2\phi_0^B + t_1\phi_2^B &= 0. \end{aligned} \quad (61)$$

In principle eqs (60) and (61) may be used to determine ϕ_n^A and ϕ_n^B for all n in terms of the three independent components ϕ_0^A, ϕ_1^A and ϕ_0^B.

In practice it is most convenient to solve eqs (60) and (61) using the method of generating functions (the discrete Laplace transform). We define,

$$f_A(w) = \sum_0^\infty \phi_n^A w^n \quad \text{and} \quad f_B(w) = \sum_0^\infty \phi_n^B w^n. \quad (62)$$

A short calculation then reveals,

$$\begin{aligned} f_A(w) &= \frac{\phi_0^A t_2}{(t_2 + \tau_2 w + \tau_1 w^2 + t_1 w^3)}\left[1 + \left(\frac{\tau_2}{t_2} + \frac{\phi_1^A}{\phi_0^A}\right) w\right], \\ f_B(w) &= \frac{\phi_0^B t_1}{(t_1 + \tau_1 w + \tau_2 w^2 + t_2 w^3)}. \end{aligned} \quad (63)$$

Eq (63) is obtained by multiplying the first line of eq (60) by w^{n+2}, the second line by w^{n+1}, then summing over n over the ranges indicated in eq (60) and making use of eq (61) to eliminate ϕ_2^A, ϕ_2^B and ϕ_1^B. The denominator of the expression for f_B coincides with the polynomial (58), previously denoted $-wz(w)$, and analyzed above in connection with the winding number. Hence its roots are α, β and γ. On the other hand one can show that the polynomial in the denominator of f_A is the dual of the denominator of f_B in the sense that its roots are $1/\alpha, 1/\beta$ and $1/\gamma$. Making use of this information allows us to rewrite the generating functions in a more transparent form,

$$\begin{aligned} f_A(w) &= \frac{\phi_0^A}{(1-\alpha w)(1-\beta w)(1-\gamma w)}\left[1 + \left(\frac{\tau_2}{t_2} + \frac{\phi_1^A}{\phi_0^A}\right) w\right] \\ f_B(w) &= \frac{\phi_0^B}{(1-\alpha^{-1}w)(1-\beta^{-1}w)(1-\gamma^{-1}w)}. \end{aligned} \quad (64)$$

In order to reconstruct the solutions it is helpful to separate the generating functions f_A and f_B into partial fractions. For example,

$$f_B(w) = \phi_0^B\left[\frac{A}{1-\alpha^{-1}w} + \frac{B}{1-\beta^{-1}w} + \frac{C}{1-\gamma^{-1}w}\right], \quad (65)$$

where the coefficients A, B and C are given by,

$$A = \frac{\beta\gamma}{(\alpha-\beta)(\alpha-\gamma)},$$
$$B = \frac{\gamma\alpha}{(\beta-\gamma)(\beta-\alpha)},$$
$$C = \frac{\alpha\beta}{(\gamma-\alpha)(\gamma-\beta)}. \quad (66)$$

It follows from eqs (62) and (65) that,

$$\phi_n^B = \phi_0^B \left[\frac{A}{\alpha^n} + \frac{B}{\beta^n} + \frac{C}{\gamma^n}\right]. \quad (67)$$

Evidently the solution eq (67) is a viable (normalizable) eigenfunction only if $|\alpha| > 1, |\beta| > 1$ and $|\gamma| > 1$. Under these conditions the winding number $Q = -1$. Thus we see that $Q = -1$ implies the existence of a zero mode that lives entirely on the B sub lattice. Conversely we see that if the winding number $Q \neq 1$ then at least one of the roots α, β or γ lies within the unit circle and hence the solution eq (67) diverges as $n \to \infty$.

Similarly f_A has a partial fraction decomposition,

$$f_A(w) = \phi_0^A \left[\frac{a}{1-\alpha w} + \frac{b}{1-\beta w} + \frac{c}{1-\gamma w}\right] \quad (68)$$

which corresponds to the wave function,

$$\phi_n^A = \phi_0^A \left[a\alpha^n + b\beta^n + c\gamma^n\right]. \quad (69)$$

Here the coefficients a, b and c depend on the ratio ϕ_1^A/ϕ_0^A. For example by choosing,

$$\frac{\tau_2}{t_2} + \frac{\phi_1^A}{\phi_0^A} = \alpha \quad (70)$$

we can choose the numerator in f_A to cancel the pole at $w = \alpha$ and hence arrange for a to vanish. In this circumstance,

$$b = \frac{\beta}{\beta-\gamma}, \quad c = \frac{\gamma}{\gamma-\beta}. \quad (71)$$

Now let us consider the viability of the solution eq (69) for different values of the winding number Q.

For $Q = 2$, we have $|\alpha| < 1, |\beta| < 1$ and $|\gamma| < 1$. Thus the solution eq (69) is normalizable regardless of the value of the ratio ϕ_1^A/ϕ_0^A and hence there are two independent zero modes on the A sub lattice for $Q = 2$. For $Q = 1$, one of the three roots, α, β or γ must lie outside the unit circle. To be definite let us assume that $|\alpha| > 1$ whilst $|\beta| < 1$ and $|\gamma| < 1$. In that case by choosing the ratio ϕ_1^A/ϕ_0^A as in eq (70) we can arrange for a to vanish and for eq (69) to still be normalizable. Thus for winding number $Q = 1$ there is only one zero mode on the A sub lattice. For $Q = 0$ and $Q = -1$ there are no zero modes on the A sub lattice because two or three terms in the wavefunction eq (69) diverge as $n \to \infty$ and by tuning the ratio ϕ_1^A/ϕ_0^A we can only cancel one of the divergent terms.

In summary we have shown that for $Q = 2$ there are two zero modes, both on the A sub lattice; for $Q = 1$, one zero mode on the A sub lattice; for $Q = 0$ there are no zero modes on either sub lattice; and for $Q = -1$ there is one zero mode on the B sub lattice.

III. PHOTONIC ANALOGUE

In order to conveniently verify the veracity of the topological index as a predictor of edge states, we exploit the known isomorphism between the Schrödinger equation and Maxwell's equations in one dimension[22] to construct a photonic analogue of the bipartite structure in the microwave regime using metamaterials.

The electromagnetic wave equation takes the form,

$$\nabla^2\phi - \frac{\varepsilon}{c^2}\frac{\partial^2\phi}{\partial t^2} = 0 \quad (72)$$

where ϕ is the electric field, c is the speed of light, and ε is the dielectric permittivity of the medium in which the wave is propagating. The plane wave solution to this is,

$$\phi = \phi_0 e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)} \quad (73)$$

allowing us to rewrite the wave equation as

$$\left[-\frac{\partial^2}{\partial x^2} - k_0^2\varepsilon\right]\phi = -k_\parallel^2\phi \quad (74)$$

where $k_0 = \omega/c$ is the wavevector in free space and $k_\parallel = \sqrt{k_x^2 + k_y^2}$ is the in-plane wavevector.

We can directly compare this wave equation with Schrödinger's equation,

$$\left[-\frac{\hbar}{2m}\frac{\partial^2}{\partial x^2} + V\right]\psi = E\psi \quad (75)$$

allowing us to identify $-k_0^2\varepsilon$ as the photonic analogue of the potential and $-k_\parallel^2$ as the analogue of the total energy. Thus a medium with a negative value of the dielectric permittivity (a metal) will constitute a potential barrier, and we can form a bipartite stack through the appropriate layering of metallic and dielectric media.

While it would be a relatively simple matter to create such a stack at optical frequencies, where the barriers would have thicknesses of the order of the skin depth (< 50 nm), metals at these frequencies are inherently lossy—the permittivity is complex—corresponding to a tunnel barrier with a complex potential. If we wish to make a direct comparison to the previous arguments this is less than ideal. At lower frequencies (microwaves) the situation is even worse. The conductivity of metals at microwave frequencies is exceptionally high, corresponding to extremely high tunnel barriers. Indeed, in order to create a barrier that reflects only 50% of incoming radiation

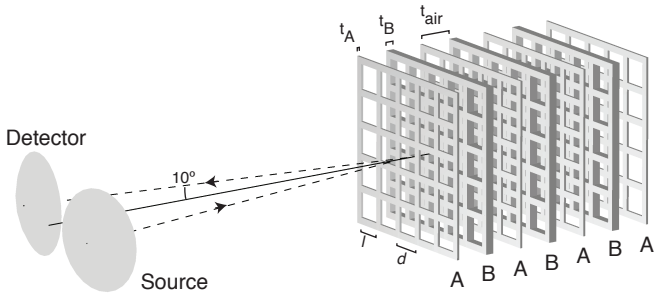

Figure 4. Experimental setup to measure reflected intensity of microwave radiation. The structure consists of metamaterial layers—aluminium plates stamped with a square array of pitch $d = 7.68$ mm and hole size $l = 6.15$ mm—of alternating thicknesses $t_A = 0.66$ mm and $t_B = 2.33$ mm and spaced by air $t_{air} = 7.6$mm.

one requires a barrier thickness of the order of 2 nm, even though the wavelengths are of the order of centimetres. These ultra-thin layers are very lossy and exceptionally difficult to fabricate. By using microwave metamaterials, artificial materials formed from sub-wavelength elements whose properties are determined not solely by the constituent materials but also their structuring[23], one can tailor the height of the tunnel barrier while maintaining an almost purely real potential.

The metamaterials used consist of aluminium sheets with thicknesses of the order of a few millimetres perforated with arrays of sub-wavelength square holes. The square holes exhibit a cut-off frequency that depends upon the size of the holes and below which only evanescently decaying fields are supported. Throughout the frequency range investigated here we are always within this regime. The transmission through the metamaterial is dominated by the fields decaying through the holes, mimicking the skin depth in real metals and effectively acting as an artificial metal with a much lower conductivity than the constituent aluminium, but with close to zero absorptive losses since almost all of the fields are confined within the perforations[24]. We note, however, that, while we can consider such a metamaterial to act as an artificial metal in this specific instance of wave propagation in the direction normal to the interface, this is not in general the case[25]. In order to control the reflectivity of this artificial metal—equivalent to changing the opacity θ in the discussion above—one can alter the size of the holes, in essence changing the effective conductivity of the medium, or alternatively change the thickness. Here we keep the structure the same and alter the thickness.

The experimental setup is illustrated in fig. 4 and details of the data collection described more fully in[26]. The experimental structure itself consists of alternating layers of dielectric (air) and metamaterials. The metamaterial layers are separated by metallic spacers around the edge of the sample yielding a nominal spacing of $t_{air} = 7.60 \pm 0.01$ mm. The metamaterial layers are made from a solid aluminium sheet perforated with a periodic square array of holes of pitch $d = 7.68 \pm 0.01$ mm and hole size of $l = 6.15 \pm 0.01$ mm; the thickness of the metamaterial in successive layers is alternated, and denoted t_A and t_B respectively, to form the required ABAB bipartite stack with 3.5 unit cells as modelled earlier.

In the experiment, the sample is illuminated by a collimated s-polarized (transverse electric) microwave beam, produced by a microwave horn placed at the focus of a spherical mirror, incident on the sample at 10°. Reflected intensity is measured as a function of frequency using a detector horn and secondary mirror connected to a scalar network analyser from which the reflected intensity was determined. The sample studied had $t_A < t_B$, specifically $t_A = 0.66 \pm 0.01$ mm and $t_B = 2.33 \pm 0.01$ mm as measured from the constructed sample. Since the opacity parameters θ of eq. (24) monotonically increase with the thicknesses t, the topological argument above predicts that the edge states should be observed in this case. The reflected intensity is plotted as a function of frequency in fig. 5(a), confirming the presence of the midgap topological mode. Due to the finite number of periods in the structure, the band is manifested as a series of reflection minima corresponding to resonant modes in the structure; if the number of unit cells were increased, the continuous sub-bands would be recovered[27]. One can make a direct qualitative comparison between the data in fig. 5(a) and fig. 2(a) for values of $\theta_2 > 1.5$.

We also simulated the response of the structure using commercial finite element software (Comsol multiphysics). We emphasize that no fitting was performed and the agreement between model and data displayed in fig. 5(a) is typical for such studies, accurately locating the position of the minima while incorrectly estimating their depth. This discrepancy is due to a number of physical effects including: finite area of the sample in the experiment; spherical aberration of the microwave source and detector; imperfections in the experimental sample such as bowing of the metamaterial layers which changes the dielectric spacing; and a radius of curvature associated with the hole edges that reduces the effective skin depth. As well as modelling the system with $t_A < t_B$ we also simulated the reversed structure, i.e. where $t_A = 2.33$ mm and $t_B = 0.66$ mm; the reflection profile of the reversed structure is plotted as a grey dashed line in fig. 5(a). No mid-gap topological mode is observed in this case in agreement with the prediction of the previous sections.

As with other experiments[28,29], modelling can also be used to visualize the electric field distribution in each mode. Plots of the norm of the electric field as a function of the distance through the sample are displayed in fig. 5(b) for each reflection minimum in the fitted profile fig. 5(a). Modes A,B,D and E are clearly distinct in character from mode C. Mode C has the predicted properties of the topological edge state: it is confined to the edge of the sample and also occurs in the middle of the band. Plots of the average electric field are also shown in fig. 5(c) for the six minima in the reversed structure. Unlike the

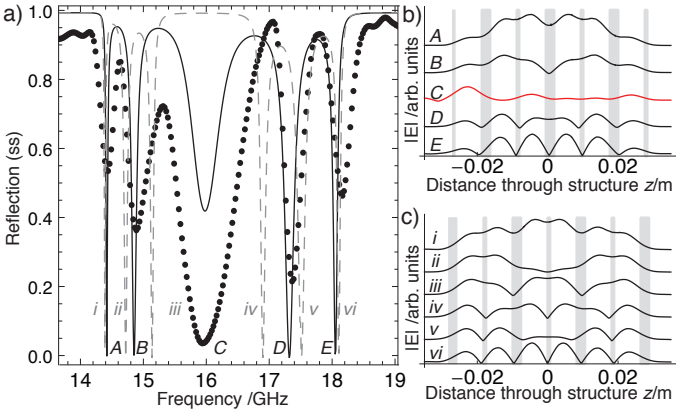

Figure 5. (a) Measured (disks) and modelled (black line) reflection profile for a 3.5 period AB bipartite stack with $t_A < t_B$; minima are labelled (A-E). The grey dashed line is the reflection profile for a structure with $t_A > t_B$ with minima labelled (i-vi). Incident radiation is from the left. (b) Modelled normalized electric field intensity profile for each observed mode A-E; metamaterial layers are indicated by gray shading. The topological mode is highlighted in red. (c) Corresponding field plots for the reversed structure i-vi.

previous case, none of the modes i—vi are localized to the edges, in agreement with the prediction of the above topological argument.

IV. CONCLUSION

A topologically protected edge state has been experimentally observed in a 1D photonic crystal with time-reversal symmetry. The mid-gap edge state might potentially be useful for the creation of a notch filter, but is of fundamental interest due to its topological orgin. The existence or otherwise of the state as a function of the design parameters of the crystal is predicted by a $\mathbb{Z}_2$ topological invariant that classifies mappings from the band structure to the space of Bloch Hamiltonians; the classification was constructed using methods previously applied to two and three dimensional topological insulators. However the crystals we consider are of a different class from conventional topological insulators in terms of their symmetries. The latter must have time reversal symmetry and strong spin-orbit interaction; in our work we use an anti-unitary charge conjugation symmetry $\mathcal{C}$ as well as the time-reversal operator $\mathcal{T}$ to develop the classification.

Our work also clarifies the distinction between topological and non-topological edge states. By analyzing a tripartite and quadripartite lattice, we found edge states that can be dissolved into the adjacent band by a continuous change of parameters; these are non-topological in origin. The tripartite lattice possesses no topological edge states while the quadripartite lattice possesses both topological and non-topological modes. By generalizing our results to arbitrary periodicities and hopping distances, we find that our $\mathbb{Z}_2$ invariant determines whether the number of topological modes is even or odd.

ACKNOWLEDGMENTS

The authors acknowledge financial support from the EPSRC through the QUEST program grant (EP/I034548/1), CAMB was supported by QinetiQ and the EPSRC through the Industrial CASE scheme (no. 08000346). HM is supported by a grant from the U.S. Department of Energy to the Particle Astrophysics Theory group at CWRU. TJA is funded by the Research Corporation for Science Advancement through a Cottrell Award.

[1] D. J. Thouless, M. Kohmoto, M. P. Nightingale, and M. den Nijs, Phys. Rev. Lett. **49**, 405 (1982).
[2] M. Z. Hasan and C. L. Kane, Rev. Mod. Phys. **82**, 3045 (2010).
[3] X.-L. Qi and S.-C. Zhang, Rev. Mod. Phys. **83**, 1057 (2011).
[4] A. H. Castro Neto, F. Guinea, N. M. R. Peres, K. S. Novoselov, and A. K. Geim, Rev. Mod. Phys. **81**, 109 (2009).
[5] M. V. Berry, Proc. Roy. Soc. A **392**, 45 (1984).
[6] J. E. Moore and L. Balents, Phys. Rev. B **75**, 121306 (2007).
[7] C. L. Kane and E. J. Mele, Phys. Rev. Lett. **95**, 146802 (2005).
[8] A. Altland and M. R. Zirnbauer, Phys. Rev. B **55**, 1142 (1997).
[9] L. Lu, J. D. Joannopoulos, and M. Soljačić, Nature Photonics (2014).
[10] Y. tuan Fang, L. kun Chen, N. Zhu, and J. Zhou, Optoelectronics, IET **7**, 9 (2013).
[11] W. P. Su, J. R. Schrieffer, and A. J. Heeger, Phys. Rev. Lett. **42**, 1698 (1979).
[12] J. Zak, Phys. Rev. Lett. **62**, 2747 (1989).
[13] If we assume that for generic values of k the Bloch Hamiltonian will not accidentally have the form eq. (4) then the map from the Brillouin zone to the complex plane can cross the x axis only at $k = 0$ and $k = \pi$; this is the reason that the winding number can only equal 0 or 1.
[14] U. Fano, Phys. Rev. **124**, 1866 (1961).
[15] A. R. Akhmerov, J. P. Dahlhaus, F. Hassler, M. Wimmer, and C. W. J. Beenakker, Phys. Rev. Lett. **106**, 057001 (2011).
[16] I. C. Fulga, F. Hassler, A. R. Akhmerov, and C. W. J. Beenakker, Phys. Rev. B **83**, 155429 (2011).
[17] D. Sprung and G. Morozov, Journal of Physics A (2007).
[18] A. Y. Kitaev, Physics-Uspekhi **44**, 131 (2001).
[19] S. Ryu, A. P. Schnyder, A. Furusaki, and A. W. W. Ludwig, New Journal of Physics **12**, 065010 (2010).

[20] A. M. Essin and V. Gurarie, Phys. Rev. B **84**, 125132 (2011).
[21] J. Song and E. Prodan, Phys. Rev. B **89**, 224203 (2014).
[22] I. R. Hooper, T. W. Preist, and J. R. Sambles, Phys. Rev. Lett. **97**, 053902 (2006).
[23] T. Cui, D. Smith, and R. Liu, *Metamaterials: Theory, Design, and Applications* (Springer US, 2009).
[24] J. S. McCalmont, M. M. Sigalas, G. Tuttle, K. Ho, and C. M. Soukolis, Appl. Phys. Lett. **68**, 2759 (1996).
[25] E. K. Stone and E. Hendry, Phys. Rev. B **84**, 035418 (2011).
[26] C. Butler, *The Microwave Response of Square Mesh Metamaterials*, Ph.D. thesis, University of Exeter (2012).
[27] D. Griffiths and C. Steinke, Am. J. Phys. **69**, 137 (2001).
[28] C. A. M. Butler, J. Parsons, J. R. Sambles, A. P. Hibbins, and P. A. Hobson, Appl. Phys. Lett. **95**, 174101 (2009).
[29] C. A. M. Butler, I. R. Hooper, A. P. Hibbins, J. R. Sambles, and P. A. Hobson, J. Appl. Phys. **109**, 013104 (2011).