Tovmasyan, Murad; Peotta, Sebastiano; Törmä, Päivi; Huber, Sebastian D.

Effective theory and emergent SU(2) symmetry in the flat bands of attractive Hubbard models

Published in:
Physical Review B (Condensed Matter and Materials Physics)

DOI:
10.1103/PhysRevB.94.245149

Published: 30/12/2016

Document Version
Publisher's PDF, also known as Version of record

Please cite the original version:
Tovmasyan, M., Peotta, S., Törmä, P., & Huber, S. D. (2016). Effective theory and emergent SU(2) symmetry in the flat bands of attractive Hubbard models. Physical Review B (Condensed Matter and Materials Physics), 94(24), 1-17. [245149]. https://doi.org/10.1103/PhysRevB.94.245149

Effective theory and emergent SU(2) symmetry in the flat bands of attractive Hubbard models

Murad Tovmasyan,[1] Sebastiano Peotta,[2] Päivi Törmä,[2,*] and Sebastian D. Huber[1,†]
[1]*Institute for Theoretical Physics, ETH Zurich, 8093 Zürich, Switzerland*
[2]*COMP Centre of Excellence, Department of Applied Physics, Aalto University School of Science, FI-00076 Aalto, Finland*
(Received 4 August 2016; revised manuscript received 14 November 2016; published 30 December 2016)

In a partially filled flat Bloch band electrons do not have a well defined Fermi surface and hence the low-energy theory is not a Fermi liquid. Nevertheless, under the influence of an attractive interaction, a superconductor well described by the Bardeen-Cooper-Schrieffer (BCS) wave function can arise. Here we study the low-energy effective Hamiltonian of a generic Hubbard model with a flat band. We obtain an effective Hamiltonian for the flat band physics by eliminating higher-lying bands via the perturbative Schrieffer-Wolff transformation. At first order in the interaction energy we recover the usual procedure of projecting the interaction term onto the flat band Wannier functions. We show that the BCS wave function is the exact ground state of the projected interaction Hamiltonian, if a simple uniform pairing condition on the single-particle states is satisfied, and that the compressibility is diverging as a consequence of an emergent SU(2) symmetry. This symmetry is broken by second-order interband transitions resulting in a finite compressibility, which we illustrate for a one-dimensional ladder with two perfectly flat bands. These results motivate a further approximation leading to an effective ferromagnetic Heisenberg model. The gauge-invariant result for the superfluid weight of a flat band can be obtained from the ferromagnetic Heisenberg model only if the maximally localized Wannier functions in the Marzari-Vanderbilt sense are used. Finally, we prove an important inequality $D \geqslant \mathcal{W}^2$ between the Drude weight D and the winding number $\mathcal{W}$, which guarantees ballistic transport for topologically nontrivial flat bands in one dimension.

DOI: 10.1103/PhysRevB.94.245149

I. INTRODUCTION

In quantum mechanics particles can localize due to the destructive interference between different classical trajectories. Such localization can result in the formation of flat bands with a diverging effective mass, or equivalently zero group velocity. Historically the first example of this phenomenon is the formation of the Landau levels of a particle in the presence of a uniform magnetic field. While this localization is a purely single-particle effect, the presence of a flat band can have a profound impact on the physics of interacting many-body systems. A prime example is the fractional quantum Hall effect [1]. Beyond the physics in a strong magnetic field, flat or nearly flat bands are a relatively common occurrence in lattice Hamiltonians where they can be realized by engineering suitable hopping matrix elements [2–17]. Given the vanishing group velocity, one could expect an electronic flat band system to be a particularly bad conductor. This is certainly true for the high-temperature phase. However, matters are less clear if a superconductor is formed in such a flat band. Here, we investigate this scenario with a special emphasis on the transport properties.

Why can we expect a superconductor to appear in a flat band to begin with? Let us recall that the BCS superconducting transition temperature scales as $T_c \propto \exp(-1/U D_\mathrm{F})$, where U is the pairing interaction strength and D_F is the density of states at the Fermi energy. On the other hand, it is also known that a flat band promotes superconductivity with a transition temperature $T_c \propto U$, which can be substantially higher than the usual exponentially small transition temperature [18–22]. Hence, we can indeed hope for an enhanced superconducting instability in a flat band. This now raises questions regarding the superfluid properties of such a flat band superconductor.

That superfluid transport can occur in the limit of strictly flat bands has been known from the experiments on the exciton condensate in quantum Hall bilayers [23–25]. However, it has been realized only recently that even flat bands that emerge from the geometry of the lattice in the presence of attractive interactions can sustain a large superfluid current. This occurs because in the flat band limit the superfluid weight is controlled by a band structure quantity called the quantum metric which is distinct from, but related to, the Chern number [26].

Despite the recent progress in the study of many-body states in the flat bands of lattice Hamiltonians, there are still many open questions. For example, it has been found in the case of bipartite lattices that the ground state in the presence of an attractive Hubbard interaction is well approximated by the BCS wave function [27]. However, it is quite unclear what are the properties of the normal state above the superconducting transition.

By definition the normal state is not a Fermi liquid since as the interaction is turned off the system becomes an insulator for any filling of the flat band. Indeed, the picture that the divergence of the effective mass implies the absence of transport is true for any noninteracting system.

Transport in a flat band is a consequence of either interaction or disorder. It is therefore interesting to characterize the properties both of the normal state and of the superconducting state at nonzero temperature as these can be accessed in current ultracold gas experiments. Moreover, the situation where the superconducting ground state is well known but the normal metallic state is less understood is reminiscent of the one in high-T_c superconductors, especially cuprates and iron pnictides, where the enigmatic pseudogap phase is

*ptorma@aalto.fi
†sebastian.huber@phys.ethz.ch

believed to be important for unveiling the pairing mechanism that gives rise to the superconducting phase. It may be possible that idealized flat band models can provide clues in this direction.

In this work we address the general problem of providing a reliable low-energy effective theory for the flat band of a multiband lattice Hamiltonian in the presence of an attractive Hubbard interaction. It is important to consider a multiband Hamiltonian since in the case of a single-band Hamiltonian, i.e., a Hamiltonian defined on a simple lattice, the only flat band that can be obtained is trivial and corresponds to the limit where all the sites are decoupled (atomic limit). Furthermore, in the multiband case the flat band can have a wide variety of properties encoded in suitable invariants constructed from the Bloch/Wannier functions. An example is the Chern number $\mathcal{C}$, a topological invariant signaling, when nonzero, that the flat band cannot be connected adiabatically to the atomic limit ($\mathcal{C} = 0$) or to a band with different Chern number. Interestingly, even a topologically trivial flat band ($\mathcal{C} = 0$) can have a nonzero quantum metric and therefore host a superconducting state [26]. The subject of multiband superconductivity has recently become important with the discovery of materials such as magnesium diboride and iron pnictides superconductors. Flat band superconductivity is an exotic example of the rich variety of phenomena encountered in multiband superconductivity.

The strategy of this work is to combine the result of Ref. [26] for the superfluid weight of a flat band in terms of the quantum metric with the general approach of Ref. [28] where flat bands are studied by projecting the interaction Hamiltonian on the Wannier functions of the flat band. This latter approach has the advantage of providing a simple low-energy effective Hamiltonian which is often accurate in predicting the properties of the ground state.

We are able to prove the useful result that, in the wide class of Hubbard models considered here, the BCS wave function is the exact ground state of the projected interaction Hamiltonian under a simple uniform pairing condition on the single-particle states. Concomitantly, the compressibility in the partially filled flat band is diverging. In fact we show how both these properties are the manifestation of an emerging SU(2) symmetry, which is due to the band flatness.

Guided by these rigorous results, we approximate the projected interaction Hamiltonian by an effective ferromagnetic Heisenberg model. Since the Wannier functions are defined up to a unitary transformation, any approximation performed on the projected interaction Hamiltonian depends on the specific choice of the Wannier functions. Therefore, the drawback of this approximation is that it depends on the specific choice of Wannier functions. Using the gauge invariant result of Ref. [26], we show that there is a preferred choice for the Wannier functions, which coincides with the maximally localized Wannier functions in the Marzari-Vanderbilt sense.

As a concrete example of our general results, we consider an attractive Hubbard model defined on a one-dimensional ladder, the Creutz ladder [3], which is sketched in Fig. 1. The band structure of the Creutz ladder consists of two perfectly flat bands. In the case of the attractive Creutz-Hubbard model the resulting effective ferromagnetic Heisenberg model takes the form of the integrable XXX chain.

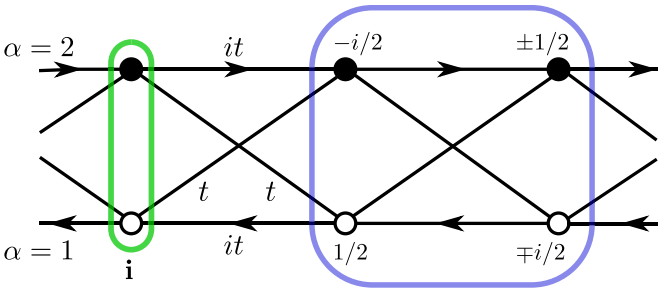

FIG. 1. The Creutz ladder: The green box indicates the **i**th unit cell containing two orbitals depicted by an empty and a full circle. The two sublattices are labeled by α. The hopping amplitudes for the spin-$\uparrow$ fermions are given on corresponding links. The arrows indicate the sense of the complex hoppings. The blue box indicates the maximally localized Wannier function for the upper/lower ($\pm$) flat band, also called plaquette states; the numbers correspond to the respective amplitudes which are nonzero only for the sites inside the blue box.

We find also that in order to account for a finite compressibility it is necessary to include the effects of interband transitions resulting in higher-order terms in the effective Hamiltonian. In the case of the Creutz-Hubbard model we provide an analytic result for the compressibility up to second order in the ratio U/E_{gap} between the interaction strength and the band gap. This result is tested against density-matrix renormalization-group (DMRG) simulations, and we find an excellent agreement.

Finally, we prove an important bound $D \geqslant \mathcal{W}^2$ between Drude weight D and the one-dimensional winding number, which extends the result for the superfluid weight $D_{\mathrm{s}} \geqslant |\mathcal{C}|$ valid in two dimensions [26]. It is shown using the Creutz-Hubbard model that the inequality is in fact optimal.

The paper is organized as follows. In Sec. II we introduce the model, the basic notations, and the perturbative Schrieffer-Wolff (SW) transformation for a generic Hubbard model. In Sec. III we derive the result that the BCS wave function is exact in the isolated flat band limit and provide the generators of the emergent SU(2) symmetry. In Sec. IV, after a suitable truncation, we map the projected interaction Hamiltonian into an effective ferromagnetic Heisenberg model. We show that this mapping is in fact exact for the Creutz-Hubbard model. In Sec. V we discuss how the result of Ref. [26] for the superfluid weight of a flat band can be recovered from the effective Heisenberg model. This is done by introducing an overlap functional for Wannier functions, whose relation with the usual Marzari-Vanderbilt functional is analyzed in detail. The proof of various results relating the two functionals is detailed in Appendix C. In Sec. VI we derive for the Creutz-Hubbard model the second-order corrections to the projected Hamiltonian using the SW transformation. In Sec. VII we use the result of the previous section to derive an analytic result for the compressibility of the Creutz-Hubbard model, which is then compared to DMRG simulations. In Sec. VIII the inequality between superfluid weight and winding number is proved. Our results and the future perspectives are discussed in the last section. Moreover, in Appendix A we discuss the relation between attractive and repulsive Hubbard models in

a flat band, and in Appendix B we prove a general identity useful for calculating the interband effects.

II. DEFINITIONS AND NOTATIONS

A. Model

Here we introduce the class of flat band models with attractive Hubbard interactions defined on a d-dimensional lattice that are considered in this work. We shall also give some useful definitions and consider a one-dimensional example, the Creutz-Hubbard model.

The kinetic part of the tight-binding Hamiltonian $\hat{\mathcal{H}} = \hat{\mathcal{H}}_{\rm kin} + \hat{\mathcal{H}}_{\rm int}$ is given by

$$\hat{\mathcal{H}}_{\rm kin} = \sum_{\sigma=\uparrow,\downarrow} \sum_{\mathbf{i}\alpha,\,\mathbf{j}\beta} t^{\sigma}_{\mathbf{i}\alpha,\,\mathbf{j}\beta}\, \hat{c}^{\dagger}_{\mathbf{i}\alpha\sigma} \hat{c}_{\mathbf{j}\beta\sigma}, \tag{1}$$

where, as usual, $\hat{c}^{\dagger}_{\mathbf{i}\alpha\sigma}$ and $\hat{c}_{\mathbf{i}\alpha\sigma}$ are the creation and the annihilation operators, respectively, of a fermion with spin σ at unit cell $\mathbf{i}$ and sublattice α. According to Eq. (1) the number of spin-$\uparrow$ and spin-$\downarrow$ particles are separately conserved. In order to preserve the time-reversal symmetry and favor the occurrence of Cooper pairs in the spirit of Anderson's theorems [29,30], we take the hopping matrix $(t^{\sigma}_{\mathbf{i}\alpha,\mathbf{j}\beta})$ such that the kinetic Hamiltonian is time-reversal invariant, namely, $t^{\downarrow}_{\mathbf{i}\alpha,\mathbf{j}\beta} = (t^{\uparrow}_{\mathbf{i}\alpha,\mathbf{j}\beta})^*$, where the star denotes the complex conjugate. Moreover, we consider models in which the kinetic Hamiltonian has an *isolated flat Bloch band* separated from the other bands by a finite energy gap $E_{\rm gap}$. Without loss of generality, we concentrate on the case where the flat band is the lowest lying band to simplify the notation.

The attractive Hubbard interaction term has the form

$$\hat{\mathcal{H}}_{\rm int} = -U \sum_{\mathbf{i}\alpha} \hat{n}_{\mathbf{i}\alpha\uparrow} \hat{n}_{\mathbf{i}\alpha\downarrow}, \tag{2}$$

where $\hat{n}_{\mathbf{i}\alpha\sigma} = \hat{c}^{\dagger}_{\mathbf{i}\alpha\sigma} \hat{c}_{\mathbf{i}\alpha\sigma}$ is the fermionic number operator, and we assume that $U > 0$.

Occasionally we comment on the corresponding repulsive Hubbard model, where by "corresponding" we mean that the spin-$\uparrow$ and spin-$\downarrow$ kinetic Hamiltonians are equal ($t^{\uparrow}_{\mathbf{i}\alpha,\mathbf{j}\beta} = t^{\downarrow}_{\mathbf{i}\alpha,\mathbf{j}\beta}$) and equal to the spin-$\uparrow$ kinetic Hamiltonian of the attractive model. Thus the corresponding repulsive Hubbard model possesses full SU(2) spin rotational symmetry. We will see that there is to some extent a duality between flat band superconductivity in the attractive model and flat band ferromagnetism in the corresponding repulsive model.

For the discussion of the interacting problem below, the Wannier states form a convenient local orthonormal basis [6,28]. The Wannier function of the flat band centered at the unit cell $\mathbf{j}$ is constructed from the Bloch functions $g_{\mathbf{k}\sigma}(\alpha)$ of the flat band according to [31,32]

$$W_{\alpha\sigma}(\mathbf{i}-\mathbf{j}) = \frac{V_{\rm c}}{(2\pi)^d} \int_{\rm B.Z.} d^d\mathbf{k}\, e^{i\mathbf{k}\cdot(\mathbf{r}_{\mathbf{i}}-\mathbf{r}_{\mathbf{j}})} g_{\mathbf{k}\sigma}(\alpha), \tag{3}$$

where $V_{\rm c}$ denotes the volume of the unit cell and $\mathbf{r}_{\mathbf{i}} = i_1\mathbf{a}_1 + i_2\mathbf{a}_2 + \cdots + i_d\mathbf{a}_d$ is the lattice vector corresponding to unit cell $\mathbf{i} = (i_1, i_2, \ldots, i_d)$. The fundamental lattice vectors of the underlying Bravais lattice are $\mathbf{a}_{j=1,2,\ldots,d}$.

It is important to note that in Eq. (3) the Wannier functions $W_{\alpha\sigma}(\mathbf{i})$ are defined in terms of the Bloch functions $g_{\mathbf{k}\sigma}(\alpha)$, which in turn are obtained from the diagonalization of the hopping matrix $t_{\mathbf{i}\alpha,\mathbf{j}\beta} = t_{\alpha,\beta}(\mathbf{i}-\mathbf{j})$. In particular, they are distinct from the Bloch functions $g_{\mathbf{k}\sigma}(\mathbf{r})$, with a continuous variable $\mathbf{r}$, of the continuous Hamiltonian from which the lattice model is derived via the tight-binding approximation. The Wannier functions corresponding to the continuous Hamiltonian can be constructed from the Wannier functions (3) of the tight-binding model given that the local orbitals used in the tight-binding approximation are known [32].

It is a well-known fact that the Wannier functions are not uniquely defined because of the gauge freedom of the Bloch functions. In fact, we can change the Bloch functions by $g_{\mathbf{k}\sigma}(\alpha) \to \exp(i\varphi_{\mathbf{k}})\, g_{\mathbf{k}\sigma}(\alpha)$, where $\varphi_{\mathbf{k}}$ is an arbitrary real function of quasimomentum $\mathbf{k}$. This transformation on the Bloch function changes in a substantial way the shape of the Wannier functions; in particular, their spatial extent [32]. Below, when necessary, we discuss the consequences of this gauge freedom in detail.

As an example we shall consider the Creutz ladder, a one-dimensional model of two cross-linked chains, depicted in Fig. 1. The hopping matrix elements for the spin-$\uparrow$ fermions are also given in Fig. 1. The band structure of this model is extremely simple consisting in two perfectly flat bands at energies $\pm 2t$ ($E_{\rm gap} = 4t$). Moreover, the Wannier functions of the Creutz ladder for both the upper and lower flat bands can be chosen to be perfectly localized on only two adjacent rungs as it is shown in Fig. 1. The lower band Wannier function for the spin-$\uparrow$ fermions correspond to the following periodic and analytic Bloch functions:

$$g_{k\uparrow}(\alpha) = \begin{cases} e^{ika/2} \sin\left(\frac{ka}{2} + \frac{\pi}{4}\right) & \text{for } \alpha = 1, \\ e^{ika/2} \cos\left(\frac{ka}{2} + \frac{\pi}{4}\right) & \text{for } \alpha = 2, \end{cases} \tag{4}$$

where a is the lattice spacing. Note that, from the time-reversal invariance of the kinetic Hamiltonian, we have $g_{\mathbf{k}\downarrow}(\alpha) = [g_{-\mathbf{k}\uparrow}(\alpha)]^*$ or, equivalently, $W_{\alpha\downarrow}(\mathbf{i}) = [W_{\alpha\uparrow}(\mathbf{i})]^*$.

The finite support of the Wannier functions of the Creutz ladder as well as the simple band structure consisting of two perfectly flat bands are convenient for analytical calculations in this demonstrative example. However, these properties are not required for the validity of any of the general results of our work.

B. Schrieffer-Wolff transformation and projected Hamiltonian

In this subsection we give the framework which we use to construct an effective low-energy theory for the class of Hubbard models introduced above. We are interested mainly in the attractive case, but this framework can be used in the case of repulsive Hubbard interactions as well.

A simple low-energy theory can be constructed by projecting the interaction term into the subspace where only the flat band states can be occupied. This is done in practice by truncating the expansion of field operators in terms of Wannier orbitals of all bands to retain only the orbitals corresponding to the chosen flat band [28], thereby restricting the Hilbert-Fock space to the flat band subspace. However, one

expects interband effects which become more relevant with increasing interaction strength U. To include these effects, we use the perturbative SW transformation that allows taking them into account by means of a low-energy SW Hamiltonian, which involves only the degrees of freedom of the lowest flat band [33].

Let us define the field operators projected in the flat band as

$$\bar{c}_{\mathbf{i}\alpha\sigma} = \sum_{\mathbf{j}} W_{\alpha\sigma}(\mathbf{i}-\mathbf{j})\,\hat{d}_{\mathbf{j}\sigma}, \tag{5}$$

where $W_{\alpha\sigma}(\mathbf{i}-\mathbf{j})$ is the Wannier function of the flat band centered at unit cell $\mathbf{j}$ and $\hat{d}_{\mathbf{j}\sigma}$ is the annihilation operator corresponding to this Wannier orbital. Hereafter, we take advantage of time-reversal symmetry and drop the spin index of the Wannier and Bloch functions, that is $W_\alpha(\mathbf{i}) \equiv W_{\alpha\uparrow}(\mathbf{i}) = [W_{\alpha\downarrow}(\mathbf{i})]^*$ and $g_{\mathbf{k}}(\alpha) \equiv g_{\mathbf{k}\uparrow}(\alpha) = [g_{-\mathbf{k}\downarrow}(\alpha)]^*$. By virtue of the orthonormality of the Wannier functions, namely $\sum_{\mathbf{i}\alpha}[W_\alpha(\mathbf{i}-\mathbf{j})]^* W_\alpha(\mathbf{i}-\mathbf{l}) = \delta_{\mathbf{j},\mathbf{l}}$, operators $\hat{d}_{\mathbf{i}\sigma}$, $\hat{d}^\dagger_{\mathbf{i}\sigma}$ satisfy the canonical anticommutation relations. It is important to note, though, that the projected field operators satisfy *modified* anticommutation relations

$$\{\bar{c}_{\mathbf{i}\alpha\uparrow}, \bar{c}^\dagger_{\mathbf{j}\beta\uparrow}\} = P_{\alpha,\beta}(\mathbf{i}-\mathbf{j}), \quad \{\bar{c}_{\mathbf{i}\alpha\downarrow}, \bar{c}^\dagger_{\mathbf{j}\beta\downarrow}\} = [P_{\alpha,\beta}(\mathbf{i}-\mathbf{j})]^*, \tag{6}$$

where we have introduced the flat band projector defined as $P_{\alpha,\beta}(\mathbf{i}-\mathbf{j}) = \sum_{\mathbf{l}} W_\alpha(\mathbf{i}-\mathbf{l})[W_\beta(\mathbf{j}-\mathbf{l})]^*$. All other anticommutators are trivial.

Despite the gauge freedom in the definition of the Wannier functions, the projected field operators given in Eq. (5) are gauge independent, i.e., they are the same for any choice of the Wannier function. Indeed, an equivalent way of defining the projected operators is via the lower band projector introduced above according to $\bar{c}_{\mathbf{i}\alpha\uparrow} = \sum_{\mathbf{j}\beta} P_{\alpha,\beta}(\mathbf{i}-\mathbf{j})\,\hat{c}_{\mathbf{j}\beta\uparrow}$ and $\bar{c}_{\mathbf{i}\alpha\downarrow} = \sum_{\mathbf{j}\beta}[P_{\alpha,\beta}(\mathbf{i}-\mathbf{j})]^*\,\hat{c}_{\mathbf{j}\beta\downarrow}$. From the last relations the gauge invariance of the projected operators becomes explicit, as the projectors are gauge independent.

Let us denote by $\hat{\mathcal{P}} = \hat{\mathcal{P}}_\uparrow\hat{\mathcal{P}}_\downarrow$ the second-quantized projection operator that projects on the subspace of the Hilbert-Fock space where only the states of the flat band can have nonzero occupancy, and $\hat{\mathcal{P}}_\sigma$ is the projector relative to spin σ. Then these operators have the property

$$\hat{\mathcal{P}}_\sigma \tilde{c}^\dagger_{\mathbf{i}\alpha\sigma} = \tilde{c}_{\mathbf{i}\alpha\sigma}\hat{\mathcal{P}}_\sigma = 0, \tag{7}$$

where $\tilde{c}_{\mathbf{i}\alpha\sigma}$ is the field operator projected into the complement of the flat band subspace, namely $\tilde{c}_{\mathbf{i}\alpha\sigma} = \hat{c}_{\mathbf{i}\alpha\sigma} - \bar{c}_{\mathbf{i}\alpha\sigma}$. The operator $\tilde{c}_{\mathbf{i}\alpha\sigma}$ can be expressed by the Wannier orbitals of the higher bands analogous to Eq. (5). The operators $\hat{\mathcal{P}}_\sigma$ commute with the flat band field operators $\bar{c}_{\mathbf{i}\alpha\sigma}$, $\bar{c}^\dagger_{\mathbf{i}\alpha\sigma}$ and with the operators of opposite spin. From the projected operator given in Eq. (5) we can define the projected number operator as $\bar{n}_{\mathbf{i}\alpha\sigma} = \bar{c}^\dagger_{\mathbf{i}\alpha\sigma}\bar{c}_{\mathbf{i}\alpha\sigma}$. It is straightforward to check that $\hat{\mathcal{P}}_\sigma\hat{n}_{\mathbf{i}\alpha\sigma}\hat{\mathcal{P}}_\sigma = \hat{\mathcal{P}}_\sigma\hat{c}^\dagger_{\mathbf{i}\alpha\sigma}\hat{c}_{\mathbf{i}\alpha\sigma}\hat{\mathcal{P}}_\sigma = \bar{c}^\dagger_{\mathbf{i}\alpha\sigma}\bar{c}_{\mathbf{i}\alpha\sigma}\hat{\mathcal{P}}_\sigma = \bar{n}_{\mathbf{i}\alpha\sigma}\hat{\mathcal{P}}_\sigma$. Then the projected interaction Hamiltonian $\hat{\mathcal{P}}\hat{\mathcal{H}}_{\rm int}\hat{\mathcal{P}} = \overline{\mathcal{H}}_{\rm int}\hat{\mathcal{P}}$ is given in terms of the operator

$$\overline{\mathcal{H}}_{\rm int} = -U\sum_{\mathbf{m}\alpha}\bar{n}_{\mathbf{m}\alpha\uparrow}\bar{n}_{\mathbf{m}\alpha\downarrow} = -U\sum_{\mathbf{i},\mathbf{j},\mathbf{k},\mathbf{l}} J_{\mathbf{ijkl}}\,\hat{d}^\dagger_{\mathbf{i}\uparrow}\hat{d}^\dagger_{\mathbf{k}\downarrow}\hat{d}_{\mathbf{l}\downarrow}\hat{d}_{\mathbf{j}\uparrow} \tag{8}$$

with

$$J_{\mathbf{ijkl}} = \sum_{\mathbf{m}\alpha}[W_\alpha(\mathbf{m}-\mathbf{i})]^* W_\alpha(\mathbf{m}-\mathbf{j})W_\alpha(\mathbf{m}-\mathbf{k})[W_\alpha(\mathbf{m}-\mathbf{l})]^*. \tag{9}$$

In the following we call $\overline{\mathcal{H}}_{\rm int}\hat{\mathcal{P}}$ the *projected Hamiltonian* to distinguish it from the ferromagnetic Heisenberg Hamiltonian which is defined below in Eq. (20). We call this latter Hamiltonian the *effective spin Hamiltonian*. Moreover, we drop the projector $\hat{\mathcal{P}}$ from the projected Hamiltonian as it is understood that $\overline{\mathcal{H}}_{\rm int}$ acts only on the flat band subspace.

Using the notation of Ref. [33], we define superoperators $\mathcal{D}(\cdot)$ and $\mathcal{O}(\cdot)$ acting on a generic operator $\hat{X}$ as

$$\mathcal{D}(\hat{X}) = \hat{\mathcal{P}}\hat{X}\hat{\mathcal{P}} + \hat{\mathcal{Q}}\hat{X}\hat{\mathcal{Q}}, \quad \mathcal{O}(\hat{X}) = \hat{\mathcal{P}}\hat{X}\hat{\mathcal{Q}} + \hat{\mathcal{Q}}\hat{X}\hat{\mathcal{P}}, \tag{10}$$

where $\hat{\mathcal{Q}} = \mathbf{1} - \hat{\mathcal{P}}$ is the complementary projector. The superoperator $\mathcal{D}(\cdot)$ extracts the diagonal part of the operator in the argument, while $\mathcal{O}(\cdot)$ extracts the off-diagonal one. Therefore, $\hat{X} = \mathcal{D}(\hat{X}) + \mathcal{O}(\hat{X})$. Let us define another superoperator $\mathcal{L}(\cdot)$

$$\mathcal{L}(\hat{X}) = \sum_{i,j}\frac{|i\rangle\langle i|\mathcal{O}(\hat{X})|j\rangle\langle j|}{E_i - E_j}, \tag{11}$$

where the labels i,j run over the eigenstates E_i, E_j of the noninteracting Hamiltonian $\hat{\mathcal{H}}_{\rm kin}$. Since the matrix element $\langle i|\mathcal{O}(\hat{X})|j\rangle$ is nonzero only if $|i\rangle$ belongs to the flat band subspace and $|j\rangle$ to the complementary subspace, or vice versa, the above sum is well defined.

With the help of the above introduced superoperators, the SW Hamiltonian up to second order in an expansion in the dimensionless coupling constant $U/E_{\rm gap}$ reads [33]

$$\hat{\mathcal{H}}_{\rm eff} \approx \hat{\mathcal{H}}_{\rm kin}\hat{\mathcal{P}} + \hat{\mathcal{P}}\hat{\mathcal{H}}_{\rm int}\hat{\mathcal{P}} + \tfrac{1}{2}\hat{\mathcal{P}}[\mathcal{L}(\hat{\mathcal{H}}_{\rm int}), \mathcal{O}(\hat{\mathcal{H}}_{\rm int})]\hat{\mathcal{P}}. \tag{12}$$

The zero-order term is the projected kinetic Hamiltonian which is a trivial constant for a flat band, while the first-order term is simply the projected interaction Hamiltonian $\overline{\mathcal{H}}_{\rm int}\hat{\mathcal{P}}$. The last term of order $(U/E_{\rm gap})^2$ allows us to incorporate effects due to interband coupling neglected in the projected Hamiltonian. In Sec. VI we evaluate explicitly the second-order term in the case of the Creutz-Hubbard model and use this latter result in Sec. VII to evaluate the compressibility.

III. EXACTNESS OF THE BCS WAVE FUNCTION AND EMERGENT SU(2) SYMMETRY

In this section we prove that the ground state of the projected attractive interaction $\overline{\mathcal{H}}_{\rm int}$ for arbitrary filling of the flat band is given exactly by the BCS wave function, provided that the uniform pairing condition introduced in Eq. (17) below is fulfilled. This statement is analogous to the well-known fact that the completely polarized ferromagnetic state is the ground state for a half-filled flat band in a repulsive Hubbard model if the flat band is the lowest lying one [34–38]. An important difference is that in the former case we known the ground state only of the *projected* Hamiltonian $\overline{\mathcal{H}}_{\rm int}$, while the latter is a statement regarding the *full* Hamiltonian $\hat{\mathcal{H}}$. Moreover, it seems that the condition in Eq. (17) is necessary in the

attractive case, but no analogous condition is required in the repulsive case.

Only the existence of a ground state given by the BCS wave function is established, but nothing is said about its uniqueness (global stability). In absence of a rigorous mathematical proof, this is a reasonable assumption for flat bands with nonzero quantum metric. Indeed this implies the superfluid weight to be nonzero and guarantees the local stability of the BCS ground state.

This result generalizes the one relative to bipartite lattices. In a bipartite lattice it is possible to relate a repulsive Hubbard model to an attractive Hubbard model by a particle-hole transformation (see Appendix A). Using this result and the Lieb theorem [39], it was shown in Ref. [27] that the BCS wave function is in fact a good approximation of the true ground state for the partially filled flat band.

To begin with, we introduce the following pseudospin operators $\hat{\mathbf{T}}_\mathbf{i} = (\hat{\mathrm{T}}^x_\mathbf{i}, \hat{\mathrm{T}}^y_\mathbf{i}, \hat{\mathrm{T}}^z_\mathbf{i})$:

$$\hat{\mathrm{T}}^z_\mathbf{i} = \tfrac{1}{2}(\hat{d}^\dagger_{\mathbf{i}\uparrow}\hat{d}_{\mathbf{i}\uparrow} + \hat{d}^\dagger_{\mathbf{i}\downarrow}\hat{d}_{\mathbf{i}\downarrow} - 1), \quad \hat{\mathrm{T}}^+_\mathbf{i} = (\hat{\mathrm{T}}^-_\mathbf{i})^\dagger = \hat{d}^\dagger_{\mathbf{i}\uparrow}\hat{d}^\dagger_{\mathbf{i}\downarrow}, \quad (13)$$

which are also used later in Sec. IV. As usual, we have $\hat{\mathrm{T}}^x_\mathbf{i} = (\hat{\mathrm{T}}^+_\mathbf{i} + \hat{\mathrm{T}}^-_\mathbf{i})/2$ and $\hat{\mathrm{T}}^y_\mathbf{i} = (\hat{\mathrm{T}}^+_\mathbf{i} - \hat{\mathrm{T}}^-_\mathbf{i})/2i$. Then we introduce the operator $\hat{b}^\dagger_0$ which creates a Cooper pair in a plane wave state with zero quasimomentum in the flat band,

$$\hat{b}^\dagger_0 = \sum_\mathbf{j} \hat{\mathrm{T}}^+_\mathbf{j} = \sum_\mathbf{j} \hat{d}^\dagger_{\mathbf{j}\uparrow}\hat{d}^\dagger_{\mathbf{j}\downarrow} = \sum_\mathbf{k} \hat{f}^\dagger_{\mathbf{k}\uparrow}\hat{f}^\dagger_{-\mathbf{k}\downarrow}. \quad (14)$$

The above introduced operator $\hat{f}^\dagger_{\mathbf{k}\sigma}$ creates a fermion in a Bloch plane-wave state with quasimomentum $\mathbf{k}$ in the flat band. Note that the pair creation operator takes the above simple form only in the flat band limit. Generally, it is given by $\hat{b}^\dagger_0 = \sum_\mathbf{k} g_\mathbf{k}\, \hat{f}^\dagger_{\mathbf{k}\uparrow}\hat{f}^\dagger_{-\mathbf{k}\downarrow}$, where $g_\mathbf{k} = v_\mathbf{k}/u_\mathbf{k}$ defines the Cooper pair wave function [40,41]. In the flat band limit the BCS coherence factors $u_\mathbf{k} = u = \sqrt{1-\nu}$ and $v_\mathbf{k} = v = \sqrt{\nu}$ are independent of the quasimomentum $\mathbf{k}$ and only in this case is Eq. (14) valid. Here $\nu = N/(2N_\mathrm{c})$ is the filling of the flat band with N_c the number of unit cells and N the total number of fermions. Hence, the BCS wave function in the grand canonical ensemble can be written in the following equivalent forms:

$$\begin{aligned}|\Omega\rangle &= u^{N_\mathrm{c}} \exp\left(\frac{v}{u}\,\hat{b}^\dagger_0\right)|\emptyset\rangle \\ &= \prod_\mathbf{j}(u + v\hat{d}^\dagger_{\mathbf{j}\uparrow}\hat{d}^\dagger_{\mathbf{j}\downarrow})|\emptyset\rangle = \prod_\mathbf{k}(u + v\hat{f}^\dagger_{\mathbf{k}\uparrow}\hat{f}^\dagger_{-\mathbf{k}\downarrow})|\emptyset\rangle, \quad (15)\end{aligned}$$

where $|\emptyset\rangle$ is the vacuum containing no fermions. In expanding the exponential in Eq. (15) we have made use of Fermi statistics, which implies $(\hat{d}^\dagger_{\mathbf{j}\sigma})^2 = (\hat{f}^\dagger_{\mathbf{k}\sigma})^2 = 0$.

As shown in Appendix B, the operator $\hat{b}^\dagger_0$ commutes with the *projected* spin operator $\overline{S}^z_{\mathbf{i}\alpha} = (\bar{n}_{\mathbf{i}\alpha\uparrow} - \bar{n}_{\mathbf{i}\alpha\downarrow})/2$, that is, $[\hat{b}^\dagger_0, \overline{S}^z_{\mathbf{i}\alpha}] = 0$. From this commutation relation one has that $\overline{S}^z_{\mathbf{i}\alpha}|\Omega\rangle = 0$. Therefore, the BCS wave function $|\Omega\rangle$ is a zero eigenvector of the positive semidefinite operator $\overline{\mathcal{H}}'_\mathrm{int} = (U/2)\sum_{\mathbf{i}\alpha}(\bar{n}_{\mathbf{i}\alpha\uparrow} - \bar{n}_{\mathbf{i}\alpha\downarrow})^2$. Note that, as a consequence of the modified anticommutation relations (6), $\bar{n}^2_{\mathbf{i}\alpha\sigma} = P_{\alpha\alpha}(\mathbf{0})\bar{n}_{\mathbf{i}\alpha\sigma}$, while the usual number operators satisfy $\hat{n}^2_{\mathbf{i}\alpha\sigma} = \hat{n}_{\mathbf{i}\alpha\sigma}$. Using this, one can expand the square

$$\overline{\mathcal{H}}'_\mathrm{int} = \frac{U}{2}\sum_{\mathbf{i}\alpha} P_{\alpha\alpha}(\mathbf{0})(\bar{n}_{\mathbf{i}\alpha\uparrow} + \bar{n}_{\mathbf{i}\alpha\downarrow}) - U\sum_{\mathbf{i}\alpha}\bar{n}_{\mathbf{i}\alpha\uparrow}\bar{n}_{\mathbf{i}\alpha\downarrow}. \quad (16)$$

The last term in the above equation is precisely the projected attractive Hubbard interaction $\overline{\mathcal{H}}_\mathrm{int}$, while the first term is in general a nontrivial orbital-resolved potential. Consider now the case in which the flat band Bloch/Wannier functions have the same "weight" on the orbitals where they are nonzero, which means

$$\begin{aligned}n_\phi &= \sum_\mathbf{i} |W_\alpha(\mathbf{i})|^2 = \frac{V_\mathrm{c}}{(2\pi)^d}\int_{\mathrm{B.Z.}} d^d\mathbf{k}\, |g_\mathbf{k}(\alpha)|^2 \\ &= P_{\alpha\alpha}(\mathbf{0}) = P_{\beta\beta}(\mathbf{0}) \quad \forall\alpha,\beta \in \mathcal{S}, \quad (17)\end{aligned}$$

for a certain subset $\mathcal{S}$ of orbitals and $P_{\gamma\gamma}(\mathbf{0}) = 0$ for $\gamma \notin \mathcal{S}$. Here $n_\phi^{-1} = |\mathcal{S}|$ is the number of orbitals on which the flat band states have nonvanishing weight and, equivalently, on which the Cooper pair wave function is uniformly delocalized. For this reason we call Eq. (17) the *uniform pairing condition*. As a consequence of the uniform pairing condition (17), the expectation values of the projected density $\langle \bar{n}_{\mathbf{i}\alpha\sigma}\rangle$ and of the pairing order parameters $\Delta_\alpha = -U\langle \hat{c}_{\mathbf{i}\alpha\downarrow}\hat{c}_{\mathbf{i}\alpha\uparrow}\rangle$ taken on the BCS wave function (15) are constant as a function of $\alpha \in \mathcal{S}$ and vanish for $\alpha \notin \mathcal{S}$. We note that the condition of constant Δ_α has been used in Ref. [26] to derive the relation between superfluid weight and quantum metric.

Using Eq. (17) the operator $\overline{\mathcal{H}}'_\mathrm{int}$ reduces to

$$\overline{\mathcal{H}}'_\mathrm{int} = \frac{n_\phi U}{2}\overline{N} + \overline{\mathcal{H}}_\mathrm{int}, \quad (18)$$

with $\overline{N} = \sum_{\mathbf{j}\sigma}\hat{d}^\dagger_{\mathbf{j}\sigma}\hat{d}_{\mathbf{j}\sigma}$ the projected particle number operator. In this case $\overline{\mathcal{H}}'_\mathrm{int}$ differs from the projected attractive Hubbard interaction $\overline{\mathcal{H}}_\mathrm{int}$ by a trivial term proportional to the particle number operator. Following the usual argument, [37] we conclude that the BCS wave function is one ground state of $\overline{\mathcal{H}}_\mathrm{int}$ if the uniform pairing condition (17) is satisfied. Moreover, the ground-state energy is

$$\frac{E_\mathrm{BCS}}{N_\mathrm{c}} = (2\varepsilon_0 - n_\phi U)\nu, \quad (19)$$

where ε_0 is the flat band energy. It is important to note that this result is only valid for small $U \ll E_\mathrm{gap}$, contrary to the repulsive case. The reason is that the pair creation operator $\hat{b}^\dagger_0$ commutes only with the projected spin operator $\overline{S}^z_{\mathbf{i}\alpha}$ and it does not commute with the full spin operator $\hat{S}^z_{\mathbf{i}\alpha} = (\hat{n}_{\mathbf{i}\alpha\uparrow} - \hat{n}_{\mathbf{i}\alpha\downarrow})/2$. In fact, one can show that $[\hat{S}^z_{\mathbf{i}\alpha}, \hat{b}^\dagger_0] = \frac{1}{2}(\tilde{c}^\dagger_{\mathbf{i}\alpha\uparrow}\bar{c}^\dagger_{\mathbf{i}\alpha\downarrow} - \bar{c}^\dagger_{\mathbf{i}\alpha\uparrow}\tilde{c}^\dagger_{\mathbf{i}\alpha\downarrow})$.

From the Cooper pair creation operator $\hat{b}^\dagger_0$, we can construct the generators of an emergent SU(2) symmetry of the projected interaction Hamiltonian which also includes the usual U(1) particle number conservation symmetry. The operators $(1/2)(\hat{b}^\dagger_0 + \hat{b}_0)$, $(-i/2)(\hat{b}^\dagger_0 - \hat{b}_0)$ and $(1/2)[\hat{b}^\dagger_0, \hat{b}_0]$ form the generators of SU(2). Using the commutator $[b^\dagger_0, [\hat{b}^\dagger_0, \hat{b}_0]] = -2b^\dagger_0$, it is straightforward to check that the three generators given above satisfy the commutation relations of the $\mathfrak{su}(2)$ algebra. Using $b_0 = \sum_\mathbf{i}\hat{\mathrm{T}}^+_\mathbf{i}$, we can also write the generators

of SU(2) in terms of pseudospin operators as $\sum_{\mathbf{i}} \hat{\mathrm{T}}^x_{\mathbf{i}}$, $\sum_{\mathbf{i}} \hat{\mathrm{T}}^y_{\mathbf{i}}$, and $\sum_{\mathbf{i}} \hat{\mathrm{T}}^z_{\mathbf{i}}$.

It follows from Eq. (19) and the definition of the inverse compressibility $\kappa^{-1} = \nu^2 \partial^2 (E_{\mathrm{BCS}}/N_{\mathrm{c}})/\partial \nu^2$ that the compressibility diverges at the level of the projected interaction Hamiltonian. Moreover, since $|\Omega\rangle$ is a zero eigenstate of $\overline{\mathcal{H}}'_{\mathrm{int}}$, one obtains the relation $n_\phi \langle \overline{n}_{\mathbf{i}\alpha\uparrow} \rangle = n_\phi \langle \overline{n}_{\mathbf{i}\alpha\downarrow} \rangle = \langle \overline{n}_{\mathbf{i}\alpha\uparrow} \overline{n}_{\mathbf{i}\alpha\downarrow} \rangle$ between density and double occupancy (expectation values are here taken on $|\Omega\rangle$). Both the result for the energy and the relation between density and double occupancy have been verified with DMRG for the Creutz-Hubbard model ($n_\phi = 1/2$).

The uniform pairing condition (17) is essential for most of our results and it is worth commenting on its significance. The case $n_\phi = 1$ implies that the flat band states are localized on the same orbital for all values of $\mathbf{k}$. This means that the flat band is composed of Wannier functions that are not overlapping and are thus completely uncoupled. A rigorous way to measure the Wannier function overlap is given in Sec. V with the overlap functional. Such a flat band is a trivial insulator and uninteresting for our purpose. Indeed under the same condition it is possible to show that the quantum metric is vanishing everywhere in the Brillouin zone resulting in a zero superfluid weight.

The first nontrivial case is $n_\phi = 1/2$ realized in the Creutz ladder presented here and in the Lieb lattice [27], while $n_\phi = 1/3$ in the kagome lattice [42], for example. The parameter n_ϕ enters in a number of important quantities such as the mean-field critical temperature $T_{\mathrm{c}} = n_\phi U/4$ and the general result (26) for the superfluid weight of a flat band.

We believe the uniform pairing condition not to be overly restrictive for two reasons: First, in many cases the lattice symmetry alone enforces Eq. (17) and, second, the intermediate case, where the diagonal elements $P_{\alpha\alpha}(\mathbf{0})$ of the projection operator are nonzero arbitrary numbers, interpolates continuously between between uniform pairing and the trivial case $n_\phi = 1$. The message is that flat band superfluidity requires the Cooper pair wave function to be spread over more than one orbital per unit cell, and only when all orbitals are equally involved is the effect the most pronounced.

IV. EFFECTIVE SPIN HAMILTONIAN

As a next step, we show that it is possible to drop a large amount of terms in the projected interaction Hamiltonian $\overline{\mathcal{H}}_{\mathrm{int}}$ at the same time preserving all of the properties mentioned above, namely the BCS ground state and the emergent SU(2) symmetry. The result of this truncation is an effective ferromagnetic Heisenberg model whose degrees of freedom are encoded in the pseudospin operators $\hat{\mathbf{T}}_{\mathbf{i}}$ introduced in Eqs. (13). The effective spin model is computationally much easier to deal with and offers an intuitive model of a flat band superconductor as a ferromagnet.

Note that in the expansion of the projected Hamiltonian given in Eq. (8) the terms for which $\mathbf{i} = \mathbf{j} = \mathbf{k} = \mathbf{l}$ take the usual Hubbard form $\propto -U \sum_{\mathbf{i}} \hat{d}^\dagger_{\mathbf{i}\uparrow} \hat{d}_{\mathbf{i}\uparrow} \hat{d}^\dagger_{\mathbf{i}\downarrow} \hat{d}_{\mathbf{i}\downarrow}$. As a consequence the energy cost for breaking a pair of particles localized on the same Wannier state is order of U. This is a large energy scale compared to the gapless long-wavelength excitations of the projected Hamiltonian, which correspond to the motion of Cooper pairs. Hence, we truncate the projected Hamiltonian to the subspace defined by the conditions $(\hat{d}^\dagger_{\mathbf{i}\uparrow} \hat{d}_{\mathbf{i}\uparrow} - \hat{d}^\dagger_{\mathbf{i}\downarrow} \hat{d}_{\mathbf{i}\downarrow}) |\psi\rangle = 0$ for every $\mathbf{i}$. In other words, only the terms in the expansion (8) that preserve and act nontrivially on this subspace are retained. In particular, terms that break $(\hat{d}^\dagger_{\mathbf{i}\uparrow} \hat{d}^\dagger_{\mathbf{j}\downarrow} \hat{d}_{\mathbf{l}\downarrow} \hat{d}_{\mathbf{l}\uparrow}$, with $\mathbf{i} \neq \mathbf{j}$) and create $(\hat{d}^\dagger_{\mathbf{l}\uparrow} \hat{d}^\dagger_{\mathbf{l}\downarrow} \hat{d}_{\mathbf{i}\downarrow} \hat{d}_{\mathbf{j}\uparrow}$ with $\mathbf{i} \neq \mathbf{j}$) a Cooper pair in a generic Wannier state $\mathbf{l}$ are dropped. After this truncation, the only remaining degrees of freedom are the presence or absence of a Cooper pair in a given Wannier state.

Using the pseudospin operators defined in Eqs. (13), we can map the projected Hamiltonian (8) after the truncation into an effective Hamiltonian which is an isotropic ferromagnetic Heisenberg model given by

$$\hat{\mathcal{H}}_{\mathrm{spin}} = -U \sum_{\mathbf{i} \neq \mathbf{j}} J(|\mathbf{i} - \mathbf{j}|)\, \hat{\mathbf{T}}_{\mathbf{i}} \cdot \hat{\mathbf{T}}_{\mathbf{j}} \tag{20}$$

with couplings $J(|\mathbf{i} - \mathbf{j}|)$ defined as

$$J(|\mathbf{i} - \mathbf{j}|) = \sum_{\mathbf{m}\alpha} |W_\alpha(\mathbf{m} - \mathbf{i})|^2\, |W_\alpha(\mathbf{m} - \mathbf{j})|^2. \tag{21}$$

The physical content of the effective spin Hamiltonian (20) is that the effective pseudospins located at unit cells $\mathbf{i}$ and $\mathbf{j}$ interact through the *density overlap* of Wannier functions centered at unit cells $\mathbf{i}$ and $\mathbf{j}$, substantiating the intuition that the density overlap of Wannier functions is responsible for the superconducting order. The ground state of the effective spin Hamiltonian is a product state where all pseudospins are aligned to the same direction. In the language of the projected attractive Hubbard model this corresponds to the BCS wave function (15) where the order parameter is given by the expectation value $\Delta_{\mathbf{i}} = -U \langle \hat{\mathrm{T}}^-_{\mathbf{i}} \rangle$. Moreover, the effective spin Hamiltonian is manifestly SU(2) invariant.

The main drawback of the effective spin Hamiltonian (20) is that it is *not gauge invariant*. The above truncation depends on the gauge choice for the Wannier states, namely on the definition of the operators $\hat{d}_{\mathbf{i}\sigma}$. Whereas the BCS ground state and the SU(2) symmetry is correctly reproduced by the effective spin Hamiltonian in any gauge, the choice of gauge affects in a substantial way the low-energy spectrum of the effective spin Hamiltonian. We show in the next section that there is a preferred gauge choice, the maximally localized Wannier functions in the Marzari-Vanderbilt sense [43]. In this basis the effective spin Hamiltonian is the best possible approximation of the projected interaction Hamiltonian $\overline{\mathcal{H}}_{\mathrm{int}}$ in the sense that the gauge invariant result for the superfluid density obtained in Ref. [26] is recovered.

The subtle point of the gauge noninvariance of the effective spin Hamiltonian is very well illustrated in the Creutz-Hubbard model. As we mentioned before, the Wannier functions of the Creutz model can be chosen to be perfectly localized on a plaquette; see Fig. 1. These plaquette states are in fact the maximally localized Wannier functions in the Marzari-Vanderbilt sense. If $\overline{\mathcal{H}}_{\mathrm{int}}$ is expanded in annihilation and creation operators of these plaquette states, pair-breaking and pair-creation terms subject to truncation are absent, and the mapping from the projected interaction Hamiltonian to $\hat{\mathcal{H}}_{\mathrm{spin}}$

is exact. However, this is not the case for arbitrary choices of the Wannier functions. Utilizing the maximally localized lower band Wannier functions of the Creutz model, we obtain the projected Hamiltonian given by

$$\begin{aligned}\overline{\mathcal{H}}_{\rm int} = &-\frac{U}{4}\sum_{\mathbf{i}}\hat{\rho}_{\mathbf{i}\uparrow}\hat{\rho}_{\mathbf{i}\downarrow} - \frac{U}{8}\sum_{\mathbf{i}}(\hat{\rho}_{\mathbf{i}-1\uparrow}\hat{\rho}_{\mathbf{i}\downarrow} + \hat{\rho}_{\mathbf{i}-1\downarrow}\hat{\rho}_{\mathbf{i}\uparrow})\\ &- \frac{U}{8}\sum_{\mathbf{i}}(\hat{d}^{\dagger}_{\mathbf{i}-1\uparrow}\hat{d}^{\dagger}_{\mathbf{i}-1\downarrow}\hat{d}_{\mathbf{i}\downarrow}\hat{d}_{\mathbf{i}\uparrow} + \text{H.c.})\\ &- \frac{U}{8}\sum_{\mathbf{i}}(\hat{d}^{\dagger}_{\mathbf{i}\uparrow}\hat{d}_{\mathbf{i}\downarrow}\hat{d}^{\dagger}_{\mathbf{i}-1\downarrow}\hat{d}_{\mathbf{i}-1\uparrow} + \text{H.c.}),\end{aligned} \tag{22}$$

where $\hat{\rho}_{\mathbf{i}\sigma} = \hat{d}^{\dagger}_{\mathbf{i}\sigma}\hat{d}_{\mathbf{i}\sigma}$. Indeed, we see that the above Hamiltonian does not mix the subspaces with different number of pairs, and hence for the Creutz-Hubbard model the above described truncation is not an approximation. Using Eqs. (13), we can map the projected Hamiltonian (22) into the ferromagnetic XXX chain with Hamiltonian

$$\hat{\mathcal{H}}_{\rm spin} = -\frac{U}{4}\sum_{\mathbf{i}}\hat{\mathbf{T}}_{\mathbf{i}}\cdot\hat{\mathbf{T}}_{\mathbf{i}+1}. \tag{23}$$

It is worth mentioning that the effective spin Hamiltonian (20) describes also the completely polarized ferromagnetic state of the corresponding repulsive Hubbard model with half-filled lowest flat band. As discussed in Sec. II A in the corresponding repulsive Hubbard model the kinetic Hamiltonian is spin isotropic. In this case it is energetically more favorable that all the Wannier states are only singly occupied, i.e., we consider the subspace defined by $(\hat{d}^{\dagger}_{\mathbf{i}\uparrow}\hat{d}_{\mathbf{i}\uparrow} + \hat{d}^{\dagger}_{\mathbf{i}\downarrow}\hat{d}_{\mathbf{i}\downarrow})|\psi\rangle = |\psi\rangle$ for all $\mathbf{i}$. As a result one obtains the same effective Hamiltonian (20) with the only difference that the pseudospin operators $\hat{\mathrm{T}}^{\pm,z}_{\mathbf{i}}$ are replaced by the true spin operators $\hat{\mathrm{S}}^{z}_{\mathbf{i}} = (\hat{d}^{\dagger}_{\mathbf{i}\uparrow}\hat{d}_{\mathbf{i}\uparrow} - \hat{d}^{\dagger}_{\mathbf{i}\downarrow}\hat{d}_{\mathbf{i}\downarrow})/2$, $\hat{\mathrm{S}}^{+}_{\mathbf{i}} = \hat{d}^{\dagger}_{\mathbf{i}\uparrow}\hat{d}_{\mathbf{i}\downarrow}$ and $\hat{\mathrm{S}}^{-}_{\mathbf{i}} = \hat{d}^{\dagger}_{\mathbf{i}\downarrow}\hat{d}_{\mathbf{i}\uparrow}$, since in this case the only remaining degree of freedom is the spin. In fact, the attractive and repulsive flat band models are exactly related by a particle-hole transformation up to first order in $U/E_{\rm gap}$ (see Appendix A), while this mapping is broken by interband transitions. This is shown in Sec. VI where the second-order SW Hamiltonian is calculated explicitly for the Creutz-Hubbard model. Therefore, the effective spin Hamiltonian (20) expresses in a concise way the duality between the ferromagnetic and BCS ground states in a flat band.

V. SUPERFLUID WEIGHT FROM THE EFFECTIVE SPIN HAMILTONIAN

In this section we discuss how the superfluid weight in a flat band can be recovered from the effective spin Hamiltonian (20) obtained in the previous section.

Recall that the superfluid weight tensor $D_{\rm s}$ is a thermodynamic quantity measuring the energy change when the complex order parameter has a plane-wave spatial modulation of the form $\Delta(\mathbf{r}) = |\Delta|e^{2i\mathbf{q}\cdot\mathbf{r}}$, imposed for example by twisted boundary conditions in a ring geometry. It can be calculated by taking the second-order derivative of the grand potential $\Omega(T,\mu,\mathbf{q})$ with respect to the modulation wave vector $\mathbf{q}$,

$$[D_{\rm s}]_{i,j} = \frac{1}{V\hbar^2}\frac{\partial^2\Omega}{\partial q_i\partial q_j}\bigg|_{\mathbf{q}=0}. \tag{24}$$

We denote by $[D_{\rm s}]_{i,j}$ the components of the tensor $D_{\rm s}$. The superfluid weight measures the ability of the system to sustain an equilibrium nondissipative current and for superfluids in the continuum is given by $D_{\rm s} = n_{\rm s}/m$ with $n_{\rm s}$ the superfluid number density. Since the mass m of particles is not well defined in lattice systems, especially in flat bands, it is preferable to use the superfluid weight instead of the superfluid density [26]. A variety of methods are used to measure the superfluid weight, for example by means of the relation $D_{\rm s} \propto \lambda_{\rm L}^{-2}$ with the London penetration length $\lambda_{\rm L}$ which characterizes the Meissner effect. Being a measure of superfluidity and of the strength of the Meissner effect, the superfluid weight is the defining property of superfluids and superconductors.

The gauge noninvariance of the effective spin Hamiltonian (20) manifests in the fact that the superfluid weight evaluated from it depends on the choice of the Wannier functions. On the other hand, in Ref. [26] a gauge invariant result for the superfluid weight of a flat band in the presence of an attractive Hubbard interaction has been derived using mean-field BCS theory. The solution of this inconsistency is that, within the effective spin Hamiltonian approximation, there exists a preferred gauge choice for the Wannier functions which allows us to obtain a result for the superfluid weight as close as possible to the gauge invariant result of Ref. [26]. Specifically, in the following we prove the inequality

$$\mathrm{Tr}D_{\rm s}^{(\rm spin)} \geqslant \mathrm{Tr}D_{\rm s} \tag{25}$$

between the traces of the superfluid weight tensor $D_{\rm s}^{(\rm spin)}$ obtained from the effective spin Hamiltonian (20) and the gauge invariant result of Ref. [26],

$$[D_{\rm s}]_{i,j} = \frac{4n_\phi U\nu(1-\nu)}{(2\pi)^d\hbar^2}\int_{\rm B.Z.} d^d\mathbf{k}\,\mathrm{Re}\,\mathcal{B}_{ij}(\mathbf{k}), \tag{26}$$

valid in arbitrary dimension d. Here the quantum geometric tensor (QGT) $\mathcal{B}_{ij}(\mathbf{k})$ is defined by

$$\mathcal{B}_{ij}(\mathbf{k}) = 2\big\langle\partial_{k_i}g_{\mathbf{k}}\big|(1-|g_{\mathbf{k}}\rangle\langle g_{\mathbf{k}}|)\big|\partial_{k_j}g_{\mathbf{k}}\big\rangle, \tag{27}$$

where $g_{\mathbf{k}}(\alpha) = \langle\alpha|g_{\mathbf{k}}\rangle$ are the flat band Bloch functions. Moreover, we show that the gauge noninvariant quantity $\mathrm{Tr}D_{\rm s}^{(\rm spin)}$ attains a global minimum if the maximally localized Wannier functions in the Marzari-Vanderbilt sense are used. In this preferred gauge the superfluid weight tensors calculated using the two different methods coincide in one dimension, while for $d \geqslant 2$ the effective spin Hamiltonian generally overestimates the superfluid weight as shown by Eq. (25).

Our argument is based on the fact that the superfluid weight calculated from the effective spin Hamiltonian is proportional to a functional, called $F_{\rm ov}$, that measures the degree of overlap between Wannier functions. This functional is similar to the Marzari-Vanderbilt functional, called $F_{\rm MV}$, which measures the spread of the Wannier functions [43]. The above results follow from some general relations between these two functionals whose detailed proof is given in the Appendix C. Our proof relies on some assumptions: the first

is the uniform pairing condition (17) which is also assumed in Ref. [26] to derive Eq. (26). Since we have proved that the BCS wave function is an exact ground state of $\overline{\mathcal{H}}_{\rm int}$, the use of mean-field BCS theory and thus the result of Eq. (26) are justified. The second assumption is the existence of a gauge in which the Bloch functions are periodic and analytic functions of quasimomentum, which is equivalent to requiring that the flat band has zero Chern number(s) [44–48]. The final assumption is Eq. (C8) which is discussed in Appendix C. All conditions are fulfilled in the case of the Creutz ladder.

The overlap functional is introduced by considering the state with finite uniform superfluid current given approximately, for small phase-winding wave vector $\mathbf{q}$, by the ansatz

$$|\Omega,\mathbf{q}\rangle = \prod_{\mathbf{j}}(u + e^{2i\mathbf{q}\cdot\mathbf{r}_\mathbf{j}} v \hat{d}^\dagger_{\mathbf{j}\uparrow}\hat{d}^\dagger_{\mathbf{j}\downarrow})|\emptyset\rangle. \tag{28}$$

It is easy to verify that $\langle\Omega,\mathbf{q}|\hat{\mathrm{T}}^-_\mathbf{j}|\Omega,\mathbf{q}\rangle = uve^{2i\mathbf{q}\cdot\mathbf{r}_\mathbf{j}} = \sqrt{\nu(1-\nu)}e^{2i\mathbf{q}\cdot\mathbf{r}_\mathbf{j}}$ and $\langle\Omega,\mathbf{q}|\hat{\mathrm{T}}^z_\mathbf{j}|\Omega,\mathbf{q}\rangle = (v^2-u^2)/2 = \nu$. Therefore, the energy change due to the finite phase-winding with wave vector $\mathbf{q}$ is given by

$$\begin{aligned}\Delta E(\mathbf{q}) &= \langle\Omega,\mathbf{q}|\hat{\mathcal{H}}_{\rm spin}|\Omega,\mathbf{q}\rangle - \langle\Omega,\mathbf{0}|\hat{\mathcal{H}}_{\rm spin}|\Omega,\mathbf{0}\rangle \\ &= N_c U\nu(1-\nu)(F_{\rm ov}(2\mathbf{q})[W] - F_{\rm ov}(\mathbf{0})[W]) \geqslant 0.\end{aligned} \tag{29}$$

Here we have defined the *overlap functional* for Wannier functions $F_{\rm ov}(\mathbf{q})[W]$ which reads

$$\begin{aligned}F_{\rm ov}(\mathbf{q})[W] &= -\sum_{\mathbf{i},\mathbf{j},\alpha}|W_\alpha(\mathbf{i})|^2\,|W_\alpha(\mathbf{j})|^2 e^{i\mathbf{q}\cdot(\mathbf{r_i}-\mathbf{r_j})} \\ &= -\sum_{\mathbf{i}-\mathbf{j}}\sum_{\mathbf{l}\alpha}|W_\alpha(\mathbf{l}-\mathbf{i})|^2\,|W_\alpha(\mathbf{l}-\mathbf{j})|^2 e^{i\mathbf{q}\cdot(\mathbf{r_i}-\mathbf{r_j})}.\end{aligned} \tag{30}$$

We call it the overlap functional since, from the second line of the above equation, it is apparent that it can be expressed as the density overlap between the Wannier functions located at unit cells $\mathbf{i}$ and $\mathbf{j}$, given by $\sum_{\mathbf{l}\alpha}|W_\alpha(\mathbf{l}-\mathbf{i})|^2|W_\alpha(\mathbf{l}-\mathbf{j})|^2$, summed over all possible relative position vectors $\mathbf{i}-\mathbf{j}$.

We also introduce the *Marzari-Vanderbilt localization functional* which differs from the overlap functional only by an additional summation $\sum_\beta$ over the orbitals

$$F_{\rm MV}(\mathbf{q})[W] = -\sum_{\mathbf{i}\alpha}\sum_{\mathbf{j}\beta}|W_\alpha(\mathbf{i})|^2|W_\beta(\mathbf{j})|^2 e^{i\mathbf{q}\cdot(\mathbf{r_i}-\mathbf{r_j})}. \tag{31}$$

Properly speaking, the Marzari-Vanderbilt localization functional is given by $\nabla^2_\mathbf{q}F_{\rm MV}$. Indeed, one has

$$\begin{aligned}&\nabla^2_\mathbf{q}F_{\rm MV}(\mathbf{q}=0)[W] \\ &= \sum_{\mathbf{i}\alpha}\sum_{\mathbf{j}\beta}|W_\alpha(\mathbf{i})|^2|W_\beta(\mathbf{j})|^2(\mathbf{r_i}-\mathbf{r_j})^2 \\ &= 2\left[\sum_{\mathbf{i}\alpha}|W_\alpha(\mathbf{i})|^2\mathbf{r}_\mathbf{i}^2 - \left(\sum_{\mathbf{i}\alpha}|W_\alpha(\mathbf{i})|^2\mathbf{r_i}\right)^2\right].\end{aligned} \tag{32}$$

Upon replacing the summations with integrals $\sum_{\mathbf{i}\alpha}\to\int d^d\mathbf{r}$ in the last equation, the usual definition of the Marzari-Vanderbilt localization functional in the continuum case [32,43] is recovered. In this sense both $F_{\rm ov}$ and $F_{\rm MV}$ are "generating functionals" whose expansion in the wave vector $\mathbf{q}$ generates various moments of the Wannier function density distribution. However, we use the same name for both (31) and (32) since the object we are referring to should be clear from the context. The same applies to the overlap functional.

The main technical results of this section are some general relations between the two functionals which are proved in Appendix C under the conditions stated above. Let us first introduce the maximally localized Wannier functions in the Marzari-Vanderbilt sense, denoted by $\overline{W}_\alpha(\mathbf{i})$, that are the global minimizers of the Marzari-Vanderbilt functional; namely, they satisfy

$$\nabla^2_\mathbf{q}F_{\rm MV}(\mathbf{q}=0)[\overline{W}] \leqslant \nabla^2_\mathbf{q}F_{\rm MV}(\mathbf{q}=0)[W] \tag{33}$$

for all Wannier functions W obtained from $\overline{W}$ by a gauge transformation. The first result is that the maximally localized Wannier functions in the Marzari-Vanderbilt sense are in fact global minimizers of the overlap functional as well; in other words,

$$\nabla^2_\mathbf{q}F_{\rm ov}(\mathbf{q}=0)[\overline{W}] \leqslant \nabla^2_\mathbf{q}F_{\rm ov}(\mathbf{q}=0)[W]. \tag{34}$$

The second result is the equality, up to the constant factor n_ϕ defined in Eq. (17), of the second derivatives of the two functionals calculated on the maximally localized Wannier functions,

$$\partial_{q_i}\partial_{q_j}F_{\rm ov}(\mathbf{q}=0)[\overline{W}] = n_\phi\partial_{q_i}\partial_{q_j}F_{\rm MV}(\mathbf{q}=0)[\overline{W}]. \tag{35}$$

Note that in general the values of the two functionals are different if they are calculated using arbitrary Wannier functions. As consequence of Eqs. (29) and (35) one has

$$\begin{aligned}\left[D_{\rm s}^{\rm (spin)}\right]_{i,j} &= \frac{1}{V\hbar^2}\partial_{q_i}\partial_{q_j}\Delta E(\mathbf{q}) \\ &= \frac{4n_\phi U\nu(1-\nu)}{V_c\hbar^2}\partial_{q_i}\partial_{q_j}F_{\rm MV}(\mathbf{q}=0),\end{aligned} \tag{36}$$

where the first equality is the superfluid weight of the effective Hamiltonian at zero temperature and the second equality holds only if the coefficients in Eq. (21) are calculated using the maximally localized Wannier functions.

The crucial point of Eq. (36) is that the superfluid weight obtained from the effective spin Hamiltonian has been related to the Marzari-Vanderbilt functional which is known to be the sum of a gauge invariant term and a gauge noninvariant one [32,43]. Not by chance, the result of Eq. (26) is recovered from the gauge invariant term of the Marzari-Vanderbilt functional. In fact one has $\partial_{q_i}\partial_{q_j}F_{\rm MV}(\mathbf{q}=0) = \Omega^I_{ij} + \widetilde{\Omega}_{ij}$ where the gauge invariant term is expressed in terms of the QGT as

$$\Omega^I_{ij} = \frac{V_c}{(2\pi)^d}\int_{\rm B.Z.} d^d\mathbf{k}\,{\rm Re}\,\mathcal{B}_{ij}(\mathbf{k}). \tag{37}$$

Using Eqs. (36) and (37) one recovers Eq. (26) from the gauge invariant part Ω^I_{ij} only, which is a positive semidefinite matrix. Also the gauge noninvariant part $\widetilde{\Omega}_{ij}$ is positive semidefinite.

Combining this fact with Eqs. (34) and (35) one obtains the chain of inequalities

$$\mathrm{Tr}[\Omega_I] \leqslant \nabla^2_{\mathbf{q}} F_{\mathrm{MV}}(\mathbf{q}=0)[\overline{W}] \leqslant n_\phi^{-1} \nabla^2_{\mathbf{q}} F_{\mathrm{ov}}(\mathbf{q}=0)[W], \quad (38)$$

that leads to Eq. (25), where $D_{\mathrm{s}}^{(\mathrm{spin})}$ is calculated with an arbitrary choice of the Wannier functions W. This means that the spin Hamiltonian overestimates the superfluid weight of the original Hubbard model if $\widetilde{\Omega}_{ij} \neq 0$ or if Wannier functions other than the maximally localized ones are used in Eq. (21). However, in one dimension $\widetilde{\Omega}_{ij} = 0$ if maximally localized Wannier functions are used. In general this is not true in $d \geqslant 2$ due to the noncommutativity of the components of the projected position operator. Also it can be shown that $\widetilde{\Omega}_{ij}$ is not vanishing in general when the Berry curvature is nonzero [43].

A concrete example of the above general results is provided by the Creutz ladder. Indeed, it is easy to verify that the plaquette states are maximally localized Wannier functions in the Marzari-Vanderbilt sense, and in this preferred gauge the number of Cooper pairs is a conserved quantity as shown by Eq. (22). As a consequence there is no approximation in going from the projected interaction Hamiltonian $\overline{\mathcal{H}}_{\mathrm{int}}$ to the spin Hamiltonian $\hat{\mathcal{H}}_{\mathrm{spin}}$ and the result $D_{\mathrm{s}}^{(\mathrm{spin})}$ for the superfluid weight coincides with Eq. (26). On the other hand, if we had chosen to perform the truncation in any other gauge, we would have obtained a superfluid weight which is strictly larger than the correct one as it can be seen from Eq. (25) and the results of Appendix C.

We now summarize the results of this section and comment on their significance. The first result is that the approximation underlying the spin effective Hamiltonian (20) is justified at least in one dimension, since the result of Eq. (26) is reproduced *exactly*. This means that the spin Hamiltonian is not only able to capture the correct ground state, but also low-lying excited states. These are the magnons of the effective spin Hamiltonian, which correspond to Cooper pairs with finite center-of-mass momentum. Indeed the state $|\Omega,\mathbf{q}\rangle$ in Eq. (28) is none else than a condensate of Cooper pairs with finite momentum whose creation operator is $\hat{b}^\dagger_{\mathbf{q}} = \sum_j e^{2i\mathbf{q}\cdot\mathbf{r}_j} \hat{d}^\dagger_{\mathbf{j}\uparrow} \hat{d}^\dagger_{\mathbf{j}\downarrow}$. In higher dimensions this approximation is justified only if $\widetilde{\Omega}$ is significantly smaller than Ω^I in some sense. In particular, we anticipate that effects due to the Berry curvature, which are absent in one dimension, may play an important role and the low-energy effective Hamiltonian may differ substantially from Eq. (20). This is reflected in that we have purposefully avoided the case of nonzero Chern number which is the result of a nonzero average Berry curvature.

As a second important result, we believe we have provided a good understanding of the reason why the quantum metric (the real part of the QGT) enters in the result of Eq. (26) for the superfluid weight of a flat band. The physical mechanism for transport in a flat band is the correlated hopping of Cooper pairs induced by the Hubbard interaction between Wannier functions with finite density overlap. This picture is nicely captured by the effective Hamiltonian (20). We emphasize, though, that the gauge invariant part of the Marzari-Vanderbilt functional measures the spread of the Wannier functions and not their overlap. As we have proved, under some conditions the two distinct functionals measuring spread and overlap of Wannier functions, respectively, are equivalent, but this may not be true in general. Better insight into this matter can be gained by studying specific models which do not satisfy the conditions required for the validity of the relations (34) and (35) between the overlap and spread functionals.

Finally, we have provided a clean example where it is necessary to employ a specific preferred gauge for the Wannier functions in order to obtain a good approximation for the low-energy Hamiltonian. Very often the maximally localized Wannier functions in the Marzari-Vanderbilt sense are used to derive simplified model Hamiltonians from electronic structure calculations. This is a generally accepted heuristic prescription, but we have shown here that there are some rigorous underlying constrains behind it, in our case the minimization of an observable quantity such as the superfluid weight.

VI. SECOND-ORDER SCHRIEFFER-WOLFF HAMILTONIAN

Contrary to the repulsive case for which the completely polarized ferromagnetic state is the exact ground state for any value of U if the half-filled flat band is the lowest band, in the attractive case the BCS wave function is the exact ground state of the SW Hamiltonian (12) only up to the first-order term. As we shall see now the second-order term in the expansion leads to a breaking of the emergent SU(2) symmetry. It is convenient to study the breaking of the emergent symmetry in the specific case of the Creutz model which is particularly simple. However, this is a general fact, since this symmetry is not present in the generic model under consideration.

The off-diagonal part, as defined in Eq. (10), of the interaction term can be written as

$$\mathcal{O}(\hat{\mathcal{H}}_{\mathrm{int}}) = \hat{A}_\uparrow + \hat{A}_\downarrow + \hat{B} + \mathrm{H.c.}, \quad (39)$$

$$\hat{A}_\sigma = -U \sum_{\mathbf{i}\alpha} \tilde{c}^\dagger_{\mathbf{i}\alpha\sigma} \bar{c}_{\mathbf{i}\alpha\sigma} \bar{n}_{\mathbf{i}\alpha\bar{\sigma}} \hat{\mathcal{P}}, \quad (40)$$

$$\hat{B} = -U \sum_{\mathbf{i}\alpha} \tilde{c}^\dagger_{\mathbf{i}\alpha\uparrow} \bar{c}_{\mathbf{i}\alpha\uparrow} \tilde{c}^\dagger_{\mathbf{i}\alpha\downarrow} \bar{c}_{\mathbf{i}\alpha\downarrow} \hat{\mathcal{P}}, \quad (41)$$

where $\bar{\sigma} = \downarrow$ if $\sigma = \uparrow$ and vice versa. The operators $\hat{A}_\sigma$, $\hat{B}$ introduced above correspond to virtual processes where one particle with spin σ or two particles with opposite spins are excited to the upper band, respectively. Then the second-order term in the expansion of the effective Hamiltonian (12) reads

$$\hat{\mathcal{H}}^{(2)}_{\mathrm{eff}} = \frac{1}{2}\hat{\mathcal{P}}[\mathcal{L}(\hat{\mathcal{H}}_{\mathrm{int}}), \mathcal{O}(\hat{\mathcal{H}}_{\mathrm{int}})]\hat{\mathcal{P}}$$
$$= -\frac{\hat{A}^\dagger_\uparrow \hat{A}_\uparrow}{E_{\mathrm{g}}} - \frac{\hat{A}^\dagger_\downarrow \hat{A}_\downarrow}{E_{\mathrm{g}}} - \frac{\hat{B}^\dagger \hat{B}}{2E_{\mathrm{g}}}. \quad (42)$$

In this last result we have made crucial use of the fact that the upper band of the Creutz ladder is also flat. For a generic lattice Hamiltonian this is not the case and the computation of the second-order term is more involved.

In order to compute the expansion of the products $\hat{A}^\dagger_\sigma \hat{A}_\sigma$, $\hat{B}^\dagger \hat{B}$ in terms of Wannier operators the following identity is useful:

$$\hat{\mathcal{P}} \tilde{c}_{\mathbf{i}\alpha\sigma} \tilde{c}^\dagger_{\mathbf{j}\beta\sigma} \hat{\mathcal{P}} = \left[\delta_{\alpha\beta}\delta_{\mathbf{i},\mathbf{j}} - P^\sigma_{\alpha,\beta}(\mathbf{i}-\mathbf{j})\right]\hat{\mathcal{P}}. \quad (43)$$

Recall that we impose time-reversal symmetry in the attractive Hubbard model which implies $P^{\uparrow}_{\alpha\beta}(\mathbf{i}-\mathbf{j}) = [P^{\downarrow}_{\alpha\beta}(\mathbf{i}-\mathbf{j})]^* = P_{\alpha\beta}(\mathbf{i}-\mathbf{j})$. Using Eq. (43) we obtain

$$\begin{aligned}\hat{A}^{\dagger}_{\sigma}\hat{A}_{\sigma} = {} & \frac{U^2}{4}\sum_{\mathbf{i},\alpha}\bar{n}_{\mathbf{i}\alpha\uparrow}\bar{n}_{\mathbf{i}\alpha\downarrow}\hat{\mathcal{P}} - \frac{U^2}{64}\sum_{\mathbf{i}}[\,\hat{\rho}_{\mathbf{i}\uparrow}\hat{\rho}_{\mathbf{i}\downarrow} + \hat{\rho}_{\mathbf{i}-1\sigma}\hat{\rho}_{\mathbf{i}-1\bar{\sigma}}\hat{\rho}_{\mathbf{i}\bar{\sigma}} + \hat{\rho}_{\mathbf{i}-1\bar{\sigma}}\hat{\rho}_{\mathbf{i}\sigma}\hat{\rho}_{\mathbf{i}+1\bar{\sigma}} + \hat{\rho}_{\mathbf{i}-1\bar{\sigma}}\hat{\rho}_{\mathbf{i}\sigma}\hat{\rho}_{\mathbf{i}\bar{\sigma}} \\ & + \hat{d}^{\dagger}_{\mathbf{i}+1\sigma}\hat{d}^{\dagger}_{\mathbf{i}+1\bar{\sigma}}\hat{d}_{\mathbf{i}-1\bar{\sigma}}\hat{d}_{\mathbf{i}-1\sigma}(1-\hat{\rho}_{\mathbf{i}\bar{\sigma}}) + \hat{d}^{\dagger}_{\mathbf{i}+1\sigma}\hat{d}^{\dagger}_{\mathbf{i}+1\bar{\sigma}}\hat{d}_{\mathbf{i}\bar{\sigma}}\hat{d}_{\mathbf{i}\sigma}(1+\hat{\rho}_{\mathbf{i}-1\bar{\sigma}}) + \hat{d}^{\dagger}_{\mathbf{i}\sigma}\hat{d}^{\dagger}_{\mathbf{i}\bar{\sigma}}\hat{d}_{\mathbf{i}-1\bar{\sigma}}\hat{d}_{\mathbf{i}-1\sigma}(1+\hat{\rho}_{\mathbf{i}+1\bar{\sigma}}) \\ & - \hat{d}^{\dagger}_{\mathbf{i}+1\sigma}\hat{d}_{\mathbf{i}-1\sigma}\hat{d}^{\dagger}_{\mathbf{i}-1\bar{\sigma}}\hat{d}_{\mathbf{i}+1\bar{\sigma}}\hat{\rho}_{\mathbf{i}\bar{\sigma}} - \hat{d}^{\dagger}_{\mathbf{i}\sigma}\hat{d}_{\mathbf{i}-1\sigma}\hat{d}^{\dagger}_{\mathbf{i}-1\bar{\sigma}}\hat{d}_{\mathbf{i}\bar{\sigma}}\hat{\rho}_{\mathbf{i}+1\bar{\sigma}} - \hat{d}^{\dagger}_{\mathbf{i}+1\sigma}\hat{d}_{\mathbf{i}\sigma}\hat{d}^{\dagger}_{\mathbf{i}\bar{\sigma}}\hat{d}_{\mathbf{i}+1\bar{\sigma}}\hat{\rho}_{\mathbf{i}-1\bar{\sigma}} + \text{H.c.}], \end{aligned} \tag{44}$$

$$\hat{B}^{\dagger}\hat{B} = \frac{U^2}{4}\sum_{\mathbf{i},\alpha}\bar{n}_{\mathbf{i}\alpha\uparrow}\bar{n}_{\mathbf{i}\alpha\downarrow}\hat{\mathcal{P}} + \frac{U^2}{64}\sum_{\mathbf{i}}(\,\hat{\rho}_{\mathbf{i}\uparrow}\hat{\rho}_{\mathbf{i}\downarrow} + 2\hat{d}^{\dagger}_{\mathbf{i}\uparrow}\hat{d}^{\dagger}_{\mathbf{i}\downarrow}\hat{d}_{\mathbf{i}-1\downarrow}\hat{d}_{\mathbf{i}-1\uparrow} + \hat{d}^{\dagger}_{\mathbf{i}+1\uparrow}\hat{d}^{\dagger}_{\mathbf{i}+1\downarrow}\hat{d}_{\mathbf{i}-1\downarrow}\hat{d}_{\mathbf{i}-1\uparrow} + \text{H.c.}). \tag{45}$$

Second-order processes give rise to three-site terms in the effective Hamiltonian, such as three-site density interaction in the first line of Eq. (44), correlated pair hopping in the second line of Eq. (44), and correlated exchange interaction in the third line of Eq. (44). It is interesting that even at second order the number of pairs is a conserved quantity, therefore the exchange terms can dropped in the attractive case and the effective Hamiltonian takes again the form of a spin Hamiltonian with pseudospin operator given by Eq. (13). We do not know if this is the case even at higher orders and if the number of pairs is an exact conserved quantity of the full model. The breaking of the SU(2) symmetry is evident in $\hat{B}^{\dagger}\hat{B}$ which produces only two-site operators such as pair hopping. The pair hopping is not compensated by any term of the form $\hat{\rho}_{\mathbf{i}-1\uparrow}\hat{\rho}_{\mathbf{i}\downarrow}$, and hence the resulting spin chain is an easy-plane ferromagnet. The most important consequence is that the divergent compressibility becomes finite as shown in the following section. Only the $\hat{B}^{\dagger}\hat{B}$ term gives rise to a finite compressibility.

VII. COMPRESSIBILITY

In Sec. III we showed that in the attractive case the BCS state is the exact ground state of the first-order term of the SW Hamiltonian (12) with energy $E^{(1)}/L = (2\varepsilon_0 - Un_{\phi})\nu$, where L denotes the number of rungs (unit cells) in the ladder, and for the Creutz ladder $n_{\phi}^{-1} = 2$. The second-order correction to the ground-state energy can be easily obtained by calculating the expectation value $\langle\Omega|\hat{\mathcal{H}}^{(2)}_{\text{eff}}|\Omega\rangle$, where $|\Omega\rangle$ is given by Eq. (15). Below all the expectation values are calculated in this state. It is straightforward to check the following expectation values of single-site operators:

$$\langle\hat{\rho}_{\mathbf{i}\uparrow}\hat{\rho}_{\mathbf{i}\downarrow}\rangle = \langle\hat{\rho}_{\mathbf{i}\sigma}\rangle = v^2 = \nu, \tag{46}$$

$$\langle\hat{d}^{\dagger}_{\mathbf{j}\uparrow}\hat{d}^{\dagger}_{\mathbf{j}\downarrow}\rangle = \langle\hat{d}_{\mathbf{j}\downarrow}\hat{d}_{\mathbf{j}\uparrow}\rangle = uv = \sqrt{\nu(1-\nu)}, \tag{47}$$

$$\langle\hat{d}^{\dagger}_{\mathbf{j}\sigma}\hat{d}_{\mathbf{j}\bar{\sigma}}\rangle = 0. \tag{48}$$

Due to the fact that the BCS wave function $|\Omega\rangle$ in the grand canonical ensemble is a product state, the expectation values of operators acting on multiple sites can be expressed in terms of the above ones. For example,

$$\langle\hat{\rho}_{\mathbf{i}-1\uparrow}\hat{\rho}_{\mathbf{i}\downarrow}\hat{\rho}_{\mathbf{i}+1\uparrow}\rangle = \langle\hat{\rho}_{\mathbf{i}-1\uparrow}\rangle\langle\hat{\rho}_{\mathbf{i}\downarrow}\rangle\langle\hat{\rho}_{\mathbf{i}+1\uparrow}\rangle = \nu^3, \tag{49}$$

$$\begin{aligned}\langle\hat{d}^{\dagger}_{\mathbf{i}-1\uparrow}\hat{d}^{\dagger}_{\mathbf{i}-1\downarrow}\hat{d}_{\mathbf{i}\downarrow}\hat{d}_{\mathbf{i}\uparrow}\hat{\rho}_{\mathbf{i}+1\downarrow}\rangle &= \langle\hat{d}^{\dagger}_{\mathbf{i}-1\uparrow}\hat{d}^{\dagger}_{\mathbf{i}-1\downarrow}\rangle\langle\hat{d}_{\mathbf{i}\downarrow}\hat{d}_{\mathbf{i}\uparrow}\rangle\langle\hat{\rho}_{\mathbf{i}+1\uparrow}\rangle \\ &= \nu^2(1-\nu). \end{aligned} \tag{50}$$

With the above expectation values we are now in a position to evaluate the expectation values of the operators $A^{\dagger}_{\sigma}A_{\sigma}$ and $B^{\dagger}B$. Either from Eq. (44) or simply by using the properties of the projected field operators (see Appendix B) we obtain that $\langle A^{\dagger}_{\sigma}A_{\sigma}\rangle = 0$. In particular, the method used in Appendix B applies to arbitrary lattices, not just the Creutz ladder. The only nonvanishing contribution comes from Eq. (45) and it is given by $\langle B^{\dagger}B\rangle/L = U^2\nu/4 - 3U^2\nu^2/32$. Hence, from Eq. (42) the ground-state energy of the system up to second order is

$$\frac{E^{(2)}}{L} = \underbrace{\left(2\varepsilon_0 - \frac{U}{2} - \frac{U^2}{8E_{\text{gap}}}\right)}_{c_1^{\text{SW}}}\nu + \underbrace{\frac{3U^2}{64E_{\text{gap}}}}_{c_2^{\text{SW}}}\nu^2. \tag{51}$$

The coefficients of the linear term and the quadratic term in ν in Eq. (51) are denoted by c_1^{SW} and c_2^{SW}, respectively. The term quadratic in the filling ensures a finite compressibility, since $\kappa^{-1} \propto \nu^2\partial^2 E/\partial\nu^2 \propto c_2^{\text{SW}}\nu^2$. Note that the above expression of the ground-state energy is valid in the thermodynamic limit, where $L \to \infty$.

We use DMRG [49–52] simulations to find the ground-state energy of the original Hubbard model on ladders of length up to $L = 26$ with periodic boundary conditions. We typically keep up to $M = 2000$ states and check the convergence with respect to M. For a fixed length L we measure the ground-state energy $E_L(\nu)$ of the system for different fillings ν and by fitting a quadratic function $E_L(\nu)/L = c_1(L)\nu + c_2(L)\nu^2$ we extract the coefficients $c_1(L)$ and $c_2(L)$. Now by applying this procedure for increasing values of L, we can see that $c_1(L)$ and $c_2(L)$ have a power-law dependence on $1/L$, which allows us to extrapolate this dependence and find the infinite length values c_1^{∞} and c_2^{∞}. Then these extrapolated values we compare to the coefficients c_1^{SW} and c_2^{SW} of Eq. (51) calculated analytically.

In Fig. 2 we plot in log-log scale the difference of $c_{1(2)}(L)$ and the extrapolated infinite system size value $c^{\infty}_{1(2)}$ versus $1/L$. These plots clearly show that we have a power-law dependence. Moreover, in the legends we give a direct comparison of the extrapolated infinite system size values of the coefficients $c^{\infty}_{1(2)}$ with the values $c^{\text{SW}}_{1(2)}$ obtained from the SW Hamiltonian. From these numbers one can see that the effective model and the DMRG simulations with a high accuracy give the same energy dependence. With the increase of the interactions strength U, the small difference between the values of c_2^{∞} and c_2^{SW} grows approximately from 2.5% (for $-U = -0.3$) to 4.5% (for $-U = -1.0$). This difference is a

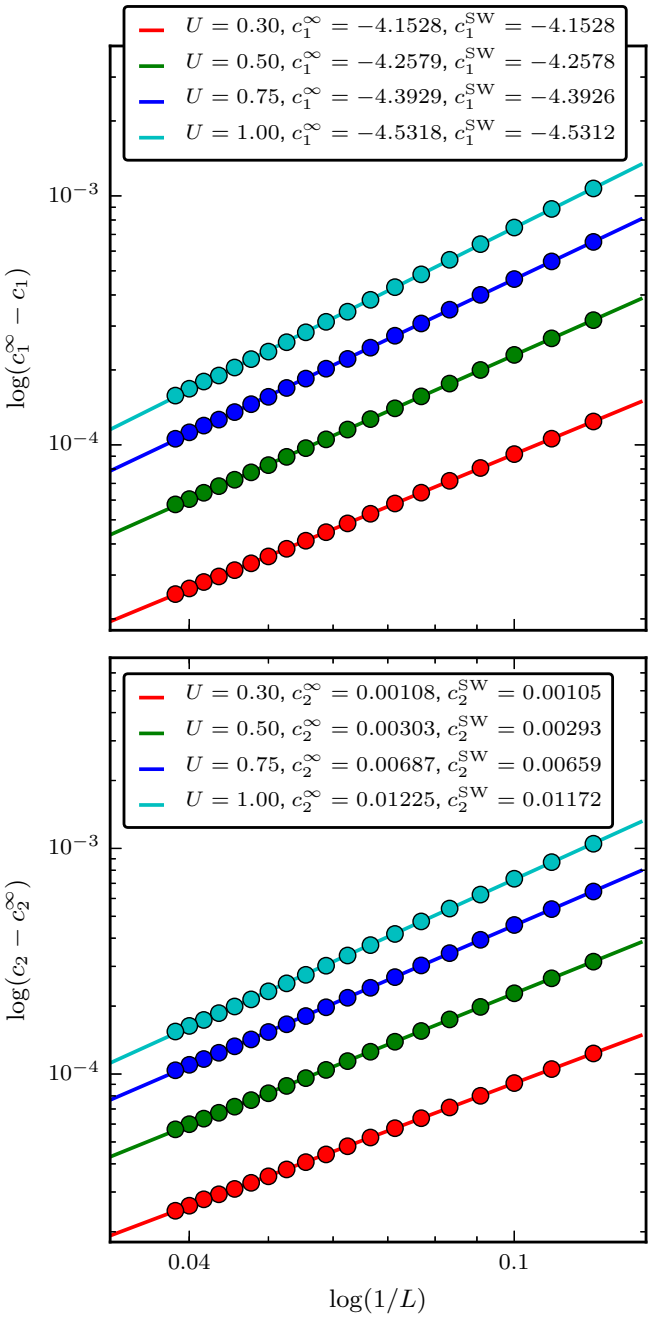

FIG. 2. The power-law dependence of $c_1(L)$ (top plot) and $c_2(L)$ (bottom plot) on $1/L$ plotted in log-log scale. The dots represent the DMRG results and the lines correspond to fitted function. Different colors correspond to different strengths of the attractive Hubbard interaction. Here, we set $t = 1$. In the legends the values of $c_{1(2)}^{\infty}$ and $c_{1(2)}^{\mathrm{SW}}$ are presented for a direct comparison.

result of higher-order interband effects, which become more relevant for a large interaction strength U.

VIII. RELATION BETWEEN WINDING NUMBER AND SUPERFLUID WEIGHT

Using Eq. (26) and the positive semidefiniteness of the QGT it has been shown in Ref. [26] that in two dimensions a flat band with nonzero Chern number is guaranteed to have a finite superfluid weight with the bound $D_s \geqslant |\mathcal{C}|$, in appropriate units. In this section we derive a similar bound in one dimension and show that a flat band with nonzero *winding number* has finite Drude weight D. In one dimension the superfluid weight is not defined and it is better to use the Drude weight instead, which measures the ability of the system to sustain ballistic transport. The Drude weight and the superfluid weight are equal in gapped systems in $d \geqslant 2$.

Let us consider a general one-dimensional system with chiral symmetry represented by a unitary operator Γ, which anticommutes with the Bloch Hamiltonian $\{\Gamma, \mathcal{H}(k)\} = 0$. The presence of a gap in the band structure allows us to deform the Bloch Hamiltonian $\mathcal{H}(k)$ adiabatically into a flat band Hamiltonian defined as $h(k) = \mathcal{Q}_k - \mathcal{P}_k$, where $\mathcal{P}_k$ and $\mathcal{Q}_k = 1 - \mathcal{P}_k$ are the projectors to the subspace of the bands below and above the gap, respectively. Then a topological invariant for the system known as the winding number is given by [53,54]

$$\mathcal{W} = \frac{i}{4\pi} \int_{\mathrm{B.Z.}} dk\, \mathrm{Tr}[\Gamma h(k) \partial_k h(k)] \tag{52}$$

$$= \frac{i}{\pi} \int_{\mathrm{B.Z.}} dk\, \mathrm{Tr}(\Gamma \mathcal{P}_k \partial_k \mathcal{P}_k), \tag{53}$$

where the last equality was obtained by using $h(k) = 1 - 2\mathcal{P}_k$ and the fact that the integral of the full derivative term $\partial_k[\mathrm{Tr}\,(\Gamma \mathcal{P}_k)]$ over the Brillouin zone vanishes.

The QGT defined in Eq. (27) can be written alternatively as $\mathcal{B}_{ij}(\mathbf{k}) = 2\,\mathrm{Tr}\,[\mathcal{P}_{\mathbf{k}} \partial_{k_i} \mathcal{P}_{\mathbf{k}} \partial_{k_j} \mathcal{P}_{\mathbf{k}}]$ and in one dimension it becomes a real valued scalar function $\mathcal{B}(k) = 2\,\mathrm{Tr}\,[\mathcal{P}_k (\partial_k \mathcal{P}_k)^2]$. Using the cyclicity of the trace, the properties $\mathcal{P}_k^2 = \mathcal{P}_k$ and $\mathcal{P}_k^{\dagger} = \mathcal{P}_k$ of the projector and by inserting $\Gamma\Gamma^{\dagger}$ one can easily verify that $\mathcal{B}(k) = 2\,\mathrm{Tr}\,[\mathcal{P}_k (\partial_k \mathcal{P}_k)^2] = 2\,\mathrm{Tr}[A_k^{\dagger} A_k]$ with $A_k = \Gamma \mathcal{P}_k \partial_k \mathcal{P}_k$. Note that the integrand of Eq. (53) is equal to $\mathrm{Tr}(A_k)$.

Let $\mathcal{M}$ be the integral of the QGT over the Brillouin zone defined by

$$\mathcal{M} = \frac{1}{a\pi} \int_{\mathrm{B.Z.}} dk\, \mathcal{B}(k), \tag{54}$$

where the factor $1/a\pi$ is for convenience and a is the lattice constant. Now we can prove the following lower bound for $\mathcal{M}$,

$$\begin{aligned} \mathcal{M} &= \frac{2}{a\pi} \int_{\mathrm{B.Z.}} dk\ \mathrm{Tr}(A_k^{\dagger} A_k) \\ &\geqslant \frac{2}{a\pi\, \mathrm{rk}(A_k)} \int_{\mathrm{B.Z.}} dk\, |\,\mathrm{Tr}(A_k)|^2 \\ &\geqslant \frac{1}{\mathrm{rk}(A_k)\,\pi^2} \left| \int_{\mathrm{B.Z.}} dk\, \mathrm{Tr}(\Gamma \mathcal{P}_k \partial_k \mathcal{P}_k) \right|^2 \geqslant \frac{\mathcal{W}^2}{\mathrm{rk}(\mathcal{P}_k)}. \end{aligned} \tag{55}$$

Above in the first step we have used the fact that for any $n \times n$ matrix A the inequality $\mathrm{Tr}(A^{\dagger} A) \geqslant |\,\mathrm{Tr}(A)|^2 / \mathrm{rk}(A)$ holds [55], where $\mathrm{rk}(A)$ denotes the rank, and then we have used a special case of Schwartz's inequality. In the last step we have used the definition of the winding number given in Eq. (53) and that $\mathrm{rk}(A_k) \leqslant \mathrm{rk}(\mathcal{P}_k)$. Note that $\mathrm{rk}(\mathcal{P}_k)$ is equal to the number of bands below the gap and obviously does not depend on k.

Since the Drude weight D is given by Eq. (26) in any dimension, including $d = 1$ where the superfluid weight is not defined, we conclude that in the flat band limit the lower bound $D \propto \mathcal{M} \geqslant \mathcal{W}^2/\,\mathrm{rk}(\mathcal{P}_k)$ holds. This bound implies that in one dimension a topological nontrivial flat band ($\mathcal{W} \neq 0$) has a guaranteed finite Drude weight for finite attractive interaction.

For the Creutz ladder the chiral symmetry is represented by the Pauli matrix σ_2. Since there is only one band below the gap we have $\mathrm{rk}(\mathcal{P}_{k,\mathrm{C}}) = 1$. This means that according to the above bound we have $\mathcal{M}_\mathrm{C} \geqslant (\mathcal{W}_\mathrm{C})^2$. Using the Bloch functions in Eq. (4), one can easily check that $\mathcal{B}_\mathrm{C}(k) = a^2/2$ and $\mathrm{Tr}\,(\sigma_2 \mathcal{P}_{k,\mathrm{C}} \partial_k \mathcal{P}_{k,\mathrm{C}}) = -ia/2$. From Eq. (53) we obtain $\mathcal{W}_\mathrm{C} = 1$ and from Eq. (54) we get $\mathcal{M}_\mathrm{C} = 1$. Hence, for the Creutz ladder the actual value of $\mathcal{M}$ coincides with the lower bound, and we have that the inequality is saturated, i.e., $\mathcal{M}_\mathrm{C} = (\mathcal{W}_\mathrm{C})^2$.

IX. DISCUSSION AND CONCLUSION

In this work we have provided some definite results that help in characterizing the low-energy physics of attractive Hubbard models in flat bands. Some results are also relevant for the repulsive case at half filling in the context of flat band ferromagnetism.

The first important conclusion is that under the condition of uniform pairing for the flat band Bloch/Wannier functions given in Eq. (17) the BCS wave function is an exact zero-temperature ground state at the level of the projected interaction Hamiltonian. This result justifies the use of mean-field theory in Ref. [26] where the relation (26) between superfluid weight and QGT has been derived and extends similar results valid only for bipartite lattices [27] and Landau levels [26]. It is an interesting question to understand what happens if the uniform pairing condition is violated. In fact in many flat band models this condition is satisfied due to symmetry reasons, such as in the Creutz ladder, the Lieb lattice, kagome lattice, and Landau levels, but it is also easy to construct realistic models where it is not.

We note that it is trivial to prove that the ferromagnetic ground state is exact in a repulsive spin-isotropic Hubbard model if the flat band is half filled and is the lowest-lying band. Our result is more subtle since the BCS ground state is not exact anymore if higher-order terms in the SW expansion are included and it is not the exact ground state of the full Hamiltonian, even if the flat band is the lowest-lying band. Indeed, at first order in U, the attractive and repulsive cases are related by a duality that takes the form of a particle-hole transformation (Appendix A). This duality is broken if interband transitions are taken into account, as we show explicitly by computing the second-order term in the SW expansion for the effective Hamiltonian.

We show that in the attractive case at the level of projected interaction Hamiltonian there is an emergent SU(2) symmetry, which corresponds to the SU(2) spin symmetry in the repulsive case possessed by the full Hamiltonian. As a manifestation of this emergent symmetry, we have that the BCS wave function is the exact ground state of $\overline{\mathcal{H}}_\mathrm{int}$, and the compressibility of the system is diverging since the ground-state energy is linear in the flat band filling ν. This is a very useful result in the context of ultracold gas experiments since the compressibility is routinely measured.

Due to the breaking of the emergent SU(2) symmetry by interband transitions, the compressibility is in fact finite and positive as we show explicitly in the case of the Creutz-Hubbard model, by calculating the ground-state energy up to second order in U. Indeed, the compressibility *decreases* if the coupling constant of the attractive interaction decreases $-U \to -\infty$. In many other fermionic systems the behavior is rather different since the compressibility *increases* if the interaction becomes more attractive. It would be interesting to understand if this persists at nonzero temperature in order to make a comparison with ultracold gas experiments. A related question is whether an intermediate behavior is possible, that is if it is possible to make the compressibility independent of the coupling constant for a suitable choice of the band energy as a function of quasimomentum.

A byproduct of our proof of the exactness of the BCS wave function is the relation $n_\phi \langle \bar{n}_{\mathbf{i}\alpha\uparrow} \rangle = n_\phi \langle \bar{n}_{\mathbf{i}\alpha\downarrow} \rangle = \langle \bar{n}_{\mathbf{i}\alpha\uparrow} \bar{n}_{\mathbf{i}\alpha\downarrow} \rangle$ between local density and double occupancy, where n_ϕ is the parameter defined in Eq. (17). This simple result can be tested in ultracold gas experiments, for example, with site-resolved imaging, and should be valid even at quite high temperatures.

We have shown how from the projected interaction Hamiltonian $\overline{\mathcal{H}}_\mathrm{int}$ many terms can be dropped resulting in an effective spin model $\hat{\mathcal{H}}_\mathrm{spin}$ (20) which is a very convenient model for computations and, furthermore, offers an intuitive model of a flat band superfluid as a ferromagnet. This approximation is justified since exact properties, such as the exactness of the BCS wave function and the emergent SU(2) symmetry, are preserved, while the main drawback is that the invariance under gauge transformations of the Wannier functions present in the projected interaction Hamiltonian is broken.

We provide at least a partial solution to this problem by noticing that in the preferred gauge of maximally localized Wannier functions the gauge invariant result for the superfluid weight of Ref. [26] is recovered in one dimension. In higher dimensions the result obtained from the effective spin model is in general larger than the gauge invariant one of Ref. [26] since the gauge noninvariant term $\widetilde{\Omega}$ of the Marzari-Vanderbilt localization functional is in general nonzero even when calculated on the maximally localized Wannier functions.

It has been recently proved that the Marzari-Vanderbilt localization functional is finite if and only if the band has zero Chern number [47]. This means that the approximation leading to the spin Hamiltonian (20) fails in the case of bands with nonzero Chern number since it leads to the unphysical result of an infinite superfluid weight. The case of nonzero Chern number is not discussed in this work and is a very interesting subject for future research. A possible approach is to use Wannier functions that mix the two components of spin [56]. Since the sum of Chern numbers of the bands with up and down spin is zero due to time-reversal symmetry, these Wannier functions can be taken exponentially localized, which leads to a finite superfluid weight. Also the case of a flat band with zero Chern number but large Berry curvature is interesting since in this case $\widetilde{\Omega}$ can be large, and the effective spin Hamiltonian (20) is probably not a very good approximation.

Our results are also relevant for the general problem of estimating the critical temperature of the superconducting

transition. Notice that the BCS wave function in a flat band is identical to the one in the strong-coupling limit $-U/t \to -\infty$ (t is the hopping energy scale) with the only difference that in the flat band the Cooper pair wave function coincides with the flat band Wannier function and is delocalized on n_ϕ^{-1} orbitals, while in the strong-coupling limit the Cooper pair is localized on a single orbital. Indeed, it is well known [57] that the strong-coupling limit can be treated with the SW transformation as an expansion in the parameter t/U which is essentially the inverse of the parameter $U/E_{\rm gap}$ employed here. Similar to the strong-coupling limit, the critical temperature is controlled by the collective fluctuations and the relevant energy scale is $\sim D_{\rm s}$. In the flat band limit one has $D_{\rm s} \propto U$, while in the strong-coupling limit $D_{\rm s} \propto t^2/U$. In both cases the effective Hamiltonian at the leading order is an isotropic Heisenberg ferromagnet with the difference that in the strong-coupling limit the range of the ferromagnetic couplings coincides with the range of the hopping terms in the kinetic Hamiltonian, while they can be long ranged in the flat band limit. If the flat band is intersected by other bands, as in the case of the Lieb lattice [27] and kagome lattice [28], or has nonzero Chern number, the Wannier functions are algebraically decaying to the extent that the Mermin-Wagner theorem may not be applicable. However it is known that the Mermin-Wagner theorem is valid for a very large class of lattice models, [58] including the ones studied here. This is another way to see that the effective spin Hamiltonian (20) fails in the case of nonzero Chern number and for models with a band intersection. Extending the effective spin Hamiltonian to these cases would be an important step towards providing a reliable estimate of the critical temperature in the flat band limit.

Some of the results provided here are of more general interest, beyond the topic of flat band superconductivity. The same effective spin Hamiltonian (20) is the first-order term in the SW expansion for a spin-isotropic repulsive Hubbard model with half-filled flat band, and is therefore the low-energy theory of a flat band ferromagnet (see Appendix A). In particular, this shows that the QGT plays a crucial role also for flat band ferromagnets, in which case the quantity corresponding to the superfluid weight is the spin stiffness. Indeed, suppose that $\Omega^I_{ij} = 0$. This implies that the QGT $\mathcal{B}_{ij}(\mathbf{k}) = 0$ vanishes everywhere in the Brillouin zone and so does $\nabla_{\mathbf{k}} P_{\mathbf{k}} = 0$ (with $P_{\mathbf{k}} = |g_{\mathbf{k}}\rangle\langle g_{\mathbf{k}}|$). Therefore, up to a gauge transformation, the Bloch functions can be taken as constants $|g_{\mathbf{k}}\rangle = |g\rangle$ and these produce maximally localized Wannier functions of the form $\overline{W}_\alpha(\mathbf{i}-\mathbf{j}) = \delta_{\mathbf{i},\mathbf{j}}\langle\alpha|g\rangle$. The spread and overlap functionals calculated on these maximally localized Wannier functions are both zero $[\nabla^2_{\mathbf{q}} F_{\rm ov}(\mathbf{q}=0)[\overline{W}] = \nabla^2_{\mathbf{q}} F_{\rm MV}(\mathbf{q}=0)[\overline{W}] = 0]$. The same set of Wannier functions form a basis for which the local connectivity condition [37] is not satisfied. The local connectivity condition is crucial in order to prove the uniqueness of the ferromagnetic ground state in a repulsive Hubbard model. An equivalent condition is the irreducibility of the projection operator [59,60], which is also not satisfied if $\Omega^I_{ij} = 0$, since it is apparent from $P_{\alpha\beta}(\mathbf{i}-\mathbf{j}) = \delta_{\mathbf{i},\mathbf{j}}\langle\alpha|g\rangle\langle g|\beta\rangle$ that the projection operator is highly reducible (block diagonal). We suspect that the local connectivity and the irreducibility of the projector are somehow related to Ω^I_{ij} being nonsingular, and this is an interesting open question for future investigations.

Other results of general interest are the inequality $\mathcal{M} \geqslant \mathcal{W}^2$ between the quantum metric and the winding number, and the various relations between the overlap functional and the Marzari-Vanderbilt localization functional, which may find applications in other contexts. In particular, we believe that inequalities between quantum metric and the rich variety of topological invariants known so far can be found, similarly to what has been done here for the one-dimensional winding number and in previous works for the Chern number.

A promising way to verify our results would be via transport measurements in cold atom experiments, as pioneered in Ref. [61]. Furthermore, for a tangible experimental realization of the Creutz ladder, we refer the reader to Refs. [62,63] and the recent advances in implementing artificial magnetic fields for neutral atoms in synthetic dimensions [64–68].

ACKNOWLEDGMENTS

We acknowledge useful discussions with L. Liang, A. Paramekanti, and J. Tempere. S.P. thanks the authors of Ref. [47] for sharing their results before publication. Our DMRG simulations are based on the ALPS libraries [69,70]. This work was supported by the Academy of Finland through its Centers of Excellence Programme (2012-2017) and under Projects No. 263347, No. 251748, No. 272490, and No. 284621, and by the European Research Council (ERC-2013-AdG-340748-CODE). This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie Grant Agreement No. 702281 (FLATOPS). M.T. and S.D.H. acknowledge support from the Swiss National Science Foundation. S.P. thanks the Pauli Center for Theoretical Studies, and M.T. thanks the COMP Centre of Excellence of Aalto University for their hospitality.

APPENDIX A: DUALITY BETWEEN ATTRACTIVE AND REPULSIVE HUBBARD MODELS

In this Appendix we explain in detail the duality between the attractive and repulsive Hubbard models. In the most general setting this duality is valid at the level of the projected interaction Hamiltonian presented in Eq. (8) and relates an attractive Hubbard model with time-reversal symmetry and symmetric with respect to rotations of the spin around a chosen axis (the z axis for instance) to a repulsive Hubbard model with SU(2) spin rotational symmetry. Our argument also clarifies why it is natural to relate Hubbard models with precisely the above symmetries.

A similar duality holds for bipartite lattices [27,71] and takes the form of a particle-hole transformation

$$\hat{c}_{\mathbf{i}\alpha\uparrow} \to \hat{c}_{\mathbf{i}\alpha\uparrow}, \quad \hat{c}_{\mathbf{i}\alpha\downarrow} \to s(\alpha)\hat{c}^\dagger_{\mathbf{i}\alpha\downarrow}. \tag{A1}$$

In a bipartite lattice the sublattices labeled by α,β can be partitioned in two subsets $\mathcal{S}_{i=1,2}$ and the hopping matrix element $t_{\mathbf{i}\alpha,\mathbf{j}\beta}$ is zero if $\alpha,\beta \in \mathcal{S}_i$, namely only hoppings connecting one subset to its complement are allowed. Correspondingly, the function $s(\alpha)$ in the above equation takes values $s(\alpha) = +1$ if $\alpha \in \mathcal{S}_1$ and $s(\alpha) = -1$ if $\alpha \in \mathcal{S}_2$. If the Hamiltonian commutes with the z component of spin, the spin up and spin down sectors of the kinetic term are decoupled and have the general form

given in Eq. (1). The spin-up term is unchanged under Eq. (A1), while for the spin down we have

$$\hat{\mathcal{H}}^{\downarrow}_{\rm kin} = \sum_{\mathbf{i}\alpha,\mathbf{j}\beta} t^{\downarrow}_{\mathbf{i}\alpha,\mathbf{j}\beta}\, \hat{c}^{\dagger}_{\mathbf{i}\alpha\downarrow}\hat{c}_{\mathbf{j}\beta\downarrow} \rightarrow \sum_{\mathbf{i}\alpha,\mathbf{j}\beta} s(\alpha)s(\beta) t^{\downarrow}_{\mathbf{i}\alpha,\mathbf{j}\beta}\hat{c}_{\mathbf{i}\alpha\downarrow}\hat{c}^{\dagger}_{\mathbf{j}\beta\downarrow}$$
$$= -\sum_{\mathbf{i}\alpha,\mathbf{j}\beta} t^{\downarrow}_{\mathbf{j}\beta,\mathbf{i}\alpha}\hat{c}_{\mathbf{j}\beta\downarrow}\hat{c}^{\dagger}_{\mathbf{i}\alpha\downarrow} = \sum_{\mathbf{i}\alpha,\mathbf{j}\beta} \left(t^{\downarrow}_{\mathbf{i}\alpha,\mathbf{j}\beta}\right)^{*} \hat{c}^{\dagger}_{\mathbf{i}\alpha\downarrow}\hat{c}_{\mathbf{j}\beta\downarrow}. \quad \text{(A2)}$$

Above we have used that $s(\alpha)s(\beta) = -1$ due to the bipartiteness and that $t^{\downarrow}_{\mathbf{j}\beta,\mathbf{i}\alpha} = (t^{\downarrow}_{\mathbf{i}\alpha,\mathbf{j}\beta})^{*}$ due to the hermiticity of the hopping matrix. Hence, the net result of this transformation is to send the hopping matrix for down spin to its time-reversal conjugate. Therefore, if initially the kinetic Hamiltonian has SU(2) spin symmetry, which means $t^{\uparrow}_{\mathbf{i}\alpha,\mathbf{j}\beta} = t^{\downarrow}_{\mathbf{i}\alpha,\mathbf{j}\beta}$, after applying the transformation (A1) one obtains a time-reversal symmetric kinetic Hamiltonian, which then satisfies the relation $t^{\uparrow}_{\mathbf{i}\alpha,\mathbf{j}\beta} = (t^{\downarrow}_{\mathbf{i}\alpha,\mathbf{j}\beta})^{*}$. Moreover the same transformation maps a repulsive Hubbard interaction to an attractive one up to a term proportional to the total number of up spin particles. Therefore for bipartite lattices one has a duality between the full Hamiltonians $\hat{\mathcal{H}} = \hat{\mathcal{H}}_{\rm kin} + \hat{\mathcal{H}}_{\rm int}$.

For general models introduced in Eq. (1) a quite similar duality holds only for the projected Hamiltonian, and it is based on the following transformations of the flat band operators:

$$\hat{d}_{\mathbf{l}\downarrow} \rightarrow \hat{d}^{\dagger}_{\mathbf{l}\downarrow}, \quad \text{(A3)}$$

$$\underbrace{\bar{c}_{\mathbf{i}\alpha\downarrow} = \sum_{\mathbf{l}} W_{\alpha}(\mathbf{i}-\mathbf{l})\,\hat{d}_{\mathbf{l}\downarrow}}_{\text{repulsive model}} \rightarrow \underbrace{\sum_{\mathbf{l}} W_{\alpha}(\mathbf{i}-\mathbf{l})\,\hat{d}^{\dagger}_{\mathbf{l}\downarrow} = \bar{c}^{\dagger}_{\mathbf{i}\alpha\downarrow}}_{\text{attractive model}}, \quad \text{(A4)}$$

while the up spin operators are unchanged. For a bipartite lattice the above transformations are simply the projected form of Eq. (A1). Indeed the zero energy flat band of a bipartite lattice has support on the orbitals belonging only to a single subset $\mathcal{S}_i$. However Eqs. (A3) and (A4) can be defined regardless of the lattice being bipartite or not. Note that in Eq. (A4) we have chosen to present the transformation with the projected creation operator of the attractive model on the right [$W_{\alpha\downarrow}(\mathbf{i}-\mathbf{l}) = W^{*}_{\alpha\uparrow}(\mathbf{i}-\mathbf{l})$], while on the left-hand side the projected annihilation operator of the repulsive model appears [$W_{\alpha\downarrow}(\mathbf{i}-\mathbf{l}) = W_{\alpha\uparrow}(\mathbf{i}-\mathbf{l})$]. Since for a flat band the projected kinetic term is just a constant, one needs only to consider the effect of the transformation (A4) when applied to the projected interaction Hamiltonian of the repulsive model,

$$\overline{\mathcal{H}}_{\rm int} = U\sum_{\mathbf{i}\alpha} \bar{n}_{\mathbf{i}\alpha\uparrow}\bar{n}_{\mathbf{i}\alpha\downarrow} \rightarrow U\sum_{\mathbf{i}\alpha} \bar{n}_{\mathbf{i}\alpha\uparrow}\bar{c}_{\mathbf{i}\alpha\downarrow}\bar{c}^{\dagger}_{\mathbf{i}\alpha\downarrow}$$
$$= -U\sum_{\mathbf{i}\alpha} \bar{n}_{\mathbf{i}\alpha\uparrow}\bar{n}_{\mathbf{i}\alpha\downarrow} + U\sum_{\mathbf{i}\alpha} \bar{n}_{\mathbf{i}\alpha\uparrow} P_{\alpha\alpha}(\mathbf{0})$$
$$= -U\sum_{\mathbf{i}\alpha} \bar{n}_{\mathbf{i}\alpha\uparrow}\bar{n}_{\mathbf{i}\alpha\downarrow} + n_{\phi}U\overline{N}_{\uparrow}. \quad \text{(A5)}$$

Note the use of the uniform pairing condition on the last step. In the case of equal numbers of up and down spins, which is dual to a half-filled band for the repulsive Hubbard model, the last line of Eq. (A5) is equivalent to Eq. (18).

Note also that under the transformation (A3) the pseudospin operators $\hat{\mathrm{T}}^{\pm,z}_{\mathbf{i}}$ in Eq. (20) are transformed to the real spin operators for the Wannier states

$$\hat{\mathrm{S}}^{z}_{\mathbf{i}} = \tfrac{1}{2}(\hat{d}^{\dagger}_{\mathbf{i}\uparrow}\hat{d}_{\mathbf{i}\uparrow} - \hat{d}^{\dagger}_{\mathbf{i}\downarrow}\hat{d}_{\mathbf{i}\downarrow}), \quad \hat{\mathrm{S}}^{+}_{\mathbf{i}} = (\hat{\mathrm{S}}^{-}_{\mathbf{i}})^{\dagger} = \hat{d}^{\dagger}_{\mathbf{i}\uparrow}\hat{d}_{\mathbf{i}\downarrow}. \quad \text{(A6)}$$

Therefore the effective (pseudo)spin Hamiltonians have the same form in the repulsive and attractive case at first order in U, but this duality breaks down already at the next order as shown explicitly in Secs. VI and VII for the Creutz-Hubbard model.

APPENDIX B: PROOF OF THE IDENTITY $A_{\sigma}|\Omega\rangle = 0$

Here we provide an alternative simple proof of the identity $\langle\Omega|\hat{A}^{\dagger}_{\sigma}\hat{A}_{\sigma}|\Omega\rangle = 0$, which has been shown explicitly and used in Sec. VII to calculate the second-order correction to the compressibility for the Creutz-Hubbard model. This proof relies only on the commutation relations of the projected fermionic operators and is valid not just for the Creutz ladder, but for arbitrary Hubbard models in the class defined in Sec. II A.

Since $\hat{A}^{\dagger}_{\sigma}\hat{A}_{\sigma}$ is a positive semidefinite operator, the vanishing of the expectation value is equivalent to $\hat{A}_{\sigma}|\Omega\rangle = 0$ and it is easier to show this last identity. In particular, we prove inductively that $\hat{A}_{\sigma}(\hat{b}^{\dagger}_{0})^{N}|\emptyset\rangle = 0$ for all N.

A consequence of time-reversal symmetry is the commutation relation $[\bar{c}_{\mathbf{i}\alpha\sigma}, \hat{b}^{\dagger}_{0}] = \sigma\bar{c}^{\dagger}_{\mathbf{i}\alpha\bar{\sigma}}$ with $\sigma = \uparrow(+)$ or $\sigma = \downarrow(-)$ and $\bar{\sigma} = -\sigma$. One immediately obtains $[\bar{n}_{\mathbf{i}\alpha\sigma}, \hat{b}^{\dagger}_{0}] = \bar{c}^{\dagger}_{\mathbf{i}\alpha\uparrow}\bar{c}^{\dagger}_{\mathbf{i}\alpha\downarrow}$, from which it follows that the operator $\hat{b}^{\dagger}_{0}$ commutes with the projected spin operator $\bar{S}^{z}_{\mathbf{i}\alpha}$, a crucial result used in Sec. III. Using the inductive hypothesis, the previous commutation relations and that $(\bar{c}^{\dagger}_{\mathbf{i}\alpha\sigma})^{2} = 0$, we have

$$\hat{A}_{\sigma}(\hat{b}^{\dagger}_{0})^{N}|\emptyset\rangle = [\hat{A}_{\sigma}, \hat{b}^{\dagger}_{0}](\hat{b}^{\dagger}_{0})^{N-1}|\emptyset\rangle$$
$$= -U\sum_{\mathbf{i}\alpha} \tilde{c}^{\dagger}_{\mathbf{i}\alpha\sigma}\bar{c}_{\mathbf{i}\alpha\sigma}\bar{c}^{\dagger}_{\mathbf{i}\alpha\uparrow}\bar{c}^{\dagger}_{\mathbf{i}\alpha\downarrow}(\hat{b}^{\dagger}_{0})^{N-1}|\emptyset\rangle$$
$$= -U\sigma\sum_{\mathbf{i}\alpha} P_{\alpha\alpha}(\mathbf{0})\tilde{c}^{\dagger}_{\mathbf{i}\alpha\sigma}\bar{c}^{\dagger}_{\mathbf{i}\alpha\bar{\sigma}}(\hat{b}^{\dagger}_{0})^{N-1}|\emptyset\rangle$$
$$+ U\sigma\sum_{\mathbf{i}\alpha} \tilde{c}^{\dagger}_{\mathbf{i}\alpha\sigma}\bar{c}^{\dagger}_{\mathbf{i}\alpha\bar{\sigma}}\bar{n}_{\mathbf{i}\alpha\sigma}(\hat{b}^{\dagger}_{0})^{N-1}|\emptyset\rangle. \quad \text{(B1)}$$

The first term in the last line vanishes due to the uniform pairing condition [$P_{\alpha\alpha}(\mathbf{0}) = n_{\phi}$] and the identity $\sum_{\mathbf{i}\alpha} \tilde{c}^{\dagger}_{\mathbf{i}\alpha\sigma}\bar{c}^{\dagger}_{\mathbf{i}\alpha\bar{\sigma}} = 0$ which can be verified by expanding in any orthonormal basis, using time-reversal symmetry and noting that the wave functions relative to different bands are orthogonal. The operator $\bar{c}^{\dagger}_{\mathbf{i}\alpha\bar{\sigma}}\bar{n}_{\mathbf{i}\alpha\sigma}$ commutes with $\hat{b}^{\dagger}_{0}$ and annihilates the vacuum state, therefore also the second term vanishes, concluding the proof.

The result $\hat{A}_{\sigma}|\Omega\rangle = 0$ is very useful since it implies that the virtual process corresponding to the transfer of a single particle to the upper band of *any* lattice model, processes encoded in the operators $\hat{A}_{\sigma}$, do not produce any second-order correction to the ground-state energy, and hence to the compressibility. This can be easily verified using Eqs. (11), (12), and (39). This is a significant simplification since only the expectation value involving the operator $\hat{B}$ needs to be evaluated.

APPENDIX C: RELATIONS BETWEEN THE OVERLAP AND MARZARI-VANDERBILT FUNCTIONALS

Before providing a proof of Eqs. (34) and (35), we present other relations between the two functionals for completeness. The two functionals can be identified up to the constant factor n_ϕ for arbitrary values of $\mathbf{q}$ if the condition

$$|W_\alpha(\mathbf{i})|^2 = |W_\beta(\mathbf{i})|^2 \quad \text{for every } \alpha,\beta,\mathbf{i}, \tag{C1}$$

is satisfied. For example, the above equation holds if the orbital $W_\alpha(\mathbf{i})$ is the plaquette state of the Creutz ladder. However, even in the Creutz ladder this condition does not hold for arbitrary gauge choices, and therefore is not useful for our purposes. On the other hand by using the elementary inequality $\sum_{\alpha=1}^n |v_\alpha|^2/n \geqslant |\sum_{\alpha=1}^n v_\alpha/n|^2$ it is easily shown that

$$F_{\rm ov}(\mathbf{q})[W] \leqslant n_\phi F_{\rm MV}(\mathbf{q})[W]. \tag{C2}$$

Here n_ϕ^{-1} is simply the number of α's for which $W_\alpha(\mathbf{i}) \neq 0$ and is not defined by the condition (17), which is not necessary for the above inequality to hold. Equation (C2) provides a strong inequality but in the wrong sense and again it is not useful in this context.

The proof of Eqs. (34) and (35) relies on the following relation between Wannier functions and Bloch functions:

$$\sum_{\mathbf{i}} |W_\alpha(\mathbf{i})|^2 e^{i\mathbf{q}\cdot\mathbf{r_i}} = \frac{V_c}{(2\pi)^d}\int_{\rm B.Z.} d^d\mathbf{k}\, g^*_{\mathbf{k}+\mathbf{q}}(\alpha) g_{\mathbf{k}}(\alpha), \tag{C3}$$

which is readily derived from the definition of Wannier functions given in Eq. (3). Accordingly, we write $F_{\rm ov}(\mathbf{q})[g]$ and $F_{\rm MV}(\mathbf{q})[g]$ to the denote that these are functionals of the Bloch functions $g_{\mathbf{k}}(\alpha)$ corresponding to the Wannier functions $W_\alpha(\mathbf{i})$. Moreover, we denote by $\bar{g}_{\mathbf{k}}(\alpha)$ the Bloch functions corresponding to the maximally localized Wannier functions in the Marzari-Vanderbilt sense $\overline{W}_\alpha(\mathbf{i})$.

In the following we assume that the Bloch functions are analytic and periodic functions of the quasimomentum $\mathbf{k}$. This allows us in particular to perform integration by parts and drop full derivatives which integrate to zero over the whole Brillouin zone. In general it is not possible to find an analytic and periodic gauge if the Chern numbers in $d \geqslant 2$ are nonzero [44,47,48], while in $d = 1$ this is always possible [72,73]. For the Creutz model the smooth and periodic Bloch functions are given in Eq. (4). In what follows we restrict to the case of bands with zero Chern numbers.

Let us first insert Eq. (C3) into the definition of the overlap and Marzari-Vanderbilt localization functionals given in Eqs. (30) and (31), respectively, and take the Laplacian $\nabla^2_{\mathbf{q}}$. After integration by parts we are left with

$$\nabla^2_{\mathbf{q}} F_{\rm ov}(\mathbf{q}=0)[g] = \frac{2V_c}{(2\pi)^d}\sum_\alpha P_{\alpha\alpha}(\mathbf{0})\int_{\rm B.Z.} d^d\mathbf{k}\, \nabla g^*_{\mathbf{k}}(\alpha) \cdot \nabla g_{\mathbf{k}}(\alpha) - 2\sum_\alpha |\boldsymbol{\theta}_\alpha|^2, \tag{C4}$$

$$\nabla^2_{\mathbf{q}} F_{\rm MV}(\mathbf{q}=0)[g] = \frac{2V_c}{(2\pi)^d}\int_{\rm B.Z.} d^d\mathbf{k}\, \langle \nabla g_{\mathbf{k}}| \cdot |\nabla g_{\mathbf{k}}\rangle - 2|\boldsymbol{\theta}|^2. \tag{C5}$$

The second terms on the right-hand sides of Eqs. (C4) and (C5) depend on the quantities

$$\boldsymbol{\theta}_\alpha = -i\frac{V_c}{(2\pi)^d}\int_{\rm B.Z.} d^d\mathbf{k}\, g^*_{\mathbf{k}}(\alpha)\nabla g_{\mathbf{k}}(\alpha), \tag{C6}$$

$$\boldsymbol{\theta} = -i\frac{V_c}{(2\pi)^d}\int_{\rm B.Z.} d^d\mathbf{k}\, \langle g_{\mathbf{k}}|\nabla g_{\mathbf{k}}\rangle = \frac{V_c}{(2\pi)^d}\int_{\rm B.Z.} d^d\mathbf{k}\, \mathbf{A}(\mathbf{k}). \tag{C7}$$

Equation (C7) is known as the Zak phase (up to normalization) while the $\boldsymbol{\theta}_\alpha$ in Eq. (C6) are orbital-resolved Zak phases whose sum give the Zak phase $\boldsymbol{\theta} = \sum_\alpha \boldsymbol{\theta}_\alpha$. The Zak phase can be expressed in terms of the Berry connection $\mathbf{A}(\mathbf{k}) = -i\langle g_{\mathbf{k}}|\nabla g_{\mathbf{k}}\rangle$ as shown in Eq. (C7). The Zak phase is gauge invariant up to shifts by an arbitrary lattice vector. Indeed, a generic gauge transformation $g_{\mathbf{k}} \to e^{i\phi_{\mathbf{k}}} g_{\mathbf{k}}$ that preserves periodicity and analyticity is given by a phase of the form $\phi_{\mathbf{k}} = \mathbf{k}\cdot\mathbf{r_l} + \bar{\phi}_{\mathbf{k}}$, where $\mathbf{r_l}$ is an arbitrary lattice vector and $\bar{\phi}_{\mathbf{k}}$ a periodic analytic function, namely $\bar{\phi}_{\mathbf{k}} = \bar{\phi}_{\mathbf{k}+\mathbf{G}_j}$ for all reciprocal-lattice vectors $\mathbf{G}_j$. The nonperiodic part of the phase $\phi_{\mathbf{k}}$ corresponds to a rigid translation of the Wannier function by the lattice vector $\mathbf{r_l}$. Under this transformation the Zak phase $\boldsymbol{\theta}$ transforms as $\boldsymbol{\theta} \to \mathbf{r_l} + \boldsymbol{\theta}$. Instead the orbital-resolved Zak phases $\boldsymbol{\theta}_\alpha$ depend on the choice of gauge for the Bloch functions.

We now restrict ourselves to the case where Eq. (17) holds and let us further assume that

$$\overline{\boldsymbol{\theta}}_\alpha = \overline{\boldsymbol{\theta}}_\beta \quad \text{for } \alpha,\beta \in \mathcal{S}, \tag{C8}$$

when the orbital-resolved Zak phases are calculated on the maximally localized Wannier/Bloch functions, namely $\overline{\boldsymbol{\theta}}_\alpha = -i\frac{V_c}{(2\pi)^d}\int_{\rm B.Z.} d^d\mathbf{k}\, \overline{g}^*_{\mathbf{k}}(\alpha)\nabla\overline{g}_{\mathbf{k}}(\alpha)$. This is the case of the Creutz ladder where $\overline{\theta}_{\alpha=1} = \overline{\theta}_{\alpha=2} = a/4$. Equation (C8) is the second requirement mentioned previously that allows us to derive Eqs. (34) and (35). Indeed, by inserting $P_{\alpha\alpha}(\mathbf{0}) = n_\phi$ and $\boldsymbol{\theta}_\alpha = n_\phi\boldsymbol{\theta}$ into Eq. (C4) and comparing with Eq. (C5), one immediately obtains

$$\nabla^2_{\mathbf{q}} F_{\rm ov}(\mathbf{q}=0)[\overline{W}] = n_\phi \nabla^2_{\mathbf{q}} F_{\rm MV}(\mathbf{q}=0)[\overline{W}]. \tag{C9}$$

The same argument can be straightforwardly adapted to prove Eq. (35). To prove that the Wannier functions are global minimizers of the overlap functional let us add and subtract a term $n_\phi\frac{2V_c}{(2\pi)^d}\int_{\rm B.Z.} d^d\mathbf{k}\,|\mathbf{A}(\mathbf{k})|^2$ in Eq. (C4) and write arbitrary Bloch functions as $g_{\mathbf{k}} = e^{i\phi_{\mathbf{k}}}\overline{g}_{\mathbf{k}}$. This leads to

$$\begin{aligned}\nabla^2_{\mathbf{q}} F_{\rm ov}(\mathbf{q}=0)[g] &= n_\phi\frac{V_c}{(2\pi)^d}\int_{\rm B.Z.} d^d\mathbf{k}\,{\rm Tr}\,\mathcal{B}(\mathbf{k}) + n_\phi\frac{2V_c}{(2\pi)^d}\int_{\rm B.Z.} d^d\mathbf{k}\,|\mathbf{A}(\mathbf{k})|^2 - 2\sum_\alpha|\boldsymbol{\theta}_\alpha|^2 \\ &= n_\phi\nabla^2_{\mathbf{q}} F_{\rm MV}(\mathbf{q}=0)[\overline{g}] + n_\phi\frac{2V_c}{(2\pi)^d}\int_{\rm B.Z.} d^d\mathbf{k}\,|\nabla\phi_{\mathbf{k}}|^2 - 2\sum_\alpha\left(\frac{V_c}{(2\pi)^d}\int_{\rm B.Z.} d^d\mathbf{k}\,|g_{\mathbf{k}}(\alpha)|^2\nabla\phi_{\mathbf{k}}\right)^2.\end{aligned} \tag{C10}$$

In order to derive the above equation we have used again the conditions in Eqs. (17) and (C8) and the fact that the Berry connection calculated on the maximally localized Bloch functions is a divergence-free vector field $\nabla \cdot \overline{\mathbf{A}}(\mathbf{k}) = 0$, as can be easily shown by varying the Marzari-Vanderbilt functional. Note that the above functional is quadratic in $\nabla\phi_{\mathbf{k}}$. Moreover, using the inequality

$$\frac{V_c}{(2\pi)^d}\int_{\text{B.Z.}} d^d\mathbf{k}\,\frac{|g_{\mathbf{k}}(\alpha)|^2}{n_\phi}|f(\mathbf{k})|^2 \geqslant \left|\frac{V_c}{(2\pi)^d}\int_{\text{B.Z.}} d^d\mathbf{k}\,\frac{|g_{\mathbf{k}}(\alpha)|^2}{n_\phi} f(\mathbf{k})\right|^2 \tag{C11}$$

valid for arbitrary functions $f(\mathbf{k})$, one can see from Eq. (C10) that the maximally localized Wannier functions in the Marzari-Vanderbilt sense correspond to a global minimum of the overlap functional, and thus Eq. (34) is proved.

One may wonder if it would be possible to generalize Eq. (34) to a matrix inequality of the form $B \geqslant A$, namely prove that $B - A$ is a positive semidefinite matrix, where the elements of B are $[B]_{i,j} = \partial_{q_i}\partial_{q_j} F_{\text{ov}}(\mathbf{q} = 0)[g]$ and the elements of A are $[A]_{i,j} = \partial_{q_i}\partial_{q_j} F_{\text{ov}}(\mathbf{q} = 0)[\bar{g}]$. This may not be the case as we explain below. In Eq. (C10) one can drop the cross term $\int d^d\mathbf{k}\,\overline{\mathbf{A}}(\mathbf{k}) \cdot \nabla\phi_{\mathbf{k}} = 0$ since $\nabla \cdot \overline{\mathbf{A}}(\mathbf{k}) = 0$. On the other hand, in order to prove the matrix inequality, one should show that $C_{ij} = \int d^d\mathbf{k}\,[\,\overline{A}_i(\mathbf{k})\partial_j\phi_{\mathbf{k}} + \overline{A}_j(\mathbf{k})\partial_i\phi_{\mathbf{k}}\,]$ is a positive semidefinite matrix. Since $\text{Tr}\,C = 0$ this would imply $C = 0$. Given that $\phi_{\mathbf{k}}$ is an arbitrary function, this is equivalent to $\partial_i\overline{A}_j(\mathbf{k}) + \partial_j\overline{A}_i(\mathbf{k}) = 0$. This is a significantly stronger condition than being divergence-free and it is probably not true in general. Therefore the trace in Eq. (25) is important.

[1] D. C. Tsui, H. L. Stormer, and A. C. Gossard, Phys. Rev. Lett. **48**, 1559 (1982).
[2] J. Motruk and A. Mielke, J. Phys. A **45**, 225206 (2012).
[3] M. Creutz, Phys. Rev. Lett. **83**, 2636 (1999).
[4] C. Wu, D. Bergman, L. Balents, and S. Das Sarma, Phys. Rev. Lett. **99**, 070401 (2007).
[5] D. L. Bergman, C. Wu, and L. Balents, Phys. Rev. B **78**, 125104 (2008).
[6] M. Tovmasyan, E. P. L. van Nieuwenburg, and S. D. Huber, Phys. Rev. B **88**, 220510 (2013).
[7] S. Takayoshi, H. Katsura, N. Watanabe, and H. Aoki, Phys. Rev. A **88**, 063613 (2013).
[8] D. Sticlet, L. Seabra, F. Pollmann, and J. Cayssol, Phys. Rev. B **89**, 115430 (2014).
[9] M. Biondi, E. P. L. van Nieuwenburg, G. Blatter, S. D. Huber, and S. Schmidt, Phys. Rev. Lett. **115**, 143601 (2015).
[10] G. Zhu, J. Koch, and I. Martin, Phys. Rev. B **93**, 144508 (2016).
[11] K. Sun, Z. Gu, H. Katsura, and S. Das Sarma, Phys. Rev. Lett. **106**, 236803 (2011).
[12] E. Tang, J.-W. Mei, and X.-G. Wen, Phys. Rev. Lett. **106**, 236802 (2011).
[13] T. Neupert, L. Santos, C. Chamon, and C. Mudry, Phys. Rev. Lett. **106**, 236804 (2011).
[14] O. Derzhko, J. Richter, and M. Maksymenko, Int. J. Mod. Phys. B **29**, 1530007 (2015).
[15] P. Müller, J. Richter, and O. Derzhko, Phys. Rev. B **93**, 144418 (2016).
[16] C. H. Lee, D. P. Arovas, and R. Thomale, Phys. Rev. B **93**, 155155 (2016).
[17] W. Maimaiti, A. Andreanov, H. C. Park, O. Gendelman, and S. Flach, arXiv:1610.02970.
[18] V. A. Khodel and V. R. Shaginyan, JETP Lett. **51**, 553 (1990).
[19] G. E. Volovik, JETP Lett. **53**, 222 (1991).
[20] V. A. Khodel, V. R. Shaginyan, and V. V. Khodel, Phys. Rep. **249**, 1 (1994).
[21] N. B. Kopnin, T. T. Heikkilä, and G. E. Volovik, Phys. Rev. B **83**, 220503 (2011).
[22] T. T. Heikkilä, N. B. Kopnin, and G. E. Volovik, JETP Lett. **94**, 233 (2011).
[23] M. Kellogg, J. P. Eisenstein, L. N. Pfeiffer, and K. W. West, Phys. Rev. Lett. **93**, 036801 (2004).
[24] E. Tutuc, M. Shayegan, and D. A. Huse, Phys. Rev. Lett. **93**, 036802 (2004).
[25] J. Eisenstein, Annu. Rev. Condens. Matter Phys. **5**, 159 (2014).
[26] S. Peotta and P. Törmä, Nat. Commun. **6**, 8944 (2015).
[27] A. Julku, S. Peotta, T. I. Vanhala, D.-H. Kim, and P. Törmä, Phys. Rev. Lett. **117**, 045303 (2016).
[28] S. D. Huber and E. Altman, Phys. Rev. B **82**, 184502 (2010).
[29] P. W. Anderson, J. Phys. Chem. Solids **11**, 26 (1959).
[30] P. W. Anderson, Phys. Rev. B **30**, 4000 (1984).
[31] G. H. Wannier, Phys. Rev. **52**, 191 (1937).
[32] N. Marzari, A. A. Mostofi, J. R. Yates, I. Souza, and D. Vanderbilt, Rev. Mod. Phys. **84**, 1419 (2012).
[33] S. Bravyi, D. P. DiVincenzo, and D. Loss, Ann. Phys. **326**, 2793 (2011).
[34] A. Mielke, J. Phys. A **24**, L73 (1991).
[35] A. Mielke, J. Phys. A **24**, 3311 (1991).
[36] H. Tasaki, Phys. Rev. Lett. **69**, 1608 (1992).
[37] A. Mielke and H. Tasaki, Commun. Math. Phys. **158**, 341 (1993).
[38] M. Maksymenko, A. Honecker, R. Moessner, J. Richter, and O. Derzhko, Phys. Rev. Lett. **109**, 096404 (2012).
[39] E. H. Lieb, Phys. Rev. Lett. **62**, 1201 (1989).
[40] J. R. Schrieffer, *Theory of Superconductivity*, Advanced Books Classics (Westview Press, Boulder, Colorado, 1999).
[41] N. Read and D. Green, Phys. Rev. B **61**, 10267 (2000).
[42] D. Green, L. Santos, and C. Chamon, Phys. Rev. B **82**, 075104 (2010).
[43] N. Marzari and D. Vanderbilt, Phys. Rev. B **56**, 12847 (1997).
[44] C. Brouder, G. Panati, M. Calandra, C. Mourougane, and N. Marzari, Phys. Rev. Lett. **98**, 046402 (2007).
[45] G. Panati, Ann. Henri Poincaré **8**, 995 (2007).
[46] G. Panati and A. Pisante, Commun. Math. Phys. **322**, 835 (2013).
[47] D. Monaco, G. Panati, A. Pisante, and S. Teufel (private communication).
[48] N. Read, arXiv:1608.04696.
[49] S. R. White, Phys. Rev. Lett. **69**, 2863 (1992).
[50] S. R. White and R. M. Noack, Phys. Rev. Lett. **68**, 3487 (1992).
[51] U. Schollwöck, Rev. Mod. Phys. **77**, 259 (2005).
[52] U. Schollwöck, Ann. Phys. **326**, 96 (2011).

[53] X. G. Wen and A. Zee, Nucl. Phys. B **316**, 641 (1989).
[54] M. Sato and S. Fujimoto, Phys. Rev. B **79**, 094504 (2009).
[55] H. Wolkowicz and G. P. Styan, Linear Algebra Appl. **29**, 471 (1980).
[56] G. W. Winkler, A. A. Soluyanov, and M. Troyer, Phys. Rev. B **93**, 035453 (2016).
[57] A. F. Ho, M. A. Cazalilla, and T. Giamarchi, Phys. Rev. A **79**, 033620 (2009).
[58] T. Koma and H. Tasaki, Phys. Rev. Lett. **68**, 3248 (1992).
[59] A. Mielke, J. Phys. A **32**, 8411 (1999).
[60] A. Mielke, Eur. Phys. J. B **85**, 184 (2012).
[61] J.-P. Brantut, J. Meineke, D. Stadler, S. Krinner, and T. Esslinger, Science **337**, 1069 (2012).
[62] L. Mazza, A. Bermudez, N. Goldman, M. Rizzi, M. A. Martin-Delgado, and M. Lewenstein, New J. Phys. **14**, 015007 (2012).
[63] S. Mugel, A. Celi, P. Massignan, J. K. Asbóth, M. Lewenstein, and C. Lobo, Phys. Rev. A **94**, 023631 (2016).
[64] M. Aidelsburger, M. Atala, M. Lohse, J. T. Barreiro, B. Paredes, and I. Bloch, Phys. Rev. Lett. **111**, 185301 (2013).
[65] H. Miyake, G. A. Siviloglou, C. J. Kennedy, W. C. Burton, and W. Ketterle, Phys. Rev. Lett. **111**, 185302 (2013).
[66] A. Celi, P. Massignan, J. Ruseckas, N. Goldman, I. B. Spielman, G. Juzeliūnas, and M. Lewenstein, Phys. Rev. Lett. **112**, 043001 (2014).
[67] 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).
[68] B. K. Stuhl, H.-I. Lu, L. M. Aycock, D. Genkina, and I. B. Spielman, Science **349**, 1514 (2015).
[69] B. Bauer, L. D. Carr, H. G. Evertz, A. Feiguin, J. Freire, S. Fuchs, L. Gamper, J. Gukelberger, E. Gull, S. Guertler, A. Hehn, R. Igarashi, S. V. Isakov, D. Koop, P. N. Ma, P. Mates, H. Matsuo, O. Parcollet, G. Pawowski, J. D. Picon, L. Pollet, E. Santos, V. W. Scarola, U. Schollwck, C. Silva, B. Surer, S. Todo, S. Trebst, M. Troyer, M. L. Wall, P. Werner, and S. Wessel, J. Stat. Mech. (2011) P05001.
[70] M. Dolfi, B. Bauer, S. Keller, A. Kosenkov, T. Ewart, A. Kantian, T. Giamarchi, and M. Troyer, Comput. Phys. Commun. **185**, 3430 (2014).
[71] V. J. Emery, Phys. Rev. B **14**, 2989 (1976).
[72] W. Kohn, Phys. Rev. **115**, 809 (1959).
[73] A. Nenciu and G. Nenciu, Commun. Math. Phys. **190**, 541 (1998).