Building flat-band lattice models from Gram matrices

Youjiang Xu and Han Pu

Department of Physics and Astronomy, and Rice Center for Quantum Materials, Rice University, Houston, Texas 77251-1892, USA

We propose a powerful and convenient method to systematically design flat-band lattice models, which overcomes the difficulties underlying the previous method. Especially, our method requires no elaborate calculations, applies to arbitrary spatial dimensions, and guarantees to result in a completely flat ground band. We use this method to generate several classes of lattice models, including models with both short- and long-range hoppings, both topologically trivial and nontrivial flat bands. Some of these models were previously known. Our method, however, provides important insight. For example, we have reproduced and generalized the Kapit-Mueller model [Kapit and Mueller, Phys. Rev. Lett. **105**, 215303 (2010)] and demonstrated a universal scaling rule between the flat-band degeneracy and the magnetic flux that was not noticed in previous studies. We show that the flat band of this model results from the (over-)completeness properties of coherent states.

I. INTRODUCTION

Tight-binding lattice models that support flat bands [1], i.e., with single-particle energy dispersion $E(\mathbf{k})$ independent of momentum $\mathbf{k}$, are of great importance. The quenched kinetic energy and the associated macroscopic degeneracy in a flat band makes the system extremely sensitive to perturbations. In particular, in a many-body setting, interaction between particles in the flat band, no matter how weak it is, can result in strong correlations and exotic quantum phases. This is exactly what happens in, for example, fractional quantum Hall systems [2–4], where the underlying single-particle spectrum features flat Landau levels. It is therefore important to understand what model Hamiltonians can support flat bands and, conversely, how one can systematically design models supporting flat bands.

At the first glance, it seems quite simple to construct a flat-band lattice model in momentum space. Given dispersion relations $\{E_1(\mathbf{k}), E_2(\mathbf{k}), ..., E_n(\mathbf{k})\}$ in which one or a few of them are set to be constant, e.g., $E_1(\mathbf{k}) = 0$, we can construct a n-band flat-band model from an arbitrary unitary matrix $U(\mathbf{k})$: $h(\mathbf{k}) := U(\mathbf{k})E(\mathbf{k})U^\dagger(\mathbf{k})$, where $E(\mathbf{k})$ is the diagonal matrix with elements $\{E_1(\mathbf{k}), E_2(\mathbf{k}), ..., E_n(\mathbf{k})\}$. However, this approach is too general to be practical. Specifically, if we want the flat-band model to possess certain properties (e.g., finite-range hopping, certain symmetry properties, etc.), it is not straightforward to put appropriate constraints on the matrices $U(\mathbf{k})$ and $E(\mathbf{k})$ to make $h(\mathbf{k})$ possess the desired properties. To solve this problem, we turn our attention from the momentum space to the real position space.

Suppose the Hilbert space is spanned by the basis $|\mathbf{R}; i = 1, 2, \ldots, n\rangle$, where $\mathbf{R} = \sum_{j=1}^d x_j \mathbf{e}_j$ ($x_j \in \mathbb{Z}$) represents the position of cells in a d-dimension lattice and i's distinguish the n sites in a cell. In general, as long as a Hamiltonian H defined on this basis satisfies the translational symmetry $U_\mathbf{R} H U_\mathbf{R}^\dagger = H$ for arbitrary $\mathbf{R}$, where $U_\mathbf{R}|\mathbf{R}'; i\rangle \equiv |\mathbf{R} + \mathbf{R}'; i\rangle$, it depicts a n-band system. Now it is not difficult to make the hopping terms in H finite ranged. This usually

comes, however, at the cost of losing control on the band dispersions.

Although it is difficult to manipulate the energy of all the n bands simultaneously in the position representation, it is still possible to set some of the bands flat. It can be done by putting the bands into the null space of the Hamiltonian. For example, the famous Lieb's lattice, whose flat bands are crucial in proving magnetization when repulsive Hubbard interaction is turned on [5], consists of two sublattices A and B with unequal numbers of sites and zero hopping magnitude between any two sites in the same sublattice. This special bipartite connectivity results in $|n_A - n_B|$ flat bands that reside in the null space of the Hamiltonian, where n_A and n_B are the number of sites in the two sublattices, respectively. The flatness originates from the fact that the Hamiltonian is a direct sum of h_{AB} and h_{BA}, consisting of hopping terms starting from sites in A and B, respectively, and the rank of h_{AB} (h_{BA}) cannot be greater than the dimension of h_{BA} (h_{AB}). Sutherland [6] studies the band structure of general bipartite systems and finds that these connectivity-induced flat bands occur at the middle of the spectrum, while there exist dispersive upper and lower bands that are reflectively symmetric about the average on-site energy of the two sublattices. It is remarkable that the flat bands of a bipartite lattice will not be lifted by an external magnetic field because the flatness is not due to fine-tuning of the hopping strengths [7]. However, a general flat-band lattice model is not necessarily bipartite.

An alternative method to put a band into the null space of the Hamiltonian is based on the following feature of any flat bands: One can superimpose Bloch states on the flat band to make localized states which remain as eigenstates on the same band. Such localized states that occupy the fewest cells are called compact localized states (CLSs). CLSs can be annihilated by their parent Hamiltonian if the band energy is shifted to zero. Conversely, as long as we find one of the CLSs, we can generate a whole family of them by translation, each centering at a different cell. Altogether, this family of

CLSs form a flat band. Thus, using CLSs as generators, one can devise a systematic way of constructing flat-band models. However, it is tricky to find appropriate CLSs. For an arbitrary localized state, its parent Hamiltonian may not exist, because the Hamiltonian is found by solving an inverse eigenvalue problem that may not have a solution. Another drawback of the CLS method is that information about the band spectrum cannot be obtained readily. In particular, one cannot know *a priori* whether the flat band is a ground or an excited band. Moreover, the inverse problem is in general computationally cumbersome, particularly for spatial dimensions larger than one. Only very recently, the complete flat-band generators on the one-dimensional (1D) lattice was found [8], but its generalization to higher dimensions is not straightforward. It is thus highly desirable to develop a more powerful and convenient method of generating flat-band models [7–11] .

In this paper, we will present such a method based on the simple mathematical properties of Gram matrices. The Gram matrix method put the ground states into the null space of the Hamiltonian. It guarantees the flatness of the lowest band by simple dimension-counting procedures without invoking complicated inverse problems. Also, generating flat-band models in high-dimensional lattices is straightforward as the method is insensitive to spatial dimensions. The remainder of the paper is organized as follows. We will describe the general principle of our method in Sec. II. In the next two sections, we will present some specific examples to demonstrate its usage. In Sec. III, we demonstrate the construction of several flat-band models with finite-range hopping. In Sec. IV, we construct the long-range Kapit-Mueller model in the two-dimensional (2D) square lattice and present its generalization to arbitrary lattice geometry. Finally, we conclude in Sec. V.

II. GENERAL PRINCIPLE OF THE GRAM MATRIX METHOD

A Gram matrix G defines a semi-inner product in a linear vector space V. Given a linear transformation $T : V \rightarrow V'$, G is the pullback of the inner product defined in V', that is, $G \equiv T^{\dagger}T$. Obviously, G is Hermitian and semipositive definite. Under a basis $\{|v_i\rangle\}_{i=1}^{N}$ of the N-dimensional space V, the ith column of the T matrix is the image of $|v_i\rangle$, $|v_i'\rangle := T|v_i\rangle$, so the matrix element of G is the inner product $\langle v_i|G|v_j\rangle = \langle v_i'|v_j'\rangle$. If the set of vectors $\{|v_i'\rangle\}_{i=1}^{N}$ are linearly independent, then G is positive definite; otherwise, G would be singular and possess zero eigenvalues, and the number of zero eigenvalues equals the dimension of the kernel of T. If $N' < N$, where N' is the dimension of V', then the set $\{|v_i'\rangle\}_{i=1}^{N}$ is necessarily linearly dependent, and the number of zero eigenvalues that G possesses is at least $N - N'$.

This simple property serves as the basic principle underlying our method. Up to a shift of energy making its ground-state energy zero, a Hamiltonian can always be written as $H \equiv T^{\dagger}T$ and can thus be interpreted as a Gram matrix. Now T is a linear transformation from the Hilbert space V to an auxiliary space V'. We take the states $|\mathbf{R}; i = 1, 2, \ldots, n\rangle$'s as the basis of V. If the auxiliary space V' is spanned by sites $|\mathbf{R}; i' = 1, 2, \ldots, n'\rangle_{\text{aux}}$ in an auxiliary lattice, where $n' < n$, then we claim that *any* n-band model with $n - n'$ flat lowest

bands is associated with a matrix T whose elements are

$$_{\text{aux}}\langle \mathbf{R}'; i'|T|\mathbf{R}; i\rangle = T_{\mathbf{R}-\mathbf{R}'}^{i',i}. \tag{1}$$

Thus we can interpret the T matrix graphically as hopping terms that connect a single cell in the real lattice to sites in the auxiliary lattice. Because $H \equiv T^{\dagger}T$, as long as T is finite ranged, so will the Hamiltonian H. We will construct some specific finite-range flat-band models in Sec. III.

The proof our statement is straightforward: Any desired flat-band Hamiltonian can be decomposed as $H = \sum_{\mathbf{k}} T_{\mathbf{k}}^{\dagger}T_{\mathbf{k}}$, where $T_{\mathbf{k}}$'s are n'-by-n matrices acting on $\mathbf{k}$-Bloch states with the usual definition $|\mathbf{k}; i\rangle \equiv \sum \frac{e^{i\mathbf{k} \cdot \mathbf{R}}}{\sqrt{N}}|\mathbf{R}; i\rangle$ in momentum space. Without loss of generality, we can identify the image of $T_{\mathbf{k}}$ with the $\mathbf{k}$-Bloch states in the auxiliary space $|\mathbf{k}; i'\rangle_{\text{aux}} \equiv \sum \frac{e^{i\mathbf{k} \cdot \mathbf{R}}}{\sqrt{N'}}|\mathbf{R}; i'\rangle_{\text{aux}}$. Thus $T = \sum_{\mathbf{k}} T_{\mathbf{k}}$ and the matrix elements of T in Eq. (1) are related by a simple Fourier transformation. We can say that the translational symmetry of the Hamiltonian is inherited by the T matrix.

To summarize our Gram matrix method, here is the protocol to generate an n-band model in d-dimensional space whose lowest few bands are degenerate.

(1) The Hilbert space is spanned by $|\mathbf{R}; i = 1, 2, \ldots, n\rangle$, where $\mathbf{R}$'s are positions of the unit cells and each unit cell contains n internal sites.

(2) Construct an auxiliary space spanned by $|\mathbf{R}; i' = 1, 2, \ldots, n'\rangle_{\text{aux}}$, where $n' < n$.

(3) A mapping T, whose matrix elements are given by Eq. (1), is constructed to map between the Hilbert and the auxiliary spaces. Note that T is completely arbitrary unless some special properties are desired for the resulting Hamiltonian.

(4) The Hamiltonian can be constructed as $H = T^{\dagger}T$. The lowest $n - n'$ bands of H are guaranteed to be flat with energy 0.

The above is the most general protocol to generate models with flat lowest bands. The choice of the auxiliary space V' is flexible in the sense that the space is not necessarily spanned by the lattice $|\mathbf{R}; i' = 1, 2, \ldots, n'\rangle_{\text{aux}}$. Especially, when the restriction on finite-range hopping is lifted, then an alternative choice of V' may better serve our purpose. In Sec. IV, we will see that if we select the auxiliary space to be spanned by a subset of coherent states, then we can elegantly reproduce the Kapit-Mueller model and its generalizations, revealing the nature of the massive degeneracy in such models, which turns out to be a universal feature as required by the (over-)completeness of the coherent states.

III. GENERATING FINITE-RANGE FLAT-BAND MODELS

Let us now demonstrate the usage of the Gram matrix method by constructing several specific models. In this section, we show the simplest models constructed by the Gram matrix method. By simplest we mean that the model Hamiltonian contains the fewest number of hopping terms in the T matrix. Considering a d-dimensional lattice, the simplest choice of the T matrix is $T|\mathbf{R}; i\rangle = |\mathbf{R}; i\rangle_{\text{aux}}$. However, it is trivial because the resulting Hamiltonian does not contain hopping terms between sites in the real lattice. To obtain a nontrivial Hamiltonian, the T matrix has to map the real lattice cell at $\mathbf{R}$ to the auxiliary lattice cell at $\mathbf{R}$ and to at least d of its
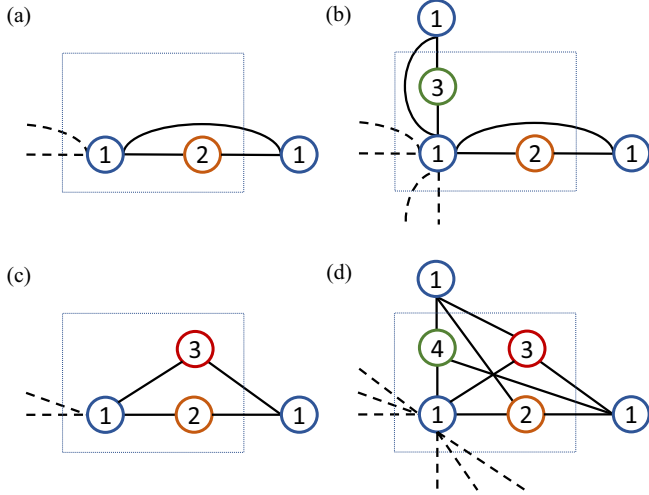
FIG. 1. Tasaki lattice in one dimension (a) and two dimensions (b) governed by Hamiltonian (2). Bipartite lattice in one dimension (c) and two dimensions (d) governed by Hamiltonian (3). Rectangular boxes represent unit cells, with internal sites labeled by circled numbers. Solid lines represent nonzero hoppings starting from a single cell. Duplicated hopping terms are marked by dashed lines.

nearest neighbor cells. For example,

$$T|\mathbf{R};1\rangle = \sum_{j=1}^{d}(a_j|\mathbf{R};j\rangle_{\text{aux}} + b_j|\mathbf{R}+\mathbf{e}_j;j\rangle_{\text{aux}}),$$

$$T|\mathbf{R};i\rangle = |\mathbf{R};i-1\rangle_{\text{aux}}, i=2,\ldots,d+1,$$

where $a_j, b_j \in \mathbb{C}$. The corresponding Hamiltonian is given by

$$H = \sum_{\mathbf{R}}\left[\sum_{i=1}^{d}(|a_i|^2+|b_i^2|)|\mathbf{R};1\rangle\langle\mathbf{R};1| + \sum_{i=2}^{d+1}|\mathbf{R};i\rangle\langle\mathbf{R};i|\right]$$

$$+ \sum_{\mathbf{R}}\sum_{i=1}^{d}[a_i|\mathbf{R};i+1\rangle\langle\mathbf{R};1|+b_i|\mathbf{R}+\mathbf{e}_i;i+1\rangle\langle\mathbf{R};1|$$

$$+ a_i^*b_i|\mathbf{R}+\mathbf{e}_i;1\rangle\langle\mathbf{R};1|+\text{H.c.}]. \quad (2)$$

The underlying lattice is the d-dimensional Tasaki's lattice [examples in one and two dimensions are presented in Figs. 1(a) and 1(b), respectively], and the hopping amplitudes in the original Tasaki's Hamiltonian [12] represent a special case of Eq. (2) with $a_i = b_i = 1/\lambda$. This Hamiltonian has $n = d + 1$ bands. Since each unit cell in the auxiliary space has $n' = d$ internal sites, H possesses $n - n' = 1$ flat band. This flat band has zero energy and represents the ground band of the system. The CLSs of the flat band can be found as

$$|\psi_{\mathbf{R}}^0\rangle = |\mathbf{R};1\rangle - \sum_{i=1}^{d}(a_i|\mathbf{R};i+1\rangle + b_i|\mathbf{R}+\mathbf{e}_i;i+1\rangle).$$

If we group the sites $|\mathbf{R};1\rangle$ as sublattice A and the remaining sites $|\mathbf{R};i=2,3,...,d+1\rangle$ as sublattice B, we find that the particle on a B site can only hop to an A site. Following the same argument about the origin of the flat bands in bipartite

lattices, we conclude that there must be $(d-1)$ additional flat bands at energy one, whose CLSs are

$$|\psi_{\mathbf{R}}^i\rangle = a_{i+1}^*|\mathbf{R};i+1\rangle + b_{i+1}^*|\mathbf{R}-\mathbf{e}_{i+1};i+1\rangle$$

$$- a_i^*|\mathbf{R};i+2\rangle - b_i^*|\mathbf{R}-\mathbf{e}_i;i+2\rangle, \quad (i=1,...,d-1).$$

Finally, there exists a dispersive top band with energy $E_{\mathbf{k}} = 1 + \sum_{i=1}^{d}|\alpha_{\mathbf{k}}^i|^2$ where $\alpha_{\mathbf{k}}^i := a_i + b_i\exp(-i\mathbf{k}\cdot\mathbf{e}_i)$. The corresponding eigenstate is

$$|\psi_{\mathbf{k}}^d\rangle = (E_{\mathbf{k}}-1)|\mathbf{k};1\rangle + \sum_{i=1}^{d}\alpha_{\mathbf{k}}^i|\mathbf{k};i+1\rangle.$$

The above example of constructing the d-dimensional Tasaki's model serves a pedagogical purpose, from which we see the power of the Gram matrix method that it takes no extra effort to generalize the model to high dimensions, whereas the conventional CLS method involves solving more and more complicated equation when d increases. By modifying the T matrix, any model with a desired number of lowest few flat bands can be readily constructed. We can also artificially put constraints on the T matrix so that the resulting model possesses desired properties. To this end, we now turn to construct a finite-range hopping lattice model which possesses a special feature: All of its bands are flat.

The basic idea rooted in the reflection symmetry of the upper and lower bands in a bipartite model. From [6] we know the middle bands in a bipartite model are flat, and if there is only one lower band (i.e., the ground band) whose flatness is guaranteed by the Gram matrix, then the only upper band must also be flat because of the reflection symmetry. So in order to build a model in which all bands are flat, we are going to design a bipartite structure which guarantees $n-2$ middle flat bands and, by choosing $n'=n-1$, there is a zero-energy flat ground band. Such a T matrix can be chosen as follows:

$$T|\mathbf{R};1\rangle = c_{d+1}|\mathbf{R};d+1\rangle_{\text{aux}} + \sum_{j=1}^{d}c_j|\mathbf{R}+\mathbf{e}_j;j\rangle_{\text{aux}},$$

$$T|\mathbf{R};i\rangle = \sum_{j=1}^{d+1}u_{i-1,j}|\mathbf{R};j\rangle_{\text{aux}}, i=2,3,\ldots,d+2,$$

where $c_j \in \mathbb{C}$ and $u_{i,j}$ is an arbitrary unitary matrix. The unitary matrix guarantees the bipartite structure: Different rows of $u_{i,j}$ are orthogonal to each other, so the hopping amplitudes between sites $|\mathbf{R};i=2,3,\ldots,d+2\rangle$ are zero, and the onsite energy of these sites is uniform because each row of the unitary matrix has the same norm. The resulting Hamiltonian is given by

$$H = \sum_{\mathbf{R}}\sum_{i=1}^{d+1}|c_i|^2|\mathbf{R};1\rangle\langle\mathbf{R};1| + \sum_{\mathbf{R}}\sum_{i=2}^{d+2}|\mathbf{R};i\rangle\langle\mathbf{R};i|$$

$$+ \sum_{\mathbf{R}}\sum_{i=1}^{d+1}\left[u_{i,d+1}^*c_{d+1}|\mathbf{R};i+1\rangle\langle\mathbf{R};1|\right.$$

$$\left. + \sum_{j=1}^{d}u_{i,j}^*c_j|\mathbf{R}+\mathbf{e}_j;i+1\rangle\langle\mathbf{R};1| + \text{H.c.}\right]. \quad (3)$$

The lattice connectivity in one and two dimensions are illustrated in Figs. 1(c) and 1(d), respectively. The bipartite nature can be easily seen if we group $|\mathbf{R}; 1\rangle$ as sublattice A with onsite energy $E_A = \sum_i^{d+1} |c_i|^2$, and $|\mathbf{R}; i = 2, .., d + 2\rangle$ as sublattice B with onsite energy $E_B = 1$. This $(d + 2)$-band model consists of a flat ground band with energy zero, d flat bands with energy $E_B = 1$, and a flat top band with energy $E_A + E_B = 1 + \sum_i^{d+1} |c_i|^2$, reflecting the reflection symmetry of bipartite lattices.

IV. THE KAPIT-MUELLER MODEL AND ITS GENERALIZATION

We have seen two examples of finite-range flat-band models. The flat bands in a finite-range Hamiltonian are necessarily nontopological with zero Chern number. This is due to a theorem [13] which states that the following three conditions concerning a band cannot be simultaneously satisfied: (1) being flat; (2) having a nonzero Chern number; (3) the Hamiltonian is finite range. To create a flat band with a finite Chern number, it is then necessary to construct a model with infinite-range hopping amplitudes. This can be done by following the same protocol and by choosing a T matrix that contains infinite-range hoppings in real space. To ensure nontrivial topology, the Fourier transform of the T matrix (i.e., the T matrix in momentum space) must possess singularities to yield finite Berry curvature. It may be more convenient to directly work in the momentum space. As we have demonstrated in the Introduction, a flat topological ground band can always be obtained by writing down a topologically nontrivial unitary matrix $U(\mathbf{k})$ to construct the Hamiltonian in the momentum space: $h(\mathbf{k}) = U(\mathbf{k})E(\mathbf{k})U^\dagger(\mathbf{k})$, where $E(\mathbf{k}) = \text{diag}(0, E_2(\mathbf{k}), E_3(\mathbf{k}), \ldots, E_n(\mathbf{k}))$. Usually, we want the constructed Hamiltonian to possess hopping amplitude whose magnitude decays fast when the hopping distance increases. It is, however, inconvenient to work in the momentum space to restrain the hopping distance. In the following, we will construct a topological flat band by slightly modifying the protocol: Instead of using orthogonal basis states $|\mathbf{R}; i' = 1, 2, \ldots, n'\rangle_{\text{aux}}$ in the auxiliary space, we choose to use nonorthogonal basis which naturally leads to long-range hopping in the resulting Hamiltonian. We will illustrate this method by constructing the Kapit-Mueller model [14].

In Ref. [14], Kapit and Mueller, working in the real space, found such a topological flat band in a 2D square lattice, and attributed the massive degeneracies in the flat band to some unrevealed symmetries. It was realized that the degenerate ground states can be regarded as discrete lowest Landau levels (LLLs), and that the degeneracy of the LLLs gives birth to the flatness. Here we find an alternative way to understand the origin of this topological flat band by reproducing the model with Gram matrices. More specifically, we reproduce the Kapit-Mueller model by a Gram matrix built upon a subset of coherent states. From this construction, the massive degeneracy of the ground band can be straightforwardly understood as a result of the linear dependency of the coherent states. And we find the generalizations of the model, which are beyond the LLL descriptions, can also share the ground band degeneracy.

A. Constructing Kapit-Mueller model on arbitrary 2D lattice

A coherent state $|z\rangle$ is an eigenstate of a bosonic annihilation operator with complex eigenvalue z. It is well known that the full set of coherent states form an overcomplete basis. Perelomov [15] studied the completeness of a countable subset of coherent states. Define

$$z_{m,n} := m\omega_1 + n\omega_2, \quad m, n \in \mathbb{Z}, \quad \omega_1, \omega_2 \in \mathbb{C}.$$

$z_{m,n}$'s form a 2D lattice on the complex plane whose unit cell area is $S := \text{Im}\,\omega_1^* \omega_2$. We collect the set of coherent states $\{|z_{m,n}\rangle\}$. Perelomov finds that if $S \leqslant \pi$, the set represents an overcomplete basis; if $S > \pi$, the set is incomplete; if $S = \pi$, we can take away any one of the $|z_{m,n}\rangle$'s from the set, and the remaining states form a complete basis.

Following the basic procedures described in Sec. II, and replacing the orthonormal basis of the auxiliary space by the lattice of coherent states, we consider a linear transformation T that maps $|m, n\rangle := |m\mathbf{e}_1 + n\mathbf{e}_2; 1\rangle$, a site on a 2D lattice containing a single state in a cell, to $|z_{m,n}\rangle$'s:

$$T|m, n\rangle = |z_{m,n}\rangle. \tag{4}$$

In other words, the columns of the T matrix are formed by $|z_{m,n}\rangle$'s. The matrix elements of the Hamiltonian $H = T^\dagger T$ are therefore the inner product of coherent states:

$$\langle m', n'|H|m, n\rangle = \langle z_{m',n'}|z_{m,n}\rangle$$
$$= e^{-|z_{m,n}-z_{m',n'}|^2/2 + i\text{Im}\,z_{m',n'}^* z_{m,n}}. \tag{5}$$

Physically, this Hamiltonian describes a fully connected 2D lattice under a magnetic field, and the flux per unit cell is $2S$. It reduces to the Kapit-Mueller model [14] when $z_{m,n}$'s form a square lattice, and to the Hofstadter model [16] by further taking the limit $S \to \infty$.

From our construction, it immediately becomes clear that the emergence of the flat ground band is guaranteed by the properties of the Gram matrix and the (over-)completeness of the coherent states. When $S > \pi$, the set of coherent states on the lattice are linearly independent, and the smallest eigenvalue of the resulting H must be positive. When $S = \pi$, the set becomes complete if we take away any one of the states, so H has a single zero eigenvalue. When $S < \pi$, we have $1/S$ states per unit area on the complex plane, while only $1/\pi$ states per unit area are needed to construct a complete basis, so a fraction of $\rho \equiv 1 - S/\pi$ eigenvalues of H must be zero. As a result, the massive degeneracy in the Hamiltonian Eq. (5) follows a universal scaling behavior, in the sense that it only depends on S and is completely independent of the lattice geometry. As we will see, it is more convenient to interpret S, instead of as the unit cell area, as the averaged area on the complex plane occupied by each coherent state, because the relation between S and the completeness extends beyond the cases that Perelomov studied. The universal degeneracy is one of the most elegant features of this model.

We carry out numerical calculation to verify the universal scaling rule between the degeneracy and S. We choose N coherent states distributed in a square region on the complex plane, where N is large but finite (typically, $N \sim 3 \times 10^3$), and numerically diagonalize the corresponding Gram matrix H to find the density of states as a function of S and the
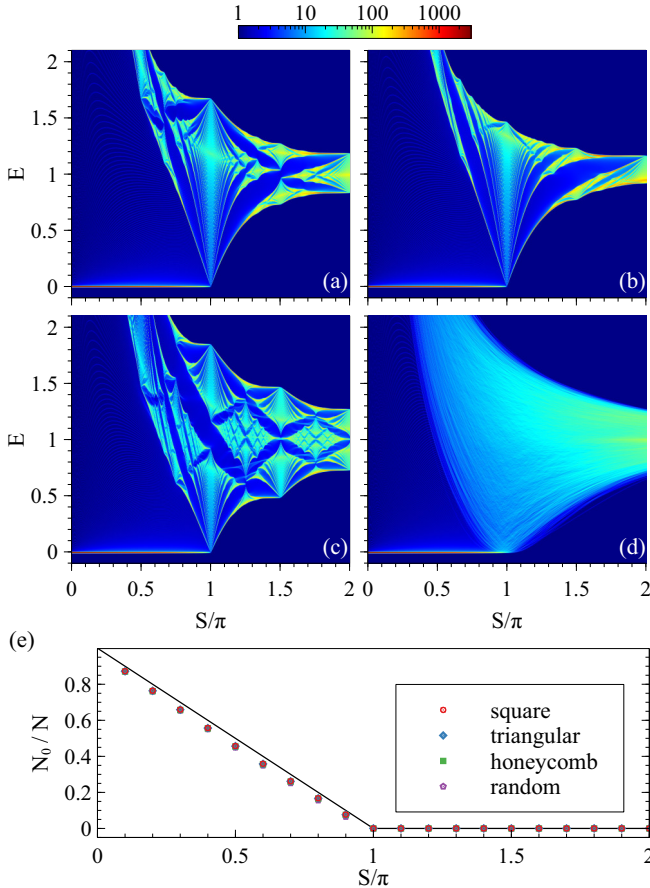
FIG. 2. (a)–(d) Spectrum of H^S for a square lattice (a), triangular lattice (b), honeycomb lattice (c), and random lattice (d). The color map represents $(D + 1)$ where $D(E) = \sum_n \delta(E - E_n)$ is the density of states, with E_n being the nth eigenvalue of H^S. In the calculation, the Dirac δ function is replaced by a smooth narrow distribution function. S is the averaged area per state. (e) We count N_0, the number of eigenenergy that is less than 10^{-5}, and compare the ratio N_0/N with the theoretical value ρ marked by the line, where $\rho = \max(1 - S/\pi, 0)$.

energy E. In Figs. 2(a)–2(c), we display the spectrum for several distinct lattice geometries (the lattice geometry refers to the geometry of the coherent states on the complex plane): square lattice, triangular lattice, and honeycomb lattice. The positive-energy part of the spectrum forms a Hofstadter butterfly, whose specific pattern depends on the lattice geometry. The universal feature for these different lattices is, however, the massively degenerate ground states at zero energy when $S < \pi$. Remarkably, this flat ground band exists even when the coherent states has a random distribution over the whole region (we specify a lower bound on the distances between sites to ensure that no two sites are too close together in order to exclude trivial zero eigenvalues of the Gram matrix) as we show in Fig. 2(d). Although in this random lattice case, the positive-energy butterfly pattern no longer exists.

As mentioned earlier, the degeneracy of the ground band is given by $N_0 = \rho N = (1 - S/\pi)N$, which should be a universal feature independent of the lattice geometry. In Fig. 2(e), we plot the numerically obtained fraction of zero-energy states

as a function for S. Results from all four lattice geometries show excellent agreement with the theoretical prediction. The small discrepancies can be attributed to the finite-size effect. In Figs. 2(a)–2(d), we see a gap between the zero-energy and the positive-energy states, which increases as S decreases and diverges when $S \to 0$. This can be easily understood as follows: Since $\text{Tr}[H] = N$ which is the sum of the positive eigenenergies, the averaged energy of the excited states should be π/S when $S < \pi$ according to the degeneracy of the ground states.

We also carry out the same numerical calculation on a small 5×5 square lattice and see that the degenerate ground states emerge again. This demonstrates that the essential physics can be observed even in such a small system, which makes its experimental realization very promising.

B. Universal ground states and the Hall dynamics

The analog between the zero-energy ground states of the model and the LLLs can be found by writing down the wave functions explicitly. Using a more general singlet sum rule [17] proved by Perelomov [15]:

$$\sum_{m,n} (-1)^{m+n+mn} e^{-\frac{|\alpha_{m,n}|^2}{2} + \alpha_{m,n}z} \equiv 0,$$

where z is an arbitrary complex number and $\alpha_{m,n}$'s build an arbitrary lattice whose unit cell area is π, we can show that the wave function of the CLS in the ground band takes the universal form ($S < \pi$),

$$\langle m, n | \psi \rangle = (-1)^{m+n+mn} e^{-\frac{\pi/S-1}{2} |z_{m,n}|^2}. \tag{6}$$

Ignoring the phase factor $(-1)^{m+n+mn}$, Eq. (6) takes the same form as the wave function of a LLL describing a particle with unit charge confined in the (x, y) plane subjected to a perpendicular magnetic field with strength B:

$$\psi_{\text{LLL}}(x, y) = e^{-B(x^2+y^2)/4\hbar}. \tag{7}$$

Also, just like the LLL, the ground band formed by $|\psi\rangle$ and its translations is topological with Chern number 1. If the nondimensionalization is done by setting $2\pi\hbar$, and the unit cell area of the lattice formed by $|m, n\rangle$'s [18] as 1, then by comparing the exponents in Eqs. (6) and (7), we find that the degeneracy of the flat band per site ρ can be regarded as the effective magnetic field for $|\psi\rangle$, which is different from the true flux $2S$. Moreover, it is this effective field ρ, rather than the true flux, that plays a role in the Hall dynamics of $|\psi\rangle$. Given a constant electric field $\mathscr{E}$, the Hall velocity of $|\psi\rangle$ is $\mathscr{E}/\rho$, because the Hall conductivity is 1. Similarly, the Hall velocity for ψ_{LLL} is $\mathscr{E}/B$.

The Hall dynamics of $|\psi\rangle$ can be verified numerically as follows. We initially prepare a wave packet in the zero-energy ground band localized at the center of an 81×81 square lattice. We then add a linear potential, with gradient U, along the y axis at $t = 0$, and the ensuing evolution of the wave packet is depicted in Fig. 3. Figure 3(a) shows the snapshots of density profiles at various times. The three columns correspond to $S = \pi/4$, $\pi/2$ and $3\pi/4$ from left to right. One can see that the wave packet disperses and moves along the x axis, perpendicular to the direction of the linear potential. The evolution
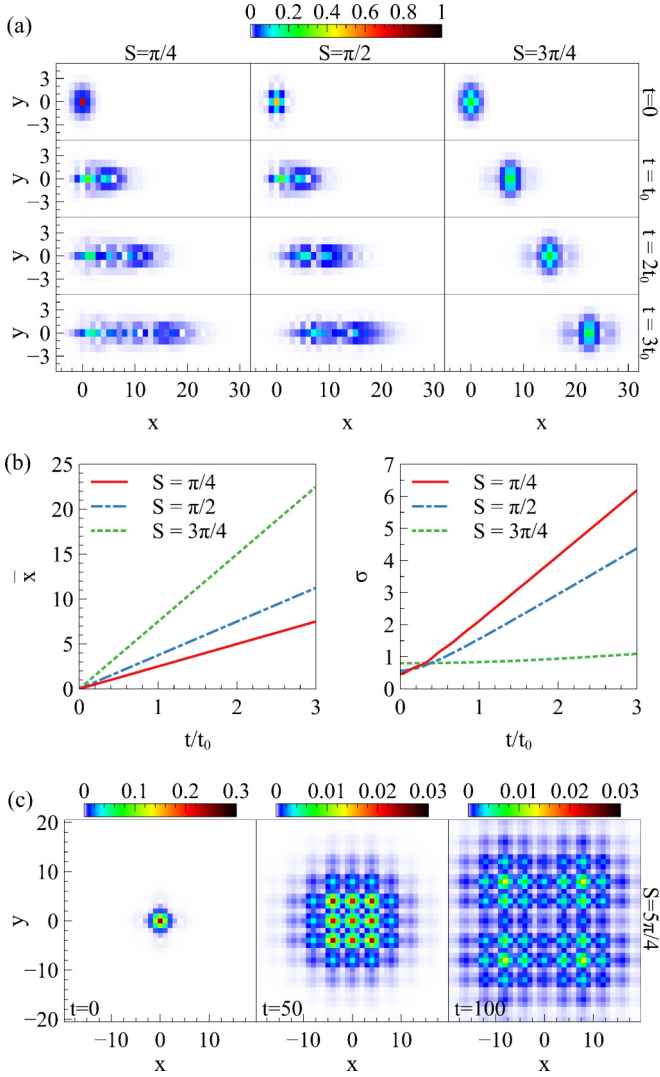
FIG. 3. (a) Evolution of the density profile of a wave packet on an $L \times L = 81 \times 81$ square lattice with a linear potential along the y axis with gradient U. The evolution operator is given by $\exp\{-2\pi it(G^S + Uy)\}$, where we take $U = 3.75 \times 10^{-4}$. The positions x and y are renormalized by the lattice constant. The initial wave packet is obtained by projecting a completely localized state to the lowest (zero-energy) band. $t_0 = 5000$. (b) The evolution of the central position $\bar{x}$ and the half width σ along the x direction of the wave packet in (a). (c) The same evolution as that in (a) except that now $S = 5\pi/4$ such that the lowest band is not a flat band.

of the center position $\bar{x}$ along the x axis is plotted in the left panel of Fig. 3(b). $\bar{x}$ is nearly linear in time, which results from the very accurate quantization of the Hall conductivity. We can readily extract the Hall velocities, $v_H = 5 \times 10^{-4}$, 7.5×10^{-4}, 1.5×10^{-3}, respectively. The Hall conductivity, given by $v_H(1 - S/\pi)/U$, is equal to 1 with a relative standard deviation on the order 10^{-7} for all three cases. As the wave packet moves, it also disperses along the x axis (but not along the y axis). The right panel of Fig. 3(b) displays the evolution of the half width σ along the x axis. The slower the wave packet moves, the faster it disperses. Note that the LLL wave packet is nondispersive under a similar situation. In Fig. 3(c), we also display the evolution of a localized wave packet for

$S = 5\pi/4$. In this case, the wave packet quickly disperses in both directions, in stark contrast against the situation depicted in Fig. 3(a) where $S < \pi$.

C. Analytic derivation of the full spectrum for special cases

Despite the numerical evidence that supports universal degeneracy presented in the generalized Kapit-Mueller model, it is helpful if we can prove the degeneracy and write down the Bloch wave functions by solving the Hamiltonian Eq. (5) directly. It can be done for some special choices of $|z_{m,n}\rangle$'s.

Consider the case when ρ is rational by taking $S = p\pi/q$ where p, q are co-prime positive integers. After a local gauge transformation $|m, n\rangle \to e^{-imnS}|m, n\rangle$ and then a Fourier transformation, H can be reduced to a q-band Bloch Hamiltonian $h^{p,q}(k, l)$, where k, l are pseudomomenta defined in the range $-\frac{1}{q} < k \leq \frac{1}{q}$, $-1 < l \leq 1$. The explicit matrix elements of $h^{p,q}$ are given by

$$
\begin{aligned}
h_{m',m}^{p,q} = \sum_{r,s} \exp\Big\{ &-|z_{qr+m-m',s}|^2/2 + i\pi\big[k(qr + m - m') \\
&+ (l + (m + m')pq)s - prs\big]\Big\}, \quad (m, m' = 1, 2, ..., q).
\end{aligned}
$$

If $z_{m,n}$'s form a rectangular lattice with aspect ratio ξ on the complex plane, then $h_{m',m}^{p,q}$ can be expressed through Jacobi θ functions (see the Appendix for definitions):

$$
\begin{aligned}
h_{m',m}^{p,q} = \exp\Big(&-\frac{S\xi(m - m')^2}{2} + i\pi k(m - m') \Big) \\
&\times \sum_{\chi=0,1} \theta_{3+\chi(p \bmod 2)}(z_1, \tau_1)\theta_{3-\chi}(z_2, \tau_2),
\end{aligned}
$$

where

$$
\tau_1 := \frac{ipq\xi}{2}, \qquad z_1 := \frac{qk}{2} + \frac{(m - m')\tau_1}{q},
$$

$$
\tau_2 := \frac{2ip}{q\xi}, \qquad z_2 := l + \frac{(m + m')p}{q}.
$$

We are going to verify that when $p = 1$, $h_{m',m}^{p,q}$ is a rank-1 matrix, so it has only one nonzero eigenvalue, which is predicted by the Gram matrix construction. Now $p = 1$, $\tau_1 = -\tau_2^{-1}$, so we can use Jacobi identities,

$$
\theta_{3+\chi}\Big(\frac{z}{\tau}, -\frac{1}{\tau}\Big) = (-i\tau)^{1/2} \exp\big(i\pi z^2/\tau\big)\theta_{3-\chi}(z, \tau),
$$

where $\chi = 0, \pm 1$, to recast $h^{1,q}$ as follows:

$$
\begin{aligned}
h_{m',m}^{1,q} = (-i\tau_2)^{1/2} &\sum_{\chi=0,1} \theta_{3-\chi}(z_1\tau_2)\theta_{3-\chi}(z_2) \\
&\times \exp\Big(-\frac{S\xi(m - m')^2}{2} + i\pi k(m - m') + i\pi z_1^2\tau_2 \Big),
\end{aligned}
$$

where, for simplicity, we have omitted the second argument τ_2 of the θ functions. Define

$$
\begin{aligned}
R_{m,m'} := \frac{i}{\tau_2}\big(&h_{m',m}^{1,q}h_{m'+1,m+1}^{1,q} - h_{m'+1,m}^{1,q}h_{m',m+1}^{1,q}\big) \\
&\times \exp\big(S\xi(m - m')^2 - 2i\pi\big(k(m - m') + z_1^2\tau_2\big)\big).
\end{aligned}
$$

$R_{m,m'} \equiv 0$ iff $h^{1,q}$ is rank-1. Representing $h^{1,q}$ by the θ functions, we have

$$R_{m,m'} = \sum_{\chi,\chi'} \theta_{3-\chi}(z_1\tau_2)\theta_{3-\chi}(z_2)\theta_{3-\chi}(z_1\tau_2)\theta_{3-\chi'}\left(z_2 + \frac{2}{q}\right)$$

$$- \theta_{3-\chi}\left(z_1\tau_2 + \frac{1}{q}\right)\theta_{3-\chi}\left(z_2 + \frac{1}{q}\right)$$

$$\times \theta_{3-\chi'}\left(z_1\tau_2 - \frac{1}{q}\right)\theta_{3-\chi'}\left(z_2 + \frac{1}{q}\right). \quad (8)$$

Using the sum rules of the θ functions, presented in the Appendix, we can prove

$$R_{m,m'} = 0.$$

So we have proved that, indeed, for rectangular lattice with aspect ratio ξ, $h^{1,q}$ has only one positive eigenvalue, whose value is given by the trace:

$$\mathrm{Tr}(h^{1,q}) = q \sum_{r,s} \exp\left(-\frac{1}{2}|z_{qr,qs}|^2 + i\pi q(kr + ls - rs)\right).$$

The corresponding Bloch wave function is given by any column vector of $h^{1,q}$.

D. Generalization beyond the LLL analogy

The Kapit-Mueller model can also be constructed using a projection method [19] which is analogous to the inverse method by treating the LLLs as CLSs. However, using the Gram matrix method, infinite variations of the model, which cannot be regarded as the parent Hamiltonians of the LLLs, can be constructed a similar protocol described in Sec. II. In the discussion above to construct the Kapit-Mueller model, we have picked a particular T matrix defined in Eq. (4). However, we can define the T matrix, which maps the real cell $|m, n\rangle$ to an arbitrary number of cells represented by the coherent states in the auxiliary space, in a more general way as follows:

$$T|m, n\rangle = \sum_{m',n'} \tau_{m',n'}^{m,n} |z_{m',n'}\rangle. \quad (9)$$

The resulting Hamiltonian is $H'^{\dagger}H\tau$, where H is the Gram matrix of the coherent states given in Eq. (5) which yields the Kapit-Mueller model. Apparently, the new Hamiltonian H' has at least the same number of zero eigenvalues as that of H. In order that the zero-energy states of H' form a flat band, H' must preserve the translational symmetry of H, which constrains the T matrix in Eq. (9). The translational symmetry is preserved as long as

$$\Phi_{\mathbf{R}}^{\dagger} U_{\mathbf{R}}^{\dagger} \tau U_{\mathbf{R}} \Phi_{\mathbf{R}} = \tau, \quad (10)$$

where $U_{\mathbf{R}}$ is the translation operator and $\Phi_{\mathbf{R}}$ a local gauge transformation as the result of the gauge flux through the lattice:

$$\Phi_{p\mathbf{e}_1+q\mathbf{e}_2}|m, n\rangle = \exp\left(-iS(np - qm)\right)|m, n\rangle,$$

where $\mathbf{e}_{1,2}$ are the two basis vectors of the lattice, and $p, q \in \mathbb{Z}$. To find τ's that preserve the translational symmetry, we
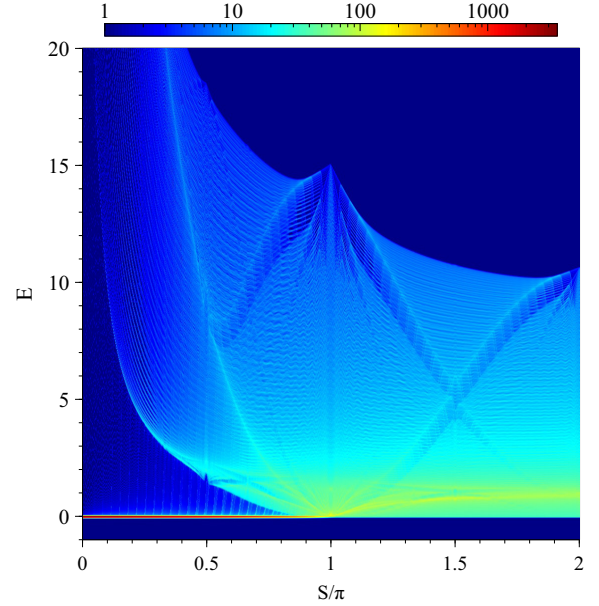

FIG. 4. The spectrum of the generalized Kapit-Mueller model given by the T matrix Eq. (13). The color map has the same meaning as that in Fig. 2(a).

recast τ into the form,

$$\tau_{m',n'}^{m,n} = \sum_{r,s} C(m, n; r, s)\delta_{m'}^{m+r}\delta_{n'}^{n+s}. \quad (11)$$

Now we have

$$(\Phi_{\mathbf{R}}^{\dagger} U_{\mathbf{R}}^{\dagger} \tau U_{\mathbf{R}} \Phi_{\mathbf{R}})_{m',n'}^{m,n}\Big|_{\mathbf{R}=p\mathbf{e}_1+q\mathbf{e}_2}$$

$$= \sum_{r,s} C(m + p, n + q; r, s) \exp\left(iS(sp - qr)\right)\delta_{m'}^{m+r}\delta_{n'}^{n+s}.$$

To make Eq. (10) hold, $C(m, n; r, s)$ must take the form,

$$C(m, n; r, s) = C(r, s) \exp\left[iS(rn - sm)\right], \quad (12)$$

where $C(r, s)$ is an arbitrary function.

For example, analogous to the 2D Tasaki's lattice constructed in Sec. III, a T matrix that maps a real cell $|m, n\rangle$ to the auxiliary cell $|z_{m,n}\rangle$ and its two nearest neighbors is given by

$$T|m, n\rangle = |z_{m,n}\rangle + e^{iSn}|z_{m+1,n}\rangle + e^{-iSm}|z_{m,n+1}\rangle. \quad (13)$$

The resulting Hamiltonian's zero-energy band has the same degeneracy as that of the Kapit-Mueller model. In Fig. 4, we illustrate the spectrum of this Hamiltonian on a square lattice by plotting the density of states as a function of the unit cell area S.

V. CONCLUSION

We have proposed a powerful method of constructing lattice models supporting flat bands. The method is based on the mathematical properties of Gram matrices. Any lattice model with the flat lowest band can be constructed through the method. The method does not require any elaborate calculations such as solving the inverse eigenvalue problem, works for arbitrary spatial dimensions, and guarantees to produce

a flat ground band. We have presented a variety of examples, including both finite- and infinite-range hopping, and topologically trivial and nontrivial flat bands. Specifically, we have constructed the d-dim Tasaki lattice, a d-dim bipartite lattice whose bands are all flat, and the generalized Kapit-Mueller lattice whose flat ground band features universal (i.e., geometry-independent) degeneracy. We study the generalized Kapit-Mueller in detail and, especially, conclude that the (over-)completeness of the coherent states is the origin of the universal degeneracy.

Finally, we want to comment on realizing flat-band models in laboratory. Over the past few years, we have witnessed rapid progress in realizing lattice models in synthetic materials, particularly synthetic dimensions, in both atomic [20–29] and photonic [30–40] systems, where the lattice sites are represented by different atomic states or photonic modes, respectively. Nearly arbitrary hopping amplitudes can be realized in such systems. Realizing the flat-band models constructed using our method with synthetic materials should therefore pose no essential difficulties.

APPENDIX: JACOBI θ FUNCTION

The Jacobi θ functions are defined by

$$\theta_1(z, \tau) := \sum_n \exp\left(\pi i \tau (n+1/2)^2 + 2i\pi \left(z - \frac{1}{2}\right)(n+1/2)\right),$$

$$\theta_2(z, \tau) := \sum_n \exp(\pi i \tau (n+1/2)^2 + 2i\pi z(n+1/2)),$$

$$\theta_3(z, \tau) := \sum_n \exp(\pi i \tau n^2 + 2i\pi z n),$$

$$\theta_4(z, \tau) := \sum_n \exp\left(\pi i \tau n^2 + 2i\pi \left(z - \frac{1}{2}\right)n\right).$$

The sum rules of the θ functions used in the paper are

$$\theta_2(v + w)\theta_2(v - w)\theta_2(x + y)\theta_2(x - y)$$
$$- \theta_2(x + w)\theta_2(x - w)\theta_2(v + y)\theta_2(v - y)$$
$$= \theta_1(v + x)\theta_1(v - x)\theta_1(y + w)\theta_1(y - w).$$

$$\theta_3(v + w)\theta_3(v - w)\theta_3(x + y)\theta_3(x - y)$$
$$- \theta_3(x + w)\theta_3(x - w)\theta_3(v + y)\theta_3(v - y)$$
$$= -\theta_1(v + x)\theta_1(v - x)\theta_1(y + w)\theta_1(y - w).$$

$$\theta_2(v + w)\theta_3(v - w)\theta_3(x + y)\theta_2(x - y)$$
$$- \theta_2(x + w)\theta_3(x - w)\theta_3(v + y)\theta_2(v - y)$$
$$= -\theta_4(v + x)\theta_1(v - x)\theta_1(y + w)\theta_4(y - w).$$

By taking

$$v = z_1\tau_2, \qquad w = 0,$$
$$x = z_2 + \frac{1}{q}, \qquad y = \frac{1}{q},$$

one can readily show that $R_{m,m'}$ in Eq. (8) indeed vanishes.

[1] D. Leykam, A. Andreanov, and S. Flach, Adv. Phys.: X **3**, 1473052 (2018).

[2] R. B. Laughlin, Phys. Rev. Lett. **50**, 1395 (1983).

[3] J. K. Jain, Phys. Rev. Lett. **63**, 199 (1989).

[4] H. L. Stormer, D. C. Tsui, and A. C. Gossard, Rev. Mod. Phys. **71**, S298 (1999).

[5] E. H. Lieb, Phys. Rev. Lett. **62**, 1927(E) (1989).

[6] B. Sutherland, Phys. Rev. B **34**, 5208 (1986).

[7] H. Aoki, M. Ando, and H. Matsumura, Phys. Rev. B **54**, R17296 (1996).

[8] W. Maimaiti, S. Flach, and A. Andreanov, Phys. Rev. B **99**, 125129 (2019).

[9] W. Maimaiti, A. Andreanov, H. C. Park, O. Gendelman, and S. Flach, Phys. Rev. B **95**, 115135 (2017).

[10] R. A. Vicencio, C. Cantillano, L. Morales-Inostroza, B. Real, C. Mejía-Cortés, S. Weimann, A. Szameit, and M. I. Molina, Phys. Rev. Lett. **114**, 245503 (2015).

[11] Y. Zong, S. Xia, L. Tang, D. Song, Y. Hu, Y. Pei, J. Su, Y. Li, and Z. Chen, Opt. Express **24**, 8877 (2016).

[12] H. Tasaki, Phys. Rev. Lett. **69**, 1608 (1992).

[13] L. Chen, T. Mazaheri, A. Seidel, and X. Tang, J. Phys. A: Math. Theor. **47**, 152001 (2014).

[14] E. Kapit and E. Mueller, Phys. Rev. Lett. **105**, 215303 (2010).

[15] A. M. Perelomov, Theor. Math. Phys. **6**, 156 (1971).

[16] D. R. Hofstadter, Phys. Rev. B **14**, 2239 (1976).

[17] R. Laughlin, Ann. Phys. **191**, 163 (1989).

[18] The distance between physical sites is not necessarily the same as the distance between $z_{m,n}$'s on the complex plane. Here we make it independent from S.

[19] H. Atakişi and M. O. Oktel, Phys. Rev. A **88**, 033612 (2013).

[20] A. Celi, P. Massignan, J. Ruseckas, N. Goldman, I. B. Spielman, G. Juzeliūnas, and M. Lewenstein, Phys. Rev. Lett. **112**, 043001 (2014).

[21] M. Mancini, G. Pagano, G. Cappellini, L. Livi, M. Rider, J. Catani, C. Sias, P. Zoller, M. Inguscio, M. Dalmonte, and L. Fallani, Science **349**, 1510 (2015).

[22] B. K. Stuhl, H.-I. Lu, L. M. Aycock, D. Genkina, and I. B. Spielman, Science **349**, 1514 (2015).

[23] T.-S. Zeng, C. Wang, and H. Zhai, Phys. Rev. Lett. **115**, 095302 (2015).

[24] S. Barbarino, L. Taddia, D. Rossini, L. Mazza, and R. Fazio, New J. Phys. **18**, 035010 (2016).

[25] B. Sundar, M. Thibodeau, Z. Wang, B. Gadway, and K. R. A. Hazzard, Phys. Rev. A **99**, 013624 (2019).

[26] B. Sundar, B. Gadway, and K. R. Hazzard, Sci. Rep. **8**, 1 (2018).

[27] F. A. An, E. J. Meier, J. Ang'ong'a, and B. Gadway, Phys. Rev. Lett. **120**, 040407 (2018).

[28] M. Lohse, C. Schweizer, H. M. Price, O. Zilberberg, and I. Bloch, Nature (London) **553**, 55 (2018).

[29] S. Sugawa, F. Salces-Carcoba, A. R. Perry, Y. Yue, and I. B. Spielman, Science **360**, 1429 (2018).

[30] E. Lustig, S. Weimann, Y. Plotnik, Y. Lumer, M. A. Bandres, A. Szameit, and M. Segev, Nature (London) **567**, 356 (2019).

[31] D. Jukić and H. Buljan, Phys. Rev. A **87**, 013814 (2013).

[32] X.-W. Luo, X. Zhou, C.-F. Li, J.-S. Xu, G.-C. Guo, and Z.-W. Zhou, Nat. Commun. **6**, 1 (2015).

[33] L. Yuan, Y. Shi, and S. Fan, Opt. Lett. **41**, 741 (2016).

[34] T. Ozawa, H. M. Price, N. Goldman, O. Zilberberg, and I. Carusotto, Phys. Rev. A **93**, 043827 (2016).

[35] L. Yuan, Q. Lin, M. Xiao, and S. Fan, Optica **5**, 1396 (2018).

[36] X.-F. Zhou, X.-W. Luo, S. Wang, G.-C. Guo, X. Zhou, H. Pu, and Z.-W. Zhou, Phys. Rev. Lett. **118**, 083603 (2017).

[37] R. Ma, C. Owens, A. LaChapelle, D. I. Schuster, and J. Simon, Phys. Rev. A **95**, 062120 (2017).

[38] S. Wang, X.-F. Zhou, G.-C. Guo, H. Pu, and Z.-W. Zhou, Phys. Rev. A **100**, 043817 (2019).

[39] B. M. Anderson, R. Ma, C. Owens, D. I. Schuster, and J. Simon, Physical Review X **6**, 041043 (2016).

[40] N. Schine, A. Ryou, A. Gromov, A. Sommer, and J. Simon, Nature (London) **534**, 671 (2016).