Temperature evolution of the phonon dynamics in the Kitaev spin liquid

Kexin Feng, Mengxing Ye, and Natalia B. Perkins

Phys. Rev. B **103**, 214416 — Published 8 June 2021

DOI: 10.1103/PhysRevB.103.214416

Temperature evolution of the phonon dynamics in the Kitaev spin liquid

Kexin Feng,[1] Mengxing Ye,[2] and Natalia B. Perkins[1]

[1]*School of Physics and Astronomy, University of Minnesota, Minneapolis, MN 55455, USA*

[2]*Kavli Institute for Theoretical Physics, University of California, Santa Barbara, CA 93106, USA*

(Dated: May 24, 2021)

Here we present a study of the phonon dynamics in the honeycomb Kitaev spin model at finite temperatures. We show that the fractionalized spin excitations of the Kitaev spin liquid, the itinerant Majorana fermions and static Z_2 fluxes, have distinct effects on the phonon dynamics, which makes the phonon dynamics a promising tool for exploring the Kitaev spin liquid candidate materials. In particular, we focus on the signature of the fractionalized excitations in the thermodynamic behaviour of the sound attenuation and the phonon Hall viscosity: The former describes the phonon decay into the fractionalized excitations, and the latter is the leading order time reversal symmetry breaking effect on the acoustic phonon. We find that the angular dependence of the attenuation coefficient and its magnitude are modified by the thermal excitation of the Z_2 fluxes. The strength of this effect strongly depends on the relative magnitude of the sound velocity and the Fermi velocity characterizing the low-energy Majorana fermions. We also show that the Hall viscosity is strongly suppressed by the increase of the density of the Z_2 fluxes at finite temperatures. All these observations reflect the effects of the emergent disorder on the Majorana fermions introduced by the Z_2 fluxes. Our analysis is based on the complementary analytical calculations in the low-temperature zero-flux sector, and numerical calculations in the inhomogeneous flux sectors at intermediate and high temperatures with stratified Monte Carlo (strMC) method.

I. INTRODUCTION

Recent years have seen a lot of progress in identifying candidate materials that can realize quantum spin liquid (QSL) phases [1], the interest to which stems from their remarkable set of collective phenomena including topological ground-state degeneracy, long-range entanglement, and fractionalized excitations [2–7]. In particular, a significant experimental and theoretical effort has been devoted to the study of magnetic properties of spin-orbit coupled 4d and 5d magnets, dubbed Kitaev materials [8–11], which can potentially realize the celebrated Kitaev honeycomb model [3].

The experimental detection and characterization of the QSL states are, however, very difficult since the absence of local magnetic orders down to zero temperatures makes them in many aspects invisible to the conventional local probes. When searching for QSL physics in the Kitaev materials, a promising route is to look for signatures of spin fractionalization into two types of quasiparticle excitations, which according to the exact solution of the Kitaev model are localized, gapped Z_2 fluxes and itinerant, gapless Majorana fermions. The dynamical probes which have been heavily exploited for this goal are inelastic neutron scattering [12–15], Raman scattering [16–23], resonant inelastic x-ray scattering [24–27], ultrafast spectroscopy [28, 29] and 2D terahertz non-linear coherent spectroscopy [30, 31].

The search for fractionalization in QSLs largely relies on their decoupling with the environment. However, the spin-lattice interaction is inevitable in real materials and is often rather strong [32–35]. In particular, since the dynamics of phonons can be modified by the coupling of phonons to fractionalized excitations [36, 37], the study of phonon dynamics can serve as an additional probe of fractionalization and provide important information about the nature of the QSL state. In the Kitaev spin liquid, the coupling of phonons to Majorana fermions is also expected to play a significant role in the thermal Hall effect [38, 39], the observation of which in the Kitaev material $\alpha-\mathrm{RuCl_3}$ [33], considered to be proximate to the Kitaev QSL state, is one of the most direct evidences of the presence of fractional excitations in this material.

Recently, some of us have shown that the phonon dynamics may serve as an indirect probe of fractionalization of spin degrees of freedom in the Kitaev QSL through the study of the sound attenuation (α_s) from the phonon decaying into a pair of Majorana fermions, and the Hall viscosity (η_H) induced by the time-reversal breaking spin Hamiltonian [37]. Specifically, the sound attenuation may be measured by the ultrasound experiment [40–45], and the Hall viscosity could be inferred from the acoustic Faraday effect [46–48], thermal Hall effect [37, 39, 49–52] and spectroscopy measurement. In

order to compute these quantities, in Ref. [37] we derived a low-energy effective spin-lattice coupling in terms of the matter Majorana fermions and the acoustic phonons and formulated a diagrammatic computation procedure. As we were focusing only on the low-temperature response, the flux degrees of freedom were totally neglected. However, since thermal excitations of fluxes at finite temperatures significantly modify the Majorana fermion spectrum, they can give rise to dynamical responses that are strikingly different from their zero-temperature counterparts [20, 26, 36].

In this work, we study the phonon dynamics at finite temperatures and focus on the effects of the Z_2 fluxes on the temperature dependence of the sound attenuation and Hall viscosity coefficients. Importantly, as the Z_2 flux excitations are static and do not couple to phonons directly [53], the renormalized phonon propagator can still be obtained from the polarization bubble whose internal states are just Majorana fermions, and the problem is still exactly solvable. However, in order to compute this polarization bubble in the presence of the fluxes when the translational symmetry for the Majorana fermions is broken, we had to derive a microscopic low-energy effective spin-lattice coupling and a diagrammatic computation procedure in the mixed representation when the Majorana fermion eigenmodes are obtained from the diagonalization in the real space and the acoustic phonons are treated in the momentum space.

We show that the sound attenuation manifests a clear sixfold symmetry with respect to the phonon momentum both in the zero-flux sector and when the sound attenuation coefficient is averaged over different thermal flux sectors. Second, the sound attenuation in the presence of fluxes shows a very different temperature evolution depending on the relative magnitude of the sound velocity and the Fermi velocity characterizing the low-energy Majoranas.

When the sound velocity is smaller than the Fermi velocity, the sound attenuation at low-temperature zero-flux sector is determined by the microscopic processes in which a Majorana fermion is excited to a higher energy fermion state (dubbed as ph-channel as we define in Sec. IV A), with the attenuation rate linear in temperature due to the vanishing density of states at the Dirac points [37]. These processes satisfy both the energy and momentum kinematic constraints. At intermediate and high temperatures, the relaxation of the momentum kinematic constraint and the modification of the fermionic spectrum due to the thermal flux lead to a significant change in the angular distribution of the attenuation coefficient in the momentum space and an overall decrease of the magnitude of the phonon decay.

When the sound velocity is larger than the Fermi velocity, in the low-temperature zero-flux sector the kinematic constraints can only be satisfied in the microscopic processes when a phonon decays into two fermions, both with positive energy (dubbed as pp-channel in Sec. IV A). These processes can happen even at zero temperature as long as a phonon has enough energy to excite a pair of particles. Consequently, both the intensity and the angular pattern of the attenuation coefficient do not strongly depend on the temperature.

In the presence of the time-reversal symmetry breaking term in the spin Hamiltonian, we study its effect (in the leading order) on the phonon dynamics through a phonon Hall viscosity [54, 55]. Our results on its temperature dependence show that while at low temperatures, when the zero-flux is a good approximation of the thermal flux, the phonon Hall viscosity remains almost temperature independent, at higher temperatures there is a significant reduction of the phonon Hall viscosity coefficient due to the presence of the Z_2 fluxes. We also show that, counter-intuitively, the phonon Hall viscosity exhibits a sizeable decrease as the strength of the time-reversal symmetry breaking perturbation in the spin Hamiltonian increases.

In order to study the phonon dynamics above the flux proliferation temperatures, we developed a method to sample the flux configurations, dubbed stratified Monte Carlo (strMC) algorithm. This algorithm stratifies the sample space according to the flux pseudo-potential energy model proposed in Ref. [56] and yields unbiased results of the thermodynamic quantities calculated throughout this paper. This algorithm has several advantages over the commonly used Markov Chain Monte Carlo (MCMC) algorithm. It fundamentally reduces the autocorrelation time to zero and accelerates the convergence. It is also convenient in implementation and parallelization, and free from the local minima problem. The details of this algorithm is presented in Appendix F.

The rest of the paper is organized as follows. In Section II A, we introduce the general aspects of the spin-phonon Hamiltonian. In Section II B, we show how to obtain the finite-temperature fermionic spectrum by solving the free-fermion problem in each flux sector exactly. We show that in the presence of fluxes the diagonalization of the Majorana fermion Hamiltonian should be performed in the real space

and can be done by using singular value decomposition. Then, we introduce the imaginary time fermion propagators which we will later use in the calculations of observable quantities. In Section II C, we derive the Majorana fermion-phonon (MFPh) coupling in the mixed representation. We then proceed to Section III and discuss the effect of the spin-lattice coupling on the phonon dynamics. Here we compute the phonon polarization bubble, which is the key quantity defining the phonon dynamics. In Section IV, we relate the imaginary part of the diagonal components of the phonon polarization bubble to the attenuation coefficients. In Section IV A, we discuss the role of the kinematic constraints on the phonon dynamics. In Section IV B, we present the analysis of the phonon dynamics in the thermal flux sectors, when the kinematic constraint of momentum conservation is relaxed. Based on this analysis, in Section IV C we present numerical results for the temperature evolution of the sound attenuation coefficient obtained by employing the stratified Monte Carlo (strMC) algorithm. The details of the strMC method, the comparison with MCMC method, and its tailored application to the Kitaev honeycomb model are discussed in Appendix F. In Section V, we study the observable consequences of the spin-lattice coupling when time reversal symmetry is broken. In particular, we study the temperature evolution of the Hall viscosity coefficient, η_H focusing on the understanding the effect of the Z_2 fluxes. We conclude with a general discussion of our results in Section F1. Technical details and other auxiliary information are provided in Appendices A-E.

II. THE MODEL

A. The spin-phonon Hamiltonian

We focus our discussion on the spin-phonon Hamiltonian:

$$\mathcal{H} = \mathcal{H}_s + \mathcal{H}_{\rm ph} + \mathcal{H}_{\rm c}, \tag{1}$$

For simplicity, we consider the isotropic Kitaev model on a two-dimensional (2d) honeycomb lattice. The model has $G = P6mm$ space group (group number 77), which is a semidirect product of the point group C_{6v} [57] and translation group P on a 2d hexagonal lattice, i.e. $G = \mathrm{C}_{6v} \ltimes P$. The space group symmetry gives constraints on the form of $\mathcal{H}$, which we discuss in detail below. Importantly, while a given flux configuration breaks the space group symmetry, the symmetry is restored after averaging over all flux configurations due to thermal fluctuations.

The *first term* in Eq. (1) represents the spin Hamiltonian given by

$$\mathcal{H}_{\rm s} = -\sum_{\alpha,\mathbf{r}\in A} J^{\alpha}\sigma^{\alpha}_{\mathbf{r}}\sigma^{\alpha}_{\mathbf{r}+\mathbf{M}_\alpha} - \kappa \sum_{\langle \mathbf{r}',\mathbf{r}'',\mathbf{r}'''\rangle_{\alpha\gamma}} \sigma^{\alpha}_{\mathbf{r}'}\sigma^{\beta}_{\mathbf{r}''}\sigma^{\gamma}_{\mathbf{r}'''}. \tag{2}$$

where J^α denotes the nearest neighbor Kitaev interaction on the corresponding bond of type $\alpha = x, y, z$, $\sigma^\alpha_{\mathbf{r}}$ are the Pauli matrices, and $\mathbf{M}_\alpha$ labels the three inequivalent bonds on the honeycomb lattice (see Fig. 1). In the isotropic Kitaev model, we consider $J^\alpha = J_K$. The second term in Eq.(2) breaks time-reversal and vertical mirror symmetries, but preserves the exact solubility of the model. It mimics the effect of an external magnetic field $\mathbf{h} = (h_x, h_y, h_z) = \frac{h}{\sqrt{3}}(1,1,1)$ perturbatively, with $\kappa \sim \frac{h_x h_y h_z}{J^2}$. The three-spin link notation $\langle \mathbf{r}', \mathbf{r}'', \mathbf{r}'''\rangle_{\alpha\gamma}$ shown in Fig. 1 labels bonds $\mathbf{r}'\mathbf{r}''$, $\mathbf{r}''\mathbf{r}'''$ of type α, γ, respectively, on three adjacent sites $\mathbf{r}', \mathbf{r}'', \mathbf{r}'''$.

The *second term* in Eq.(1) is the bare Hamiltonian for the acoustic phonons on the honeycomb lattice,

$$\mathcal{H}_{\rm ph} = \mathcal{H}^{kinetic}_{\rm ph} + \mathcal{H}^{elastic}_{\rm ph}. \tag{3}$$

Here, $\mathcal{H}^{kinetic}_{\rm ph} = \sum_{\mathbf{q}} \frac{\mathbf{P}_{-\mathbf{q}}\cdot\mathbf{P}_{\mathbf{q}}}{2\rho\delta_V}$, where $\mathbf{P}_{\mathbf{q}} = \rho\,\partial_{\mathbf{t}}\mathbf{u}_{\mathbf{q}}$ is the momentum of the phonon. δ_V is the area enclosed in one unit cell, ρ is the mass density of the lattice ion, and $\mathbf{u} = \{u_x, u_y\}$ is the lattice displacement vector. The elastic energy is determined by the point group of the crystal. For the C_{6v} group of our interest, it is expressed in terms of the strain tensor $\epsilon_{\alpha\beta} = \frac{1}{2}(\partial_\alpha u_\beta + \partial_\beta u_\alpha)$, as

$$\mathcal{H}^{elastic}_{\rm ph} = \int d^2x \left[C_1(\epsilon_{xx}+\epsilon_{yy})^2 + C_2(\epsilon_{xx}-\epsilon_{yy}+2i\epsilon_{xy})(\epsilon_{xx}-\epsilon_{yy}-2i\epsilon_{xy})\right]. \tag{4}$$

Here, $\epsilon_{xx}+\epsilon_{yy}$ and $\{\epsilon_{xx}-\epsilon_{yy}, 2\epsilon_{xy}\}$ are from the basis of A_1^{ph} and E_2^{ph} irreducible representations (IRRs) of the C_{6v} point group, respectively, and C_1 and C_2 are the only two independent nonzero elastic module tensor coefficients [37]. The lattice symmetry holds even in the presence of fluxes, so we can still use symmetry considerations to describe phonon modes at finite temperatures. Based on the analysis of Ref. [37], we can write the longitudinal/transverse acoustic phonon spectrum and the polarization vectors (defined through $\mathbf{u}_{\mathbf{q}} =$

$\sum_\nu \hat{e}^\nu_{\mathbf{q}} \tilde{u}^\nu_{\mathbf{q}}$) as

$$\Omega^{\parallel}_{\mathbf{q}} = v^{\parallel}_s q = \sqrt{\frac{C_1 + C_2}{\rho}}\, q, \quad \hat{e}^{\parallel}_{\mathbf{q}} = \{\cos\theta_{\mathbf{q}}, \sin\theta_{\mathbf{q}}\}$$
$$\Omega^{\perp}_{\mathbf{q}} = v^{\perp}_s q = \sqrt{\frac{C_2}{\rho}}\, q, \quad \hat{e}^{\perp}_{\mathbf{q}} = \{-\sin\theta_{\mathbf{q}}, \cos\theta_{\mathbf{q}}\}, \quad (5)$$

where $q = \sqrt{q_x^2 + q_y^2}$ and $\theta_{\mathbf{q}}$ is defined as the angle between $\mathbf{q}$ and $\hat{x}$ axis,

Knowing the acoustic phonon dispersion relations, we can now determine the free phonon propagator in terms of lattice displacement field $\tilde{u}_{\mathbf{q},\nu}$ as

$$D^{(0)}_{\nu\nu}(\mathbf{q}, t) = -i\langle T_t \tilde{u}^\nu_{-\mathbf{q}}(t)\tilde{u}^\nu_{\mathbf{q}}(0)\rangle^{(0)}, \quad (6)$$

where T_t is time ordering operator, $\nu = \parallel, \perp$ labels the polarization, and the quantized displacement field is given by the standard expression:

$$\tilde{u}^\nu_{\mathbf{q}}(t) = i\sqrt{\frac{\hbar}{2\rho\delta_V \Omega^\nu_{\mathbf{q}}}}\left(a_{\mathbf{q},\nu}e^{-i\Omega^\nu_{\mathbf{q}}t} + a^\dagger_{-\mathbf{q},\nu}e^{i\Omega^\nu_{\mathbf{q}}t}\right). \quad (7)$$

In the rest of the discussions we set $\hbar = 1$. Also, since in the following we will focus on the finite-temperature physics, it is convenient for us to rewrite the phonon propagator using the Matsubara formalism [58]:

$$D^{(0)}_{\nu\nu}(\mathbf{q}, i\Omega_n) = -\frac{1}{\rho\delta_V}\frac{1}{(i\Omega_n)^2 - (\Omega^\nu_{\mathbf{q}})^2}. \quad (8)$$

The *third term* term in Eq.(1) denotes the magneto-elastic coupling that arises from the change in the Kitaev coupling J_K due to the lattice vibrations. In the long wavelength limit for the acoustic phonons, assuming that J_K only depends on the distance r between the atoms, the coupling Hamiltonian can be written in a differential form:

$$\mathcal{H}_{\rm c} = -\lambda\sum_{\mathbf{r},\alpha}\mathbf{M}_\alpha\cdot(\mathbf{u}(\mathbf{r}) - \mathbf{u}(\mathbf{r}+\mathbf{M}_\alpha))\,\sigma^\alpha_{\mathbf{r}}\sigma^\alpha_{\mathbf{r}+\mathbf{M}_\alpha} \quad (9)$$
$$= \lambda\sum_{\mathbf{r},\alpha}\mathbf{M}_\alpha\cdot[(\mathbf{M}_\alpha\cdot\nabla)\,\mathbf{u}(\mathbf{r})]\,\sigma^\alpha_{\mathbf{r}}\sigma^\alpha_{\mathbf{r}+\mathbf{M}_\alpha},$$

where $\lambda \sim \left(\frac{{\rm d}J_K}{{\rm d}r}\right)_{eq} l_a$ is the strength of the spin-phonon interaction and l_a is the lattice constant. Note that since the strength of the three-spin interaction term κ depends J_K, it will also change due to the lattice vibrations. However the contribution from this term to the magneto-elastic coupling will appear at higher orders in h/J_K, and thus will be neglected.

The general symmetry allowed spin-lattice coupling can be expressed in terms of A_1 and E_2 channels independently [37].

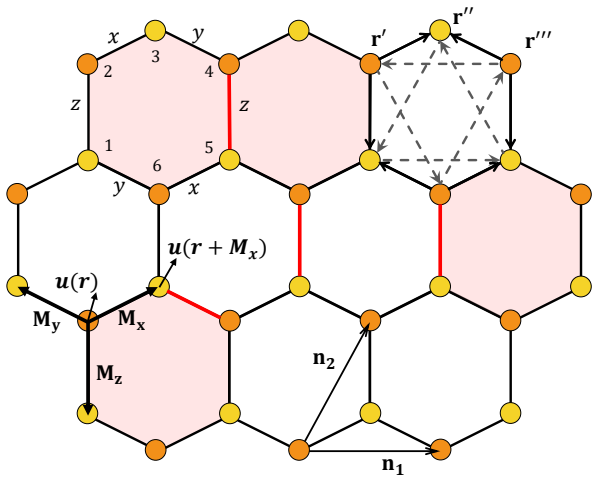

FIG. 1. Kitaev model on the honeycomb lattice with unit vectors $\mathbf{n}_1 = (1, 0)$ and $\mathbf{n}_2 = (\frac{1}{2}, \frac{\sqrt{3}}{2})$ shown in a random thermal flux sector. The sites on the A and B sublattices of the honeycomb lattice are marked by yellow and orange circles. The three vectors $\mathbf{M}_{x,y} = \left(\pm\frac{\sqrt{3}}{2}, \frac{1}{2}\right)$, and $\mathbf{M}_z = (0, -1)$ connect nearest neighbors (NN) by x, y, and z bonds, respectively. The black arrow represent the displacement vector $\mathbf{u}(\mathbf{r}) = \{u_x, u_y\}$. The red hexagon on the top left plaquette shows a π flux, i.e. the eigenvalue of $\hat{W}_p = \sigma^x_1\sigma^y_2\sigma^z_3\sigma^x_4\sigma^y_5\sigma^z_6$ is equal to -1. The red thick lines represent the link gauge variable with $\eta_{lm} = -1$, which creates or annihilates fluxes on its two adjacent plaquettes. The arrows on the top right plaquette show the convention of the signs of the NN couplings and NNN couplings, e.g. an arrow pointing from $\mathbf{r}'$ to $\mathbf{r}''$ means that the value of the NN bond $\eta_{\mathbf{r}'\mathbf{r}''}$ is by default positive in Eq. (11). The sites $\langle\mathbf{r}'\mathbf{r}''\mathbf{r}'''\rangle$ is an example of the NNN triplet used in Eq. (2) and corresponds to $\langle lnm\rangle$ in Eq. (11).

In Ref.[37], it was shown that the dominant contribution to the magneto-elastic coupling comes from the E_2 symmetry channel, and in the following we will focus only on it [59]. Then, the magnetoelastic coupling is given by

$$\mathcal{H}^{E_2}_{\rm c} = \lambda_{E_2}\sum_{\mathbf{r}\in A}\Big[2\sqrt{3}\epsilon_{xy}\left(\sigma^x_{\mathbf{r}}\sigma^x_{\mathbf{r}+\mathbf{M}_x} - \sigma^y_{\mathbf{r}}\sigma^y_{\mathbf{r}+\mathbf{M}_y}\right)$$
$$+ (\epsilon_{xx} - \epsilon_{yy})\left(\sigma^x_{\mathbf{r}}\sigma^x_{\mathbf{r}+\mathbf{M}_x} + \sigma^y_{\mathbf{r}}\sigma^y_{\mathbf{r}+\mathbf{M}_y} - 2\sigma^z_{\mathbf{r}}\sigma^z_{\mathbf{r}+\mathbf{M}_z}\right)\Big]. \quad (10)$$

B. Real space diagonalization of the Majorana fermion Hamiltonian and fermionic propagators

Using the Kitaev four Majorana fermion representation of spin [3], $\sigma_l^\alpha = ib_l^\alpha c_l$, we rewrite the spin Hamiltonian as

$$\begin{aligned}\mathcal{H}_s =& J_K \sum_{\langle lm\rangle} \frac{i}{2}\eta_{lm} c_l c_m + \kappa \sum_{\langle lnm\rangle} \frac{i}{2}\eta_{ln}\eta_{mn} c_l c_m \\ =& \frac{i}{4}\sum_{\langle lm\rangle} h_{lm} c_l c_m, \end{aligned} \tag{11}$$

where $\eta_{lm} = ib_l^\alpha b_m^\alpha$ is the Z_2 link operator on α-bond between sites l and m of the honeycomb lattice, and $h_{lm} \equiv 2J_K\eta_{lm} + 2\kappa\,\eta_{ln}\eta_{mn}$ denote matrix elements of the coupling matrix $\hat{h}$.

The Hilbert space of the fermionic model Eq. (11) is larger than that of the spin model; the latter is recovered by imposing the gauge condition, $b_l^x b_l^y b_l^z c_l|\psi_s\rangle = |\psi_s\rangle$, where $|\psi_s\rangle$ is a physical spin state, which is obtained by a projection of a Majorana fermion state $|\Psi_{\rm MF}\rangle$ solved from the Hamiltonian Eq. (11) to the physical spin state $|\psi_s\rangle$ [3]. The bond operators η_{lm} commute with the Hamiltonian Eq. (11) and take eigenvalues $\eta_{lm} = \pm 1$. A configuration of $\{\eta_{lm}\}$ can be understood as $\mathbb{Z}_2$ gauge fields, which corresponds to different gauge choices. The physical spin states, obtained after projecting out the gauge degrees of freedom, is denoted by the eigenvalue of the loop operator $\hat{W}_p$, which in the fermionic representation is expressed as $\hat{W}_p = \prod_{(l,m)\in \text{edge}(p)} \eta_{lm}$. It can be understood as a gauge invariant Wilson loop operator around a single plaquette p. It takes eigenvalue $W_p = \pm 1$. $W_p = -1$ corresponds to a (gapped) π-flux excitation on the plaquette in question. In the following we will also use notation $\{\phi_p\}$ to denote a particular flux sector, described by choosing a particular configuration of the bond values $\{\eta_{lm}\}$.

A quantitative description of the finite-temperature fermionic spectrum is obtained by solving the free-fermion problem (11) in each flux sector exactly. To this end, we first specify the link variable for each bond and then determine the physically relevant flux sector. After the flux sector is determined, the Hamiltonian (11) can be solved exactly as a tight-binding model of Majorana fermions [3].

At $\kappa = 0$, the Hamiltonian matrix $\hat{h}$ is anti-symmetric, i.e. $\hat{h}^\top = -\hat{h}$. Thus, it can be written in a block off-diagonal form $\hat{h} = \begin{bmatrix} 0 & \hat{M} \\ -\hat{M}^\top & 0 \end{bmatrix}$. It can be block-diagonalized by the orthogonal matrix $\hat{Q} \in O(2N)$, where N is the number of unit cells, given by $\hat{Q} = \begin{bmatrix} \hat{u} & 0 \\ 0 & \hat{v} \end{bmatrix}$, which can be obtained by singular value decomposition. Then, the orthonormal Majorana modes are given by

$$\begin{aligned}\gamma_{A,i} &= \sum_{\mathbf{r}} u_{\mathbf{r}i} c_{A,\mathbf{r}}\,, \\ \gamma_{B,i} &= \sum_{\mathbf{r}} v_{\mathbf{r}i} c_{B,\mathbf{r}}\,, \end{aligned} \tag{12}$$

and $\hat{M} = \hat{u}\,\hat{\Omega}\,\hat{v}^\top$, where $\hat{\Omega} = \text{diag}(\epsilon_1, \epsilon_2, ..., \epsilon_N)$ is the diagonal energy matrix. In terms of complex fermions through

$$\beta_i = \left(\gamma_{A,i} + i\gamma_{B,i}\right)/2, \tag{13}$$

the Hamiltonian can be written in the canonical form:

$$\mathcal{H}_s = \sum_{i=1}^{N} \epsilon_i \left(\beta_i^\dagger \beta_i - 1/2\right), \tag{14}$$

where β_i are complex matter fermions which label the eigenmodes with the fermion energies ϵ_i for a given flux sector. From here we can obtain the expressions for the fermion propagators. Since we are interested in finite-temperature picture, we use the Matsubara formalism [58]:

$$\begin{aligned} g_i(i\omega_n) &= -\langle T_\tau \beta_i(\tau)\beta_i^\dagger(0)\rangle_{\omega_n} = \frac{1}{i\omega_n - \epsilon_i}, \\ \bar{g}_i(i\omega_n) &= -\langle T_\tau \beta_i^\dagger(\tau)\beta_i(0)\rangle_{\omega_n} = \frac{1}{i\omega_n + \epsilon_i}, \end{aligned} \tag{15}$$

where $\langle \hat{O}\rangle_{\omega_n} = \int_0^\beta \mathrm{d}\tau e^{i\omega_n\tau}\langle \hat{O}(\tau)\rangle$, and T_τ is imaginary time ordering operator. Next, to describe the diagonalization of the Hamiltonian Eq. (11), we introduce the following matrix form notations. We denote the vector of Majorana fermions as $\mathbf{C}^\top = [c_{A,1}, \cdots c_{A,N}, c_{B,1}\cdots c_{B,N}]$, and the vector of the Bogoliubov quasiparticles as $\mathbf{B}^\dagger = \left[\beta_1^\dagger, \cdots \beta_N^\dagger, \beta_1, \cdots \beta_N\right]$. Then the transformations Eq. (12-13) can be written as a unitary matrix: $\mathbf{B}^\dagger = \frac{1}{\sqrt{2}}\mathbf{C}^\dagger \hat{U}$, where $\hat{U}$ is the unitary matrix that diagonalizes the Hamiltonian matrix $\hat{h}$, and the factor $\frac{1}{\sqrt{2}}$ is used to recover the correct anti-commutation relation of the complex fermions. The vector $\mathbf{B}^\dagger$ has the dimension $2N$ and is naturally divided into the space of creation operators corresponding to the positive energy spectrum and the space of annihilation operators corresponding to the negative energy spectrum.

C. Mixed representation of the Majorana fermion-phonon (MFPh) coupling

Now we can rewrite the coupling Hamiltonian (10) using the Majorana fermion representation of spin as

$$\mathcal{H}_{\rm c}^{E_2} = -i\lambda_{E_2}\sum_{\mathbf{r}\in A}\Big[2\sqrt{3}\,\epsilon_{xy}\left(\eta_{\mathbf{r},\mathbf{r}+\mathbf{M}_x}c_{\mathbf{r}}c_{\mathbf{r}+\mathbf{M}_x}-\eta_{\mathbf{r},\mathbf{r}+\mathbf{M}_y}c_{\mathbf{r}}c_{\mathbf{r}+\mathbf{M}_y}\right)+(\epsilon_{xx}-\epsilon_{yy}) \left(\eta_{\mathbf{r},\mathbf{r}+\mathbf{M}_x}c_{\mathbf{r}}c_{\mathbf{r}+\mathbf{M}_x}+\eta_{\mathbf{r},\mathbf{r}+\mathbf{M}_y}c_{\mathbf{r}}c_{\mathbf{r}+\mathbf{M}_y}-2\eta_{\mathbf{r},\mathbf{r}+\mathbf{M}_z}c_{\mathbf{r}}c_{\mathbf{r}+\mathbf{M}_z}\right)\Big]. \tag{16}$$

Next we will use a mixed representation, i.e. we will transform the strain tensor into the momentum space using the long wavelength limit, $\epsilon_{\alpha\beta}(\mathbf{r}) = \frac{1}{\sqrt{N}}\sum_{\mathbf{q}}\frac{i}{2}\left(q_\alpha u_{\mathbf{q},\beta}+q_\beta u_{\mathbf{q},\alpha}\right)e^{i\mathbf{q}\cdot\mathbf{r}_l}$, where α,β denote x or y, but keep Majorana fermions in the real space. The coupling Hamiltonian (16) then reads

$$\mathcal{H}_{\rm c} = \frac{1}{\sqrt{N}}\sum_{\mathbf{q}}V_{\mathbf{q}},$$
$$V_{\mathbf{q}} = -\frac{i}{2}\sum_{\langle l\in A,m\in B\rangle}c_l c_m\left(\hat{\lambda}^{\parallel}_{\mathbf{q},lm}\tilde{u}^{\parallel}_{q}+\hat{\lambda}^{\perp}_{\mathbf{q},lm}\tilde{u}^{\perp}_{\mathbf{q}}\right)e^{i\mathbf{q}\cdot\mathbf{r}_l}, \tag{17}$$

where the MFPh coupling vertices are given by

$$\begin{aligned}\hat{\lambda}^{\parallel}_{\mathbf{q},lm} =& 2i\lambda\Big\{\delta_{\langle lm\rangle_x}\eta_{lm}\Big[(c_{\mathbf{q}}+\sqrt{3}s_{\mathbf{q}})q_x+(\sqrt{3}c_{\mathbf{q}}-s_{\mathbf{q}})q_y\Big] \\ &+\delta_{\langle lm\rangle_y}\eta_{lm}\Big[(c_{\mathbf{q}}-\sqrt{3}s_{\mathbf{q}})q_x+(-\sqrt{3}c_{\mathbf{q}}+s_{\mathbf{q}})q_y\Big] \\ &+\delta_{\langle lm\rangle_z}\eta_{lm}\,2\left(-c_{\mathbf{q}}q_x+s_{\mathbf{q}}q_y\right)\Big\}, \end{aligned}\tag{18}$$
$$\begin{aligned}\hat{\lambda}^{\perp}_{\mathbf{q},lm} =& 2i\lambda\Big\{\delta_{\langle lm\rangle_x}\eta_{lm}\Big[(\sqrt{3}c_{\mathbf{q}}-s_{\mathbf{q}})q_x+(-c_{\mathbf{q}}-\sqrt{3}s_{\mathbf{q}})q_y\Big] \\ &+\delta_{\langle lm\rangle_y}\eta_{lm}\Big[(-\sqrt{3}c_{\mathbf{q}}-s_{\mathbf{q}})q_x+(-c_{\mathbf{q}}+\sqrt{3}s_{\mathbf{q}})q_y\Big] \\ &+\delta_{\langle lm\rangle_z}\eta_{lm}\,2\left(s_{\mathbf{q}}q_x+c_{\mathbf{q}}q_y\right)\Big\}. \end{aligned}\tag{19}$$

Here, for the compactness of the equations we denote $c_{\mathbf{q}} = \cos\theta_{\mathbf{q}}$ and $s_{\mathbf{q}} = \sin\theta_{\mathbf{q}}$, and $\delta_{\langle lm\rangle_\alpha} = 1$ when $\langle lm\rangle$ is the nearest neighboring link of type $\alpha\in\{x,y,z\}$, and zero otherwise. Since l and m always belong to different A and B sublattices, we can write the coupling Hamiltonian Eq. (17) as

$$V_{\mathbf{q}} = -\frac{i}{4}\sum_{l,m}c_l c_m\Lambda^{\mu}_{\mathbf{q},lm}\tilde{u}^{\mu}_{\mathbf{q}} \tag{20}$$

where the coupling matrices $\Lambda^{\mu}_{\mathbf{q},lm}$, $\mu=\parallel,\perp$, in the sublattice matrix representation are given by

$$\Lambda^{\mu}_{\mathbf{q},lm} = \begin{bmatrix} O & \hat{\lambda}^{\mu}_{\mathbf{q}} \\ \hat{\lambda}^{\mu\dagger}_{\mathbf{q}} & O \end{bmatrix}_{lm} \odot \begin{bmatrix} O & \cdots & \cdots & \cdots \\ & \cdots & e^{i\mathbf{q}\cdot\mathbf{r}_l} & \cdots \\ & \cdots & \cdots & \cdots \\ \vdots\ \vdots\ \vdots & & & \\ \vdots\ e^{i\mathbf{q}\cdot\mathbf{r}_m}\ \vdots & & O & \\ \vdots\ \vdots\ \vdots & & & \end{bmatrix}_{lm}, \tag{21}$$

where $\odot$ is element-wise multiplication. Notice that in the spin-lattice coupling Eq. (17), the plane wave phase factor is applied only on the A sites. To make it invariant under the C_{6v}, a further symmetrization of the coupling matrix between between A and B sublattices is required. By performing this symmetrization, we obtained the symmetrized spin-lattice coupling between the eigenmodes of the Hamiltonian:

$$\begin{aligned}V_{\mathbf{q}} &= -\frac{i}{2}\sum_{\langle l\in A,m\in B\rangle}c_l c_m\hat{\lambda}^{\mu}_{\mathbf{q},lm}\tilde{u}^{\mu}_{\mathbf{q}}\frac{1}{2}(e^{i\mathbf{q}\cdot\mathbf{r}_l}+e^{i\mathbf{q}\cdot\mathbf{r}_m}) \\ &= -\frac{i}{4}\sum_{l,m}c_l c_m\Lambda^{s,\mu}_{\mathbf{q},lm}\tilde{u}^{\mu}_{\mathbf{q}} = -\frac{i}{2}\mathbf{B}^{\dagger}\tilde{\Lambda}^{\mu}_{\mathbf{q}}\mathbf{B}\,\tilde{u}^{\mu}_{\mathbf{q}}\end{aligned} \tag{22}$$

where $\tilde{\Lambda}^{\mu}_{\mathbf{q}} \equiv \hat{U}^{\dagger}\Lambda^{s,\mu}_{\mathbf{q}}\hat{U}$ is the symmetrized coupling matrix, whose entries are the coupling vertices between two fermion eigenmodes. The coupling matrices $\tilde{\Lambda}^{\mu}_{\mathbf{q}}$ can also be divided into four blocks according to the division into the creation and annihilation subspace:

$$\tilde{\Lambda}^{\mu}_{\mathbf{q}} \equiv \begin{bmatrix} \tilde{\Lambda}^{\mu}_{\mathbf{q},11} & \tilde{\Lambda}^{\mu}_{\mathbf{q},12} \\ \tilde{\Lambda}^{\mu}_{\mathbf{q},21} & \tilde{\Lambda}^{\mu}_{\mathbf{q},22} \end{bmatrix}. \tag{23}$$

These coupling matrices will be used in the final expression of the polarization bubble, which we will derive in the next section.

III. PHONON POLARIZATION BUBBLE

Next we discuss the effect of the spin-lattice coupling on the phonon dynamics. By calculating the one-loop phonon self-energy $\Pi^{\mu\nu}_{ph}(\mathbf{q},\Omega)$, the renormalization to the sound velocity, mixing of the transverse and longitudinal phonon modes, and attenuation/absorption of sound wave may be obtained. Using Eq. (22), the phonon self-energy or, equivalently, the polarization bubble, can be expressed as

$$\Pi^{\mu\nu}(\mathbf{q},\tau) = \left\langle T_\tau\left(\mathbf{B}^{\dagger}\tilde{\Lambda}^{\mu}_{\mathbf{q}}\mathbf{B}\right)(\tau)\left(\mathbf{B}^{\dagger}\tilde{\Lambda}^{\nu}_{-\mathbf{q}}\mathbf{B}\right)(0)\right\rangle, \tag{24}$$

where the factor $\frac{1}{2}$ from the Eq. (22) has been absorbed into the definition of $\tilde{\Lambda}_{\mathbf{q}}$. By using the Wick's theorem, the polarization bubble can be explicitly written as:

$$\Pi^{\mu\nu}(\mathbf{q},\tau) = \Big\langle T_\tau \mathbf{B}_k^\dagger(\tau)\,\mathbf{B}_m(0)\Big\rangle \Big\langle T_\tau \mathbf{B}_l(\tau)\,\mathbf{B}_n^\dagger(0)\Big\rangle \tilde{\Lambda}^\mu_{kl,\mathbf{q}} \tilde{\Lambda}^\nu_{nm,-\mathbf{q}} - \Big\langle T_\tau \mathbf{B}_k^\dagger(\tau)\,\mathbf{B}_n^\dagger(0)\Big\rangle \Big\langle T_\tau \mathbf{B}_l(\tau)\,\mathbf{B}_m(0)\Big\rangle \tilde{\Lambda}^\mu_{kl,\mathbf{q}} \tilde{\Lambda}^{\nu\,\top}_{mn,-\mathbf{q}}, \tag{25}$$

where the indices k, l, m, n range from 1 to $2N$. Performing the Fourier transform and using the Eq. (23), we then get

$$\begin{aligned}\Pi^{\mu\nu}(\mathbf{q}, i\Omega_m) = \mathrm{Tr}\,[\\ &\bar{g}(i\omega_{n_1})\,\tilde{\Lambda}^\mu_{\mathbf{q},11}\,g(i\omega_{n_2})\,\tilde{\Lambda}^\nu_{\mathbf{q},11} + g(i\omega_{n_1})\,\tilde{\Lambda}^\mu_{\mathbf{q},21}\,g(i\omega_{n_2})\,\tilde{\Lambda}^\nu_{\mathbf{q},12}\\ +&\bar{g}(i\omega_{n_1})\,\tilde{\Lambda}^\mu_{\mathbf{q},12}\,\bar{g}(i\omega_{n_2})\,\tilde{\Lambda}^\nu_{\mathbf{q},21} + g(i\omega_{n_1})\,\tilde{\Lambda}^\mu_{\mathbf{q},22}\,\bar{g}(i\omega_{n_2})\,\tilde{\Lambda}^\nu_{\mathbf{q},22}\\ -&\bar{g}(i\omega_{n_1})\,\tilde{\Lambda}^\mu_{\mathbf{q},11}\,g(i\omega_{n_2})\,\tilde{\Lambda}^{\nu\top}_{\mathbf{q},22} - g(i\omega_{n_1})\,\tilde{\Lambda}^\mu_{\mathbf{q},21}\,g(i\omega_{n_2})\,\tilde{\Lambda}^{\nu\top}_{\mathbf{q},12}\\ -&\bar{g}(i\omega_{n_1})\,\tilde{\Lambda}^\mu_{\mathbf{q},12}\,\bar{g}(i\omega_{n_2})\,\tilde{\Lambda}^{\nu\top}_{\mathbf{q},21} - g(i\omega_{n_1})\,\tilde{\Lambda}^\mu_{\mathbf{q},22}\,\bar{g}(i\omega_{n_2})\,\tilde{\Lambda}^{\nu\top}_{\mathbf{q},11}\Big],\end{aligned} \tag{26}$$

where $\mathrm{Tr}[...]$ now sums over the Matsubara frequencies $i\omega_n$ as $T\sum_n$, and the energy conservation constraint $\Omega_m - \omega_{n_1} - \omega_{n_2} = 0$ is imposed. Here, $g(i\omega_n) = \mathrm{diag}(\cdots g_i(i\omega_n)\cdots)$ and $\bar{g}(i\omega_n) = \mathrm{diag}(\cdots \bar{g}_i(i\omega_n)\cdots)$ are the diagonal matrices. $\Pi^{\mu\nu}(\mathbf{q}, i\Omega_m)$ can also be conveniently written in a matrix form:

$$\begin{aligned}\Pi^{\mu\nu}(\mathbf{q}, i\Omega_m) = -\,\mathrm{Tr}\,[\\ G_1(i\omega_{n_1})\tilde{\Lambda}^\mu_{\mathbf{q}} G_1^*(i\omega_{n_2})\tilde{\Lambda}^\nu_{-\mathbf{q}} + G_2(i\omega_{n_1})\tilde{\Lambda}^\mu_{\mathbf{q}} G_2(i\omega_{n_2})\tilde{\Lambda}^{\nu\top}_{-\mathbf{q}}\Big],\end{aligned} \tag{27}$$

where

$$G_1(i\omega_n) \equiv \begin{bmatrix} \bar{g}(i\omega_n) & O \\ O & g(i\omega_n)\end{bmatrix},$$
$$G_2(i\omega_n) \equiv \begin{bmatrix} O & g(i\omega_n) \\ \bar{g}(i\omega_n) & O\end{bmatrix}, \tag{28}$$

and $\bar{g}(i\omega_n) = -g^*(i\omega_n)$. From Eq. (26), it is clear that the frequency dependence only appears in the denominator of the fermion propagators. Thus, we can explicitly sum over the Matsubara frequences, then take analytical continuation: $i\Omega_m \to \Omega + i\delta$ and obtain the final expression of phonon polarization bubble:

$$\begin{aligned}\Pi^{\mu\nu}(\mathbf{q},\Omega) = \frac{1}{N}\sum_{ij}\\ &\Big[P_{ij}^{\bar{g}g}\Big[\tilde{\Lambda}^\mu_{\mathbf{q},11}\Big]_{ij}\Big[\tilde{\Lambda}^\nu_{\mathbf{q},11}\Big]_{ij} + P_{ij}^{gg}\Big[\tilde{\Lambda}^\mu_{\mathbf{q},21}\Big]_{ij}\Big[\tilde{\Lambda}^\nu_{\mathbf{q},12}\Big]_{ij}\\ &+ P_{ij}^{\bar{g}\bar{g}}\Big[\tilde{\Lambda}^\mu_{\mathbf{q},12}\Big]_{ij}\Big[\tilde{\Lambda}^\nu_{\mathbf{q},21}\Big]_{ij} + P_{ij}^{g\bar{g}}\Big[\tilde{\Lambda}^\mu_{\mathbf{q},22}\Big]_{ij}\Big[\tilde{\Lambda}^\mu_{\mathbf{q},22}\Big]_{ij}\\ &- P_{ij}^{\bar{g}g}\Big[\tilde{\Lambda}^\mu_{\mathbf{q},11}\Big]_{ij}\Big[\tilde{\Lambda}^\nu_{\mathbf{q},22}\Big]_{ij} - P_{ij}^{gg}\Big[\tilde{\Lambda}^\mu_{\mathbf{q},21}\Big]_{ij}\Big[\tilde{\Lambda}^\mu_{\mathbf{q},12}\Big]_{ij}\\ &-P_{ij}^{\bar{g}\bar{g}}\Big[\tilde{\Lambda}^\mu_{\mathbf{q},12}\Big]_{ij}\Big[\tilde{\Lambda}^\nu_{\mathbf{q},21}\Big]_{ij} - P_{ij}^{g\bar{g}}\Big[\tilde{\Lambda}^\mu_{\mathbf{q},22}\Big]_{ij}\Big[\tilde{\Lambda}^\nu_{\mathbf{q},11}\Big]_{ij}\Big],\end{aligned} \tag{29}$$

where $P_{ij}^{g\bar{g}}, P_{ij}^{\bar{g}g}, P_{ij}^{gg}$ and $P_{ij}^{\bar{g}\bar{g}}$ are obtained from the Matsubara summation over the frequencies and are explicitly given in Eq. (A1) of App. A. We will see in the next section that since $P_{ij}^{g\bar{g}}, P_{ij}^{\bar{g}g}, P_{ij}^{gg}$ and $P_{ij}^{\bar{g}\bar{g}}$ encode the information of the fermionic energy spectrum, they determine the intensity of the Majorana fermion-phonon scattering and its dependence on temperature. We will also see that the matrix elements $\tilde{\Lambda}^\mu_{\mathbf{q},ij}$, describing the coupling between the i-th and j-th fermionic eigenmodes and the acoustic phonon with wavevector $\mathbf{q}$ and the polarization μ, are responsible for the angular dependence of the Majorana fermion-phonon scattering.

IV. THE SOUND ATTENUATION COEFFICIENT

The quantitative description of the attenuation process can be obtained through the lossy acoustic wave function which decays with distance away from the driving source as

$$\mathbf{u}(\mathbf{r},t) = \mathbf{u}_0 e^{-\alpha_s(\mathbf{q})r} e^{i(\Omega t - \mathbf{q}\cdot\mathbf{r})}, \tag{30}$$

where $\alpha_s(\mathbf{q})$ is the attenuation coefficient defined as the inverse of the phonon mean free path, which can be computed from the imaginary part of the diagonal components of the phonon self-energy [37, 60]:

$$\alpha_s^\mu(\mathbf{q}) \propto -\frac{1}{v_s^2 q}\,\mathrm{Im}\Big[\Pi^{\mu\mu}_{\mathrm{ph}}(\mathbf{q},\Omega)\Big]_{\Omega=v_s q}, \tag{31}$$

where μ is the polarization component. In Ref. [37] we have shown that at low temperatures below the energy scale of the flux gap, the sound attenuation is determined by the decay of a phonon into a pair of Majorana fermions, with the attenuation rate linear in temperature due to the density of states that scales linearly in energy at the Dirac points. [Note that this only applies when $v_s < v_F$.] The question on how the presence of thermally excited fluxes modifies this picture is addressed in this section.

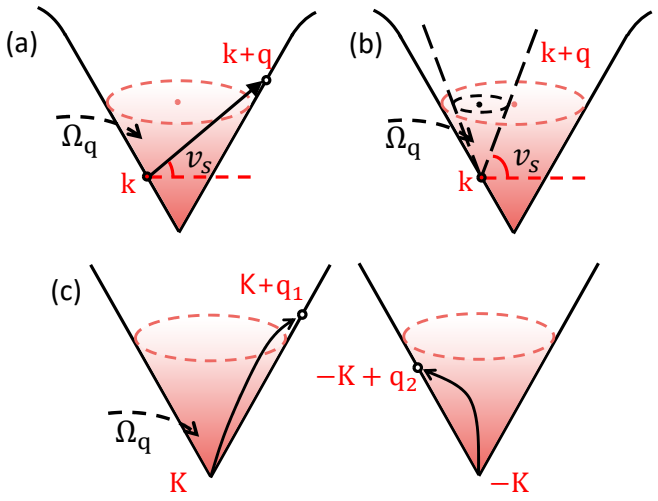

FIG. 2. The illustration of different scattering channels near the bottom of the Dirac cone with different acoustic phonon velocities v_s: (a) ph-channel with $v_s < v_F$, (b) ph-channel with $v_s > v_F$, (c) pp-channel with $v_s > v_F$. Here, v_F is the Fermi velocity visualized by the slope of the Dirac cone. The phonon energy is $\Omega_{\mathbf{q}} = v_s|\mathbf{q}|$. In scenario (a), the kinematic constraint is satisfied for a hopping from a fermionic mode $\mathbf{k}$ to $\mathbf{k}+\mathbf{q}$. In scenario (b), the dashed cone illustrating a hopping from a fermionic mode $\mathbf{k}$ to any of the possible final states that satisfies the kinematic constraint has no intersection with the fermionic cone (solid line) since $v_s > v_F$. Thus, this scenario is not possible.

A. The kinematic constraints in phonon dynamics

Existence of the coupling between the Majorana fermions and phonons is necessary but not a sufficient condition for the decay of the phonon into the low-energy Majorana fermions modes. In the case of the translationally invariant system, the decay rate is governed by the kinematic conditions reflecting the conservation of energy and momentum. At low temperatures, when the Kitaev spin liquid can be considered in its ground state zero-flux sector, both energy and momenta constraints are present. At finite temperatures, the thermally excited fluxes destroy the translational symmetry such that a fermionic momentum is not any more a good quantum number. Thus, at finite temperatures the decay rate is determined by a weaker kinematic conditions. In addition, the disorder from the thermal fluxes destroys the Dirac cones of the Majorana fermions and flattens their density of states over the whole energy range, so that the scattering of the phonons is happening on a very different manifold of low-energy fermionic states.

In this section, we analyze the phonon decay in the zero-flux sector, which dominates at low temperatures. The relative strength of the sound velocity, v_s, and the Fermi velocity, v_F, that characterizes the slope of the low-energy Majorana Dirac cone defines the phase space for the decay and determines whether the decay happens in the particle-hole (ph-) or in the particle-particle (pp-) channel. Here, by particle and hole, we mean if the state i with energy $\epsilon_i > 0$ in Eq. (14) is occupied or empty. In other words, the particle number refers to that of the complex fermion β_i.

The kinematic constraint for a ph-process, in which a phonon mode with $\Omega_{\mathbf{q}} = v_s|\mathbf{q}|$ scatters a fermion from the state at $\mathbf{k}$ to the state at $\mathbf{k}+\mathbf{q}$ is shown schematically in Fig. 2 (a). As the maximum energy difference between the two states is $v_F|\mathbf{q}|$, the ph-process is allowed kinematically only when $v_s < v_F$. Moreover, since in order for this process to happen, some positive energy state, e.g. $\beta^\dagger_{\mathbf{k}}|0\rangle$ shown in Fig. 2 (a), must be occupied, finite temperature is required.

For a pp-process, illustrated in Fig. 2 (c), in which a phonon decays into two fermions with positive energies, the kinematic constraints require:

$$\begin{aligned}\Omega_{\mathbf{q}} &= \epsilon_{\mathbf{K}+\mathbf{q}_1} + \epsilon_{-\mathbf{K}+\mathbf{q}_2} = v_F(|\mathbf{q}_1| + |\mathbf{q}_2|) \\ &\geq v_F|\mathbf{q}_1 + \mathbf{q}_2| = v_F|\mathbf{q}|, \end{aligned} \tag{32}$$

where $\mathbf{K}+\mathbf{q}_1$, $-\mathbf{K}+\mathbf{q}_2$ are, respectively, the momenta of the pair of fermion particles produced by the phonon, and we have expanded the energy of a single fermion $\epsilon_{\mathbf{K}+\mathbf{q}} = 2J_K|\sum_{\alpha=x,y,z} e^{i(\mathbf{K}+\mathbf{q})\cdot\mathbf{M}_\alpha}| \simeq v_F|\mathbf{q}|$ near the corresponding Dirac points. Thus, the pp-process is lower bounded by $v_F|\mathbf{q}| = \epsilon_{K+\mathbf{q}}$. Clearly, when $v_s < v_F$ the phonon doesn't have enough energy to produce a pair of fermion particles, and, resultantly, in this limit the low-temperatures sound attenuation is mainly caused by the ph-scattering. On contrary, when $v_s \geqslant v_F$, the constraint Eq. (32) is easily satisfied, leading to the decay of the phonon into a pair of Majorana fermions. Unlike the ph-process, the pp-process does not require a finite occupation of the fermionic states and, in principle, can happen at zero temperature (when phonons are externally pumped into the system) as long as the incident phonon has enough energy to excite a pair of particles. As illustrated in Fig. 2 (b), when $v_s \geqslant v_F$ the kinematic constraints for the ph-process can not be satisfied, an thus the sound attenuation coefficient is entirely determined by the pp-scattering.

Fig. 3 presents the imaginary part of the longitudinal com-

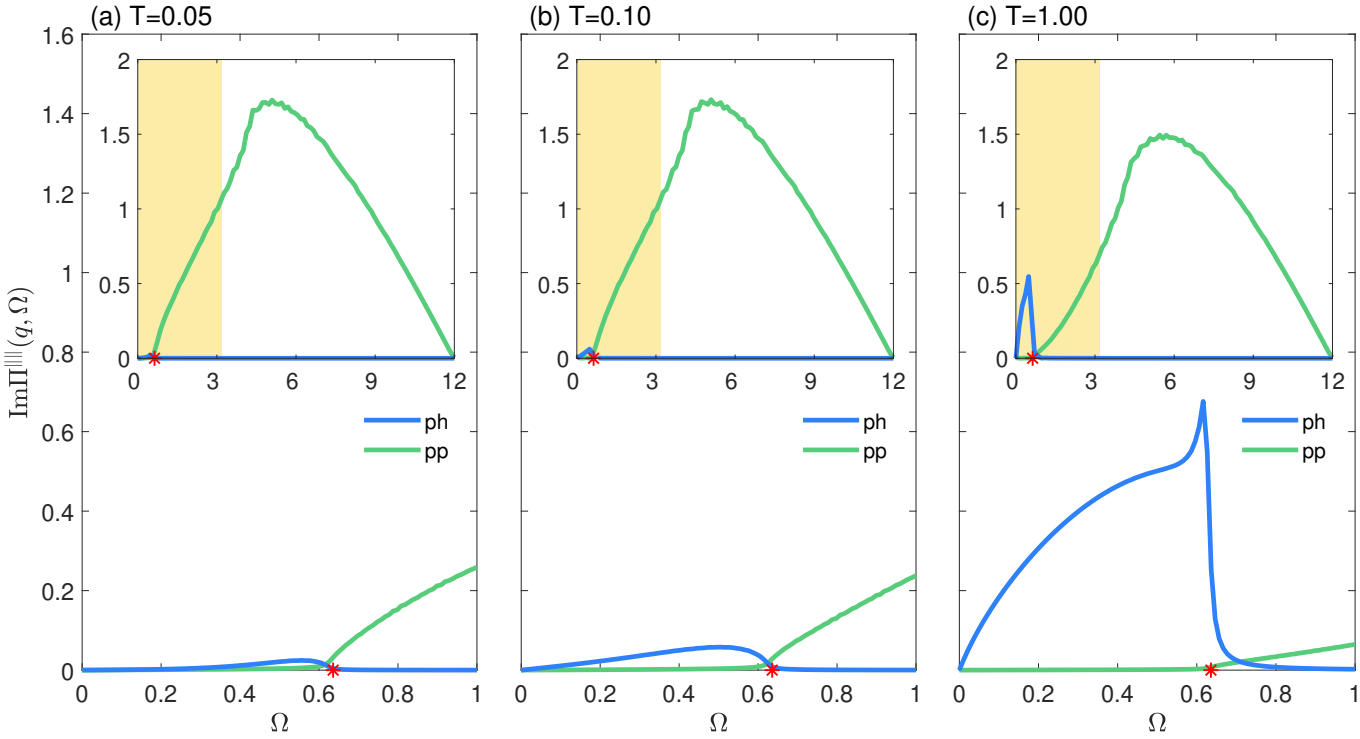

FIG. 3. The imaginary part of the phonon self-energy, $\mathrm{Im}\Pi^{|||}(\mathbf{q},\Omega)$ as a function of the phonon frequency Ω and fixed phonon momentum $\mathbf{q}=(0.1\pi,0)$ computed assuming the zero-flux sector for various temperatures. The ph-channel and pp-channel are shown as blue and green curves, respectively. The calculations are performed on the finite size lattice with $N_1=N_2=1000$. We use $\delta=0.01$ for imaginary energy broadening. The red star at the frequency $\Omega^*=\epsilon_{\mathbf{K}+\mathbf{q}}=v_F|\mathbf{q}|$, where $\mathbf{K}$ denoting one of the Dirac points, corresponds to the upper bound for the ph-continuum and the lower bound of the pp-continuum. The inset in each of the panel shows the ph- and pp-contributions to $\mathrm{Im}\Pi^{|||}(\mathbf{q},\Omega)$ in a wide range of the phonon frequencies. The shaded region in each of the insets highlights the range of the phonon frequencies in which the long wavelength description of the phonon media is applicable. The frequency, Ω, and temperature, T, are measured in units of the Kitaev interaction, J_K.

ponent of the phonon self-energy, $\mathrm{Im}\Pi^{|||}(\mathbf{q},\Omega)$, as a function of phonon frequency in the zero-flux sector for (a) $T=0.05$, (b) $T=0.1$ and (c) $T=1.0$. The contributions from the ph-channel and pp-channel computed at fixed phonon momentum $\mathbf{q}=(0.1\pi,0)$ are plotted with blue and green curves, respectively.

The main panels of Fig. 3 show the ph- and pp-contributions to $\mathrm{Im}\Pi^{|||}(\mathbf{q},\Omega)$ computed in the low-frequency region, while the inset in each of the panel shows them in a wider range of the phonon frequencies. The red star at the frequency $\Omega^*=v_F|\mathbf{q}|=0.1\pi v_F$ corresponds to the upper bound for the ph- and the lower bound of the pp-continuum of scattering. By comparing the panels (a), (b) and (c) of Fig. 3, we can see that the magnitude of the pp-channel contribution into $\mathrm{Im}\Pi^{|||}(\mathbf{q},\Omega)$ is almost independent on the temperature. This is because the pp-channel does not require a finite particle population and the decay of the phonon into a pair of Majorana fermions can happen even at zero temperature, as long as the incident phonon has enough energy to excite a pair of particles. On the other hand, the ph-channel requires a finite population of the fermionic states, and its intensity increases with increasing temperature leading to larger population of the low-energy with fermionic states. Note, however, that with further increase of temperature, there will be a critical temperature at which the particle and the hole population become balanced, and thus above this temperature the scattering probabilities will be suppressed. Clearly, at $T\to\infty$, the intensities of both the ph- and pp-channels will decay to zero due to the Pauli exclusion principle. Similar plots for the zero-flux sector and the random flux sectors were shown in Ref. [36].

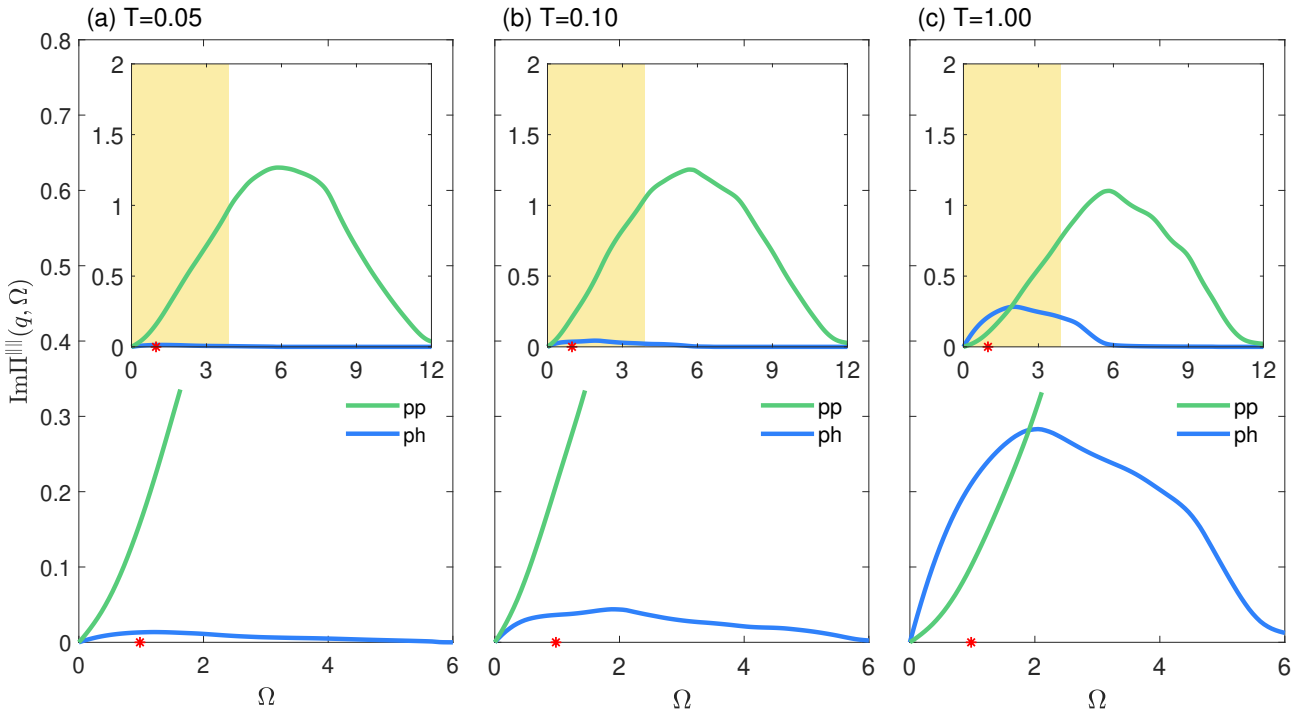

FIG. 4. The imaginary part of the phonon self-energy $\mathrm{Im}\Pi^{\parallel\parallel}(\mathbf{q},\Omega)$ as a function of the phonon frequency Ω and fixed phonon momentum $\mathbf{q}=(0.1\pi,0)$ computed using Eq. (29) for (a) $T=0.05$, (b) $T=0.1$ and (c) $T=1.0$. For every temperature, 200 inhomogeneous flux configurations are sampled with the help of the strMC algorithm (see Appendix F) implemented on the finite size lattice with $N_1=N_2=32$. The ph-channel and pp-channel are shown as blue and green curves, respectively. We use $\delta=0.2$ for the imaginary energy broadening. The red star at the frequency $\Omega^*=\epsilon_{\mathbf{K}+\mathbf{q}}=v_F|\mathbf{q}|$, where $\mathbf{K}$ denoting one of the Dirac points, corresponds to the upper bound for the ph-continuum and the lower bound of the pp-continuum in the zero-flux sector. The inset in each of the panel shows the ph- and pp-contributions to $\mathrm{Im}\Pi^{\parallel\parallel}(\mathbf{q},\Omega)$ in a wide range of the phonon frequencies. The shaded region in each of the insets highlights the range of the phonon frequencies in which the long wavelength description of the phonon media is applicable. The frequency, Ω, and temperature, T, are measured in units of the Kitaev interaction, J_K.

B. Phonon dynamics in the presence of thermal fluxes

In the inhomogeneous thermal flux sectors, which are relevant at elevated temperatures, the constraint of momentum conservation is relaxed. Therefore, at finite temperatures when flux proliferates, the decay of a phonon involves both the ph- and pp-processes. In order to illustrate this, in Fig. 4 we plot the imaginary part of the phonon self-energy computed with Eq. (29) as a function of phonon energy in the thermal flux sectors for (a) $T=0.05$, (b) $T=0.1$ and (c) $T=1.0$. Again we focus on the diagonal longitudinal component of the polarization bubble $\mathrm{Im}\Pi^{\parallel\parallel}(\mathbf{q},\Omega)$. The sampling over flux configurations is performed with a help of the stratified Monte Carlo (strMC) algorithm (see Appendix F for details) which exploits the fact that the energy of each flux configuration can be computed exactly by diagonalizing the quadratic Majorana Hamiltonian (14). The results are computed at fixed phonon momentum $\mathbf{q}=(0.1\pi,0)$ and are averaged over 200 inhomogeneous flux configurations for each data point. In all panels, the red star at the phonon frequency $\Omega^*=v_F|\mathbf{q}|=0.1\pi v_F$ marks the critical value of the phonon frequency below which in the zero-flux sector the phonon decays into the ph-continuum and above which into the pp-continuum.

Fig. 4 shows the effects of the relaxation of the kinematic constraints and the modifications of the fermionic spectrum due to the thermal flux. In particular, we can clearly see that the contribution from the phonon decay into the ph-continuum significantly increases with temperature, while the contribution from the decay into the pp-continuum remains almost unchanged.

Fig. 4 (a) shows $\mathrm{Im}\Pi^{\parallel\parallel}(\mathbf{q},\Omega)$ computed at temperature $T=0.05$, which is well below the flux gap. At this temperature, the MC sampling is done predominantly in the zero-flux

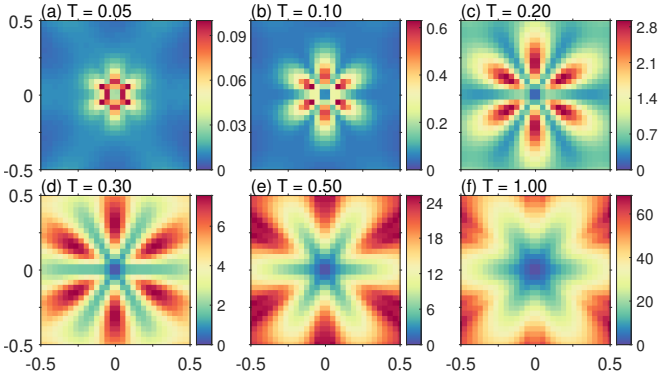

FIG. 5. Case $v_s < v_F$: The temperature evolution of the sound attenuation coefficient $\alpha_s^{||}(\mathbf{q})$ computed in the zero-flux sector. The diagonalization of the Majorana fermion Hamiltonian Eq. (11) is done in the momentum space on the lattice with $N_1 = N_2 = 500$. The phonon momentum $\mathbf{q}$ belongs to the region $(q_x, q_y) \in [-0.5\pi, 0.5\pi]^2$. In the calculations, we set $v_s = 0.1v_F$ and the imaginary energy broadening $\delta = 0.2$.Temperature is measured in units of the Kitaev interaction, J_K.

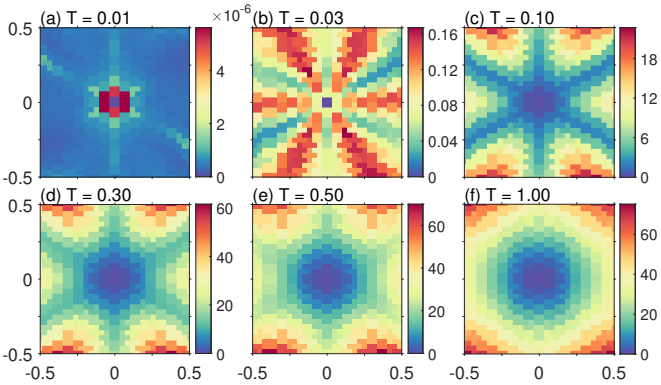

FIG. 6. Case $v_s < v_F$: The temperature evolution of the sound attenuation coefficient $\alpha_s^{||}(\mathbf{q})$ computed in the inhomogeneous flux sectors sampled by the strMC method. The lattice size is $N_1 = N_2 = 32$, and each data point is obtained by averaging over 200 flux realizations. The phonon momentum $\mathbf{q}$ is in region $(q_x, q_y) \in [-0.5\pi, 0.5\pi]^2$. In the calculations, we set $v_s = 0.1v_F$ and the imaginary energy broadening $\delta = 0.2$. Temperature is measured in units of the Kitaev interaction, J_K.

sector or in a very low-density flux sectors. Thus, the pp- and the ph-contribution to $\mathrm{Im}\Pi^{|||}(\mathbf{q}, \Omega)$ should be almost identical to those shown in Fig. 3 (a). Nevertheless, we can see in Fig. 4 (a) that the contribution from the pp-channel at frequencies below Ω^* is not zero and is notably larger than the contribution from the ph-channel, which violates the kinematic constraints discussed above. This unphysical nonzero pp-channel can be attributed to a finite-size effect, since here we consider much smaller system with $N_1 = N_2 = 32$ in order to perform MC calculation, while in Fig. 3 we have used the system with $N_1 = N_2 = 1000$. At higher temperatures shown in Fig. 4 (b) and (c), the flux proliferates, and the ph-continuum is smeared and further extends into the higher frequency region beyond Ω^* due to the relaxation of the kinematic constraints. For the same reason, the pp-continuum could also extend to the lower frequency region below Ω^*. Notice, however, large finite-size effects in this region. Therefore, in the region $\Omega < \Omega^*$, the ph-channel of the scattering is the dominant over pp-channel. Thus, when $v_s < v_F$ we will consider only the contribution from the ph-channel. When $v_s \geqslant v_F$, which corresponds to $\Omega \geqslant \Omega^*$ (but not too large that the long-wavelength approximation remains valid), both the ph- and the pp-channels contribute similarly into the phonon scattering, so we will account for both contributions.

C. Numerical results for the sound attenuation coefficient

In this section, we will present numerical results for the temperature evolution of the sound attenuation coefficient given by Eq.(31). We will consider two cases: $v_s < v_F$ and $v_s \geqslant v_F$ corresponding, respectively, to $\Omega < \Omega^*$ and $\Omega > \Omega^*$, both in the zero-flux and in the thermal flux sectors.

1. *The sound attenuation coefficient: $v_s < v_F$*

We will first evaluate $\alpha_s^{||}(\mathbf{q})$ assuming the zero-flux sector at all temperatures. Fig. 5 shows $\alpha_s^{||}(\mathbf{q})$ for various temperatures computed in the system with $N_1 = N_2 = 500$. At all considered temperatures, the sound attenuation coefficient displays the six-fold symmetry and the pattern of the magnitude, which agrees with the analytical results of [37] (see Fig. 13 in App. B for an explicit comparison of the angular positions of maximum and minimum values).

As the temperature increases, the overall intensity of the sound attenuation increases, which can be clearly seen from the increasing range of the colorbars. This is because the phase space of the Majorana fermion scattering in the ph-channel scales with temperature, thus at higher temperature, the intensity of the ph-channel scattering increases. It can also be observed that the area in the phonon momentum space

that actively contributes to the phonon's decay grows, which is a reflection that the scattering of the incident phonons with larger momentum become allowed by the kinematic constraints. When temperature reaches $T = 1$ and goes beyond, the overall magnitude starts to decrease, as analyzed above in Sec. IV A.

Above the flux onset temperature, the fluxes begin to proliferate. The evolution of the flux density with increasing temperature obtained with the strMC method is shown in Fig. 7. For the case of unbroken time reversal symmetry, $\kappa = 0$, the flux onset temperature is around $T^* \simeq 10^{-1.5} \approx 0.03$. Around $T_{\max} \simeq 10^{0.5} \approx 3$, the flux density reaches its maximum concentration, and the distribution of fluxes becomes completely random.

Fig. 6 shows the sound attenuation coefficient $\alpha_s^{||}(\mathbf{q})$ computed at various temperatures in the inhomogeneous flux sectors. $\alpha_s^{||}(\mathbf{q})$ is obtained by averaging over 200 flux configurations at a given temperature. Already from a first glance comparison of Fig. 5 and Fig. 6, we can see that the attenuation coefficient pattern in the presence of fluxes is very different from the pattern in the zero-flux sector shown in Fig. 5. Indeed, in the inhomogeneous flux sector, the peak and valley angular position are rotated compared to those in the zero-flux sector leading to the appearance of the star-like pattern. Also, the overall magnitude of the phonon decay increases with temperature remarkably faster than in the zero-flux sector.

There are three factors that contribute into these modifications of the sound attenuation coefficient. First, the overall magnitude of $\alpha_s^{||}(\mathbf{q})$ increases with temperature. This happens due to the combined effect of the increasing with temperature fermionic population and the appearance of additional low-energy fermionic modes associated with Z_2 flux sectors [61]. Second, the fermionic eigen functions are strongly modified in the presence of the fluxes, which leads to the modifications of the coupling matrix elements of the phonon scattering (See Eq. (22)). Consequently, this change leads to the modification of the sound attenuation coefficient pattern. We also notice that in the temperature interval $T = 0.03 \sim 0.1$, the star-like structure shows the 6-fold symmetry. The slight asymmetry of the pattern displayed near the flux onset temperature can be, perhaps, attributed to a finite-size effect, which gives rise to the inequivalence among the four topological sectors [56]. On the other hand, at $T \simeq 1$ an isotropic circular pattern begins to emerge, which is a combined effect of high population of fermion excitations and fully random flux configurations.

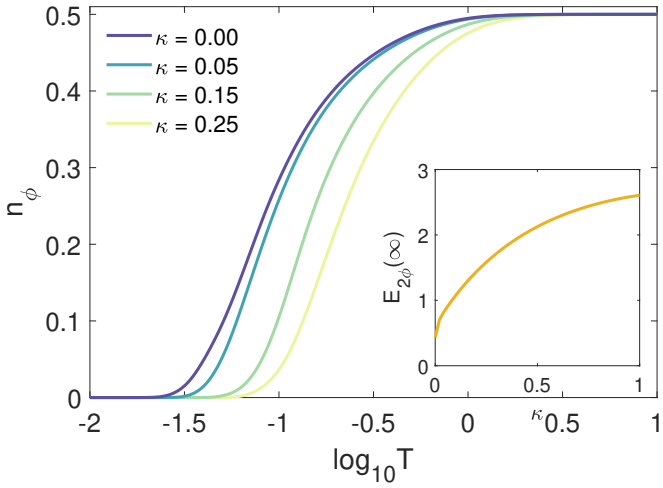

FIG. 7. The temperature dependence of the flux density for various κ obtained by the strMC method on a lattice of size $N_1 = N_2 = 32$. Each data point samples 50,000 flux configurations for each temperature. The inset shows the two-flux gap energy's dependence on κ which is adapted from Ref. [56]. See the definition of the two-flux gap energy therein. The curve shows that the flux gap energy increases with κ. So, in the main panel, the flux onset temperature increases as κ increases. Temperature and the two-flux gap energy are measured in units of the Kitaev interaction, J_K.

Third, as we discussed above, in the absence of the translational invariance, the kinematic constraint related to the conservation of momentum is relaxed and only the energy conservation constraint remains. This allows for more fermionic modes to contribute to the ph-channel scattering. This can be seen that at low temperatures the pattern of the inhomogeneous flux sectors has much larger active area in the phonon momentum space than that of the zero-flux sector. To further analyze the influence of the flux background on the pattern of the sound attenuation coefficient, in App. C we calculate the average $\alpha_s(\mathbf{q})$ over uniformly sampled flux configurations with fixed flux density.

2. *The sound attenuation coefficient: $v_s \geqslant v_F$*

We now turn to the analysis of the sound attenuation coefficient for the case of $v_s \geqslant v_F$. Fig. 8 presents $\alpha_s^{||}(\mathbf{q})$ in the zero-flux sector obtained by the calculation in the momentum space of the finite-size system with $N_1 = N_2 = 500$. Similarly to the case of $v_s < v_F$, at all considered temperatures, $\alpha_s^{||}(\mathbf{q})$ displays the 6-fold symmetry. However, the pattern of the decay intensity distribution pattern is quite different from

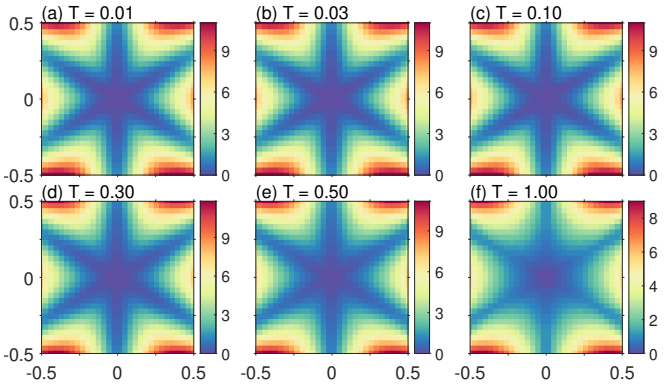

FIG. 8. Case $v_s \geqslant v_F$: The temperature evolution of the sound attenuation coefficient $\alpha_s^{||}(\mathbf{q})$ computed in the zero-flux sector. The diagonalization of the Majorana fermion Hamiltonian Eq. (11) is done in the momentum space on the lattice with $N_1 = N_2 = 500$. The contributions from the pp- and ph-channels of the scattering are summed up. The phonon momentum $\mathbf{q}$ belongs to the region $(q_x, q_y) \in [-0.5\pi, 0.5\pi]^2$. In the calculations, we set $v_s = 1.1v_F$ and the imaginary energy broadening $\delta = 0.2$. Temperature is measured in units of the Kitaev interaction, J_K.

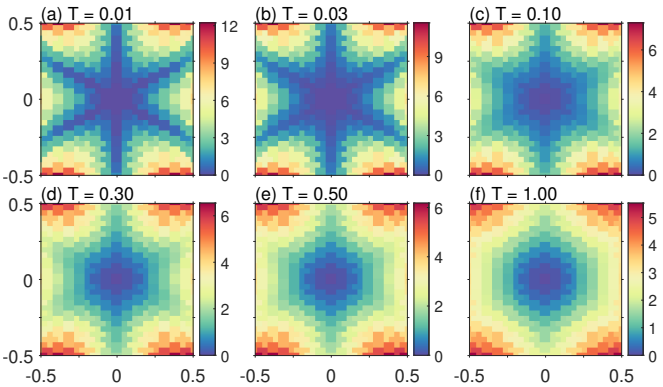

FIG. 9. Case $v_s \geqslant v_F$: The temperature evolution of the sound attenuation coefficient $\alpha_s^{||}(\mathbf{q})$ computed in the inhomogeneous flux sectors sampled by the strMC method. The lattice size is $N_1 = N_2 = 32$, and each data point is obtained by averaging over 200 flux realizations. The phonon momentum $\mathbf{q}$ is in region $(q_x, q_y) \in [-0.5\pi, 0.5\pi]^2$. The imaginary energy broadening is $\delta = 0.2$ and $v_s = 1.1v_F$. Temperature is measured in units of the Kitaev interaction, J_K.

the case of $v_s < v_F$ – the peak's and valley's angular positions have been rotated compared to Fig. 5. Also the active phonon momentum region fills the whole region shown in the plot for all temperatures, while in the case of $v_s < v_F$, the active phonon momentum region grows with temperature. Another major difference lies in the overall magnitude of the attenuation coefficient and its temperature dependence. In Fig. 8, as temperature increases, the overall magnitude of $\alpha_s^{||}(\mathbf{q})$ remains almost unchanged until $T = 1$, then begins to decrease, and the isotropic pattern appears. Recall, that in this case the dominant contribution into the phonon's decay comes from the pp-channel, in which the kinematic constraints can be always satisfied at $v_s \geqslant v_F$. Thus, already at the lowest temperature $T = 0.05$, almost all fermionic modes are involved in the scattering unlike in the ph-channel where the scattering is limited by the available fermionic particle population. Consequently, at $T = 0.05$ (see Fig. 8 (a)) $\alpha_s^{||}(\mathbf{q})$ already displays a star-like pattern. With increasing temperature, the evolution of the pattern and the overall magnitude is mainly decided by the Fermi-Dirac function n_F, and is in agreement with the temperature evolution of the pp-channel scattering shown in Fig. 3. This also verifies that even though both the ph- and pp-scattering channels have been summed up in Fig. 8, the dominant one is the pp-channel.

Fig. 9 shows the sound attenuation for various temperatures computed in the inhomogeneous flux sectors. At each temperature, $\alpha_s^{||}(\mathbf{q})$ is obtained by averaging over 200 random flux configurations sampled by the strMC method. Similarly to the behavior of sound attenuation in the zero-flux sector, the increasing of temperature doesn't change the pattern and the magnitude of the sound attenuation coefficient as dramatically as in the case of $v_s < v_F$ (see Fig. 6). This can again be attributed to the difference in the underlying dynamics between the two scattering channels. With increasing temperature and the corresponding increase of the flux density, the star-like pattern becomes blurry and smears into the isotropic pattern.

Similarly to the case of $v_s < v_F$, we also analyze the influence of the flux background by disentangling its effect from the effect of thermal population of the fermionic states. To this end, in App. C we present the results of $\alpha_s(\mathbf{q})$ averaged over uniformly sampled flux configurations with fixed flux density. Fig. 16 shows that the star-like pattern remains almost unchanged when the density of fluxes increases. Fig. 16 also shows that the star-like pattern persists even at $n_\phi = 0.5$, when the flux configurations are totally random.

V. FINITE TEMPERATURE EFFECTS ON THE PHONON DYNAMICS WITHOUT TIME REVERSAL SYMMETRY

In this section, we will study the observable consequences of the spin-lattice coupling when time reversal symmetry is

broken by the three-spin interaction κ-term. Recall that this three-spin interaction term changes the energetics of the Kitaev model (2) by simultaneously gapping out the fermionic spectrum and introducing localized zero-energy Majorana modes in the presence of isolated fluxes [3]. The spin-lattice coupling Hamiltonian Eq. (9), however, does not contain a contribution from the three-spin interaction κ-term to the leading order (see Sec. II C for more details). Also this term does not break the 6-fold rotation symmetry of the model. As a result, the decomposition of the spin-lattice coupling Hamiltonian into A_1 and E_2 irreducible representations remains valid. Therefore, the sound attenuation coefficient is modified only due to the changes in the fermion spectrum at low energy. Our calculations in App. E indeed show that as κ increases, the temperature evolution of the pattern of the sound attenuation $\alpha_s^{\parallel}(\mathbf{q})$ is similar to that of $\kappa = 0$; the only main difference is that the phonon decay starts at a higher temperature because of the increased fermionic gap energy as κ increases.

Importantly, the κ-term breaks time-reversal and vertical mirror symmetries. What are the consequences to the phonon system? Here, we show that the phonon system acquires the Berry curvature induced by the κ-term due to the spin-lattice coupling, and study its evolution with temperature and the magnitude of κ. The Berry curvature effect can be described by the Hall viscosity term in the phonon effective action [37, 54, 55]:

$$\mathcal{S}_{\rm ph}^{(a)} = \int \mathrm{d}^2x\mathrm{d}t\, \eta_{ijlk}^{(a)} \epsilon_{ij}\dot{\epsilon}_{ik}, \tag{33}$$

where the viscosity tensor is anti-symmetric, i.e. $\eta_{ijlk}^{(a)} = -\eta_{lkij}^{(a)}$. It was also shown in Ref. [37] that by symmetry constraints, the Hall viscosity coefficient has only one non-zero component η_H.

Using the linear response theory [37, 55], we can relate η_H to the off-diagonal component of the phonon polarization bubble:

$$\eta_H = \frac{1}{q^2}\eta^{\perp\parallel} = \frac{1}{q^2\Omega l_a^d}\,\mathrm{Im}\,\Pi^{\parallel\perp}(\mathbf{q},\Omega)|_{\Omega\to 0}, \tag{34}$$

where l_a is the lattice constant. Note that in the contribution to the Hall viscosity η_H, the off-diagonal component $\mathrm{Im}\,\Pi^{\perp\parallel}(\mathbf{q},\Omega)$ must be anti-symmetric w.r.t. exchanging the polarization indices $\perp, \parallel$, which is required by the fact that the Hall viscosity is non-dissipative. Also, it is shown in Ref. [37] that the nonzero contribution to $\mathrm{Im}\,\Pi^{\perp\parallel}(\mathbf{q},\Omega)$ comes from the off-shell processes, in contrast to the dissipative on-shell processes involving the poles in $\mathrm{Im}\,\Pi^{\perp\parallel}(\mathbf{q},\Omega)$.

In the following, we will discuss the temperature evolution of the Hall viscosity coefficient η_H, focusing on understanding of the Z_2 fluxes effects. We will show that due to the distinct difference underlying the off-shell scattering processes compared with the on-shell scattering processes considered in previous sections, the temperature evolution of the Hall viscosity coefficient is very different from that of the sound attenuation coefficient. We will also show that the magnitude of the Hall viscosity coefficient η_H decreases rapidly with increasing density of the Z_2 fluxes.

A. The Hall viscosity coefficient in the zero-flux sector

We start by analyzing the temperature evolution of the Hall viscosity coefficient in the zero-flux sector, when translation invariance allows to perform all the calculations in the momentum space. The analytical derivation of the Hall viscosity coefficient at $T = 0$ was performed in Ref. [37], and the explicit expression of the off-diagonal component $\mathrm{Im}\,\Pi^{\perp\parallel}(\mathbf{q}, i\Omega)$ that contributes to the Hall viscosity is given by

$$\Pi_{\rm ph}^{\perp\parallel}(\mathbf{q},i\Omega) = -\,\mathrm{Tr}\left[\hat{\lambda}^{\parallel}_{\mathbf{q},\mathbf{k}}\mathcal{G}(\mathbf{k},i\omega)\hat{\lambda}^{\perp}_{\mathbf{q},\mathbf{k}}\mathcal{G}(\mathbf{k}-\mathbf{q},i(\omega+\Omega))\right] = 2q^2\Omega\int\frac{\mathrm{d}^2k}{\mathcal{A}_{\rm BZ}}\frac{1}{\Omega^2+4\left|\vec{\mathcal{F}}_{\mathbf{k}}\right|^2}\frac{\Delta_{\mathbf{k}}}{\left|\vec{\mathcal{F}}_{\mathbf{k}}\right|}\left(\lambda^{\parallel}_{x,\mathbf{k},\mathbf{q}}\lambda^{\perp}_{y,\mathbf{k},\mathbf{q}} - \lambda^{\parallel}_{y,\mathbf{k},\mathbf{q}}\lambda^{\perp}_{x,\mathbf{k},\mathbf{q}}\right), \tag{35}$$

where $\mathcal{G}(\mathbf{k}, i\omega)$ is the Majorana fermion Green function, $\hat{\lambda}_{\mathbf{q},\mathbf{k}}$ denotes the Majorana fermion-phonon coupling matrix, which is expanded by Pauli matrices: $\hat{\lambda} = \sum_\alpha \lambda_\alpha \hat{\tau}^\alpha$, and $\vec{\mathcal{F}}_{\mathbf{k}} = \{-\,\mathrm{Im}\,f_{\mathbf{k}}, -\,\mathrm{Re}\,f_{\mathbf{k}}, \Delta_{\mathbf{k}}\}$ with $f_{\mathbf{k}} = 2J_K(1+$

$2\cos\sqrt{3}k_x/2\, e^{i3k_y/2}$) and $\Delta_{\mathbf{k}} = 4\kappa\big(\sin\mathbf{k}\cdot\mathbf{n}_1 - \sin\mathbf{k}\cdot\mathbf{n}_2 + \sin\mathbf{k}\cdot(\mathbf{n}_1-\mathbf{n}_2)\big)$ [see [37] for details of the derivation]. Note that $\vec{\mathcal{F}}_{\mathbf{k}}$ denotes the coefficients of decomposing the Majorana fermion Bloch Hamiltonian Eq. (11) onto the basis of the Pauli matrices.

From Eq. (35) one clearly sees that the contribution to $\Pi^{\mu\nu}_{\rm ph}(\mathbf{q}, i\Omega)$ vanishes for $\mu = \nu$, and that $\Pi^{\parallel\perp}_{\rm ph}(\mathbf{q}, i\Omega) = -\Pi^{\perp\parallel}_{\rm ph}(\mathbf{q}, i\Omega)$. Also, since $\Pi^{\parallel\perp}_{\rm ph}(\mathbf{q}, i\Omega) \propto \Delta_{\mathbf{k}} \propto \kappa$, it is nonzero only when time-reversal symmetry is broken. More formally, the key factor of yielding $\Pi^{\parallel\perp}_{\rm ph}(\mathbf{q}, i\Omega) \neq 0$ lies in the fact that it is proportional to purely imaginary and anti-symmetric quantity $\mathrm{Tr}\,[\hat{\tau}^x\hat{\tau}^y\hat{\tau}^z] = 2i\epsilon^{xyz}$, where ϵ^{xyz} is Levi-Civita tensor. To get nonzero result, the three Pauli matrices must be a permutation of $\hat{\tau}^x, \hat{\tau}^y$ and $\hat{\tau}^z$. Two of them are contributed from the coupling matrix $\hat{\lambda}$, which are proportional to a combination of $\hat{\tau}^x$ and $\hat{\tau}^y$ matrices. The third Pauli matrix comes from one of the propagators $\mathcal{G}(\mathbf{k}, i\omega)$, which comes with a coefficient $\Delta_{\mathbf{k}}$. Consequently, $\mathrm{Im}\,\Pi^{\parallel\perp}_{\rm ph}(\mathbf{q}, i\Omega)$ is obtained by taking the real part of the dynamical projection operators, whose equivalent real-space expressions are shown in Eq. (A1). Since the real part of these operators is taken, the energy conservation constraint is absent, showing explicitly that the non-zero Hall viscosity originates from the off-shell processes, in which all Majorana fermions contribute. Therefore, it is expected that the change in the fermionic population will not have any significant affect on the Hall viscosity at low temperatures (below any other energy scales in the problem, e.g. κ, J_K), in contrast to the on-shell processes contributing to the sound attenuation coefficient. In the high temperature region, where the particle and hole population becomes balanced, it is expected that the Hall viscosity coefficient will be suppressed to zero due to the Pauli exclusion principle.

This analysis is confirmed by Fig. 11 (a), where the Hall viscosity coefficient η_H computed in the zero-flux sector (purple line) at $\kappa = 0.05$ is plotted as a function of temperature. We can see that below $T = 10^{-0.5} \sim 0.3$, η_H remains almost unchanged. Above $10^{-0.5}$, the majority of Majorana fermions are excited and η_H starts to decay to zero. The results for the non-zero flux densities will be analyzed in the next section.

The κ dependence of the Hall viscosity η_H in the zero-flux sector is shown in Fig. 10. The dashed curves are obtained from the analytical calculations [37], while the solid curves are obtained by the numerical evaluation of Eq. (29). It shows that for all values of κ, the Hall viscosity η_H starts to decrease with temperature above $T = 0.5$, which is consistent

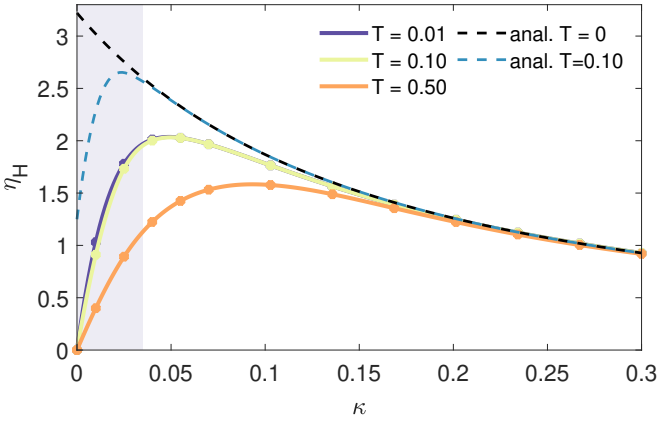

FIG. 10. The Hall viscosity coefficient η_H as of function of κ computed for various temperatures in the zero-flux sector. The solid curves are obtained from the numerical evaluation of Eq. (29). The dashed curves are based on the analytical calculation of Eq. (35) in Ref.[37]. The lattice size used for the numerical calculation is $N_x = N_y = 500$. The phonon momentum and energy used in the calculation are $\mathbf{q} = [0, 0.1]\pi$ with $|\mathbf{q}| = 0.36$ and $\Omega = 0.2$. The imaginary energy broadening $\delta = 0.2$. The shaded area shows the region of small value of κ, where the deviation of the numerical results for $\eta_H(T)$ from the zero temperature analytical result $\eta_H(T=0)$ is the most apparent. T, Ω and κ are measured in units of J_K.

with the analysis above. The comparison between the zero-temperature analytical curve (black dashed line) and the numerical curve computed at $T = 0.01$ (blue solid line) shows that they agree well with each when $\kappa > 0.04$. However, when $\kappa < 0.04$ (shaded region), the numerical result for η_H deviates significantly from the η_H computed analytically. To clarify the origin of this discrepancy, we plotted another η_H-κ curve at $T = 0.1$ (blue dashed line) obtained from the analytical calculation, where the contribution to η_H from each fermionic mode is weighed by a factor $\tanh(\epsilon_k/T)$ calculated from the Fermi-Dirac distribution function, in the integration over the whole Brillouin zone. So as temperature increases, the contribution from the modes whose energies are just above the bulk gap introduced by κ will be suppressed by this factor. As shown in Fig. 10, the Hall viscosity η_H in the shaded region decreases significantly at $T = 0.1$. This indicates that the low-energy fermionic modes has a dominant contribution to η_H at small values of κ. Since our numerical calculations are unable to capture the contribution from all low-energy modes due to the discreteness of energy levels in the finite size calculations, our results show significant decrease of η_H when κ is very small. Therefore, the difference between the analyt-

ical results and the zero-temperature numerical results in the limit of small κ can be attributed to the finite-size effects. The shaded region in Fig. 10, Fig. 11 (b) and Fig. 12 (b) shows where such deviation is the most apparent. Note also that in the numerical calculation, we have to keep Ω small but finite, and in our calculations we fix it as $\Omega = 0.2$. In App. D, we show the numerical verification that for $\kappa = 0.05$ the frequency $\Omega = 0.2$ belongs to the region where $\mathrm{Im}\,\Pi^{\parallel\perp}(\mathbf{q}, i\Omega)$ is linear in Ω.

B. The Hall viscosity coefficient in the inhomogeneous flux sectors

In this section, we explore the effect of the Z_2 fluxes on the temperature evolution of the Hall viscosity coefficient. In the presence of fluxes, the translational symmetry is broken and momentum is not a good quantum number. This means that the block diagonal matrix structure of $\mathrm{Im}\,\Pi^{\parallel\perp}(\mathbf{q}, i\Omega)$ will be mixed up by non-zero off-block-diagonal entries. Consequently, the proportionality of $\mathrm{Im}\,\Pi^{\parallel\perp}(\mathbf{q}, i\Omega)$ to $\mathrm{Tr}\left[\hat{\tau}^\alpha\hat{\tau}^\beta\hat{\tau}^\gamma\right] = 2i\epsilon^{\alpha\beta\gamma}$, that is required for a nonzero Hall viscosity coefficient, is not guaranteed to be valid anymore. Thus, we expect that the Hall viscosity coefficient will be reduced. Moreover, now $\mathrm{Im}\,\Pi^{\parallel\perp}(\mathbf{q}, i\Omega)$ has to be anti-symmetrized w.r.t. the exchange of two polarizations $\perp$ and $\parallel$.

Fig. 11 presents the Hall viscosity η_H computed for various flux densities. Each point is obtained by averaging over 200 randomly sampled flux configuration for a given flux density. Both η_H-T curves (Fig. 11 (a)) and η_H-κ curves (Fig. 11 (b)) show that the Hall viscosity decreases when the density of fluxes increases. Also, this decrease is significant for a small increase of the flux density, which indicates that the Hall viscosity is sensitive to the matrix structure mixing introduced by the fluxes. Note that the time reversal symmetry is still broken which endows the elastic medium with nonzero Berry curvature, but due to the proliferation of the Z_2 fluxes, the Hall viscosity is reduced to basically zero.

Fig. 11 (b) shows η_H-κ curves for various n_ϕ computed at $T = 0.01$, and therefore it extends the zero-flux sector results shown in Fig. 10. In Fig. 11(b), the η_H-κ curves of small flux densities first increase in small κ region and then decreases in large κ region. The behaviour in the shaded region has been analyzed above. The behavior in a region of larger κ in the presence of small flux densities is consistent with the zero flux

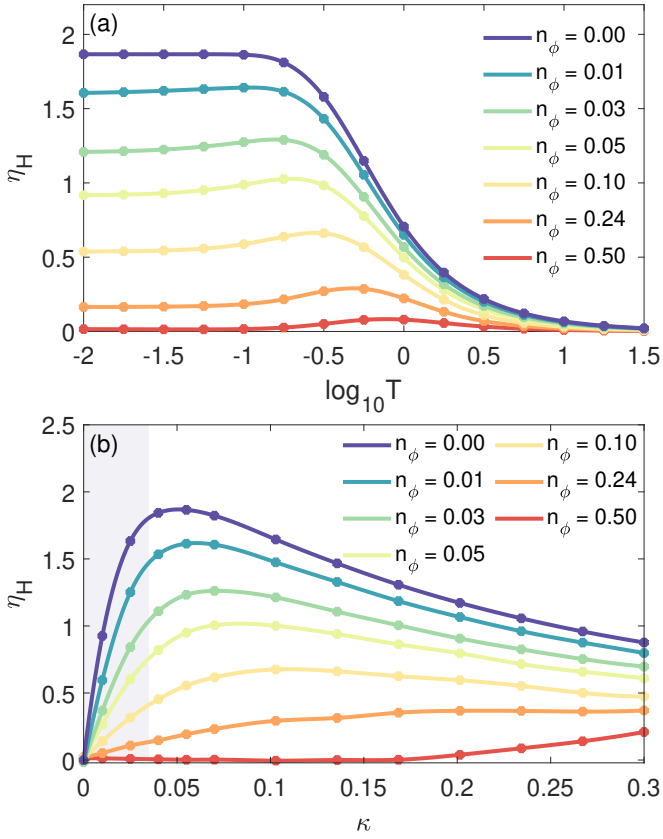

FIG. 11. (a) The temperature dependence of η_H for various flux densities computed at $\kappa = 0.05$. (b) The κ dependence of η_H for various flux densities at fixed $T = 0.01$, which extends the $T = 0$ result shown in Fig. 10. Each curve is obtained by taking the average over 200 flux configurations uniformly sampled for each flux density. $\mathrm{Im}\,\Pi^{\parallel\perp}$ has been antisymmetrized w.r.t. exchanging the polarization indices. In calculations we set $\mathbf{q} = [0, 0.1]\pi$ with $|\mathbf{q}| = 0.36$ and $\Omega = 0.2$, which are chosen based on the validation shown in Fig. 17. The lattice size is $N_1 = N_2 = 32$. The imaginary energy broadening $\delta = 0.2$. The shaded area in (b) shows the region of small value of κ, where the deviation of the numerical results for $\eta_H(T)$ from the zero temperature analytical result $\eta_H(T = 0)$ is the most apparent. T, Ω and κ are measured in units of J_K.

case [37]. Namely, since the fermion energy gap Δ_κ increases with κ [3], the contribution of the low-energy modes into the Hall viscosity is suppressed. Consequently, η_H decreases with increasing κ. On the other hand, at large flux densities n_ϕ, the fermionic spectrum is significantly modified and part of the spectral weight near the van Hove singularity is shifted to lower energies [62]. This may give rise to a small residual Hall viscosity even at relatively large κ.

Next we study the evolution of the Hall viscosity coefficient in the presence of fluxes in a more realistic approach when in-

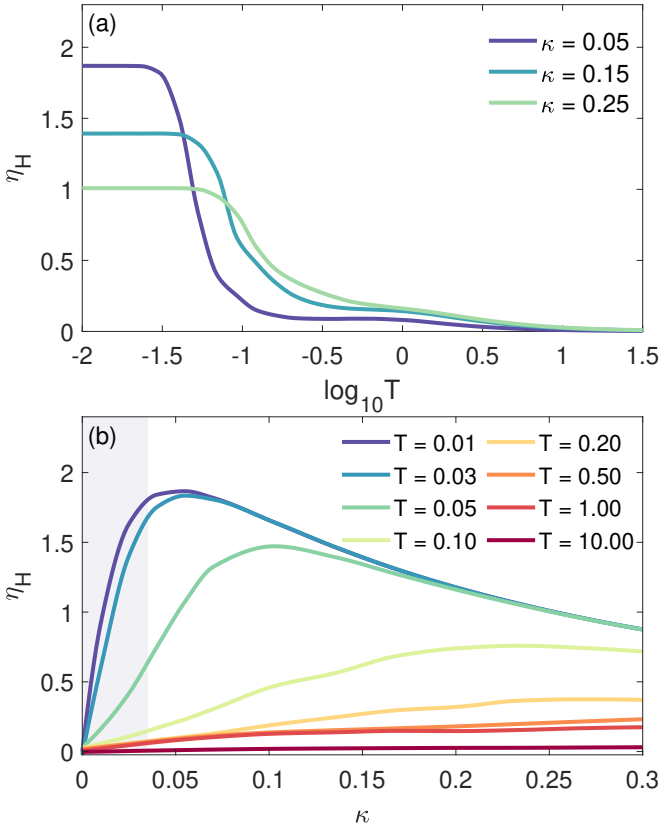

FIG. 12. (a) The temperature dependence of the Hall viscosity η_H computed for various values of κ. (b) The dependence of η_H on the strength of κ computed for various temperatures. Each curve is obtained by the strMC method, which samples independent 100,000 flux configurations for each data point. $\mathrm{Im}\,\Pi^{\parallel\perp}$ has been antisymmetrized w.r.t. exchanging the polarization indices. In calculations we set $\mathbf{q} = [0, 0.1]\pi$ with $|\mathbf{q}| = 0.36$ and $\Omega = 0.2$, which are chosen based on the validation shown in Fig. 17. The lattice size is $N_1 = N_2 = 32$. The imaginary energy broadening is $\delta = 0.2$. The shaded area in (b) shows the region of small value of κ, where the deviation of the numerical results for $\eta_H(T)$ from the zero temperature analytical result $\eta_H(T = 0)$ is the most apparent.T, Ω and κ are measured in units of J_K.

homogeneous flux configurations are sampled by the strMC method (see details in App.F). Fig. 12 shows the η_H-$\log T$ and η_H-κ curves, obtained by the strMC algorithm which samples $100,000$ flux configurations for each point. In Fig. 12(a), the Hall viscosity coefficients at the lowest temperature agree with the η_H-κ curve at $T = 0.01$ shown in Fig. 11 (b). As temperature increases, flux density increases, so the η_H-T curves start to decrease at $T = 10^{-1.5} \sim 10^{-1}$, which is earlier than in Fig. 11(a). This again reflects the reduction the Hall viscosity coefficient by the Z_2 fluxes. Fig. 12 (a) also shows that the η_H for larger values of κ starts to decrease at higher temperatures. This is because the flux gap energy $E_{2\phi}$ increases with κ (see inset in Fig. 7), so they proliferate at higher temperatures as κ increases.

In Fig. 12 (b), we present the η_H-κ curves computed at various temperatures. We can see that the η_H-κ curves display an overall downshift as temperature increases, which is mainly due to increased flux density. Also, since at the same fixed temperature the flux density in the small κ region is larger than that in the large κ region (see Fig. 7), the Hall viscosity coefficient is suppressed stronger by the fluxes in the small κ region than in the large κ region. When temperature reaches $T = 0.5$, where the flux density is close to saturation, the η_H-κ curve becomes almost linear and diminishes to zero.

VI. SUMMARY

In this paper, we proposed that the study of the temperature evolution of the acoustic phonon dynamics can be used as potential probe of spin fractionalization in the Kitaev materials. In our study we focused on two experimental observables – the sound attenuation coefficient (α_s) and the Hall viscosity coefficient (η_H). In particular, we explored how the sound attenuation and the Hall viscosity changes in the presence of the thermally excited Z_2 fluxes. We showed that since the Z_2 fluxes do not couple to the acoustic phonons directly, their effect comes mainly from providing the disorder potential for the itinerant Majorana fermions and, thus, renormalizing their spectrum, and relaxing the kinematic constraints for the scattering processes.

We computed the sound attenuation and the Hall viscosity coefficients by relating them to the imaginary part of the phonon polarization bubble. To compute the bubble in the presence of the thermally excited Z_2 fluxes, when the translational symmetry for the Majorana fermions is broken, we derived a microscopic low-energy effective spin-lattice coupling Hamiltonian and formulated a diagrammatic computation procedure in the mixed representation treating the Majorana fermions in the real space and the acoustic phonons in the momentum space.

We found that the Z_2 fluxes can significantly change the the sound attenuation, making it very different from the one at the zero-flux sector. We demonstrated that both the angular dependence of $\alpha_s(\mathbf{q})$ and its magnitude show characteris-

tic changes when temperature is increasing and various inhomogenious flux sectors are being populated. We showed that the strength of the flux effect on the phonon decay strongly depends on the ratio of the sound velocity and the Fermi velocity characterizing the low-energy Majorana fermions. Namely, we found that when $v_s < v_F$, the thermal excitation of the Z_2 fluxes increases the overall intensity of the sound attenuation and changes the sound attenuation pattern from the flower-like pattern in the zero-flux sector into the star-like pattern in the thermal flux sector. When $v_s > v_F$, both the overall intensity and the angular distribution of maxima and minima of the sound attenuation coefficient are changing only slightly with temperature. These differences reflect different scattering processes contributing into the phonon decay in these two cases. They also show the combined effect of the proliferation of the Z_2 fluxes on the Majorana fermion-phonon coupling vertices, modified fermionic spectrum and relaxed kinematic conditions.

We found that the Hall viscosity coefficient, which is non-zero when time-reversal symmetry is broken, e.g., when external magnetic field is applied, decreases rapidly with increasing density of the Z_2 fluxes. Predominantly, this happens because in the presence of fluxes, the block diagonal matrix structure of the imaginary part of the off-diagonal component of the polarization bubble is mixed up by the non-zero off-block-diagonal entries.

To explore the phonon dynamics above the flux proliferation temperatures, we developed the stratified Monte Carlo (strMC) algorithm to sample the flux configurations. One of the main advantages of this method is that it has zero autocorrelation time, since the samples in the strMC method are independent. This leads to faster convergence and less amount of samples, which is particularly important in the calculations of the phonon dynamics.

Finally, we note that our study was performed for the pure Kitaev model. Of course, real Kitaev materials feature additional weak time-reversal-invariant non-Kitaev interactions, which generally give rise to a wealth of nontrivial phases competing with the quantum spin liquids. While the Kitaev model looses its exact solubility in the presence of additional non-Kitaev terms, we believe that the temperature evolution of the sound attenuation and the Hall viscosity will remain similar to the one in the pure Kitaev model as long these perturbations are small enough to leave the system in the Kitaev-like spin liquid phase. In this case, the effect of the non-Kitaev terms on the dispersion of the Majorana fermions will be small, and their dominant effect will be in adding a dispersion to otherwise localized dispersionless modes corresponding to the flux excitations. However, as long as these terms will be small, the dispersion of the flux modes will be small too, and the fluxes will be excited at more or less the same temperature window as in the pure Kitaev model. Thus, the effect of thermal excitation of the fluxes on the Majorana fermions will remain similar and, consequently, our results on the temperature evolution of the sound attenuation and the Hall viscosity will be robust with respect to a generic time-reversal-invariant perturbation.

Acknowledgments: M.Y. and N.B.P. thank Fiona Burnell, Rafael Fernandes and Wen-Han Kao for valuable discussions. K.F. and N.B.P. were supported by the U.S. Department of Energy, Office of Basic Energy Sciences under Award No. DE-SC0018056. N.B.P. acknowledges the hospitality of Kavli Institute for Theoretical Physics and the National Science Foundation under Grant No. NSF PHY-1748958. M.Y. was supported in part by the Gordon and Betty Moore Foundation through Grant GBMF8690 to UCSB and by the National Science Foundation under Grant No. NSF PHY-1748958. The authors thank the Minnesota Supercomputing Institute for providing computing resources, with which the numerical calculations in this paper were performed.

Appendix A: The Matsubara frequency summation in the evaluation of the polarization bubble Eq. (27)

We can evaluate the Matsubara frequency summation in using the standard residue method [63], and get the following

expressions:

$$
\begin{aligned}
P_{kl}^{g\bar{g}} &= T\sum_{iw_n}\frac{1}{i\omega_n-\epsilon_k}\frac{1}{(i\Omega_m-iw_n)+\epsilon_l}\\
&=\frac{n_F(\epsilon_k)-n_F(\epsilon_l)}{i\Omega_m-\epsilon_k+\epsilon_l},\\
P_{kl}^{\bar{g}g} &= T\sum_{iw_n}\frac{1}{i\omega_n+\epsilon_k}\frac{1}{(i\Omega_m-i\omega_n)-\epsilon_l}\\
&=\frac{n_F(-\epsilon_k)-n_F(-\epsilon_l)}{i\Omega_m+\epsilon_k-\epsilon_l},\\
P_{kl}^{\bar{g}\bar{g}} &= T\sum_{iw_n}\frac{1}{i\omega_n+\epsilon_k}\frac{1}{(i\Omega_m-i\omega_n)+\epsilon_l}\\
&=\frac{n_F(-\epsilon_k)-n_F(\epsilon_l)}{i\Omega_m+\epsilon_k+\epsilon_l},\\
P_{kl}^{gg} &= T\sum_{iw_n}\frac{1}{i\omega_n-\epsilon_k}\frac{1}{(i\Omega_m-i\omega_n)-\epsilon_l}\\
&=\frac{n_F(\epsilon_k)-n_F(-\epsilon_l)}{i\Omega_m-\epsilon_k-\epsilon_l},
\end{aligned}
\tag{A1}
$$

where $P_{kl}^{g\bar{g}}, P_{kl}^{\bar{g}g}$ contribute to the particle-hole channel, P_{kl}^{gg} the particle-particle channel and $P_{kl}^{\bar{g}\bar{g}}$ the hole-hole channel. $n_F(\epsilon)$ is the Fermi-Dirac distribution function.

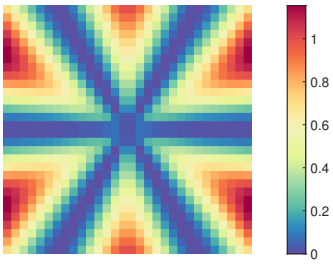

FIG. 13. The plot of the attenuation coefficient $\alpha_s^{||}(\mathbf{q})$ in the zero-flux sector computed from the analytical expression Eq. (B1).The result is shown in the $(q_x, q_y) \in [-0.5\pi, 0.5\pi]^2$ region of the Brillouin zone. To get a sense of how large is this region, one can compare the boundaries of this region with the positions of the high-symmetry points: $\Gamma = [0,0], M \approx [0,1.2]\pi, K \approx [0.66,1.2]\pi, K' \approx [1.3,0]\pi$. In this plot, we set $v_s = 0.1v_F$.

Appendix B: The sound attenuation coefficient in the zero-flux sector

Fig. 13 shows the magnitude of the sound attenuation coefficient $\alpha_s^{||}(\mathbf{q})$ in the region of the Brillouin zone with $(q_x, q_y) \in ([-0.5, 0.5]\pi)^2$ computed in the momentum space according to the analytical expression $\alpha_s^{||}(\mathbf{q}) \sim \frac{1}{q}\,\mathrm{Im}\,\Pi_{\mathrm{ph}}^{|||}(\mathbf{q},\Omega)$ with

$$
\mathrm{Im}\,\Pi_{\mathrm{ph}}^{|||}(\mathbf{q},\Omega) \approx -\frac{36\pi\lambda_{E_2}^2 q|\Omega|}{v_F^3\mathcal{A}_{BZ}}T\,(1-\cos 6\theta_{\mathbf{q}})\ln 2, \tag{B1}
$$

where only the leading order of the q term is kept [37]. Here, only the ph-channel of phonon decay contributes. The spatial distribution of $\alpha_s^{||}(\mathbf{q})$ has the six-fold rotational symmetry originated from the $q(1-\cos(6\theta_{\mathbf{q}}))$ angular dependence of $\mathrm{Im}\,\Pi_{\mathrm{ph}}^{|||}(\mathbf{q},\Omega)$. We will use this result as a benchmark for the numerical calculations of the sound attenuation coefficient.

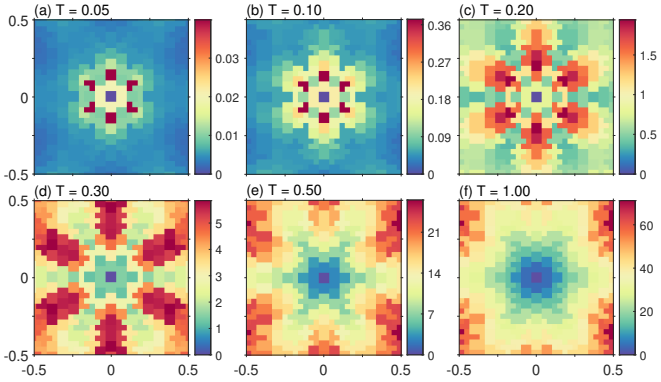

FIG. 14. The temperature evolution of sound attenuation coefficient $\alpha_s^{||}(\mathbf{q})$ for $v_s = 0.1v_F$, computed in the zero-flux sector with Majorana fermion eigenmodes obtained by the real space diagonalization of the Hamiltonian on $N_1 = N_2 = 32$ lattice. The phonon momentum $\mathbf{q}$ is shown in region $(q_x, q_y) \in [-0.5\pi, 0.5\pi]^2$. The imaginary energy broadening $\delta = 0.2$. Temperature is measured in units of J_K.

The sound attenuation coefficient in the zero-flux sector can also be computed from Eq. (24) by using the eigenmodes of the Majorana fermions obtained by the diagonalization of the Hamiltonian in the real space instead of the momentum space. This is our approach to the calculations in the inhomogeneous flux sectors with broken translational symmetry, when the momentum is not a good quantum number. Because we sample different flux sectors with the Monte Carlo approach, simulations on large systems are costly and we have to restrain ourselves to simulations on the $N_1 = N_2 = 32$ lattice. In order to separate the finite size effects from the effect of fluxes, here we compare the sound attenuation computed in the zero-flux sector on the $N_1 = N_2 = 32$ lattice and shown in Fig. 14 with Fig. 5, where diagonalization in the zero-flux sector was performed in the momentum space with $N_1 = N_2 = 500$. This comparison tells that apart from the finite size effects, which are rather significant in this case, Fig. 14 and Fig. 5 basically agree with each other. At $T \geq 0.3$ (panels (d)-(f)

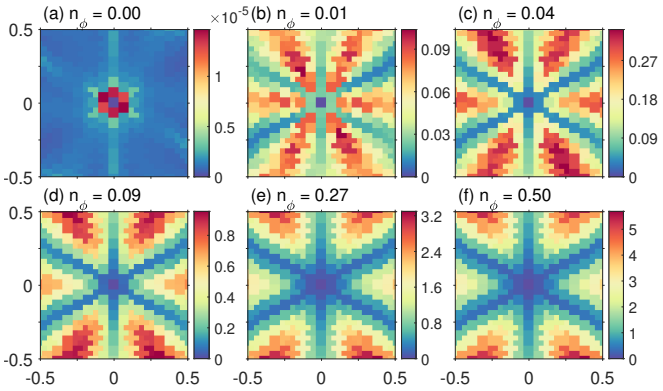

FIG. 15. The sound attenuation coefficient $\alpha_s^{\parallel}(\mathbf{q})$ for $v_s < v_F$ obtained by taking an average over 50 flux configurations uniformly sampled from each flux density at fixed temperature $T = 0.01$. The choice of such a low temperature is made in order to show how the fluxes modifies the low-energy fermionic wave function and its observable effect into the attenuation coefficient. It shows that the star-shape pattern seen in Fig. 6 is due to the presence of the flux background at finite temperatures. The phonon momentum $\mathbf{q}$ is shown in region $(q_x, q_y) \in [-0.5\pi, 0.5\pi]^2$. The lattice size is $N_1 = N_2 = 32$. The imaginary energy broadening $\delta = 0.2$ and we set $v_s = 0.1 v_F$. Temperature is measured in units of J_K.

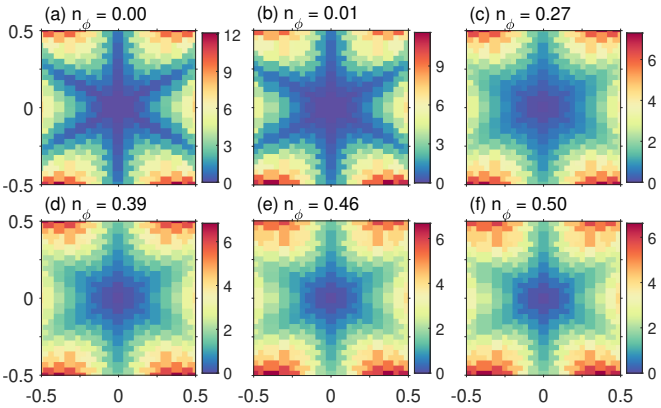

FIG. 16. The sound attenuation coefficient $\alpha_s^{\parallel}(\mathbf{q})$ for $v_s \geqslant v_F$ obtained by taking an average over 50 flux configurations uniformly sampled from each flux density at fixed temperature $T = 0.01$. The contributions from both the pp- and ph-channels of scattering have been summed up. The choice of such a low temperature is made in order to show how the fluxes modifies the low-energy fermionic wave function and its observable effect into the attenuation coefficient. It shows that increasing flux density does not dramatically change the pattern of $\alpha_s^{\parallel}(\mathbf{q})$ but only smears it as the flux density increases. This happens because of the relaxed kinematic constraint in the presence of fluxes. As temperature increases, the fermionic modes are excited, and the result converges to the strMC results in Fig. 9. The phonon momentum $\mathbf{q}$ is shown in region $(q_x, q_y) \in [-0.5\pi, 0.5\pi]^2$. The lattice size is $N_1 = N_2 = 32$. The imaginary energy broadening $\delta = 0.2$ and we set $v_s = 1.1 v_F$. Temperature is measured in units of J_K.

on both plots), they show especially good consistency both in terms of the pattern of the angular distribution of the scattering probability and in the magnitude. However at low temperatures (compare (a)-(b) panels of the corresponding plots), the finite size effects are stronger and, therefore, the discrepancies between the real space calculation and the momentum space calculations are more pronounced.

Appendix C: The sound attenuation coefficient $\alpha_s(\mathbf{q})$ averaged over uniformly sampled flux configurations with fixed flux density.

To further analyze the influence of the flux background on the pattern of the sound attenuation coefficient, here we present $\alpha_s(\mathbf{q})$ computed at very low temperature $T = 0.01$ and averaged over uniformly sampled flux configurations. This allows to practically exclude the effect of thermal population of the fermionic states and analyze the effect of fluxes on the pattern and the magnitude of the sound attenuation coefficient. In Fig. 15 and Fig. 16, we present the results computed for $v_s = 0.1\, v_F$ and $v_s = 1.1\, v_F$, respectively.

Appendix D: The validation of the numerical approach for the computation of the Hall viscosity coefficient

The calculation of the Hall viscosity coefficient η_H is based on Eq. (34), which is obtained from the linear response theory through the Kubo formula. This formula gives proper energy and momentum independent expression for η_H if the off-diagonal component of the polarization bubble $\mathrm{Im}\,\Pi^{\parallel\perp}(\mathbf{q}, \Omega)$ is linear in Ω and quadratic in $|\mathbf{q}|$. The analytical $T = 0$ expression for $\mathrm{Im}\,\Pi^{\parallel\perp}(\mathbf{q}, \Omega)$ given by Eq. (35) clearly satisfies this requirement [37]. However, when we numerically compute $\mathrm{Im}\,\Pi^{\parallel\perp}(\mathbf{q}, \Omega)$ according to Eq. (26), we need to verify that $\mathrm{Im}\,\Pi^{\parallel\perp}(\mathbf{q}, \Omega)$ is indeed linear in Ω and quadratic in $|\mathbf{q}|$, and confine our computations only to those Ω and $|\mathbf{q}|$, for which it is satisfied.

The validation results for $\kappa = 0.05$ are shown in Fig. 17, where for various flux densities we plot (a) $\mathrm{Im}\,\Pi^{\parallel\perp}$-$\Omega$ curves in a low-energy region and (b) $\mathrm{Im}\,\Pi^{\parallel\perp}$-$\mathbf{q}$ curves in small mo-

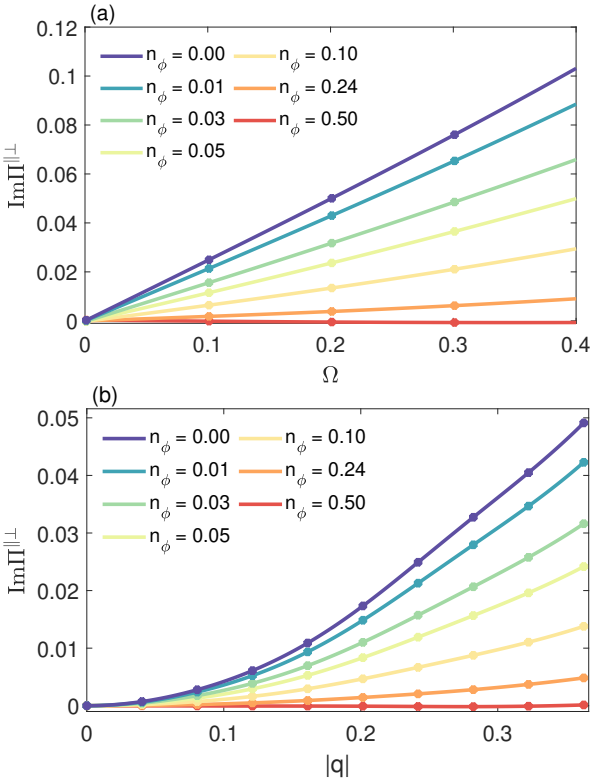

FIG. 17. The imaginary part of the polarization bubble $\mathrm{Im}\,\Pi^{\parallel\perp}(\mathbf{q},\Omega)$ plotted as a function of (a) phonon energy Ω and (b) norm of phonon momentum $|\mathbf{q}|$, for various flux densities. Each curve is obtained by averaging over 200 flux configurations uniformly sampled for a given flux density. $\mathrm{Im}\,\Pi^{\parallel\perp}$ has been antisymmetrized w.r.t. exchanging the polarization indices. Panel (a) shows the linear-Ω behavior of $\mathrm{Im}\,\Pi^{\parallel\perp}$ in the low-energy region, and (b) shows the quadratic-$|\mathbf{q}|$ behavior in the small momentum region. The values of $\Omega = 0.2$ and $\mathbf{q} = [0, 0.1]\pi$ with $|\mathbf{q}| = 0.36$ are within the required region of the linear-Ω and the quadratic-$|\mathbf{q}|$ behavior. The temperature is set to $T = 0.01$ and we use $\kappa = 0.05$. The lattice size is $N_1 = N_2 = 32$. The imaginary energy broadening is $\delta = 0.2$. T and Ω are measured in units of J_K.

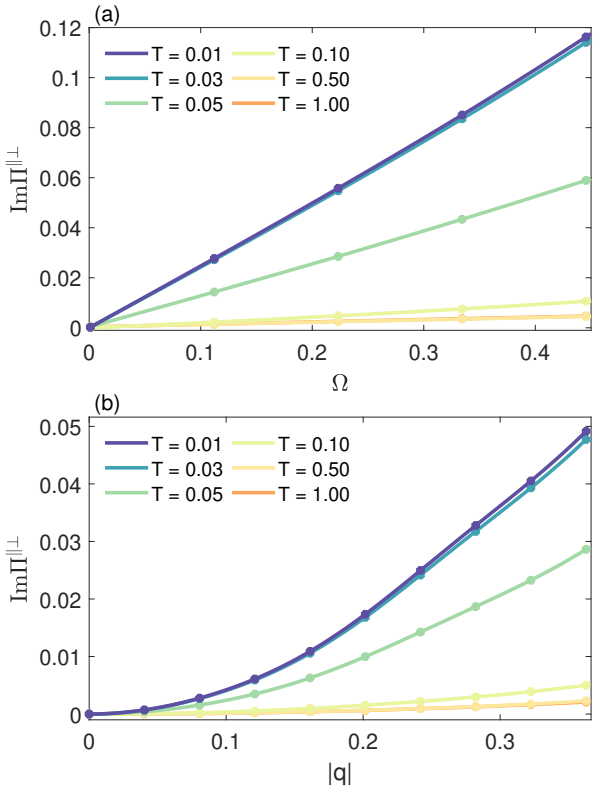

FIG. 18. The imaginary part of the polarization bubble $\mathrm{Im}\,\Pi^{\parallel\perp}(\mathbf{q},\Omega)$ plotted as a function of (a) phonon energy Ω and (b) norm of the phonon momentum $|\mathbf{q}|$, for different temperatures. Each curve is obtained by averaging over 1000 flux configurations sampled by the strMC method. $\mathrm{Im}\,\Pi^{\parallel\perp}$ has been antisymmetrized w.r.t. exchanging the polarization indices. Panel (a) shows the linear-Ω behavior of $\mathrm{Im}\,\Pi^{\parallel\perp}$ in the low-energy region, and (b) shows the quadratic-$|\mathbf{q}|$ behavior in a small momentum region. The values of $\Omega = 0.2$ and $\mathbf{q} = [0, 0.1]\pi$ with $|\mathbf{q}| = 0.36$ are within the required region of the linear-Ω and the quadratic-$|\mathbf{q}|$ behavior. The lattice size is $N_1 = N_2 = 32$. The imaginary energy broadening is $\delta = 0.2$. T and Ω are measured in units of J_K.

mentum region. Each curve is obtained by averaging over 200 flux configurations uniformly sampled for a given flux density. It shows that $\mathrm{Im}\,\Pi^{\parallel\perp}$ is linearly dependent on Ω and quadratically dependent on $|\mathbf{q}|$ for the flux densities shown in the figures. Our numerical calculation of η_H for various fixed flux densities shown in Fig. 11 uses $\Omega = 0.2$, which is within the linear-Ω region, and $\mathbf{q} = [0, 0.1]\pi$ with $|\mathbf{q}| = 0.36$, which is small for the long wavelength approximation and within the quadratic-$|\mathbf{q}|$ region.

Fig. 18 shows that the polarization bubble $\mathrm{Im}\,\Pi^{\parallel\perp}(\mathbf{q},\Omega)$ remains linear in Ω and quadratic in $|\mathbf{q}|$ in the same low-energy and small-momentum region when fluxes are sampled by the strMC method. Each curve shown in Fig. 18 is obtained by averaging over 1000 flux configurations sampled by strMC method. $\mathrm{Im}\,\Pi^{\parallel\perp}$ has also been antisymmetrized w.r.t. exchanging the polarization indices.

Appendix E: The sound attenuation coefficient $\alpha_s(\mathbf{q})$ with time reversal symmetry breaking

In order to verify that the sound attenuation coefficient when time reversal symmetry is broken, i.e. $\kappa > 0$, is qualitatively the same as that when $\kappa = 0$, we calculate $\alpha_s(\mathbf{q})$ for the

case of $v_s < v_F$ for $\kappa = 0.05$ in Fig. 19 and $\kappa = 0.15$ in Fig. 20. These plots are again obtained by strMC method and take the effects of the Z_2 fluxes into consideration. Comparing with Fig. 6, we find indeed that the pattern of $\alpha_s(\mathbf{q})$ remains almost the same with increased κ. So the analysis in for $\kappa = 0$ still applies. The only difference is in the overall magnitude of $\alpha_s(\mathbf{q})$ for temperatures below $T = 0.1$: it decreases when κ increases. This is because the fermionic energy gap Δ_κ increases with κ, so the low-energy modes starts to contribute to phonon's decay at higher temperature. For temperatures above $T = 0.1$, the overall magnitude of $\alpha_s(\mathbf{q})$ remains almost unchanged. This is consistent with the fact that the overall fermionic band height remains almost unchanged for this range of κ [56].

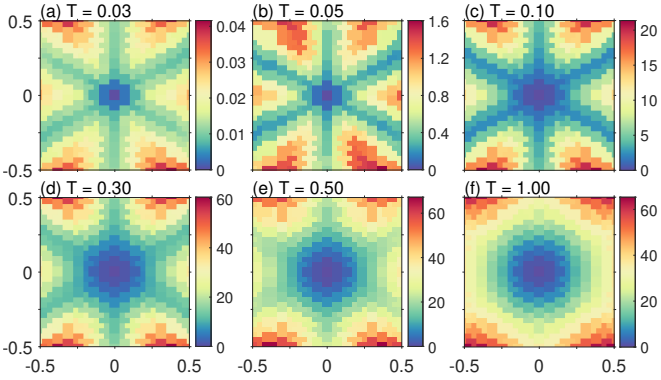

FIG. 19. The temperature evolution of the sound attenuation coefficient $\alpha_s^{\parallel}(\mathbf{q})$ computed in inhomogeneous flux sectors sampled by the strMC method for the case $v_s < v_F$ and $\kappa = 0.05$. The other parameters are the same as in Fig. 6.

Appendix F: Stratified Monte Carlo method

The calculations that include the flux degrees of freedom at finite temperatures are done with the stratified Monte Carlo method (strMC) [64, 65]. Different from a more generally applied Markov Chain Monte Carlo (MCMC) algorithm, this method is designed specifically for the Kitaev honeycomb model and is based on the flux energy model proposed in the recent work [56]. As will shown below, this method has several advantages over the MCMC method.

Stratified sampling [64] is a variance reduction technique in the Monte Carlo simulations. It consists of dividing the sample space to strata and then estimating the weighted average of the yield from each stratum. In this approach, the samples are concentrated on the high probability regions, and the variance is thus reduced. It also offers an unbiased estimator to thermodynamic quantities, and is in principle more efficient than the MCMC method, since the samples are independent from each other and have zero autocorrelation time.

Specifically to the Kitaev model, this method is designed to solve the following problem. For a given observable $O(\{\phi_p\})$ whose value is determined by a certain flux configuration $\{\phi_p\}$ (with p denoting the p-th plaquette and $\{\phi_p\}$ shorted as ϕ_p hereafter), the task is to find the thermodynamic expectation value of this observable over the ensemble of all flux configurations $\langle O\rangle_{\phi_p}$. We denote $I \equiv \langle O\rangle_{\phi_p}$, and it can be generally computed as the following [56]:

$$I = \sum_{\phi_p} p(\phi_p) O(\phi_p), \quad p(\phi_p) = \frac{1}{Z} e^{-\beta F(\phi_p)}, \qquad \text{(F1)}$$

where $\beta = 1/T$ and we assume $k_B = 1$. $p(\phi_p)$ is the probability of the realization of a flux configuration ϕ_p, which is decided by the free energy $F(\phi_p) = E_{\text{flux}}(\phi_p) - TS_F(\phi_p)$. Here, $E_{\text{flux}}(\phi_p) = -\frac{1}{2}\sum_k \epsilon_k(\phi_p)$, $S_F(\phi_p) = \sum_k \ln(1 + e^{-\beta\epsilon_k}(\phi_p))$ and $Z = \sum_{\phi_p} e^{-\beta F(\phi_p)}$ is, respectively, the flux energy, fermionic entropy and partition function of the system in a given flux sector ϕ_p. $\epsilon_k(\phi_p)$ denote the fermionic energy levels in the flux configuration ϕ_p. The central goal of this method is to generate a large set of sample flux configurations $\{\phi_p^i\}$ (where i denotes the i-th sample), and use them to estimate the thermodynamic expectation I. The function evaluated from these samples takes a form of sample average, which is called an estimator of I and denoted as $\hat{I}$. In the following introduction to the strMC method, the focus will be on showing the way of generating the sample flux configurations $\{\phi_p^i\}$, the construction of the estimator $\hat{I}$, and the validation that $\hat{I}$ converges to I in the large sample limit.

The summation in Eq. (F1) is naturally stratified into the groups according to the number of fluxes denoted as n:

$$I = \frac{1}{Z}\sum_n \sum_{\phi_{p,n}} O(\phi_{p,n})\, e^{-\beta F(\phi_{p,n})} = \sum_n \frac{Z_n}{Z}\cdot I_n, \qquad \text{(F2)}$$

where $\phi_{p,n}$ is a flux configuration restricted within an n-flux sector, and $I_n = \sum_{\phi_{p,n}} O(\phi_{p,n}) \frac{e^{-\beta F(\phi_{p,n})}}{Z_n}$, $Z_n = \sum_{\phi_{p,n}} e^{-\beta F(\phi_{p,n})}$ are respectively the thermal expectation and the partition function within the group denoted by n. It will be shown that the integral I_n can be easily estimated. The number of samples allocated to each group is proportional to the group's probability $\frac{Z_n}{Z}$, so the sampling will be concentrated on the group with larger probability. Then after obtain-

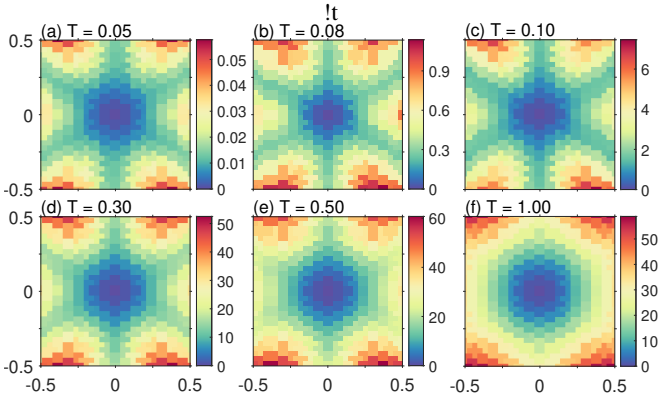

FIG. 20. The temperature evolution of the sound attenuation coefficient $\alpha_s^{\parallel}(\mathbf{q})$ computed in inhomogeneous flux sectors sampled by the strMC method for the case $v_s < v_F$ and $\kappa = 0.15$. The other parameters are the same as in Fig. 6.

ing the estimation of I_n within each group, the final estimation of I is obtained by a weighted sum over I_n, with weight proportional to the group's probability $\frac{Z_n}{Z}$.

To decide the number of samples allocated to each group, the probability of each group $\frac{Z_n}{Z}$ has to be known beforehand. To this end, we find that this probability can be approximated by a phenomenological model, which describes the distribution of the flux energies as a function of flux density. For sufficiently large lattices, this distribution is sharply peaked and is approximately independent of lattice size [56]. Therefore, we can model flux thermodynamics by numerically fitting the flux energy with a polynomial curve. The resultant best-fit flux energy as a function of flux density, so called pseudo-potential energy (PPE) [56], is shown in the inset of Fig. 21. Basically, this best-fit flux energy $\bar{E}(n)$ and $W_n = \binom{N_p}{n}$, the number of flux configurations for a fixed flux number with N_p denoting the total number of plaquettes, decide the approximate probability of each group by $Z_n \approx W_n e^{-\beta \bar{E}(n)}$, and allocate the samples to each group accordingly. Note that this approximate weight $W_n e^{-\beta \bar{E}(n)}$ is only used to allocate the samples. When calculating the weighted average across the groups, this weight will be modified by a factor $\frac{\hat{Z}_n/W_n}{e^{-\beta \bar{E}(n)}}$, which will be shown next.

To estimate the thermal expectation value I_n, we rewrite it as the following:

$$I_n = \frac{1}{Z_n/W_n} \sum_{\phi_{p,n}} O\left(\phi_{p,n}\right) \frac{e^{-\beta F(\phi_{p,n})}}{W_n}. \tag{F3}$$

Based on this expression, we can construct an estimator $\hat{I}_n$:

$$\hat{I}_n = \frac{1}{\hat{Z}_n/W_n} \frac{1}{N_n} \sum_{i=1}^{N_n} O\left(\phi_{p,n}^i\right) e^{-\beta F\left(\phi_{p,n}^i\right)}, \tag{F4}$$

where $\hat{Z}_n/W_n = \frac{1}{N_n} \sum_i^{N_n} e^{-\beta F(\phi_{p,n}^i)}$ is relevant to the estimation of group partition function, and N_n is the number of samples allocated to group n. $\phi_{p,n}^i$ denotes the i-th sample flux configuration that is generated from the uniform distribution within the flux sector n and thus, are totally independent from each other. This is fundamentally different from the MCMC algorithm, where each running sample is probabilistically dependent on the previous samples.

Now combining the estimations from all groups, we can obtain an estimator of the integral I, denoted by $\hat{I}$:

$$\hat{I} = \sum_n \frac{\hat{Z}_n}{Z} \hat{I}_n = \frac{1}{Z} \sum_n W_n \frac{1}{N_n} \sum_{i=1}^{N_n} O\left(\phi_{p,n}^i\right) e^{-\beta F\left(\phi_{p,n}^i\right)}, \tag{F5}$$

where the independently generated samples $\{\phi_{p,n}^i\}$ are used to estimate a thermodynamic average of a quantity I. But this expression is not convenient for translating to algorithm, so we rewrite it as

$$\hat{I} = \frac{1}{Z} \sum_n W_n e^{-\beta \bar{E}(n)} \frac{\hat{Z}_n/W_n}{e^{-\beta \bar{E}(n)}} \hat{I}_n, \tag{F6}$$

where the approximate weight $W_n e^{-\beta \bar{E}(n)}$, which determines the number of samples allocated into each group, has been modified by the factor $\frac{Z_n/W_n}{e^{-\beta \bar{E}(n)}}$, which is evaluated from the sampled data. To get a sense of this modification factor, note that $\hat{Z}_n/W_n = \frac{1}{N_n} \sum_i^{N_n} e^{-\beta F(\phi_{p,n}^i)}$ is the sample averaged Boltzmann weight of a flux sector, where $F(\phi_p)$ is mainly decided by $E_{\text{flux}}(\phi_p)$. So this modification factor is of the same order of 1, which is understood as a tuning of the approximate weight. After this tuning, the estimation $\hat{I}$ will be shown to be accurate, even though initially the approximate weight was used.

It remains to show that the estimator Eq. (F5) does not include systematic error, i.e. is unbiased. If an estimator of thermodynamic quantity is not constructed wisely, the sample average may not converge to the true thermal expectation value at large sample number limit. This validation done by computing the expectation value of $\hat{I}$ on the random variable $\phi_{p,n}^i$ obeying uniform distribution $\mathrm{p}(\phi_{p,n}^i) = \frac{1}{W_n}$ within a flux sec-

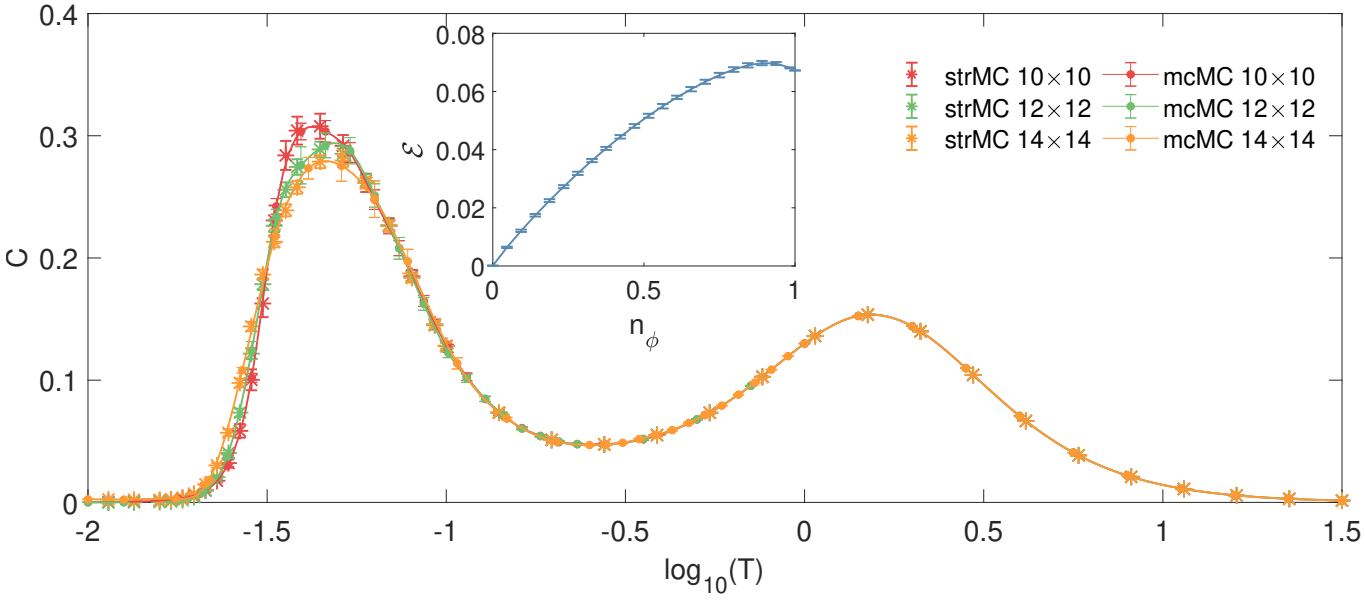

FIG. 21. The temperature dependence of the specific heat of the Kitaev honeycomb model. The dashed curve is obtained by the MCMC algorithm, described in Ref. [56], while the stars are obtained by the strMC algorithm. The total number of sample is 1,000,000 for each data point. The agreement between the two sets of results validates the strMC algorithm. The inset shows the plot of the best-fit flux energy density $\mathcal{E}(n_\phi)$ as a function of the flux density n_ϕ for $\kappa = 0$, which is adapted from Ref. [56]. The flux energy $\bar{E}(n_\phi) = \mathcal{E}(n_\phi)N_p$, where N_p is total plaquette number. For the explicit expressions of the fit polynomial function and the result of other values of κ, see Ref. [56]. Temperature and energies are measured in units of J_K.

tor n:

$$\begin{aligned}\mathbb{E}_{\phi^i_{p,n}}[\hat{I}] &= \frac{1}{Z}\sum_n W_n \frac{1}{N_n}\sum_{i=1}^{N_n} \mathbb{E}_{\phi^i_{p,n}} O\left(\phi^i_{p,n}\right) e^{-\beta F\left(\phi^i_{p,n}\right)} \\ &= \frac{1}{Z}\sum_n W_n \frac{1}{N_n}\sum_{i=1}^{N_n}\sum_{\phi^i_{p,n}} \frac{1}{W_n} O\left(\phi^i_{p,n}\right) e^{-\beta F\left(\phi^i_{p,n}\right)} \\ &= \frac{1}{Z}\sum_n \sum_{\phi_{p,n}} O\left(\phi_{p,n}\right) e^{-\beta F(\phi_{p,n})} = I. \end{aligned} \tag{F7}$$

So this estimator is indeed unbiased, which guarantees that as the number of samples N goes to infinity, the sample average converge to the thermodynamic average [66].

In Fig. 21 we plot the specific heat as a function of temperature (in a semi-logarithmic scale) calculated by both the MCMC and the strMC methods. Fig. 21 displays a good agreement between the results from the two algorithms. Note that the low-temperature specific heat peak is rather sensitive to the thermal fluctuations as shown in Ref. [56]. Therefore, the good agreement between the two algorithms especially at the low-temperature peak validates the strMC algorithm we proposed.

To conclude this Appendix, let us underline the differences and similarities between the MCMC and the strMC algorithms. In essence, both the MCMC and the strMC use the empirical probability distributions. The probability distribution in the MCMC is given by

$$p_{\text{MCMC}}(\phi_p) = \frac{1}{N}\sum_{i=1}^{N}\delta(\phi_p - \phi^i_{p'}), \tag{F8}$$

where $\{\phi^i_{p'}\}$ are sampled according to their Boltzmann weights $p(\phi_{p'}) \sim e^{-\beta F(\phi_{p'})}$, while the empirical distribution of the strMC is

$$p_{\text{strMC}}(\phi_p) = \frac{1}{Z}\sum_n W_n \frac{1}{N_n}\sum_{i=1}^{N_n} e^{-\beta F(\phi^i_{p',n})}\delta(\phi_p - \phi^i_{p',n}), \tag{F9}$$

where $\{\phi^i_{p',n}\}$ are sampled from uniform distribution within flux sectors of a fixed flux number n, and then modulated by their Boltzmann weights to get an unbiased estimation.

An immediate advantage of the strMC over the MCMC is that it fundamentally reduces the autocorrelation time to zero, since the samples in the strMC are independent. In the recent works, where the machine learning aided MCMC algorithms were applied to condensed matter systems [67–69], the deep neural networks have been employed to fit the system's free

energy and to generate the MCMC updates that have an acceptance probability of nearly one. These works achieved a significant reduction on the autocorrelation time. The autocorrelation time in the strMC algorithm is reduced to zero, leading to faster convergence and less amount of samples. This is particularly important in the calculations of the phonon dynamics, where the main computation cost comes from the matrix production $\mathbf{B}^\dagger \tilde{\Lambda}^\mu_{\mathbf{q}} \mathbf{B}$ in Eq. (22). Other advantages of the strMC method include the convenience in the implementation and parallelization, and the exemption from the annealing process (initial thermalization) and the local minima problems, which is common in the MCMC algorithms. Note, however, that in the application of the strMC in this paper, the division of the groups is specific to the Kitaev honeycomb model, where the flux number is naturally used to denote the group. For other models, the strMC method needs to be specifically designed.

In a nutshell, in this strMC algorithm, the number of samples allocated to each group is based on the flux PPE model which is approximate, while the actual superposition weight across the groups is fine-tuned by the accurate probability distribution. Thus the estimation of the thermal expectation value is unbiased. This algorithm offers a good example of utilizing a pre-trained model to accelerate the Monte Carlo simulation, which is the idea shared by many machine learning algorithms [67–69]. It is promising to combine this algorithm with the machine learning techniques to further improve the performance.

[1] P. W. Anderson, Materials Research Bulletin **8**, 153 (1973).

[2] X.-G. Wen, Phys. Rev. B **65**, 165113 (2002).

[3] A. Kitaev, Annals of Physics **321**, 2 (2006).

[4] L. Balents, Nature **464**, 199 (2010).

[5] L. Savary and L. Balents, Rep. Prog. Phys. **80**, 016502 (2017).

[6] J. Knolle and R. Moessner, Annual Review of Condensed Matter Physics **10**, 451 (2019).

[7] C. Broholm, R. J. Cava, S. A. Kivelson, D. G. Nocera, M. R. Norman, and T. Senthil, Science **367** (2020), 10.1126/science.aay0668.

[8] G. Jackeli and G. Khaliullin, Phys. Rev. Lett. **102**, 017205 (2009).

[9] S. Trebst, arXiv:1701.07056 (2017).

[10] H. Takagi, T. Takayama, G. Jackeli, G. Khaliullin, and S. E. Nagler, Nat. Rev. Phys. **1**, 264 (2019).

[11] Y. Motome and J. Nasu, J. Phys. Soc. Jpn **89**, 012002 (2020).

[12] J. Knolle, D. L. Kovrizhin, J. T. Chalker, and R. Moessner, Phys. Rev. B **92**, 115127 (2015).

[13] A. Banerjee, C. A. Bridges, J.-Q. Yan, A. A. Aczel, L. Li, M. B. Stone, G. E. Granroth, M. D. Lumsden, Y. Yiu, J. Knolle, S. Bhattacharjee, D. L. Kovrizhin, R. Moessner, D. A. Tennant, M. D. G., and S. E. Nagler, Nature materials (2016), 10.1038/nmat4604.

[14] A. Banerjee, J. Yan, J. Knolle, C. A. Bridges, M. B. Stone, M. D. Lumsden, D. G. Mandrus, D. A. Tennant, R. Moessner, and S. E. Nagler, Science **356**, 1055 (2017).

[15] A. Banerjee, P. Lampen-Kelley, J. Knolle, C. Balz, A. A. Aczel, B. Winn, Y. Liu, D. Pajerowski, J. Yan, C. A. Bridges, et al., npj Quantum Materials **3**, 8 (2018).

[16] J. Knolle, G.-W. Chern, D. L. Kovrizhin, R. Moessner, and N. B. Perkins, Phys. Rev. Lett. **113**, 187201 (2014).

[17] L. J. Sandilands, Y. Tian, K. W. Plumb, Y.-J. Kim, and K. S. Burch, Phys. Rev. Lett. **114**, 147201 (2015).

[18] B. Perreault, J. Knolle, N. B. Perkins, and F. J. Burnell, Phys. Rev. B **92**, 094439 (2015).

[19] B. Perreault, J. Knolle, N. B. Perkins, and F. J. Burnell, Phys. Rev. B **94**, 104427 (2016).

[20] I. Rousochatzakis, S. Kourtis, J. Knolle, R. Moessner, and N. B. Perkins, Phys. Rev. B **100**, 045117 (2019).

[21] A. Sahasrabudhe, D. A. S. Kaib, S. Reschke, R. German, T. C. Koethe, J. Buhot, D. Kamenskyi, C. Hickey, P. Becker, V. Tsurkan, A. Loidl, S. H. Do, K. Y. Choi, M. Grüninger, S. M. Winter, Z. Wang, R. Valentí, and P. H. M. van Loosdrecht, Phys. Rev. B **101**, 140410 (2020).

[22] Y. Wang, G. B. Osterhoudt, Y. Tian, P. Lampen-Kelley, A. Banerjee, T. Goldstein, J. Yan, J. Knolle, H. Ji, R. J. Cava, J. Nasu, Y. Motome, S. E. Nagler, D. Mandrus, and K. S. Burch, npj Quantum Materials **5**, 14 (2020).

[23] D. Wulferding, Y. Choi, S.-H. Do, C. H. Lee, P. Lemmens, C. Faugeras, and K.-Y. Gallais, Yann andChoi, Nat. Commun. **11**, 1603 (2020).

[24] G. B. Halász, N. B. Perkins, and J. van den Brink, Phys. Rev. Lett. **117**, 127203 (2016).

[25] G. B. Halász, B. Perreault, and N. B. Perkins, Phys. Rev. Lett. **119**, 097202 (2017).

[26] G. B. Halász, S. Kourtis, J. Knolle, and N. B. Perkins, Phys. Rev. B **99**, 184417 (2019).

[27] A. Ruiz, N. P. Breznay, M. Li, I. Rousochatzakis, A. Allen, I. Zinda, V. Nagarajan, G. Lopez, M. H. Upton, J. Kim, A. H. Said, X.-R. Huang, T. Gog, D. Casa, R. J. Birgeneau, J. D. Koralek, J. G. Analytis, N. B. Perkins, and A. Frano, arXiv:2102.02714.

[28] Z. Alpichshev, F. Mahmood, G. Cao, and N. Gedik, Phys. Rev. Lett. **114**, 017203 (2015).

[29] H. Zhang, S. Kim, Y.-J. Kim, H.-Y. Kee, and L. Yang, arXiv:1908.04807 (2019).

[30] A. Little, L. Wu, P. Lampen-Kelley, A. Banerjee, S. Patankar, D. Rees, C. A. Bridges, J.-Q. Yan, D. Mandrus, S. E. Nagler, and J. Orenstein, Phys. Rev. Lett. **119**, 227201 (2017).

[31] Y. Wan and N. P. Armitage, Phys. Rev. Lett. **122**, 257401 (2019).

[32] R. Hentrich, A. U. B. Wolter, X. Zotos, W. Brenig, D. Nowak, A. Isaeva, T. Doert, A. Banerjee, P. Lampen-Kelley, D. G. Mandrus, S. E. Nagler, J. Sears, Y.-J. Kim, B. Büchner, and C. Hess, Phys. Rev. Lett. **120**, 117204 (2018).

[33] Y. Kasahara, T. Ohnishi, Y. Mizukami, O. Tanaka, S. Ma, K. Sugii, N. Kurita, H. Tanaka, J. Nasu, Y. Motome, T. Shibauchi, and Y. Matsuda, Nature **559**, 227 (2018).

[34] S. Pal, A. Seth, P. Sakrikar, A. Ali, S. Bhattacharjee, D. V. S. Muthu, Y. Singh, and A. K. Sood, arXiv:2011.00606 (2020).

[35] H. Li, T. T. Zhang, A. Said, G. Fabbris, D. G. Mazzone, J. Q. Yan, D. Mandrus, G. B. Halasz, S. Okamoto, S. Murakami, M. P. M. Dean, H. N. Lee, and H. Miao, arXiv:2011.07036 (2020).

[36] A. Metavitsiadis and W. Brenig, Phys. Rev. B **101**, 035103 (2020).

[37] M. Ye, R. M. Fernandes, and N. B. Perkins, Phys. Rev. Research **2**, 033180 (2020).

[38] M. Ye, G. B. Halász, L. Savary, and L. Balents, Phys. Rev. Lett. **121**, 147201 (2018).

[39] Y. Vinkler-Aviv and A. Rosch, Phys. Rev. X **8**, 031032 (2018).

[40] A. Pippard, The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science **46**, 1104 (1955).

[41] A. I. Akhiezer, M. I. Kaganov, and G. Y. Lyubarskyi, Sov. Phys. JETP **5**, 685 (1957).

[42] E. I. Blount, Phys. Rev. **114**, 418 (1959).

[43] T. Tsuneto, Phys. Rev. **121**, 402 (1961).

[44] B. Batlogg, D. Bishop, B. Golding, C. M. Varma, Z. Fisk, J. L. Smith, and H. R. Ott, Phys. Rev. Lett. **55**, 1319 (1985).

[45] H. Won and K. Maki, Phys. Rev. B **49**, 1397 (1994).

[46] M. Boiteux, P. Doussineau, B. Ferry, J. Joffrin, and A. Levelut, Phys. Rev. B **4**, 3077 (1971).

[47] A. Sytcheva, U. Löw, S. Yasin, J. Wosnitza, S. Zherlitsyn, P. Thalmeier, T. Goto, P. Wyder, and B. Lüthi, Phys. Rev. B **81**, 214415 (2010).

[48] T. I. Tuegel and T. L. Hughes, Phys. Rev. B **96**, 174524 (2017).

[49] T. Qin, J. Zhou, and J. Shi, Phys. Rev. B **86**, 104305 (2012).

[50] X. Zhang, Y. Zhang, S. Okamoto, and D. Xiao, Phys. Rev. Lett. **123**, 167202 (2019).

[51] J.-Y. Chen, S. A. Kivelson, and X.-Q. Sun, Phys. Rev. Lett. **124**, 167601 (2020).

[52] M. Ye, L. Savary, and L. Balents, arXiv:2103.04223 (2021).

[53] This statement assumes that the spin-lattice couplings are restricted to the Kitaev interaction form, i.e. it only modifies the strength of the spin interaction but does not generate additional forms.

[54] J. Avron, R. Seiler, and P. G. Zograf, Physical review letters **75**, 697 (1995).

[55] M. Barkeshli, S. B. Chung, and X.-L. Qi, Physical Review B **85**, 245107 (2012).

[56] K. Feng, N. B. Perkins, and F. J. Burnell, Phys. Rev. B **102**, 224402 (2020).

[57] Y.-Z. You, I. Kimchi, and A. Vishwanath, Physical Review B **86**, 085145 (2012).

[58] Mahan, G. D., Many-Particle Physics (Plenum Press, New York, 1990).

[59] We have checked numerically that the contribution from the A_1 channel remains parametrically small even in the presence of the Z_2 fluxes.

[60] Y. Zhou and P. A. Lee, Phys. Rev. Lett. **106**, 056402 (2011).

[61] V. Lahtinen, New Journal of Physics **13**, 075009 (2011).

[62] W.-H. Kao and N. B. Perkins, Annals of Physics.

[63] Altland, Alexander and Simons, Ben D., Condensed Matter Field Theory, 2nd ed. (Cambridge University Press, 2010).

[64] R. Y. Rubinstein and D. P. Kroese, Simulation and the Monte Carlo method, Vol. 10 (John Wiley & Sons, 2016).

[65] P. Glasserman, Monte Carlo methods in financial engineering, Vol. 53 (Springer Science & Business Media, 2013).

[66] Careful readers may notice that the partition function Z in $\hat{I}$ is also stochastic, since it is fine-tuned by $\frac{\hat{Z}_n/W_n}{e^{-\beta\bar{E}(n)}}$. Here, we have ignored this randomness in the expectation calculation Eq. (F7), since it can be shown by central limit theory that, in the large sample limit $N \to \infty$, Z converges to the non-stochastic partition function [70]. Therefore, our approximation does not change the asymptotic behaviour of the estimator $\hat{I}$, which still converges to its thermodynamic expectation value.

[67] L. Huang and L. Wang, Physical Review B **95**, 035105 (2017).

[68] Y. Nagai, H. Shen, Y. Qi, J. Liu, and L. Fu, Phys. Rev. B **96**, 161102 (2017).

[69] D. Alcalde Puente and I. M. Eremin, Phys. Rev. B **102**, 195148 (2020).

[70] G. Casella and R. L. Berger, Statistical inference, 2nd ed. (Cengage Learning, 2021) Chap. 10.