

## Universality of Abelian and non-Abelian Wannier functions in generalized one-dimensional Aubry-André-Harper models

Kiryl Piasotski,<sup>1</sup> Mikhail Pletyukhov<sup>1</sup>, Clara S. Weber,<sup>1</sup> Jelena Klinovaja<sup>2</sup>, Dante M. Kennes<sup>1,3,\*</sup> and Herbert Schoeller<sup>1</sup>

<sup>1</sup>*Institut für Theorie der Statistischen Physik, RWTH Aachen University and*

*JARA—Fundamentals of Future Information Technology, 52056 Aachen, Germany*

<sup>2</sup>*Department of Physics, University of Basel, Klingelbergstrasse 82, CH-4056 Basel, Switzerland*

<sup>3</sup>*Max Planck Institute for the Structure and Dynamics of Matter, Center for Free Electron Laser Science (CFEL), Luruper Chaussee 149, 22761 Hamburg, Germany*



(Received 17 May 2021; accepted 3 August 2021; published 19 August 2021)

Within a Dirac model in 1+1 dimensions, a prototypical model to describe low-energy physics for a wide class of lattice models, we propose a field-theoretical version for the representation of Wannier functions, the Zak-Berry connection, and the geometric tensor. In two natural Abelian gauges we present universal scaling of the Dirac Wannier functions in terms of four fundamental scaling functions that depend only on the phase  $\gamma$  of the gap parameter and the charge correlation length  $\xi$  in an insulator. The two gauges allow for a universal low-energy formulation of the surface charge and surface fluctuation theorem, relating the boundary charge and its fluctuations to bulk properties. Our analysis describes the universal aspects of Wannier functions for the wide class of one-dimensional generalized Aubry-André-Harper lattice models. In the low-energy regime of small gaps we demonstrate universal scaling of all lattice Wannier functions and their moments in the corresponding Abelian gauges. In particular, for the quadratic spread of the lattice Wannier function, we find the universal result  $Za\xi/8$ , where  $Za$  is the length of the unit cell. This result solves a long-standing problem providing further evidence that an insulator is only characterized by the two fundamental length scales  $Za$  and  $\xi$ . Finally, we discuss also non-Abelian lattice gauges and find that lattice Wannier functions of maximal localization show universal scaling and are uniquely related to the Dirac Wannier function of the lower band. In addition, via the winding number of the determinant of the non-Abelian transformation, we establish a bulk-boundary correspondence for the number of edge states up to the bottom of a certain band, which requires no symmetry constraints. Our results present evidence that universal aspects of Wannier functions and of the boundary charge are uniquely related and can be elegantly described within universal low-energy theories.

DOI: [10.1103/PhysRevResearch.3.033167](https://doi.org/10.1103/PhysRevResearch.3.033167)

### I. INTRODUCTION

Wannier functions [1] have matured to an invaluable tool in various fields of solid-state physics. They provide a useful basis set of exponentially localized single-particle states [2] integral to density functional theory [3], are used in the definition of tight-binding models with short-ranged hoppings, can characterize the topological properties of crystals [4–6], and are the basis for the modern theory of polarization and localization [7–9]. However, Wannier functions are gauge dependent, begging the following question: Which gauges are particularly useful? One of them has been identified as the gauge of maximally localized Wannier functions [10], which we call the ML gauge in the following. In this gauge the quadratic spread  $\langle \Delta x^2 \rangle^{1/2}$  of the Wannier function is related

to the fluctuations of the bulk polarization and to the quasimomentum integral of the geometric tensor [10,11] quantifying important topological hallmarks of the system. Furthermore, the first moment of the Wannier function (its average) can be related to the Zak-Berry phase or the bulk polarization [9], a quantity widely used to determine Chern numbers in two-dimensional topological systems.

Complementary to the bulk polarization the macroscopic boundary (or surface) charge  $Q_B$  for a half-infinite insulating system has been studied extensively in a series of recent works [12–18]. Interestingly, the boundary charge [19] and its fluctuations [18] can be related to the bulk polarization and its fluctuations by the so-called *surface charge* and *surface fluctuation* theorem. The surface charge theorem has first been discussed in Ref. [19] relating  $Q_B$  to the Zak-Berry phase of all occupied bands, up to an unknown integer, for noninteracting, nondisordered systems. Recently, this theorem was further analyzed in Ref. [15] for the wide class of generalized Aubry-André-Harper models [20–26]. Here even the unknown integer of the surface charge theorem was determined fixing the precise gauge in the Wannier functions needed to relate its first moment to the Zak-Berry phase. In the small gap limit a low-energy Dirac model in 1+1 dimensions [13,17,27] can be used to reveal a universal form of the

\*dante.kennes@mpsd.mpg.de

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/). Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI. Open access publication funded by the Max Planck Society.

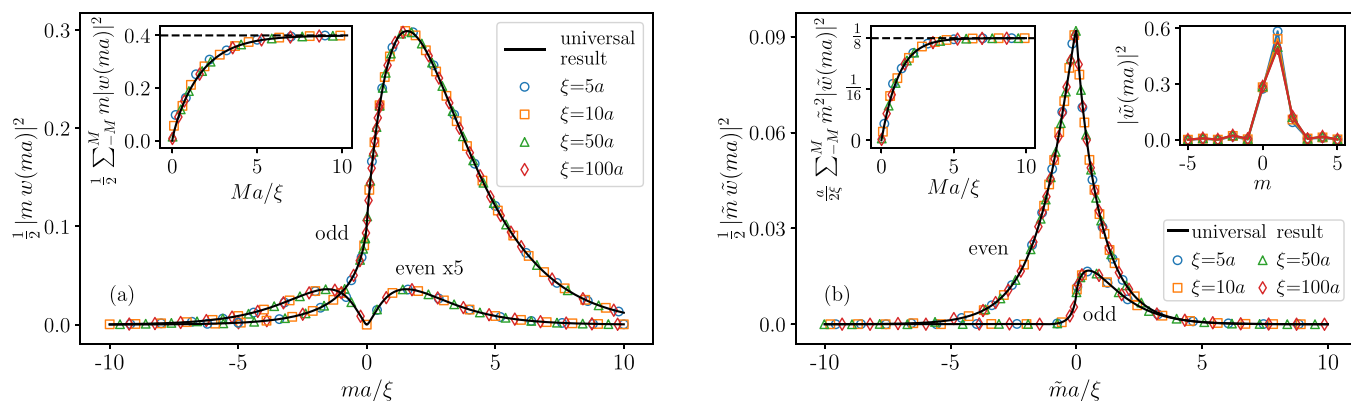


FIG. 1. Universal scaling of the lattice Wannier functions  $w(ma)$  and  $\tilde{w}(ma)$  for the lower band  $\alpha = 1$  of the Rice-Mele model in (a) AF gauge and (b) ML gauge, respectively, for different correlation lengths  $\xi = ta/\Delta$  determined by the gap size  $2\Delta$  and the average hopping  $t$ . The gap parameter is given by  $\Delta e^{i\gamma} = -\delta v - 2i\delta t$ , where  $v_{1,2} = \pm\delta v$  and  $t_{1,2} = t \pm \delta t$  are the alternating onsite energies and hopping amplitudes, respectively. The phase of the gap parameter is chosen as  $\gamma = 0.2\pi$ . Main panels:  $\frac{1}{2}|m w(ma)|^2$  (AF gauge) and  $\frac{1}{2}|\tilde{m} \tilde{w}(ma)|^2$  (ML gauge); with  $\tilde{m} = m - \gamma_1/\pi$  and  $\gamma_1 = -\gamma + \pi \operatorname{sgn}\gamma = 0.8\pi$  denoting the Zak-Berry phase as function of  $ma/\xi$  for different  $\xi = 5a, 10a, 50a, 100a$ . All discrete lattice points fall on top of the fundamental universal scaling functions of the Dirac model,  $|F_{A/B}(ma/\xi)|^2$  (AF gauge) and  $|\tilde{F}_{A/B}(\tilde{m}a/\xi)|^2$  (ML gauge), for  $m$  odd ( $F_A$  and  $\tilde{F}_A$ ) and  $m$  even ( $F_B$  and  $\tilde{F}_B$ ) (black solid lines; the even points in the AF gauge are multiplied by a factor of 5 for visibility). Left inset of (a): the scaling of the first moment  $C_1(Ma)/(2a) = \frac{1}{2} \sum_{m=-M}^M m |w(ma)|^2$  as function of  $Ma/\xi$  for different  $\xi = 5a, 10a, 50a, 100a$ . All lattice points fall on top of a universal scaling function  $\frac{1}{2} \int_{-Ma/\xi}^{Ma/\xi} dy [F_A(y)^2 + F_B(y)^2]/y$  (black solid line). They converge smoothly to the value  $\gamma_1/(2\pi) = C_1/(2a) = 0.4$ . Left inset of (b): scaling of the quadratic spread  $\langle \Delta x^2 \rangle (Ma)/(2a\xi) = \frac{a}{2\xi} \sum_{m=-M}^M \tilde{m}^2 |\tilde{w}(ma)|^2$  as function of  $Ma/\xi$  for different  $\xi = 5a, 10a, 50a, 100a$ . All lattice points fall on top of the universal result  $\int_0^{Ma/\xi} dy [\tilde{F}_A(y)^2 + \tilde{F}_B(y)^2]$  (black solid line) and converge smoothly to the universal value  $\frac{1}{8}$ . Right inset of (b):  $|\tilde{w}(ma)|^2$  for  $\xi = 5a, 10a, 50a, 100a$  shown as function of  $m$ , giving rise to the misleading visual impression of a localized state with  $\xi$ -independent spread of order  $a$ . A similar form appears for  $|w(ma)|^2$  (not shown).

boundary charge in terms of the phase of the gap parameter [17]. Complementing the surface charge theorem, recently the surface fluctuation theorem [18] was established as a relationship between the boundary charge fluctuations and the second moment of the density-density correlation function or, equivalently, the fluctuations of the bulk polarization. The boundary charge is of particular interest since it is a measurable observable, e.g., via density measurements in cold-atom systems [28,29] or via scanning single-electron transistor techniques [30–34] in carbon nanotubes [35], graphene nanoribbons [36], and Rashba nanowires [37,38].

These recent developments raise the important question whether Wannier functions themselves exhibit universal properties beyond their first two moments, in particularly fundamental gauges. To address this question we cover the wide class of generalized Aubry-André-Harper models. We do so by considering a two-band Dirac model, which requires a field-theoretical generalization of Wannier functions, the Zak-Berry connection, and the geometric tensor, which we provide. Any of such field theories must always be complemented by appropriate asymptotic conditions at high energies to provide a well-defined relation to the lattice models falling into the universality class of the respective field theory. In addition to the above-mentioned ML gauge, we define the asymptotically free (AF) gauge, which is fixed by requiring that the eigenfunctions turn into conventional plane waves at large momenta. Here, we show that, for a microscopic lattice model, this gauge corresponds precisely to the one used in Ref. [15] to fix the unknown integer in the surface charge theorem. As a result, the AF gauge is useful to formulate a universal field-theoretical version of the surface charge theorem,

relating the boundary charge to the first moment of the Dirac Wannier function and to the phase  $\gamma$  of the gap parameter, which in turn is related to the field-theoretical version of the Zak-Berry phase. In contrast, we show that the fundamental ML gauge is useful for the universal formulation of the field-theoretical version of the surface fluctuation theorem, relating the boundary charge fluctuations to the second moment of the Wannier function or the momentum integral over the field-theoretical version of the geometric tensor.

Returning to the question of universality of Wannier functions beyond the first and second moments we find that strikingly the entire line shape of Wannier functions exhibits universal behavior in the small gap limit. Surprisingly, we show that for the whole class of generalized Aubry-André-Harper models all lattice Wannier functions display universal scaling on the length scale  $\xi$ , which is the fundamental length scale of insulators on which charge correlations decay exponentially.  $\xi$  is proportional to the inverse gap and is defined here by the exponential decay length of the square of the Wannier functions. We show that all scaling functions can be related to four fundamental functions of the Dirac model which are fully characterized by the phase  $\gamma$  of the gap parameter.

In Figs. 1(a) and 1(b) we illustrate these four fundamental scaling functions, denoted by  $F_{A,B}$  and  $\tilde{F}_{A,B}$  in the following, for the lower-band Wannier functions in the AF and ML gauges of the Rice-Mele model, a fundamental one-dimensional lattice model with a two-site unit cell put forward in the context of topological insulators [39]. In Fig. 1(a) we show that for the AF gauge, all Wannier functions multiplied with the lattice site index  $m \rightarrow ma/\xi$

for different correlation lengths  $\xi$  (here,  $a$  denotes the lattice constant). For the ML gauge, as shown in Fig. 1(b), the same applies but it turns out to be important to include a shift of the lattice site index  $m$  by the first moment of the Wannier function. Surprisingly, the universal scaling holds even outside the strict range of low-energy theories, i.e., it persists for rather large gaps [where  $\xi \sim O(a)$ ] and also for rather small length scales  $|ma| \sim O(a)$ . The inset of Fig. 1(a) and the left inset of Fig. 1(b) show the universality of the surface charge and surface fluctuation theorem discussed above, respectively, which follow from the universality of the Wannier functions. We find that the first moment converges to the universal result  $C_1/(Za) = \gamma_1/(2\pi)$ , where  $Za$  is the length of the unit cell and  $\gamma_1$  is the Zak-Berry phase of the lower band, which is related to the phase of the gap parameter and to the boundary charge. For the quadratic spread in the ML gauge we find the universal result  $\langle \Delta x^2 \rangle / (Za) = \xi/8$ , which is in full agreement with the universal result for the boundary charge fluctuations [18]. Interestingly, we obtain the scaling  $\langle \Delta x^2 \rangle / (Za) \sim O(\xi)$  in any gauge and for any band for arbitrary unit-cell length  $Za$ . This shows that the spread is not an independent length scale but is universally linked to the exponential decay length  $\xi$ . In the small gap limit this solves a long-standing problem since in the previous literature only the inequality  $\langle \Delta x^2 \rangle \lesssim Za\xi$  has been stated [8,11], leaving it open whether the so-called localization length  $\lambda = \langle \Delta x^2 \rangle^{1/2}$  defines an independent length scale [40].

Our analysis shows that universal aspects of Wannier functions and the boundary charge are ultimately related to generic effects arising at band anticrossing points with a small gap defining a corresponding universal length scale clearly separated from the lattice spacing. The reason why universal scaling was easily overlooked before is at least twofold: On the one hand, density functional theory calculations for specific materials are often applied to systems where the lattice spacing  $a$  and the exponential decay length  $\xi$  are not clearly separated. On the other hand, even in cases of clearly separated  $\xi \gg a$ , the visual impression of the square  $|w(ma)|^2$  of the Wannier function does not reveal any universal scaling. As shown in the right inset of Fig. 1(b), the visual impression is that of a rather boring Lorentzian with broadening  $\sim a$ . This form is only important for the correct normalization of the Wannier function but it does not reveal any universal scaling on length scale  $\xi$ . Universal scaling is only visible when multiplying the Wannier function by the spatial coordinate and plotting the square  $|m w(ma)|^2$ , such that the asymptotic  $(1/m)^2$  behavior of the Lorentzian is canceled. As a consequence, it is roughly the product of a Lorentzian on scale  $a$  and a universal scaling function on scale  $\xi$  which characterizes the subtle and universal line shape of Wannier functions revealed in this work.

Furthermore, we show that the fully universal scaling exemplified for the Rice-Mele model persists for the Wannier functions of all bands for any size  $Z$  of the unit cell. It turns out that the lattice Wannier functions of a given band with two gaps at the bottom and top of the band show universal scaling in terms of two different correlation length scales referring to these two energetically adjacent gaps. Additionally, we show that each Wannier function can be naturally split in two additive contributions, one corresponding to the upper and one

to the lower half of the band, each of them scaling only with a single length scale. Moreover, we also discuss a non-Abelian gauge of maximal localization (NA-ML) mixing the Wannier functions of all bands including and below a given one, which has been proposed in Ref. [10] to maximally localize the total sum of the quadratic spreads of a certain set of Wannier functions. Interestingly, we find that all Wannier functions in the NA-ML gauge reveal universal scaling on the same length scale  $\xi$  referring to the band gap between that given band and the next. Additionally, these Wannier functions can be uniquely related to the lower band Wannier function of the Dirac model. Furthermore, we will present an explicit construction of the NA-ML gauge in terms of the Wilson propagator and show that the winding number associated with the unitary non-Abelian gauge transformation allows for a formulation of a bulk-boundary correspondence (BBC) determining the number of edge states up to the band chosen to define the NA-ML gauge via bulk properties. This form of the BBC is of particular interest since it is not restricted to zero-energy edge states or to particular symmetry properties of the system. We would like to additionally emphasize that this formulation of the BBC has nothing to do with a relation of the boundary charge to some winding number. In general case the boundary charge (as the Zak-Berry phase) can take any value and is only quantized in the case of special symmetries or filling factors.

Our work is organized as follows. In Sec. II we start with the description of the lattice models under consideration and the relation of their eigenstates to the eigenstates of the Dirac model. The low-energy theory of the Zak-Berry connection, the Zak-Berry phase, and the geometric tensor, together with their relation to the corresponding lattice quantities is the subject of Sec. III. In this section we will also define the AF and ML gauges and formulate the low-energy version of the surface charge and surface fluctuation theorem in terms of the low-energy version of the Zak-Berry phase and the momentum integral of the geometric tensor, respectively. The definition of Wannier functions in Dirac theory and their precise relation to the lattice Wannier functions will be analyzed in Sec. IV. Furthermore, we will describe in this section the relation of all moments of the Wannier functions in Dirac and lattice theory and explain how the surface charge and surface fluctuation theorem can be formulated in terms of the moments of the Dirac Wannier functions. The universal scaling of the Dirac and lattice Wannier functions is the central subject of Sec. V. We will calculate the Dirac Wannier functions together with their moments analytically in the AF and ML gauges, and state the definitions and properties of the fundamental scaling functions. The universal scaling of all lattice Wannier functions and their moments is then stated via their relation to the Dirac Wannier functions, and is demonstrated for two explicit examples  $Z = 2$  and  $3$ . Finally, in Sec. VI we construct explicitly the NA-ML gauge and demonstrate that all non-Abelian Wannier functions show universal scaling corresponding to the Dirac theory of the highest gap defining the set of mixed bands. In addition, we set up an interesting bulk-boundary correspondence relating the winding number of the determinant of the non-Abelian gauge transformation to the number of edge states present in all gaps up to a certain one. We close with a summary and outlook in Sec. VII. Some

more involved technical details are presented in a series of appendices.

We use units  $e = \hbar = 1$ .

## II. MODEL AND EIGENSTATES

In this section we will state the class of lattice models under consideration together with their relation to the Dirac field theory. In particular, we will show the representation of the Bloch eigenstates in lattice theory and the eigenfunctions of the Dirac Hamiltonian, together with their precise relationship. In summary it will turn out that the Dirac field theory is an elegant way for a formulation of degenerate perturbation theory in terms of slowly varying right and left movers, applicable to lattice models with arbitrary size of the unit cell.

### A. Lattice model

We consider the following class of so-called generalized Aubry-André-Harper lattice models in one dimension, characterized by nearest-neighbor hopping and a single orbital per site, together with any periodic onsite and hopping modulation:

$$H = H_0 + H', \quad (2.1)$$

$$H_0 = -t \sum_m (|m+1\rangle\langle m| + \text{H.c.}), \quad (2.2)$$

$$H' = \sum_m \delta v_m |m\rangle\langle m| - \sum_m \delta t_m (|m+1\rangle\langle m| + \text{H.c.}). \quad (2.3)$$

Here,  $m$  is the lattice site index,  $t > 0$ , and all  $\delta v_m = \delta v_{m+Z}$ ,  $\delta t_m = \delta t_{m+Z}$  are periodic and real (possible phases of the hoppings can always be gauged away in one dimension). The number of sites in a unit cell is denoted by  $Z$ . We relate  $m = Z(n-1) + j$  to the index  $n$  labeling the unit cells and the index  $j = 1, \dots, Z$  labeling the sites within a unit cell.

We assume the condition

$$|\delta v_m|, |\delta t_m| \ll t, \quad (2.4)$$

such that  $H'$  can be considered as a small perturbation. Due to the presence of  $H'$  there will be  $Z$  bands labeled by the band index  $\alpha = 1, \dots, Z$  (from bottom to top), with  $Z-1$  gaps in-between labeled by the gap index  $\nu = 1, \dots, Z-1$  (see the sketch of the band structure shown in Fig. 2).

In standard solid-state notation of the reduced zone scheme, the Bloch eigenfunctions on the lattice are written as

$$\psi_{k\alpha}(ma) = \sqrt{\frac{Za}{2\pi}} u_{k\alpha}(ma) e^{ikma}, \quad (2.5)$$

where  $\alpha = 1, \dots, Z$  is the band index (labeled from bottom to top),  $ma$  denotes the lattice space position, and  $a$  is the lattice spacing. The length  $Za$  defines the size of the unit cell. The quasimomentum  $k$  is defined within the first Brillouin zone  $-\frac{\pi}{Za} < k < \frac{\pi}{Za}$  and

$$u_{k\alpha}(ma) = u_{k\alpha}(ma + Za) \quad (2.6)$$

is periodic with the unit-cell size. We assume that the gauge is always chosen such that the total Bloch function is periodic

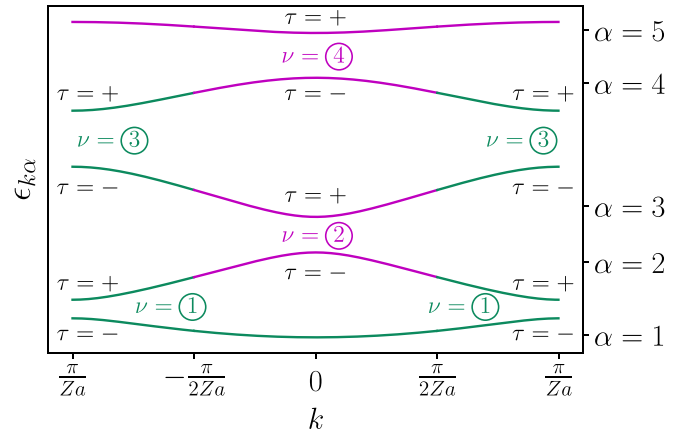


FIG. 2. Sketch of the band structure for  $Z = 5$ . We find  $Z - 1 = 4$  gaps labeled by  $\nu = 1, \dots, 4$  opening up at  $k = 0$  for  $\nu = 2, 4$  even, and at  $k = \pm \frac{\pi}{Za}$  for  $\nu = 1, 3$  odd. All bands labeled by  $\alpha = 1, \dots, 5$  are then naturally split in two halves, the upper and lower halves labeled by  $(\nu, \tau) = (\alpha - 1, +)$ , respectively. An exception are the lowest  $\alpha = 1$  and highest  $\alpha = 5$  bands. They are not split in two halves and labeled by  $(\nu, \tau) = (1, -)$  and  $(\nu, \tau) = (4, +)$ . The Dirac theory for gap  $\nu$  is a good description for all eigenstates referring to a pair of band parts labeled by  $(\nu, \pm)$  (indicated by violet and green color for  $\nu$  even or odd, respectively). The momentum  $q$  in Dirac theory is the difference of  $k$  to the point where the gap opens, i.e.,  $k = q$  for  $\nu$  even, and  $k = \pm \frac{\pi}{Za} + q$  for  $\nu$  odd and  $\text{sgn}q = \mp$ . The cutoff for  $|q|$  is given by  $\Lambda_\alpha = \frac{\pi}{2Za}$  for  $\alpha = 2, 3, 4$  and by  $\Lambda_\alpha = \frac{\pi}{Za}$  for  $\alpha = 1, 5$ .

in  $k$ :

$$\psi_{k\alpha}(ma) = \psi_{k+\frac{2\pi}{Za}, \alpha}(ma). \quad (2.7)$$

Note that this means that  $u_{k\alpha}(ma)$  is not periodic in  $k$  but fulfills the condition

$$u_{k+\frac{2\pi}{Za}, \alpha}(ma) = u_{k\alpha}(ma) e^{-i\frac{2\pi}{Z}m}. \quad (2.8)$$

At fixed  $k$  in the first Brillouin zone, the states  $u_{k\alpha}(ja) = \langle j | u_{k\alpha} \rangle$  in unit-cell space  $j = 1, \dots, Z$  are eigenfunctions of a Hamiltonian  $h_k$  defined on the periodic continuation of the unit cell

$$h_k |u_{k\alpha}\rangle = \epsilon_{k\alpha} |u_{k\alpha}\rangle, \quad (2.9)$$

$$h_k = \sum_{j=1}^Z \delta v_j |j\rangle\langle j| - \sum_{j=1}^Z t_j (|j+1\rangle\langle j| e^{-ika} + |j\rangle\langle j+1| e^{ika}), \quad (2.10)$$

where  $t_j = t + \delta t_j$ , and we identify  $|Z+1\rangle \equiv |1\rangle$ . In contrast to  $h_k$  and  $|u_{k\alpha}\rangle$ , the dispersion relation is periodic in  $k$ :  $\epsilon_{k+\frac{2\pi}{Za}, \alpha} = \epsilon_{k\alpha}$ .

We note the following normalization and completeness relations:

$$\sum_{\alpha} \psi_{k\alpha}(ma)^* \psi_{k'\alpha'}(ma) = \delta_{\alpha\alpha'} \delta(k - k'), \quad (2.11)$$

$$\sum_{\alpha} \int_{-\pi/Za}^{\pi/Za} dk \psi_{k\alpha}(ma) \psi_{k\alpha}(m'a)^* = \delta_{mm'}. \quad (2.12)$$

This corresponds to the following relations for the periodic Bloch part  $u_{k\alpha}(ja)$  at fixed quasimomentum  $k$  and  $j = 1, \dots, Z$ :

$$\sum_{j=1}^Z u_{k\alpha}(ja)^* u_{k\alpha'}(ja) = \delta_{\alpha\alpha'}, \quad (2.13)$$

$$\sum_{\alpha} u_{k\alpha}(ja) u_{k\alpha}(j'a)^* = \delta_{jj'}. \quad (2.14)$$

Introducing the following scalar product in unit-cell space

$$\langle u_{k\alpha} | u_{k\alpha'} \rangle = \sum_{j=1}^Z u_{k\alpha}(ja)^* u_{k\alpha'}(ja), \quad (2.15)$$

we can write (2.13) and (2.14) in compact form as

$$\langle u_{k\alpha} | u_{k\alpha'} \rangle = \delta_{\alpha\alpha'}, \quad (2.16)$$

$$\sum_{\alpha} |u_{k\alpha}\rangle \langle u_{k\alpha}| = \mathbb{1}_Z. \quad (2.17)$$

### B. Dirac model

For a small perturbation  $H'$  compared to the typical bandwidth set by the average hopping  $t$ , the states in the middle of the bands are nearly untouched and only close to the gaps the eigenstates of  $H_0$  are coupled significantly by  $H'$ . We denote the single-particle eigenstates of  $H_0$  in the extended zone scheme as

$$\langle m | k \rangle = \sqrt{\frac{a}{2\pi}} e^{ikma}, \quad (2.18)$$

with energy  $\epsilon_k = -2t \cos(ka)$  and  $-\pi/a < k < \pi/a$ . Since  $H'$  is periodic with the size  $Za$  of the unit cell, the gap openings will happen in general via higher-order processes close to the two Fermi points  $\pm k_F^{(v)}$  where

$$k_F^{(v)} = \frac{\pi v}{Za}, \quad v = 1, \dots, Z-1. \quad (2.19)$$

As outlined in detail in Refs. [13,17,27] the coupling of the low-energy states close to gap  $v$  can be described via Brillouin-Wigner perturbation theory by the effective Hamiltonian

$$H_{\text{eff}}^{(v)} = P^{(v)} \left( H + H' Q^{(v)} \frac{1}{\epsilon_F^{(v)} - Q^{(v)} H Q^{(v)}} Q^{(v)} H' \right) P^{(v)}, \quad (2.20)$$

where  $P^{(v)}$  projects on the low-energy space,  $Q^{(v)} = 1 - P^{(v)}$ , and  $\epsilon_F^{(v)} = -2t \cos(k_F^{(v)} a)$  is the Fermi energy where the gap opens. This coupling leads to a complex gap parameter  $\Delta^{(v)} e^{i\gamma^{(v)}}$  defined by

$$\langle k_F^{(v)} + k | H_{\text{eff}}^{(v)} | -k_F^{(v)} + k' \rangle \approx \Delta^{(v)} e^{i\gamma^{(v)}} \delta(k - k'). \quad (2.21)$$

In standard convention of low-energy theories it is very useful to split off the strongly oscillating parts  $e^{\pm ik_F^{(v)} ma}$  of the eigenfunctions (2.5) and parametrize the Bloch states  $\psi_{k\alpha}(ma)$  in terms of slowly varying right and left movers  $\psi_{q\tau p}^{(v)}(x)$  in continuum notation as

$$\psi_{k\alpha}(ma) = \sqrt{a} \psi_{q\tau}^{(v)}(ma), \quad (2.22)$$

$$\psi_{q\tau}^{(v)}(x) = \sum_{p=\pm} \psi_{q\tau p}^{(v)}(x) e^{ipk_F^{(v)} x}, \quad (2.23)$$

$$\psi_{q\tau p}^{(v)}(x) = \chi_{q\tau p}^{(v)} \frac{1}{\sqrt{2\pi}} e^{iqx}. \quad (2.24)$$

Here,  $p = \pm$  is the index for right and left movers, and  $\tau = \pm$  is the index describing the bands above and below the gap  $v$  (see Fig. 2):

$$\tau = \begin{cases} + & \text{for } \alpha = v + 1, \\ - & \text{for } \alpha = v. \end{cases} \quad (2.25)$$

Thereby, the band part labeled by  $(v, \tau)$  corresponds to the lower (upper) half of band  $\alpha = v + 1$  ( $\alpha = v$ ) for  $\tau = \pm$  (see Fig. 2). An exception are the lowest  $\alpha = 1$  and highest  $\alpha = Z$  bands which are not split in two halves and labeled by  $(v, \tau) = (1, -)$  or  $(v, \tau) = (Z - 1, +)$ , respectively.

The momentum  $q$  in Eq. (2.22) corresponds to the difference of the quasimomentum  $k$  in lattice theory to the quasimomentum defining the band bottom or band top at which the gap appears. This means that for even  $v$  one expands around  $k = 0$  such that  $q \equiv k$ , and for odd  $v$  one expands around  $k = \pm \frac{\pi}{Za}$  such that  $q \equiv k \mp \frac{\pi}{Za}$  (see Fig. 2). In both cases the momentum  $q$  has to fulfill the condition

$$|q| < \Lambda_{\alpha} = \begin{cases} \frac{\pi}{2Za} & \text{for } \alpha = 2, \dots, Z-1, \\ \frac{\pi}{Za} & \text{for } \alpha = 1, Z. \end{cases} \quad (2.26)$$

By linearizing the dispersion relation  $\epsilon_k$  of  $H_0$  around  $\pm k_F^{(v)}$ , and using (2.21) one can set up an effective Dirac Hamiltonian  $H_D^{(v)}$  corresponding to gap  $v = 1, \dots, Z-1$  with gap parameter  $\Delta^{(v)} e^{i\gamma^{(v)}}$ , which has the vectors  $\underline{\psi}_{q\tau}^{(v)}(x) = (\psi_{q\tau+}^{(v)}(x), \psi_{q\tau-}^{(v)}(x))^T$  as eigenstates

$$H_D^{(v)} = -iv_F^{(v)} \partial_x \sigma_z + \Delta^{(v)} \cos(\gamma^{(v)}) \sigma_x - \Delta^{(v)} \sin(\gamma^{(v)}) \sigma_y, \quad (2.27)$$

where  $v_F^{(v)} = 2ta \sin(k_F^{(v)} a)$  is the Fermi velocity. The energy of the Dirac eigenstates is given by

$$\epsilon_{q\tau}^{(v)} = \tau \epsilon_q^{(v)}, \quad \epsilon_q^{(v)} = \sqrt{(v_F^{(v)} q)^2 + (\Delta^{(v)})^2}, \quad (2.28)$$

which provides a very good approximation to the true dispersion when  $|q| \ll \frac{\pi}{2Za}$ .

In contrast to the dispersion, we note that the Dirac model provides a very good approximation to the exact eigenstates for all  $|q| < \Lambda_{\alpha}$ . For large  $|q| \sim 1/(Za) \gg \Delta^{(v)}/v_F^{(v)}$ , the eigenstates of the Dirac model are (up to a gauge factor) given by

$$\chi_{q\tau p}^{(v)} \approx \delta_{p, \text{sgn}(q\tau)}. \quad (2.29)$$

Inserting this form in (2.22)–(2.24), one obtains the correct eigenfunctions of the lattice model in the absence of a gap (see Appendix A). This means that for small gaps  $\Delta^{(v)} \ll t$ , the Dirac theory is a useful approximation for the description of all eigenstates of the lattice model. Therefore, all Wannier functions can be fully described by Dirac theory on all length scales. In contrast, dynamical quantities like Green's and correlation functions can only be described for low energies since the energy dispersion enters into these quantities.

When comparing the Bloch eigenstates (2.5) with the parametrization (2.22)–(2.24) one finds straightforwardly the following relations between  $u_{k\alpha}(ja)$  and  $\chi_{q\tau p}^{(v)}$  [note that  $\alpha$  and  $(\nu, \tau)$  are related via (2.25)]:

$$\begin{aligned} \nu \text{ even : } & k = q \\ u_{k\alpha}(ja) &= \frac{1}{\sqrt{Z}} \sum_{p=\pm} \chi_{q\tau p}^{(v)} e^{i\pi p \frac{\nu}{Z} j}, \end{aligned} \quad (2.30)$$

$$\begin{aligned} \nu \text{ odd : } & k = \pm \frac{\pi}{Za} + q, \quad \text{sgn}(q) = \mp \\ u_{k\alpha}(ja) &= \frac{1}{\sqrt{Z}} \sum_{p=\pm} \chi_{q\tau p}^{(v)} e^{i\pi p \frac{\nu}{Z} j} e^{\mp i \frac{\pi}{Z} j}. \end{aligned} \quad (2.31)$$

We note that both conditions (2.6) and (2.8) are respected by (2.30) and (2.31). For the special case  $j = Z$  we get for arbitrary  $\nu$

$$u_{k\alpha}(Za) = \frac{1}{\sqrt{Z}} \sum_{p=\pm} \chi_{q\tau p}^{(v)}. \quad (2.32)$$

Using this relation one can easily check whether the two local gauges of lattice and Dirac theory have been chosen identical.

Since the Dirac theory explicitly exhibits the slowly varying parts  $\psi_{q\tau p}^{(v)}(x)$ , one can interpret it as a continuum theory for all  $x$  on the real axis and not only for  $x = ma$ , where the lattice theory is reproduced. Therefore, although the Dirac theory has only a physical meaning for  $|q| < \Lambda_\alpha$ , one takes all values for the momentum  $q$  into account and writes the normalization and completeness relations as

$$\sum_p \int dx \psi_{q\tau p}^{(v)}(x)^* \psi_{q'\tau' p}^{(v)}(x) = \delta_{\tau\tau'} \delta(q - q'), \quad (2.33)$$

$$\sum_\tau \int dq \psi_{q\tau p}^{(v)}(x) \psi_{q\tau p}^{(v)}(x')^* = \delta_{pp'} \delta(x - x'). \quad (2.34)$$

This is equivalent to taking the following relations for the parts  $\chi_{q\tau p}^{(v)}$  at fixed  $q$ :

$$\sum_p (\chi_{q\tau p}^{(v)})^* \chi_{q\tau' p}^{(v)} = \delta_{\tau\tau'}, \quad (2.35)$$

$$\sum_\tau \chi_{q\tau p}^{(v)} (\chi_{q\tau' p}^{(v)})^* = \delta_{pp'}. \quad (2.36)$$

Introducing the following scalar product in right and left space

$$\langle \chi_{q\tau}^{(v)} | \chi_{q\tau'}^{(v)} \rangle = \sum_{p=\pm} (\chi_{q\tau p}^{(v)})^* \chi_{q\tau' p}^{(v)}, \quad (2.37)$$

we can write (2.35) and (2.36) in compact form as

$$\langle \chi_{q\tau}^{(v)} | \chi_{q\tau'}^{(v)} \rangle = \delta_{\tau\tau'}, \quad (2.38)$$

$$\sum_{\tau=\pm} |\chi_{q\tau}^{(v)} \rangle \langle \chi_{q\tau}^{(v)}| = \mathbb{1}_2. \quad (2.39)$$

It can be checked that the relations (2.35) and (2.36) for  $\chi_{q\tau p}^{(v)}$ , together with (2.30) and (2.31), are consistent with the normalization condition (2.13) and the completeness relation (2.14) for  $u_{k\alpha}(ja)$  (see Appendix A). The normalization condition follows exactly from (2.30) and (2.31), and is a special case of the useful relation (valid for any integers  $r, s \geq 0$ )

$$\langle \partial_k^r u_{k\alpha} | \partial_k^s u_{k\alpha'} \rangle = \delta_{\nu\nu'} \langle \partial_k^r \chi_{q\tau}^{(v)} | \partial_k^s \chi_{q\tau'}^{(v)} \rangle. \quad (2.40)$$

The completeness relation is more subtle and uses in addition the fact that the Dirac theory can reproduce all eigenstates of the original lattice model under the perturbative condition (2.4).

The fact that the Dirac theory can capture all eigenstates of the lattice model leads to the following rigorous rule when considering any  $k$  integral over a function  $f_{k\alpha}$  depending on the eigenfunctions of band  $\alpha$  of the lattice theory (like projectors on Bloch states for the completeness relation, Zak-Berry connection for the Zak-Berry phase, geometric tensor for boundary charge fluctuations, or Bloch states for Wannier functions):

$$\int_{|k| < \pi/Za} dk f_{k\alpha} \approx \int_{|q| < \Lambda_\alpha} dq (f_{q-}^{(\alpha)} + f_{q+}^{(\alpha-1)}), \quad (2.41)$$

where  $f_{q\tau}^{(v)}$  is the corresponding function in Dirac theory obtained by the replacements (2.22), (2.30), or (2.31). Thereby, for  $\alpha = 1$  or  $Z$ , we omit implicitly the terms with  $\nu = 0$  or  $Z$  on the right-hand side, respectively. The approximate sign in this relation is meant in a perturbative sense that all higher-order corrections are of relative order  $O(\Delta^{(v)}/t)$  and negligible for  $\Delta^{(v)} \ll t$ . However, we note that (2.41) is only valid when the gauges in lattice and Dirac theory have been chosen identical such that the relation (2.22) between the eigenfunctions holds globally for all quasimomenta in a certain band  $\alpha$ .

Since a field theory does not know anything about scales of the order of the lattice spacing, the integrals over  $q$  on the right-hand side of (2.41) are conveniently extended to infinity in the field-theoretical version and properly regularized in case of divergences. Since the regime of large momenta refers to the physics in the absence of a gap their contribution can be analytically analyzed quite easily. For the various physical quantities discussed in this work we will show that the field-theoretical contributions beyond the cutoff either vanish for each individual term on the right-hand side of (2.41) (as for Zak-Berry connection and geometric tensor) or, after a proper regularization, the contributions of high momenta cancel between the two terms (as for Wannier functions in certain gauges). In this way a full equivalence between lattice and Dirac theory can be set up, provided that the gaps are small compared to the typical bandwidth. Furthermore, for the bands  $\alpha = 2, \dots, Z$  the two terms on the right-hand side will be shown to refer to the description of the universal behavior of the two halves of the bands, the upper one corresponding to the Dirac theory for  $\nu = \alpha$  and the lower one to  $\nu = \alpha - 1$ . However, for this interpretation to make sense we will see that nonuniversal terms arising from cutting the spectrum in the middle of the band have to be removed by extending the lattice theory for the two halves to infinity via an asymptotically free theory far away from the gap.

When eigenfunctions are compared between the lattice and Dirac theories and identified via (2.22) one always has to guarantee that the gauges are chosen in precisely the same way (at least locally for a certain region in quasimomentum space). Therefore, it will turn out to be important to state specific gauges of particular interest. One of them is the so-called asymptotically free (AF) gauge, which is characterized by a

real and positive value of  $u_{k\alpha}(Za)$ :

$$\text{AF gauge: } u_{k\alpha}(Za) > 0 \Leftrightarrow \sum_p \chi_{q\tau p}^{(\nu)} > 0. \quad (2.42)$$

Here we used (2.32) to formulate the equivalent condition in Dirac theory. This gauge has the particular advantage that it leads precisely to the asymptotic condition (2.29) without any further gauge factors. As explained later in Sec. III C 1, the AF gauge leads to a unique relation between the Zak-Berry phase and the boundary charge both in lattice and Dirac theories. Another gauge will be introduced in Secs. III A and III B which is called the maximally localized (ML) gauge. In the ML gauge the Wannier functions in lattice and Dirac theories have minimal spread which can be related to the boundary charge fluctuations. It will turn out that the ML gauges in lattice and Dirac theories are *not* the same, such that the relation (2.22) involves additional phase factors depending on the quasimomentum. Finally, in Sec. VI we will also discuss non-Abelian lattice gauges where a mixing of the Bloch states from a set of bands  $\alpha = 1, \dots, \nu$  is involved, such that the non-Abelian Wannier functions turn out to have maximal localization. This gauge is called the non-Abelian gauge of maximal localization (NA-ML). It is a special lattice gauge which turns out to be constructed in such a way that the non-Abelian Wannier functions can be directly related to the Wannier function of the upper half of the highest band  $\nu$  or, equivalently, to the Wannier function of the lower band of the Dirac model corresponding to the gap between the highest band  $\nu$  and band  $\nu + 1$ .

### III. ZAK-BERRY CONNECTION, GEOMETRIC TENSOR, AND BOUNDARY CHARGE

As a prerequisite for the analysis of universal aspects of Wannier functions, we develop in this section the low-energy theory for the Zak-Berry connection, the Zak-Berry phase, and the geometric tensor in Dirac theory, together with the relation to the corresponding objects in lattice theory. In this connection, we will also introduce the definition of the ML gauge in Dirac theory and relate it to the corresponding definition in lattice theory. Finally, we will show how an important physical observable, the boundary charge and its fluctuations, can be related via the surface charge and surface fluctuation theorem to the Zak-Berry phase and the momentum integral of the geometric tensor in low energy.

#### A. Zak-Berry connection and geometric tensor for lattice model

For the lattice the Zak-Berry connection and geometric tensor are defined by

$$(A_k)_{\alpha\beta} = (A_k)_{\beta\alpha}^* = i\langle u_{k\alpha} | \partial_k u_{k\beta} \rangle, \quad (3.1)$$

$$(\mathcal{Q}_k)_{\alpha\beta} = (\mathcal{Q}_k)_{\beta\alpha} \quad (3.2)$$

$$= \langle \partial_k u_{k\alpha} | \partial_k u_{k\beta} \rangle \delta_{\alpha\beta} - |\langle u_{k\alpha} | \partial_k u_{k\beta} \rangle|^2 \quad (3.3)$$

$$= \sum_{\alpha' \neq \alpha} |\langle u_{k\alpha'} | \partial_k u_{k\alpha} \rangle|^2 (\delta_{\alpha\beta} - \delta_{\alpha'\beta}), \quad (3.4)$$

where we made use of the normalization and completeness relations (2.16) and (2.17) to derive the last equality. We note that the diagonal component  $\mathcal{Q}_{k\alpha} \equiv (\mathcal{Q}_k)_{\alpha\alpha}$  of the geometric

tensor can also be written as

$$\mathcal{Q}_{k\alpha} = \sum_{\alpha' \neq \alpha} |\langle u_{k\alpha'} | \partial_k u_{k\alpha} \rangle|^2 \quad (3.5)$$

$$= \partial_k \langle u_{k\alpha} | \partial_k u_{k\alpha} \rangle + \langle u_{k\alpha} | (i\partial_k)^2 | u_{k\alpha} \rangle - \langle u_{k\alpha} | i\partial_k | u_{k\alpha} \rangle^2. \quad (3.6)$$

Taking a gauge transformation

$$\tilde{u}_{k\alpha} = e^{i\varphi_{k\alpha}} u_{k\alpha}, \quad (3.7)$$

we find from (3.3) that the geometric tensor is gauge invariant and from (3.1) that the Zak-Berry connection is gauge invariant for  $\alpha \neq \beta$ , whereas  $A_{k\alpha} \equiv (A_k)_{\alpha\alpha}$  transforms as

$$\tilde{A}_{k\alpha} = A_{k\alpha} - \partial_k \varphi_{k\alpha}. \quad (3.8)$$

Furthermore,

$$\gamma_\alpha = \int_{-\pi/Za}^{\pi/Za} dk A_{k\alpha} \quad (3.9)$$

denotes the Zak-Berry phase (which should be distinguished from the notation  $\gamma^{(\nu)}$  for the phase of the gap parameter in Dirac theory). According to (3.8), the Zak-Berry phase transforms under a gauge transformation by the winding number of the phase  $\varphi_{k\alpha}$ :

$$\tilde{\gamma}_\alpha = \gamma_\alpha - \varphi_{\pi/Za,\alpha} + \varphi_{-\pi/Za,\alpha}. \quad (3.10)$$

In the AF gauge defined by (2.42) it is shown in Ref. [15] that

$$u_{-k,\alpha} = (u_{k\alpha})^* \Rightarrow A_{-k,\alpha} = A_{k\alpha}. \quad (3.11)$$

The ML gauge is defined by a constant Zak-Berry connection

$$\text{Lattice ML gauge: } \tilde{A}_{k\alpha} = \frac{Za}{2\pi} \tilde{\gamma}_\alpha. \quad (3.12)$$

According to (3.8) this can be achieved by the gauge factor  $e^{i\varphi_{k\alpha}}$  with

$$\varphi_{k\alpha} = \int_{-\pi/Za}^k dk' A_{k'\alpha} - k \frac{Za}{2\pi} \gamma_\alpha - \frac{1}{2} \gamma_\alpha, \quad (3.13)$$

which fulfils

$$\varphi_{k\alpha} = \varphi_{k+2\pi/Za,\alpha}, \quad (3.14)$$

$$\varphi_{\pm\pi/Za,\alpha} = 0, \quad (3.15)$$

$$\varphi_{-k,\alpha} = -\varphi_{k\alpha}, \quad (3.16)$$

where we used (3.11) for the last equality. From (3.15) we find that the winding number is zero leaving the Zak-Berry phase invariant

$$\tilde{\gamma}_\alpha = \gamma_\alpha. \quad (3.17)$$

Since the AF and ML gauges are the relevant gauges used in this work, we will implicitly indicate the AF gauge by symbols without a tilde and the ML gauge by a tilde symbol.

**B. Zak-Berry connection and geometric tensor for Dirac model**

In the Dirac theory we define the Zak-Berry connection and the geometric tensor by

$$(A_q^{(v)})_{\tau\tau'} = i\langle\chi_{q\tau}^{(v)}|\partial_q\chi_{q\tau'}^{(v)}\rangle, \tag{3.18}$$

$$(Q_q^{(v)})_{\tau\tau'} = \langle\partial_q\chi_{q\tau}^{(v)}|\partial_q\chi_{q\tau'}^{(v)}\rangle\delta_{\tau\tau'} - |\langle\chi_{q\tau}^{(v)}|\partial_q\chi_{q\tau'}^{(v)}\rangle|^2 \tag{3.19}$$

$$= \text{sgn}(\tau\tau')|\langle\chi_{q\bar{\tau}}^{(v)}|\partial_q\chi_{q\tau}^{(v)}\rangle|^2, \tag{3.20}$$

where we defined  $\bar{\tau} = -\tau$  and used  $\delta_{\tau\tau'} - \delta_{\bar{\tau}\tau'} = \text{sgn}(\tau\tau')$  to get the last equation. This gives the important property

$$\sum_{\tau=\pm} (Q_q^{(v)})_{\tau\tau} = \sum_{\tau'=\pm} (Q_q^{(v)})_{\tau\tau'} = 0. \tag{3.21}$$

Analog to (3.6), the diagonal component  $Q_{q\tau}^{(v)} \equiv (Q_q^{(v)})_{\tau\tau}$  can also be written as

$$Q_{q\tau}^{(v)} = \partial_q\langle\chi_{q\tau}^{(v)}|\partial_q\chi_{q\tau}^{(v)}\rangle + \langle\chi_{q\tau}^{(v)}|(i\partial_q)^2|\chi_{q\tau}^{(v)}\rangle - \langle\chi_{q\tau}^{(v)}|i\partial_q|\chi_{q\tau}^{(v)}\rangle^2. \tag{3.22}$$

Taking a gauge transformation

$$\tilde{\chi}_{q\tau p}^{(v)} = e^{i\phi_{q\tau}^{(v)}}\chi_{q\tau p}^{(v)}, \tag{3.23}$$

we find analog to the lattice that the geometric tensor and the nondiagonal components of the Zak-Berry connection are gauge invariant, whereas  $A_{q\tau}^{(v)} = (A_q^{(v)})_{\tau\tau}$  transforms as

$$\tilde{A}_{q\tau}^{(v)} = A_{q\tau}^{(v)} - \partial_q\phi_{q\tau}^{(v)}. \tag{3.24}$$

In the following we allow only for gauges where the Zak-Berry connection vanishes for  $|q| \rightarrow \infty$ ,

$$\lim_{|q|\rightarrow\infty} A_{q\tau}^{(v)} = 0, \tag{3.25}$$

such that the Zak-Berry phase in the Dirac theory, defined by

$$\gamma_\tau^{(v)} = \int dq A_{q\tau}^{(v)}, \tag{3.26}$$

is a well-defined quantity. We note that this is fulfilled for the AF gauge (2.42) where we get from (2.29) for momenta beyond the cutoff  $\Lambda_c$

$$\chi_{q\tau p}^{(v)} \approx \begin{cases} \delta_{\tau p} & \text{for } q > \Lambda_c, \\ \delta_{p,-\tau} & \text{for } q < -\Lambda_c \end{cases} \tag{3.27}$$

leading to a vanishing Zak-Berry connection for  $|q| > \Lambda_c$ .

To get the relation for the Zak-Berry connection and the geometric tensor between the lattice and Dirac definition, we use the identity (2.40) and find

$$(A_k)_{\alpha\alpha'} = \delta_{\nu\nu'}(A_q^{(v)})_{\tau\tau'}, \tag{3.28}$$

$$(Q_k)_{\alpha\alpha'} = \delta_{\nu\nu'}(Q_q^{(v)})_{\tau\tau'}. \tag{3.29}$$

Since the geometric tensor is gauge invariant in both the lattice and Dirac theories, Eq. (3.29) holds independent of the gauge choice in lattice and Dirac theories (they can be even different). The relation (3.28) for the Zak-Berry connection holds within any choice for the local gauge such that the relations (2.30) and (2.31) hold between the eigenfunctions.

Concerning the Zak-Berry phase of band  $\alpha$  we find from (3.9), (3.28), and (2.41)

$$\gamma_\alpha \approx \gamma_-^{(\alpha)} + \gamma_+^{(\alpha-1)}. \tag{3.30}$$

As discussed above, this holds only when the Zak-Berry connection in Dirac theory vanishes beyond the cutoff  $\Lambda_c$  which is fulfilled for the AF gauge. In the AF gauge we will furthermore show below via the explicit eigenstates of the Dirac model (see Sec. V A) that  $\gamma_\tau^{(v)}$  is related to the phase  $\gamma^{(v)}$  of the gap parameter  $\Delta^{(v)}e^{i\gamma^{(v)}}$  by

$$\gamma_+^{(v)} = \gamma^{(v)}, \quad \gamma_-^{(v)} = -\gamma^{(v)} + \pi s_\gamma, \tag{3.31}$$

where we abbreviated the sign function by

$$s_\gamma = \text{sgn}\gamma, \tag{3.32}$$

and assumed  $-\pi < \gamma^{(v)} < \pi$  with periodic continuation to the other regimes.

In the following, we denote the eigenstates of lattice and Dirac theories in the AF gauge by  $\psi_{k\alpha}(ma)$  and  $\psi_{q\tau}^{(v)}(x)$ , respectively, which are connected by the relation (2.22). Correspondingly, we use the notation  $u_{k\alpha}$  and  $\chi_{q\tau p}^{(v)}$  in this gauge which are related by (2.30) and (2.31). From (3.11) we note that we get in the AF gauge

$$A_{-q,\tau}^{(v)} = A_{q\tau}^{(v)}. \tag{3.33}$$

The ML gauge in Dirac theory is defined by a vanishing Zak-Berry connection

$$\text{Dirac ML gauge : } \tilde{A}_{q\tau}^{(v)} = 0, \tag{3.34}$$

such that the corresponding Zak-Berry phase is also zero

$$\tilde{\gamma}_\tau^{(v)} = 0. \tag{3.35}$$

According to (3.24) this corresponds to the choice of a gauge factor  $e^{i\phi_{q\tau}^{(v)}}$  with

$$\phi_{q\tau}^{(v)} = \int_0^q dq' A_{q'\tau}^{(v)} + \phi_{0,\tau}^{(v)}. \tag{3.36}$$

Note that this is not a contradiction to the corresponding condition (3.12) for the lattice theory where it is not possible to remove the Zak-Berry phase via a global gauge transformation. The Dirac model for a given gap  $\nu$  describes only those states in  $k$  space which belong to the two halves of the bands separated by gap  $\nu$ . Therefore, a particular global gauge chosen in Dirac theory corresponds to a certain *local* gauge in lattice theory defined for all quasimomenta lying in one-half of a given band. Such a gauge can not necessarily be extended to a global gauge for a certain band which is continuous and periodic in  $k$ .

In order to show the explicit difference between the ML gauges in lattice and Dirac theories we start from the AF gauge and define a gauge transformation to the ML gauge by the transformed quantities

$$\tilde{u}_{k\alpha} = e^{i\phi_{k\alpha}}u_{k\alpha}, \quad \tilde{\chi}_{q\tau p}^{(v)} = e^{i\phi_{q\tau}^{(v)}}\chi_{q\tau p}^{(v)}. \tag{3.37}$$

From the conditions (3.12), (3.13), (3.34), and (3.36), defining the ML gauge in lattice and Dirac theories, we get

$$\varphi_{k\alpha} = \int_{-\pi/Za}^k dk' A_{k'\alpha} - \left(k + \frac{\pi}{Za}\right)\frac{Za}{2\pi}\gamma_\alpha \tag{3.38}$$

$$= \int_0^k dk' A_{k'\alpha} - k\frac{Za}{2\pi}\gamma_\alpha, \tag{3.39}$$

$$A_{q\tau}^{(v)} = \partial_q\phi_{q\tau}^{(v)}, \tag{3.40}$$



where we made use of (3.11) in the AF gauge to get the form (3.39). Using (3.28) we find from the two forms for  $\varphi_{k\alpha}$  that, for both the lower and upper halves of band  $\alpha$ , we get the following relation between the ML gauges of lattice and Dirac theories:

$$\varphi_{k\alpha} = \int_0^q dq' A_{q'\tau}^{(v)} - q \frac{Za}{2\pi} \gamma_\alpha = \phi_{q\tau}^{(v)} - \phi_{0,\tau}^{(v)} - q \frac{Za}{2\pi} \gamma_\alpha, \quad (3.41)$$

leading to the following relation between the gauge factors:

$$e^{i\varphi_{k\alpha}} = e^{i\phi_{q\tau}^{(v)}} e^{-i\phi_{0,\tau}^{(v)}} e^{-iq \frac{Za}{2\pi} \gamma_\alpha}. \quad (3.42)$$

The last two factors on the right-hand side define the difference between the ML gauges in lattice and Dirac theories, which have to be added to (2.30) and (2.31) to get the precise relation between the eigenstates of lattice and Dirac theories in the ML gauge. Furthermore, they show that the ML gauge in Dirac theory is not smooth when crossing over the middle of a band  $\alpha$  since the indices ( $\nu\tau$ ) and the relation between  $q$  and  $k$  change.

### C. Boundary charge

#### 1. Surface charge theorem

The charge accumulated at the boundary, or the boundary charge [12–18], is defined by restricting the tight-binding Hamiltonian (2.1) to the half-infinite space  $m > 0$  and averaging the excess charge with a macroscopic envelope function  $f_m$  via

$$Q_B^{(v)} = \sum_{m=1}^{\infty} (\rho_m^{(v)} - \bar{\rho}^{(v)}) f_m. \quad (3.43)$$

Here,  $\rho_m^{(v)}$  is the average charge at lattice site  $m$  for a half-infinite system if the chemical potential is located in gap  $\nu$ , including all edge states up to this energy. The average charge per site in the bulk is denoted by  $\bar{\rho}^{(v)} = \nu/Z$ . The macroscopic and probe-specific envelope function is denoted by  $f_m$  which must have certain properties. In particular, it has the value 1 in the first range of  $m$ , which far exceeds the localization length  $\xi$ , and then it smoothly crosses over to 0 over the second range of  $m$ , which is also  $\gg \xi$ . The importance of including these features into the definition of  $f_m$  for the universality of the boundary charge properties as well as their experimental relevance has been elucidated in Refs. [12–15]. If only a single band  $\alpha$  is occupied we denote the corresponding boundary charge by  $Q_{B,\alpha}$ , such that the total boundary charge can be decomposed as

$$Q_B^{(v)} = \sum_{\alpha=1}^{\nu} Q_{B,\alpha} + Q_E^{(v)} \quad (3.44)$$

$$= \sum_{\alpha=1}^{\nu} (Q_{B,\alpha} + Q_{E,\alpha}), \quad (3.45)$$

where  $Q_E^{(v)}$  denotes the integer contribution from all occupied edge states up to the chemical potential, and  $Q_{E,\alpha}$  denotes the contribution of an edge state present in gap  $\alpha$ . The latter is either unity or zero and we assume for the validity of (3.45)

that, for the last gap  $\alpha = \nu$ , the chemical potential is located at the bottom of the conduction band  $\alpha = \nu + 1$ .

As shown in Ref. [15], one finds in the AF gauge of the lattice theory that the Zak-Berry phase is related to the boundary charge of a single band in the following way:

$$Q_{B,\alpha} = -\frac{\gamma_\alpha}{2\pi} - \frac{Z-1}{2Z}. \quad (3.46)$$

Inserting (3.30) and (3.31) in the expression (3.46) for the boundary charge we get

$$Q_{B,\alpha} = \frac{\gamma^{(\alpha)} - \gamma^{(\alpha-1)}}{2\pi} + \frac{1}{2Z} - \theta(\gamma^{(\alpha)}). \quad (3.47)$$

Noting that an edge state appears in gap  $\nu$  for  $0 < \gamma^{(\nu)} < \pi$  (see Ref. [17]), we find that the sum of the boundary charge of band  $\alpha$  and the charge  $Q_{E,\alpha}$  of the edge state in gap  $\alpha$  is given by

$$Q_{B,\alpha} + Q_{E,\alpha} = \frac{\gamma^{(\alpha)} - \gamma^{(\alpha-1)}}{2\pi} + \frac{1}{2Z}. \quad (3.48)$$

Assuming that the chemical potential is located at the bottom of the conduction band, this gives the following result for the boundary charge  $Q_B^{(v)}$  when the lowest  $\nu$  bands are filled:

$$Q_B^{(v)} = \frac{\gamma^{(v)}}{2\pi} + \frac{\nu}{2Z}. \quad (3.49)$$

Here we have used the fact that the lower half of the lowest band gives a negligible contribution to the Zak-Berry phase since we can approximate the eigenstates by the free ones in this regime leading to a vanishing Zak-Berry connection (see Appendix A). The result (3.49) states the low-energy version of the *surface charge theorem* relating the boundary charge to the phase  $\gamma^{(v)}$  of the gap parameter which is related via (3.31) to the low-energy version of the Zak-Berry phase. We note that the result (3.49) is fully consistent with the direct calculation of the boundary charge for a half-infinite Dirac model as shown in Ref. [17].

An essential point for the formulation of the surface charge theorem in low energy is the fact that the sum of the two Zak-Berry phases  $\gamma_\tau^{(v)}$  over  $\tau = \pm$  gives in the AF gauge

$$\sum_{\tau=\pm} \gamma_\tau^{(v)} = \pi s_{\gamma^{(v)}}. \quad (3.50)$$

This gives rise to the effect that all other gaps  $1, \dots, \nu - 1$  contribute only constants to the boundary charge  $Q_B^{(v)}$  and the final result (including the edge states) can be written in terms of the Zak-Berry phase  $\gamma_-^{(v)}$  of the top of the valence band alone. We note that the result  $\sum_\tau \gamma_\tau^{(v)} = \pi \text{ mod}(2\pi)$  can be shown on quite general principles without using the explicit forms of the eigenfunctions of the Dirac model (see Appendix B).

#### 2. Surface fluctuation theorem

In lattice theory the  $k$  integral over the geometric tensor is of particular interest since it can be related to the boundary charge fluctuations [18]

$$l_p(\Delta Q_B^{(v)})^2 = \sum_{\alpha,\beta=1}^{\nu} \int_{-\pi/Za}^{\pi/Za} \frac{dk}{2\pi} (Q_k)_{\alpha\beta}, \quad (3.51)$$

where  $(\Delta Q_B^{(\nu)})^2$  are the fluctuations when the chemical potential lies in gap  $\nu$  (the precise position is unimportant). Here,  $l_p \gg \xi$  is a length scale on which the charge measurement probe loses smoothly the contact to the sample (see Ref. [18] for details). Since the geometric tensor is gauge invariant, this relation holds in any gauge, in particular in the AF and ML gauges. Summing the geometric tensor  $(Q_k)_{\alpha\beta}$  over  $\alpha$  or  $\beta$  up to some  $\nu$  in the low-energy regime, we find from (3.29) and (3.21) that all contributions vanish where two bands are separated by a gap. Therefore, we obtain

$$\sum_{\alpha=1}^{\nu} (Q_k)_{\alpha\beta} = \delta_{\beta\nu} Q_{q-}^{(\nu)} \begin{cases} \delta_{\nu,\text{even}} & \text{for } k \approx 0, \\ \delta_{\nu,\text{odd}} & \text{for } k \approx \pm \frac{\pi}{Za} \end{cases} \quad (3.52)$$

and

$$\sum_{\alpha,\beta=1}^{\nu} (Q_k)_{\alpha\beta} = Q_{q-}^{(\nu)} \begin{cases} \delta_{\nu,\text{even}} & \text{for } k \approx 0, \\ \delta_{\nu,\text{odd}} & \text{for } k \approx \pm \frac{\pi}{Za}. \end{cases} \quad (3.53)$$

These relations are also independent of the gauge choice in Dirac theory since the geometric tensor is gauge invariant in Dirac theory as well. Using (2.41) and the fact that the geometric tensor vanishes in Dirac theory in the asymptotically free region  $|q| > \Lambda_c$ , we obtain in the low-energy regime for the fluctuations

$$l_p (\Delta Q_B^{(\nu)})^2 \approx \int \frac{dq}{2\pi} Q_{q-}^{(\nu)} = \int \frac{dq}{2\pi} |(A_q^{(\nu)})_{+-}|^2 \quad (3.54)$$

$$= \int \frac{dq}{2\pi} \left\{ |\chi_{q-}^{(\nu)}|^2 |(\partial_q)^2 \chi_{q-}^{(\nu)}|^2 - (A_{q-}^{(\nu)})^2 \right\}, \quad (3.55)$$

where we made use of (3.20), (3.22), and (3.25) for the last two steps.

The result (3.54) is the low-energy form of the *surface fluctuation theorem* relating the boundary charge fluctuations to the momentum integral of the geometric tensor. As shown, only the geometric tensor of the valence band enters. This is physically intuitive since one expects no fluctuations of the charge from the low-lying bands. As will be shown below in Sec. V we obtain explicitly via the eigenfunctions of the Dirac model

$$\int \frac{dq}{2\pi} Q_{q-}^{(\nu)} = \xi^{(\nu)}/8, \quad \xi^{(\nu)} = \frac{v_F^{(\nu)}}{2\Delta^{(\nu)}}, \quad (3.56)$$

showing that the result (3.54) is fully consistent with the direct calculation of the boundary charge fluctuations within the Dirac model via the second momentum of the correlation function [18].

Most importantly, we note that the fluctuations of the total boundary charge can *not* be written as the sum over the fluctuations of the individual bands. If only band  $\alpha$  is occupied, the fluctuations are given by

$$l_p \Delta(Q_{B,\alpha})^2 = \int_{-\pi/Za}^{\pi/Za} \frac{dk}{2\pi} Q_{k\alpha}. \quad (3.57)$$

When summing this expression over  $\alpha$  one does not obtain the fluctuations  $l_p (\Delta Q_B^{(\nu)})^2$  when all bands  $\alpha = 1, \dots, \nu$  are occupied, as given by Eq. (3.51) involving the summation over all matrix elements of the geometric tensor. In particular, one loses the important cancellation property (3.21) which is the central ingredient that all contributions from gaps between

occupied bands do not contribute to the fluctuations of the total boundary charge, rendering the fluctuations to depend only on the low-energy properties of the model.

#### IV. WANNIER FUNCTIONS

In this section we present the definition of the central objects of our work, the Wannier functions in Dirac theory and their corresponding moments, together with their precise relation to the lattice Wannier functions. In Sec. IV A we present a short summary of the definition of Wannier functions in lattice theory. The Dirac Wannier functions are defined in Sec. IV B and it is shown how their first and second moments can be related to the Zak-Berry phase and the momentum integral of the geometric tensor as defined within the Dirac theory in Sec. III B. At the end of this section we show how the surface charge and surface fluctuation theorem can be formulated very elegantly in terms of the moments of the Dirac Wannier functions. The precise relation of the Dirac Wannier functions and their moments to the corresponding lattice quantities is the subject of Secs. IV C 1 and IV C 2 in the AF and ML gauge, respectively.

##### A. Wannier functions for lattice model

On the lattice the dimensionless Wannier functions for band  $\alpha$  are defined by

$$w_{R,\alpha}(ma) = \sqrt{\frac{Za}{2\pi}} \int_{-\pi/Za}^{\pi/Za} dk \psi_{k\alpha}(ma) e^{-ikR}, \quad (4.1)$$

where  $R = Zan$ , with  $n$  integer, denotes a lattice vector. Using (2.11) and (2.12) one finds that the Wannier functions labeled by  $\alpha$  and  $R$  form an orthonormal and complete set of states

$$\sum_m w_{R,\alpha}(ma)^* w_{R',\alpha'}(ma) = \delta_{\alpha\alpha'} \delta_{RR'}, \quad (4.2)$$

$$\sum_{\alpha} \sum_R w_{R,\alpha}(ma) w_{R,\alpha}(m'a)^* = \delta_{mm'}. \quad (4.3)$$

Defining by  $w_{\alpha}(ma) \equiv w_{0,\alpha}(ma)$  the Wannier function centered at zero

$$w_{\alpha}(ma) = \sqrt{\frac{Za}{2\pi}} \int_{-\pi/Za}^{\pi/Za} dk \psi_{k\alpha}(ma) \quad (4.4)$$

$$= Za \int_{-\pi/Za}^{\pi/Za} \frac{dk}{2\pi} u_{k\alpha}(ma) e^{ikma}, \quad (4.5)$$

we find with the Bloch form (2.5) and the periodicity property (2.6) that the Wannier function  $w_{R,\alpha}(ma)$  follows from shifting  $w_{\alpha}(ma)$  by the lattice vector  $R$ :

$$w_{R,\alpha}(ma) = w_{\alpha}(ma - R). \quad (4.6)$$

Therefore, all properties of the Wannier functions follow from studying the properties of  $w_{\alpha}(ma)$ .

The Wannier functions depend on the gauge in a non-trivial way. If (4.5) is the Wannier function in the AF gauge, we obtain in the ML gauge

$$\tilde{w}_{\alpha}(ma) = Za \int_{-\pi/Za}^{\pi/Za} \frac{dk}{2\pi} e^{i\varphi_{k\alpha}} u_{k\alpha}(ma) e^{ikma}, \quad (4.7)$$

where the phase  $\varphi_{k\alpha}$  is given by (3.13). We note that due to the properties (3.11) and (3.16), both the Wannier functions in AF and ML gauges are real.

For the bands  $\alpha = 2, \dots, Z-1$  we show in Appendix C that the Wannier functions in the AF or ML gauge can be split into two contributions corresponding to the upper and lower halves of the band,

$$w_\alpha(ma) = w_{u,\alpha}(ma) + w_{d,\alpha}(ma), \quad (4.8)$$

$$w_{u/d,\alpha}(ma) = w'_{u/d,\alpha}(ma) \pm \delta w_\alpha(ma), \quad (4.9)$$

where  $w'_{u/d,\alpha}(ma)$  denotes the part from the corresponding integration regions

$$w'_{u/d,\alpha}(ma) \quad (4.10)$$

$$= \frac{Za}{2\pi} \begin{cases} \int_{|k| < \frac{\pi}{2Za}} dk u_{k\alpha}(ma) e^{ikma} & \text{for } \alpha \text{ even/odd,} \\ \int_{\frac{\pi}{2Za} < |k| < \frac{\pi}{Za}} dk u_{k\alpha}(ma) e^{ikma} & \text{for } \alpha \text{ odd/even} \end{cases} \quad (4.11)$$

and  $\delta w_\alpha(ma)$  arises from the extension of the quasimomentum integrations to  $\pm\infty$  taking the free solutions of  $u_{k\alpha}(ma)$  in the gapless case (see Appendix C for details). In the AF and ML gauges one obtains explicitly

$$\delta w_\alpha(ma) = \frac{\sqrt{Z}}{\pi m} \sin \left\{ \frac{\pi}{Z} \left( \alpha - \frac{1}{2} \right) m \right\}, \quad (4.12)$$

$$\begin{aligned} \delta \tilde{w}_\alpha(ma) \\ = \frac{\sqrt{Z}}{\pi \tilde{m}_\alpha} \sin \left\{ \frac{\pi}{Z} \left( \alpha - \frac{1}{2} \right) m + \frac{1}{4} (\gamma_+^{(\alpha-1)} - \gamma_-^{(\alpha)}) \right\}, \end{aligned} \quad (4.13)$$

where we defined the shifted variable

$$\tilde{m}_\alpha = m - \frac{Z}{2\pi} \gamma_\alpha. \quad (4.14)$$

Whereas the corrections  $\delta w_\alpha(ma)$  cancel out when considering the total Wannier function  $w_\alpha(ma)$  of a band, we will see in Sec. V that the Wannier functions  $w_{u/d,\alpha}(ma)$  and  $\tilde{w}_{u/d,\alpha}(ma)$  show only universal scaling if the corrections  $\pm\delta w_\alpha(ma)$  and  $\pm\delta \tilde{w}_\alpha(ma)$  are taken into account.

For the bands  $\alpha = 1, Z$  the splitting in upper and lower parts makes no sense since the lower and upper halves of the band  $\alpha = 1/Z$  are already described by a free theory for small gap. Therefore, we use the convention

$$w_{d,1}(ma) = w_{u,Z}(ma) = 0, \quad \delta w_{1/Z}(ma) = 0. \quad (4.15)$$

The moments of the Wannier functions are defined by

$$C_{r\alpha} = \langle x^r \rangle_\alpha = \sum_m (ma)^r |w_\alpha(ma)|^2. \quad (4.16)$$

A corresponding definition is used for the moments  $C_{u/d,r\alpha}$  of  $w_{u/d,\alpha}$ . Inserting (4.4) and (2.5) we find after some straightforward manipulations

$$C_{r\alpha} = \frac{Za}{2\pi} \int_{-\pi/Za}^{\pi/Za} dk \langle u_{k\alpha} | (i\partial_k)^r | u_{k\alpha} \rangle. \quad (4.17)$$

For  $r = 1$  and 2 we find with (3.1) and (3.6)

$$C_{1\alpha} = \frac{Za}{2\pi} \int_{-\pi/Za}^{\pi/Za} dk A_{k\alpha} = \frac{Za}{2\pi} \gamma_\alpha, \quad (4.18)$$

$$C_{2\alpha} = \frac{Za}{2\pi} \int_{-\pi/Za}^{\pi/Za} dk \{ Q_{k\alpha} + (A_{k\alpha})^2 \}, \quad (4.19)$$

and for the quadratic spread

$$\begin{aligned} \langle \Delta x^2 \rangle_\alpha &= C_{2\alpha} - (C_{1\alpha})^2 \\ &= Za \int_{-\pi/Za}^{\pi/Za} \frac{dk}{2\pi} \left\{ Q_{k\alpha} + \left( A_{k\alpha} - \frac{Za}{2\pi} \gamma_\alpha \right)^2 \right\}. \end{aligned} \quad (4.20)$$

Since the geometric tensor is gauge invariant, the condition for maximal localization is given by a constant Zak-Berry connection corresponding to the ML gauge (3.12). Since, according to (3.17), the Zak-Berry phase does not change in this gauge, the first moment stays invariant

$$\tilde{C}_{1\alpha} = C_{1\alpha} = \frac{Za}{2\pi} \gamma_\alpha. \quad (4.21)$$

In the ML gauge we get for the minimal quadratic spread  $\langle \Delta x^2 \rangle_{\alpha, \min}$  from (4.20) and (3.51) the result

$$\langle \Delta x^2 \rangle_{\alpha, \min} = Za \int_{-\pi/Za}^{\pi/Za} \frac{dk}{2\pi} Q_{k\alpha} = Za l_p \Delta (Q_{B,\alpha})^2, \quad (4.22)$$

where  $l_p \Delta (Q_{B,\alpha})^2$  are the fluctuations when only the band  $\alpha$  is occupied.

So, in summary we can formulate the surface charge and surface fluctuation theorem for a singly occupied band in terms of the first and second moments of the Wannier function as

$$Q_{B,\alpha} = -\frac{C_{1\alpha}}{Za} - \frac{Z-1}{2Z}, \quad (4.23)$$

$$l_p \Delta (Q_{B,\alpha})^2 = \frac{1}{Za} \tilde{D}_{2\alpha}, \quad (4.24)$$

where  $C_{r\alpha}$  refers to the AF gauge, and  $\tilde{D}_{r\alpha}$  are the moments in the ML gauge defined relative to the first moment  $C_{1\alpha} = \frac{Za}{2\pi}$ :

$$\tilde{D}_{r\alpha} = \sum_m (\tilde{m}_\alpha a)^r |\tilde{w}_\alpha(ma)|^2, \quad (4.25)$$

with  $\tilde{m}_\alpha$  defined in (4.14). These moments are related to the moments  $\tilde{C}_r^{(\alpha)}$ . For example, for  $r = 1, 2$ , one obtains by using the normalization  $\sum_m |\tilde{w}_\alpha(ma)|^2 = 1$  and (4.21)

$$\tilde{D}_{1\alpha} = \tilde{C}_1^{(\alpha)} - \frac{Za}{2\pi} \gamma_\alpha = 0, \quad (4.26)$$

$$\tilde{D}_{2\alpha} = \tilde{C}_2^{(\alpha)} - (\tilde{C}_1^{(\alpha)})^2. \quad (4.27)$$

As discussed at the end of Sec. III C 2 the fluctuations  $l_p \Delta (Q_{B,\alpha})^2$  of the individual bands are not sufficient to calculate the fluctuations  $l_p (\Delta Q_B^{(v)})^2$  when all bands  $\alpha = 1, \dots, \nu$  are occupied. However, as we will see in the next two sections for the low-energy regime of small gaps, the surface charge and surface fluctuation theorem for  $\nu$  occupied bands can be written entirely in terms of the moments of the upper

components  $w_{u,v}$  and  $\tilde{w}_{u,v}$  of the Wannier functions for the highest valence band as

$$Q_B^{(v)} \approx -\frac{1}{Za} C_{u,1v} + \frac{1}{2} s_{\gamma^{(v)}} + \frac{\nu}{2Z}, \quad (4.28)$$

$$l_p(\Delta Q_B^{(v)})^2 \approx \frac{1}{Za} \tilde{C}_{u,2v}. \quad (4.29)$$

This physically intuitive result reflects the fact that the boundary charge is a low-energy property only which is insensitive to the properties of low-lying and occupied bands. Alternatively, as we will discuss in Sec. VI, it is also possible to define maximally localized Wannier functions in a non-Abelian gauge where all occupied bands are mixed. As we will see, these Wannier functions are closely related to the Wannier function  $\tilde{w}_{u,v}$  in the Abelian ML gauge and show precisely the same universal scaling as the Dirac Wannier function of the valence band.

### B. Wannier functions for Dirac model

Within the Dirac theory we define the Wannier functions by

$$w_{y,\tau}^{(v)}(x) = \sum_{p=\pm} w_{y,\tau p}^{(v)}(x) e^{ipk_F^{(v)}(x-y)} = w_{\tau}^{(v)}(x-y), \quad (4.30)$$

where

$$w_{y,\tau p}^{(v)}(x) = \int \frac{dq}{\sqrt{2\pi}} \psi_{q\tau p}^{(v)}(x) e^{-iqy} = w_{\tau p}^{(v)}(x-y) \quad (4.31)$$

and

$$w_{\tau}^{(v)}(x) = w_{y=0,\tau}^{(v)}(x), \quad w_{\tau p}^{(v)}(x) = w_{y=0,\tau p}^{(v)}(x), \quad (4.32)$$

$$w_{\tau p}^{(v)}(x) = \int \frac{dq}{\sqrt{2\pi}} \psi_{q\tau p}^{(v)}(x) = \int \frac{dq}{2\pi} \chi_{q\tau p}^{(v)} e^{iqx}. \quad (4.33)$$

Correspondingly, for the ML gauge, we define  $\tilde{w}_{\tau p}^{(v)}(x)$  by

$$\tilde{w}_{\tau p}^{(v)}(x) = \int \frac{dq}{2\pi} \tilde{\chi}_{q\tau p}^{(v)} e^{iqx}. \quad (4.34)$$

In contrast to the lattice Wannier functions, the Dirac Wannier functions are complex quantities both in the AF and ML gauges.

Since the Wannier functions in the Dirac model contain a continuous shift  $y$  their dimension is given by inverse length and the normalization and completeness relations follow from (2.33) and (2.34) as

$$\sum_p \int dx w_{y,\tau p}^{(v)}(x) w_{y',\tau' p}^{(v)*}(x) = \delta_{\tau\tau'} \delta(y-y'), \quad (4.35)$$

$$\sum_{\tau} \int dy w_{y,\tau p}^{(v)}(x) w_{y,\tau p}^{(v)*}(x') = \delta_{pp} \delta(x-x'). \quad (4.36)$$

The moments  $C_{r\tau}^{(v)}$  of the Dirac Wannier functions are defined by

$$C_{r\tau}^{(v)} = \int dx x^r \sum_{p=\pm} |w_{\tau p}^{(v)}(x)|^2 \quad (4.37)$$

and, correspondingly, we denote the moments in the ML gauge by  $\tilde{C}_{r\tau}^{(v)}$ . Note that these moments have dimension  $(\text{length})^{r-1}$  in contrast to the moments  $C_{r\alpha}$  defined within the

lattice theory. This is due to the different normalization in the continuum Dirac theory. Nevertheless, we will see below that the first and second moments are related to the boundary charge and the fluctuations, respectively, in a similar way as in (4.23) and (4.24) but without the denominator  $Za$  (see below). Note that the first moment does *not* play the role of the position of the Wannier function in Dirac theory since it is dimensionless. A finite value of the first moment  $C_{1\tau}^{(v)}$  indicates an asymmetry of  $\sum_{p=\pm} |w_{\tau p}^{(v)}(x)|^2$  for positive and negative  $x$ .

Using the form (4.33) we find after some straightforward manipulations

$$C_{r\tau}^{(v)} = \int \frac{dq}{2\pi} \langle \chi_{q\tau}^{(v)} | (i\partial_q)^r | \chi_{q\tau}^{(v)} \rangle. \quad (4.38)$$

Using (3.26), (3.20), and (3.22) we can thus write for the first and second moments

$$C_{1\tau}^{(v)} = \frac{\gamma_{\tau}^{(v)}}{2\pi}, \quad (4.39)$$

$$C_{2\tau}^{(v)} = \int \frac{dq}{2\pi} |(A_q^{(v)})_{\tau\tau}|^2 + \int \frac{dq}{2\pi} (A_{q\tau}^{(v)})^2. \quad (4.40)$$

Since the first term on the right-hand side of (4.40) is gauge invariant and the second one has a positive integrand it follows that the gauge of maximally localized Wannier functions is given by the ML gauge (3.34) where the Zak-Berry connection vanishes. We note that there is a fundamental difference to the lattice theory where the quadratic spread is defined by  $C_{2\alpha} - (C_{1\alpha})^2$ . The analog formula is not possible in low energy since the dimension of  $C_{r\tau}^{(v)}$  is  $(\text{length})^{r-1}$ .

Taking (4.39) and (4.40) together with (3.49), (3.31), and (3.55), we can formulate the surface charge and surface fluctuation theorem in low energy via the first and second moments of the Dirac Wannier functions as

$$Q_B^{(v)} \approx -C_{1-}^{(v)} + \frac{1}{2} s_{\gamma^{(v)}} + \frac{\nu}{2Z}, \quad (4.41)$$

$$l_p(\Delta Q_B^{(v)})^2 \approx \tilde{C}_{2-}^{(v)}. \quad (4.42)$$

Here, the moments  $C_{r-}^{(v)}$  refer to the AF gauge, whereas  $\tilde{C}_{r-}^{(v)}$  are the moments in the ML gauge. Note that, in contrast to (4.23) and (4.24) (where a single band has been considered), we consider here the total boundary charge and its fluctuations when the lowest  $\nu$  bands are filled and the chemical potential is located at the bottom of the conduction band (which is only important to calculate the edge-state contribution for the boundary charge in the last gap). This result shows that the boundary charge and its fluctuations are quantities probing only low-energy features and, therefore, can be related in a universal way to the first and second moments of the Dirac Wannier functions corresponding to the top of the highest valence band.

### C. Relation between Wannier functions in lattice and Dirac theories

We now consider the precise relation of the Dirac Wannier functions to the Wannier function defined within the lattice theory. We start with the AF gauge where the relation (2.22) between eigenfunctions of the lattice and the Dirac theories

hold globally for all quasimomenta in a certain band  $\alpha$  since the gauge is the same. For the ML gauge it is more subtle since there is difference between the gauges in lattice and Dirac theories [see Eq. (3.42)]. Therefore, one has to add gauge factors depending on whether one considers the contribution of the upper or lower half of the band to the Wannier function.

### 1. AF gauge

For the AF gauge, we can insert (2.22) in (4.4) and use (2.41) to get

$$w_\alpha(ma) = a\sqrt{Z} \int_{|q| < \Lambda_\alpha} \frac{dq}{\sqrt{2\pi}} \times \{ \psi_{q_-}^{(\alpha)}(ma) + \psi_{q_+}^{(\alpha-1)}(ma) \}, \quad (4.43)$$

where as usual we omit the terms with  $\psi_{q_+}^{(0)}$  and  $\psi_{q_-}^{(Z)}$  if  $\alpha = 1$  or  $Z$ , respectively. As shown in Appendix C the region  $|q| > \Lambda_\alpha$  contributes  $\pm \delta w_{u/d,\alpha}(ma)$  to the first and second terms on the right-hand side, i.e., precisely the same contribution (4.12) we used to extend half of a band for  $\alpha = 2, \dots, Z-1$  (for  $\alpha = 1, Z$  there is no contribution from  $|q| > \frac{\pi}{Za}$ ). Therefore, after inserting (2.23) and (4.33), we obtain the following relation between the Wannier functions in lattice and Dirac theories:

$$w_{u,\alpha}(ma) = a\sqrt{Z} \sum_p w_{-,p}^{(\alpha)}(ma) e^{ipk_F^{(\alpha)} ma}, \quad (4.44)$$

$$w_{d,\alpha}(ma) = a\sqrt{Z} \sum_p w_{+,p}^{(\alpha-1)}(ma) e^{ipk_F^{(\alpha-1)} ma}. \quad (4.45)$$

Thereby, we will show in Appendix C that the momentum integral to define the Wannier functions  $w_{\tau p}^{(v)}(x)$  via (4.33) has to be regularized in such a way that unphysical contributions  $\sim \delta(x)$  are absent (see the explicit calculation in Sec. V).

To understand the relation of the various moments in the AF gauge, we need to evaluate  $|w_\alpha(ma)|^2$  using the Wannier function from the sum of (4.44) and (4.45). The simplest cases are  $\alpha = 1$  or  $Z$ , where only one term contributes:  $w_1(ma) = w_u^{(1)}(ma)$  and  $w^Z(ma) = w_d^{(1)}(ma)$  [see (4.15)]. For  $\alpha = 1$  we get with  $k_F^{(1)} a = \pi/Z$ :

$$|w_1(ma)|^2 = a^2 Z \sum_p |w_{-,p}^{(1)}(ma)|^2 + a^2 Z \sum_p w_{-,p}^{(1)}(ma) (w_{-,p}^{(1)}(ma))^* e^{i2\pi pm/Z}, \quad (4.46)$$

and a similar result for  $\alpha = Z$  with  $w_{-,p}^{(1)} \rightarrow w_{+,p}^{(Z-1)}$  and  $e^{i2\pi pm/Z} \rightarrow e^{-i2\pi pm/Z}$ . When inserting this formula into the definition (4.16) of the moments and neglecting the variation of the slowly varying function  $w_{-,p}^{(1)}(ma)$  over the unit cell, we find that the strongly oscillating terms of (4.46) can be neglected when averaging them over a unit cell. This leads to the result

$$C_{r1} \approx a^2 Z \sum_m (ma)^r \sum_p |w_{-,p}^{(1)}(ma)|^2 \approx Za \int dx x^r \sum_p |w_{-,p}^{(1)}(x)|^2 = Za C_{r-}^{(1)}, \quad (4.47)$$

$$C_{rZ} \approx a^2 Z \sum_m (ma)^r \sum_p |w_{+,p}^{(Z-1)}(ma)|^2 \approx Za \int dx x^r \sum_p |w_{+,p}^{(Z-1)}(x)|^2 = Za C_{r+}^{(Z-1)}. \quad (4.48)$$

For the bands  $\alpha = 2, \dots, Z-1$ , both terms (4.44) and (4.45) contribute to the Wannier function. This leads to more oscillating terms occurring for the moments involving also  $e^{i\pi p(2\alpha-1)m/Z}$  and  $e^{i\pi pm/Z}$ . Such terms give again a negligible contribution to the moments when averaging them over two unit cells, leading to the general result

$$C_{r\alpha} \approx Za \{ C_{r-}^{(\alpha)} + C_{r+}^{(\alpha-1)} \}, \quad (4.49)$$

together with

$$C_{u,r\alpha} \approx Za C_{r-}^{(\alpha)}, \quad C_{d,r\alpha} \approx Za C_{r+}^{(\alpha-1)}. \quad (4.50)$$

Inserting (4.50) in (4.41) we find (4.28), which states the surface charge theorem in terms of the upper component  $w_{u,v}$  of the lattice Wannier function for the highest valence band.

As discussed later in all detail in Sec. VE, we note that the divergence of the Dirac Wannier functions  $|w_{\tau p}^{(v)}(x)|^2 \sim 1/x^2$  for  $|x| \ll \xi^{(v)}$  is not important for the calculation of the moments for  $r > 0$  since the contributions from the region  $|m| \sim O(1)$  can be neglected in the universal limit  $\xi \gg a$ . The only exception is  $r = 0$  where the normalization of the lattice Wannier function is fully determined by the region  $|m| \sim O(1)$ .

### 2. ML gauge

In the ML gauge we have shown via the two last factors on the right-hand side of (3.42) that there is a difference between the gauges of maximally localized Wannier functions in lattice and Dirac theories. The last factor leads to a trivial shift of the position of the Wannier function  $w_{\tau p}^{(v)}(x)$ :

$$w_{\tau p}^{(v)}(x) \rightarrow w_{\tau p}^{(v)}\left(x - \frac{Za}{2\pi} \gamma_\alpha\right). \quad (4.51)$$

Including this shift and the phase factor  $e^{-i\phi_{0,\tau}^{(v)}}$  we have to modify (4.44) and (4.45) in the following way (see Appendix C for details) to get the correct relationship between the maximally localized Wannier functions in lattice and Dirac theories:

$$\tilde{w}_{u,\alpha}(ma) = a\sqrt{Z} \sum_p e^{-i\phi_{0,-}^{(\alpha)}} \tilde{w}_{-,p}^{(\alpha)}(\tilde{m}_\alpha a) e^{ipk_F^{(\alpha)} ma}, \quad (4.52)$$

$$\tilde{w}_{d,\alpha}(ma) = a\sqrt{Z} \sum_p e^{-i\phi_{0,+}^{(\alpha-1)}} \tilde{w}_{+,p}^{(\alpha-1)}(\tilde{m}_\alpha a) e^{ipk_F^{(\alpha-1)} ma}, \quad (4.53)$$

where  $\tilde{m}_\alpha$  is defined in (4.14).

Using (4.52) and (4.53), we obtain analog to (4.49) from the definition (4.25) for the moments in the ML gauge

$$\tilde{D}_{r\alpha} \approx Za \{ \tilde{C}_{r-}^{(\alpha)} + \tilde{C}_{r+}^{(\alpha-1)} \}, \quad (4.54)$$

together with

$$\tilde{D}_{u,r\alpha} \approx Za \tilde{C}_{r-}^{(\alpha)}, \quad \tilde{D}_{d,r\alpha} \approx Za \tilde{C}_{r+}^{(\alpha-1)}. \quad (4.55)$$

Inserting (4.55) in (4.42) we find (4.29), which states the surface fluctuation theorem in terms of the upper component  $\tilde{w}_{u,v}$  of the lattice Wannier function for the highest valence band.

As discussed in more detail in Sec. V E, the relations (4.54) and (4.55) should be used only for even values of  $r$ . As shown in Sec. V C the Dirac moments  $\tilde{C}_{r\tau}^{(v)}$  are exactly zero for all odd values of  $r$ . However, this does not mean that the lattice moments  $\tilde{D}_{r\alpha}$  and  $\tilde{D}_{w/d,r\alpha}$  are exactly zero for odd values of  $r$  (except for  $\tilde{D}_{1,\alpha} = 0$  which is zero by definition). The smallness of the odd moments in lattice theory has to be understood by an order of magnitude analysis compared to the even moments. The odd moments are of subleading order  $O(a^2\xi^{r-2})$  as compared to the even moments which are of  $O(a\xi^{r-1})$ . Therefore, in the limit of small gaps  $\xi \gg a$ , the odd moments are of no interest in the ML gauge and can be neglected.

We note that corresponding relations of  $\tilde{C}_{r\alpha}$  or  $\tilde{C}_{w/d,r\alpha}$  to the Dirac theory are not possible since the Dirac Wannier functions have a divergence  $|\tilde{w}_{\tau p}^{(v)}(x)|^2 \sim 1/x^2$  for  $x \rightarrow 0$  (see Sec. V A). Therefore, it is important to define the moments of the Dirac Wannier functions around the reference point  $x = 0$ , otherwise they contain a divergence.

## V. UNIVERSALITY OF WANNIER FUNCTIONS

This section is devoted to the most important result of this work that, in the case of small gaps, all lattice Wannier functions in AF or ML gauge show universal scaling in terms of a small set of universal scaling functions determining the shape of the Dirac Wannier functions in the corresponding gauges. To obtain this result we first calculate all eigenstates and the Wannier functions of the Dirac model analytically in Sec. V A. The fundamental universal scaling functions in the Dirac theory are then introduced in Sec. V B, together with a summary of all their symmetry properties and asymptotic forms. The moments of the Dirac Wannier functions are discussed analytically in Sec. V C, both in the AF and ML gauges. In Sec. V D we derive the universal scaling form of all lattice Wannier functions and their moments in terms of the fundamental scaling functions of the Dirac theory. The universal scaling is demonstrated explicitly for two examples with unit cells of sizes  $2a$  and  $3a$  in Secs. V D 1 and V D 2, respectively. Finally, Sec. V E contains a qualitative discussion of the properties of lattice Wannier functions and the scaling of their moments on all length scales, in particular including the one at small scales of the order of the lattice spacing. This analysis shows clearly that the visual impression of the square of lattice Wannier functions reveals only the trivial gapless limit, leading to the misleading visual impression of a localized wave function with spread determined by the lattice spacing. In contrast, the whole universal scaling properties on length scale  $\xi$  are only visible when multiplying the lattice Wannier function with the spatial coordinate and then squaring it.

### A. Explicit evaluation of Dirac Wannier functions

The eigenfunctions (2.23) of the Dirac model (2.27) in the AF and ML gauges are given by

$$|\tilde{\chi}_{q\tau}^{(v)}\rangle = e^{i\phi_{q\tau}^{(v)}} |\chi_{q\tau}^{(v)}\rangle = \frac{1}{\sqrt{2\epsilon_q^{(v)}}} \begin{pmatrix} g_{q\tau}^{(v)} e^{i\gamma^{(v)}} \\ \tau g_{q\bar{\tau}}^{(v)} \end{pmatrix}, \quad (5.1)$$

where  $\bar{\tau} = -\tau$ ,

$$\epsilon_q^{(v)} = \sqrt{(v_F^{(v)} q)^2 + (\Delta^{(v)})^2}, \quad g_{q\tau}^{(v)} = \sqrt{\epsilon_q^{(v)} + \tau v_F^{(v)} q}, \quad (5.2)$$

and the gauge factor is given by

$$e^{i\phi_{q\tau}^{(v)}} = \frac{g_{q\tau}^{(v)} e^{i\gamma^{(v)}} + \tau g_{q\bar{\tau}}^{(v)}}{|g_{q\tau}^{(v)} e^{i\gamma^{(v)}} + \tau g_{q\bar{\tau}}^{(v)}|}, \quad (5.3)$$

where

$$|g_{q\tau}^{(v)} e^{i\gamma^{(v)}} + \tau g_{q\bar{\tau}}^{(v)}| = \sqrt{2(\epsilon_q^{(v)} + \tau \Delta^{(v)} \cos \gamma^{(v)})}. \quad (5.4)$$

We note the helpful properties

$$g_{q\tau}^{(v)} g_{q\bar{\tau}}^{(v)} = \Delta^{(v)}, \quad g_{-q,\tau}^{(v)} = g_{q\bar{\tau}}^{(v)}. \quad (5.5)$$

As required for the AF gauge by (2.42), we find with (2.32) a positive  $Z$  component for  $u_{k\alpha}(Za)$ :

$$u_{k\alpha}(Za) = \frac{1}{\sqrt{Z}} \sum_{p=\pm} \chi_{qp}^{(v)} = \frac{|g_{q\tau}^{(v)} e^{i\gamma^{(v)}} + \tau g_{q\bar{\tau}}^{(v)}|}{\sqrt{2Z\epsilon_q^{(v)}}} > 0. \quad (5.6)$$

We note that the gauge factor (5.3) allows for a unique and analytic definition of the phase  $\phi_{q\tau}^{(v)}$  since

$$\text{sgn}\{\text{Im}(g_{q\tau}^{(v)} e^{i\gamma^{(v)}} + \tau g_{q\bar{\tau}}^{(v)})\} = s_{\gamma^{(v)}}. \quad (5.7)$$

Therefore, the phase  $\phi_{q\tau}^{(v)}$  can be chosen to have the same sign as  $\gamma^{(v)}$  for all  $q$  (note that we take  $-\pi < \gamma^{(v)} < \pi$ ). Using

$$e^{i\phi_{q+}^{(v)}} \rightarrow \begin{cases} e^{i\gamma^{(v)}} & \text{for } q \rightarrow \infty, \\ 1 & \text{for } q \rightarrow -\infty, \end{cases} \quad (5.8)$$

$$e^{i\phi_{q-}^{(v)}} \rightarrow \begin{cases} -1 & \text{for } q \rightarrow \infty, \\ e^{i\gamma^{(v)}} & \text{for } q \rightarrow -\infty, \end{cases} \quad (5.9)$$

we get from

$$A_{q\tau}^{(v)} = \partial_q \phi_{q\tau}^{(v)} = \frac{v_F^{(v)} \Delta^{(v)} \sin \gamma^{(v)}}{2\epsilon_q^{(v)} (\epsilon_q^{(v)} + \tau \Delta^{(v)} \cos \gamma^{(v)})} \quad (5.10)$$

the result (3.31) for the Zak-Berry phases.

We note that  $\sin \gamma^{(v)} = 0$  is a special point where the Dirac Zak-Berry connection is zero. For a half-infinite system with  $x > 0$  it can be shown [17] that an edge state is present in the gap for  $\gamma^{(v)} > 0$  with energy  $\epsilon_E^{(v)} = -\Delta^{(v)} \cos \gamma^{(v)}$ . Therefore, at this special point, the edge-state energy touches either the higher (for  $\gamma^{(v)} = \pm\pi$ ) or the lower band (for  $\gamma^{(v)} = 0$ ), and the Wannier functions of the corresponding bands have very special properties, they do not decay exponentially [for  $w_{\tau p}^{(v)}(x)$  in AF gauge] or change discontinuously [for  $\tilde{w}_{\tau p}^{(v)}(x)$  in ML gauge]. Therefore, we exclude the cases  $\gamma^{(v)} = 0$ ,  $\tau = -$  and  $\gamma^{(v)} = \pm\pi$ ,  $\tau = +$  in the following and discuss them separately in Appendix G.

For the phase factors  $e^{-i\phi_{0,\tau}^{(v)}}$  occurring in (4.52) and (4.53) we get from (5.3) the result

$$\begin{aligned} e^{-i\phi_{0,\tau}^{(v)}} &= \frac{e^{-i\gamma^{(v)}} + \tau}{|e^{-i\gamma^{(v)}} + \tau|} \\ &= e^{-i\gamma^{(v)}/2} \begin{cases} 1 & \text{for } \tau = +, \\ -i s_{\gamma^{(v)}} & \text{for } \tau = -, \end{cases} \end{aligned} \quad (5.11)$$

where we used the convention  $-\pi < \gamma^{(v)} < \pi$ .

Using (5.1), (5.3), and (5.4), we find for  $\chi_{q\tau p}^{(v)}$  the form

$$\chi_{q\tau p}^{(v)} = \frac{\epsilon_q^{(v)} + p\tau v_F^{(v)} q + \tau \Delta^{(v)} e^{ip\gamma^{(v)}}}{2\sqrt{\epsilon_q^{(v)}(\epsilon_q^{(v)} + \tau \Delta^{(v)} \cos \gamma^{(v)})}}. \quad (5.12)$$

Inserting the forms (5.1) and (5.12) for  $\tilde{\chi}_{q\tau}^{(v)}$  and  $\chi_{q\tau}^{(v)}$  in (4.33), and introducing the dimensionless variables

$$\bar{q} = 2\xi^{(v)} q, \quad \bar{x} = \frac{x}{2\xi^{(v)}}, \quad \bar{\epsilon}_q = \frac{\epsilon_q^{(v)}}{\Delta^{(v)}} = \sqrt{1 + \bar{q}^2}, \quad (5.13)$$

with  $\xi^{(v)} = v_F^{(v)}/(2\Delta^{(v)})$ , we find for the Wannier functions the integral representation

$$w_{\tau p}^{(v)}(x) = \frac{1}{4\pi\xi^{(v)}} \int d\bar{q} e^{-\bar{\eta}|\bar{q}|} e^{i\bar{q}\bar{x}} \chi_{\tau p}^{(v)}(\bar{q}), \quad (5.14)$$

$$\tilde{w}_{\tau p}^{(v)}(x) = \frac{1}{4\pi\xi^{(v)}} \int d\bar{q} e^{-\bar{\eta}|\bar{q}|} e^{i\bar{q}\bar{x}} \tilde{\chi}_{\tau p}^{(v)}(\bar{q}), \quad (5.15)$$

where

$$\chi_{\tau p}^{(v)}(\bar{q}) = \frac{\bar{\epsilon}_q + p\tau\bar{q} + \tau e^{ip\gamma^{(v)}}}{2\sqrt{\bar{\epsilon}_q}\sqrt{\bar{\epsilon}_q + \tau \cos \gamma^{(v)}}}, \quad (5.16)$$

$$\tilde{\chi}_{\tau p}^{(v)}(\bar{q}) = \frac{\sqrt{\bar{\epsilon}_q + p\tau\bar{q}}}{\sqrt{2\bar{\epsilon}_q}} \begin{cases} e^{i\gamma^{(v)}} & \text{for } p = +, \\ \tau & \text{for } p = -. \end{cases} \quad (5.17)$$

We have included a convergence factor  $e^{-\bar{\eta}|\bar{q}|}$ , with  $\bar{\eta} \sim a/\xi^{(v)}$ , since the integrals diverge for large  $|\bar{q}|$ . This divergence occurs since a low-energy theory can only describe the universal regime  $|x| \gg a$ . As shown in Appendix E, one can also find a more convenient integral representation by closing the integration contour for  $s_x = \text{sgn}(x) = \pm$  in the upper and lower half, respectively, leading to convergent integrals such that the limit  $\bar{\eta} \rightarrow 0$  can be performed under the integral. This regularization is possible for all  $x \neq 0$  and removes an unphysical contribution  $\sim \delta(x)$  from the Dirac Wannier function (see Appendices C and D). In this way we obtain the universal representation

$$w_{\tau p}^{(v)}(x) = \frac{is_x}{4\pi\xi^{(v)}} \int_0^\infty d\kappa e^{-|\bar{x}|\kappa} \delta\chi_{\tau p}^{(v)}(is_x\kappa), \quad (5.18)$$

$$\tilde{w}_{\tau p}^{(v)}(x) = \frac{is_x}{4\pi\xi^{(v)}} \int_0^\infty d\kappa e^{-|\bar{x}|\kappa} \delta\tilde{\chi}_{\tau p}^{(v)}(is_x\kappa), \quad (5.19)$$

where we defined  $s_x = \text{sgn}(x)$ , and

$$\delta\chi_{\tau p}^{(v)}(is_x\kappa) = \chi_{\tau p}^{(v)}(is_x\kappa + 0^+) - \chi_{\tau p}^{(v)}(is_x\kappa - 0^+), \quad (5.20)$$

$$\delta\tilde{\chi}_{\tau p}^{(v)}(is_x\kappa) = \tilde{\chi}_{\tau p}^{(v)}(is_x\kappa + 0^+) - \tilde{\chi}_{\tau p}^{(v)}(is_x\kappa - 0^+) \quad (5.21)$$

describe the jump of the integrands across the branch cut along the imaginary axis which emerge from the various square roots. Alternatively, we show in Appendix D how the Wannier functions can be represented via convergent momentum integrals on the real axis.

The results (5.18) and (5.19) provide the form of the Wannier functions in the universal regime  $|x|, \xi^{(v)} \gg a$ . However, as long as  $\xi^{(v)} \gg a$ , the lattice Wannier functions at  $x = ma$  can be reproduced for *all*  $m$  (even for  $m = 0$ ) from the Dirac Wannier functions via the sum of (4.44) and (4.45) (in the AF gauge) or the sum of (4.52) and (4.53) (in the ML gauge).

As shown in Appendix C, this arises from the fact that the high-momentum region  $|q| > \Lambda_\alpha$  does not contribute to the total Wannier function of a certain band  $\alpha$ .

## B. Universal scaling functions

It is convenient to express the complex Dirac Wannier functions  $w_{\tau p}^{(v)}(x)$  and  $\tilde{w}_{\tau p}^{(v)}(x)$  in terms of real functions  $w_{A/B,\tau}^{(v)}(x)$  and  $\tilde{w}_{A/B,\tau}^{(v)}(x)$  by writing

$$w_{\tau p}^{(v)}(x) = \frac{1}{2} [w_{B,\tau}^{(v)}(x) + ip w_{A,\tau}^{(v)}(x)], \quad (5.22)$$

$$\tilde{w}_{\tau p}^{(v)}(x) = e^{i\phi_{0,\tau}^{(v)}} \frac{1}{2} [\tilde{w}_{B,\tau}^{(v)}(x) + ip \tilde{w}_{A,\tau}^{(v)}(x)], \quad (5.23)$$

which correspond to the definitions

$$w_{A,\tau}^{(v)}(x) = -i \sum_p p w_{\tau p}^{(v)}(x) = w_{A,\tau}^{(v)}(x)^*, \quad (5.24)$$

$$w_{B,\tau}^{(v)}(x) = \sum_p w_{\tau p}^{(v)}(x) = w_{B,\tau}^{(v)}(x)^*, \quad (5.25)$$

$$\tilde{w}_{A,\tau}^{(v)}(x) = -i e^{-i\phi_{0,\tau}^{(v)}} \sum_p p \tilde{w}_{\tau p}^{(v)}(x) = \tilde{w}_{A,\tau}^{(v)}(x)^*, \quad (5.26)$$

$$\tilde{w}_{B,\tau}^{(v)}(x) = e^{-i\phi_{0,\tau}^{(v)}} \sum_p \tilde{w}_{\tau p}^{(v)}(x) = \tilde{w}_{B,\tau}^{(v)}(x)^*. \quad (5.27)$$

The fundamental universal and dimensionless scaling functions are then defined by

$$F_{A/B,\tau}(x/\xi^{(v)}; \gamma^{(v)}) = x w_{A/B,\tau}^{(v)}(x), \quad (5.28)$$

$$\tilde{F}_{A/B,\tau}(x/\xi^{(v)}; \gamma^{(v)}) = x \tilde{w}_{A/B,\tau}^{(v)}(x), \quad (5.29)$$

in terms of which the moments (4.37) can be expressed as

$$C_{r\tau}^{(v)} = \frac{1}{2} (\xi^{(v)})^{r-1} \int dy y^{r-2} \{F_{A,\tau}(y)^2 + F_{B,\tau}(y)^2\}, \quad (5.30)$$

$$\tilde{C}_{r\tau}^{(v)} = \frac{1}{2} (\xi^{(v)})^{r-1} \int dy y^{r-2} \{\tilde{F}_{A,\tau}(y)^2 + \tilde{F}_{B,\tau}(y)^2\}. \quad (5.31)$$

The explicit form of the scaling functions in terms of integrations along the real momentum axis are derived in Appendix D and are explicitly presented in (D7)–(D12). These useful forms allow for a direct numerical evaluation.

In Appendix F we have listed a number of properties of the Dirac Wannier functions  $w_{\tau p}^{(v)}(x)$  and  $\tilde{w}_{\tau p}^{(v)}(x)$  under inversion of  $p$ ,  $\tau$ , or  $x$  and certain transformations of  $\gamma^{(v)}$ , together with corresponding properties of  $w_{A/B,\tau}^{(v)}(x)$  and  $\tilde{w}_{A/B,\tau}^{(v)}(x)$ . As a consequence we note the following useful properties of the universal scaling functions  $F_{A/B,\tau}(y; \gamma)$  and  $\tilde{F}_{A/B,\tau}(y; \gamma)$  under inversion of  $y$ ,  $\tau$ , or  $\gamma$  and under the transformation  $\gamma \rightarrow -\gamma + \pi s_\gamma$  [alternatively, these relations follow directly from the explicit forms (D7)–(D12) of the scaling functions]

$$F_{A,\tau}(-y; \gamma) = F_{A,\tau}(y; -\gamma), \quad (5.32)$$

$$F_{B,\tau}(-y; \gamma) = -F_{B,\tau}(y; \gamma) = -F_{B,\tau}(y; -\gamma), \quad (5.33)$$

$$\tilde{F}_{A,\tau}(-y; \gamma) = \tilde{F}_{A,\tau}(y; -\gamma) = -\tau s_\gamma \tilde{F}_{B,\tau}(y; -\gamma + \pi s_\gamma) \quad (5.34)$$

$$= \tau \{\cos \gamma \tilde{F}_{A,\tau}(y; \gamma) - \sin \gamma \tilde{F}_{B,\tau}(y; \gamma)\}, \quad (5.35)$$

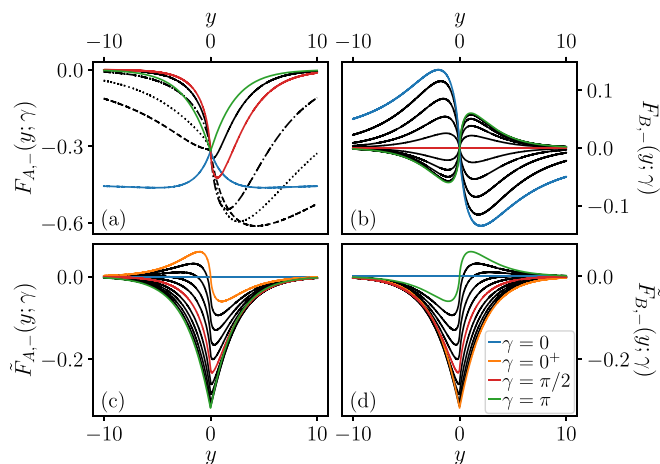


FIG. 3. The scaling functions  $F_{A/B,-}(y; \gamma)$  and  $\tilde{F}_{A/B,-}(y; \gamma)$  for various phases  $\gamma$ . We show only  $\tau = -$  and consider  $0 \leq \gamma \leq \pi$  since all other cases follow from the transformation laws (5.32)–(5.41) as  $F_{A/B,+}(y; \gamma) = \mp F_{A/B,-}(y; \pi - \gamma)$ ,  $\tilde{F}_{A/B,+}(y; \gamma) = \mp \tilde{F}_{A/B,-}(y; \pi - \gamma)$ ,  $F_{A/B,\tau}(y; -\gamma) = \pm F_{A/B,\tau}(y; \gamma)$ , and  $\tilde{F}_{A/B,\tau}(y; -\gamma) = \pm \tilde{F}_{A/B,\tau}(y; \gamma)$ . According to (5.34), we note the relation  $\tilde{F}_{B,-}(y; \gamma) = \tilde{F}_{A,-}(y; \pi - \gamma)$ . All scaling functions are exponentially decaying for  $|y| \gg 1$ , except for  $F_{A,-}(y; 0) \rightarrow -\sqrt{2}/\pi$  and  $F_{B,-}(y; 0) \rightarrow -\sqrt{2}/(\pi y)$  [see (5.55) and (5.56)]. The scaling functions  $\tilde{F}_{A/B,-}(y; \gamma)$  change discontinuously at  $\gamma = 0$  by a sign change [see (5.57) and (5.58)]. The cases  $\gamma = 0, 0^+, \pi/2, \pi$  are indicated by blue, orange, red, and green, respectively. The grid for the values of  $\gamma$  is defined by  $0.1\pi n$  with  $n = 0, 1, \dots, 10$  in (b)–(d). In (a) the black curves show  $\gamma = 0.05\pi$  (dashed line),  $0.1\pi$  (dotted),  $0.2\pi$  (dashed-dotted), and  $0.75\pi$  (solid).

$$\tilde{F}_{B,\tau}(-y; \gamma) = -\tilde{F}_{B,\tau}(y; -\gamma) = -\tau s_\gamma \tilde{F}_{A,\tau}(y; -\gamma + \pi s_\gamma) \quad (5.36)$$

$$= -\tau \{ \sin \gamma \tilde{F}_{A,\tau}(y; \gamma) + \cos \gamma \tilde{F}_{B,\tau}(y; \gamma) \}, \quad (5.37)$$

$$F_{A,-\tau}(y; \gamma) = -F_{A,\tau}(y; -\gamma + \pi s_\gamma), \quad (5.38)$$

$$F_{B,-\tau}(y; \gamma) = F_{B,\tau}(y; -\gamma + \pi s_\gamma), \quad (5.39)$$

$$\tilde{F}_{A,-\tau}(y; \gamma) = \tau s_\gamma \tilde{F}_{B,\tau}(-y; \gamma), \quad (5.40)$$

$$\tilde{F}_{B,-\tau}(y; \gamma) = -\tau s_\gamma \tilde{F}_{A,\tau}(-y; \gamma). \quad (5.41)$$

For the special case  $\sin \gamma = 0$  we list all properties in Appendix G.

The typical forms of the four fundamental scaling functions  $F_{A/B,-}(y; \gamma)$  and  $\tilde{F}_{A/B,-}(y; \gamma)$  are illustrated in Fig. 3. We show only the lower band  $\tau = -$  and positive values of the phase  $\gamma$  since a sign change of  $\tau$  or  $\gamma$  is covered by the above properties. A special case is  $\cos \gamma = 0$ , where

$$F_{B,\tau}(y; \gamma) = 0 \quad \text{for } \cos \gamma = 0. \quad (5.42)$$

The asymptotic forms for small  $|y| \ll 1$  and large  $|y| \gg 1$  can be analyzed analytically (see Appendix E), where we derive the corresponding asymptotic behavior of the Dirac Wannier functions in the AF and ML gauges for small  $|x| \ll \xi^{(v)}$  and large  $|x| \gg \xi^{(v)}$ . For  $y = 0$  we get from (E9) and (E10),

$$F_{A,\tau}(0; \gamma) = \frac{\tau}{\pi}, \quad (5.43)$$

$$F_{B,\tau}(0; \gamma) = 0, \quad (5.44)$$

$$\tilde{F}_{A,\tau}(0; \gamma) = \frac{\tau}{2\pi} |e^{i\gamma} + \tau|, \quad (5.45)$$

$$\tilde{F}_{B,\tau}(0; \gamma) = -\frac{s_\gamma}{2\pi} |e^{i\gamma} - \tau|. \quad (5.46)$$

The asymptotic behavior for large  $|y| \gg 1$  follows from (E14), (E16), and (E18) as

$$\tau = \text{sgn}(\cos \gamma):$$

$$F_{A,\tau}(y; \gamma) \rightarrow \frac{\tau(1 + s_y \sin \gamma) \Gamma(\frac{3}{4})}{2\pi \sqrt{|\cos \gamma|}} |y|^{\frac{1}{4}} e^{-\frac{|y|}{2}}, \quad (5.47)$$

$$F_{B,\tau}(y; \gamma) \rightarrow \frac{\Gamma(\frac{3}{4}) \sqrt{|\cos \gamma|}}{2\pi} s_y |y|^{\frac{1}{4}} e^{-\frac{|y|}{2}}, \quad (5.48)$$

$$\tau = -\text{sgn}(\cos \gamma):$$

$$F_{A,\tau}(y; \gamma) \rightarrow \frac{\tau}{\sqrt{2\pi} |\sin \gamma|} |y|^{-\frac{1}{2}} e^{-|\sin \gamma| \frac{|y|}{2}} + \tau(1 + s_y s_\gamma) \sqrt{\frac{|\sin \gamma|}{2\pi}} |y|^{\frac{1}{2}} e^{-|\sin \gamma| \frac{|y|}{2}}, \quad (5.49)$$

$$F_{B,\tau}(y; \gamma) \rightarrow -\frac{\sqrt{|\sin \gamma|}}{\sqrt{2\pi} |\cos \gamma|} s_y |y|^{-\frac{1}{2}} e^{-|\sin \gamma| \frac{|y|}{2}}, \quad (5.50)$$

$$\cos \gamma = 0:$$

$$F_{A,\tau}(y; \gamma) \rightarrow \frac{\tau}{2\sqrt{\pi}} (|y|^{-\frac{1}{2}} + (1 + s_y s_\gamma) |y|^{\frac{1}{2}}) e^{-\frac{|y|}{2}}, \quad (5.51)$$

$$F_{B,\tau}(y; \gamma) = 0. \quad (5.52)$$

We note that the last result (5.52) holds for all  $y$  in the special case  $\cos \gamma = 0$  [see (5.42)]. From (E20) we get for all cases

$$\tilde{F}_{A,\tau}(y; \gamma) \rightarrow \frac{\Gamma(\frac{3}{4})}{4\pi} |y|^{\frac{1}{4}} e^{-\frac{|y|}{2}} \tau (s_y s_\gamma |e^{i\gamma} - \tau| + |e^{i\gamma} + \tau|), \quad (5.53)$$

$$\tilde{F}_{B,\tau}(y; \gamma) \rightarrow \frac{\Gamma(\frac{3}{4})}{4\pi} |y|^{\frac{1}{4}} e^{-\frac{|y|}{2}} (s_y |e^{i\gamma} + \tau| - s_\gamma |e^{i\gamma} - \tau|). \quad (5.54)$$

As discussed after (5.10), we remind that the cases  $\gamma = 0, \tau = -$  and  $\gamma = \pm\pi, \tau = +$  are excluded. At these points the scaling functions  $F_{A/B,\tau}(y; \gamma)$  are not exponentially decaying,

$$F_{A,-}(y; 0) = -F_{A,+}(y; \pi) \rightarrow -\frac{\sqrt{2}}{\pi}, \quad (5.55)$$

$$F_{B,-}(y; 0) = F_{B,+}(y; \pi) \rightarrow -\frac{\sqrt{2}}{\pi} \frac{1}{y}, \quad (5.56)$$

and the scaling functions  $\tilde{F}_{A/B,\tau}(y; \gamma)$  change discontinuously by a sign change (with vanishing value exactly at  $\gamma = 0, \pi$ )

$$\tilde{F}_{A,-}(y; 0^\pm) = \mp \tilde{F}_{B,+}(y; 0), \quad (5.57)$$

$$\tilde{F}_{B,-}(y; 0^\pm) = \mp \tilde{F}_{A,+}(y; 0), \quad (5.58)$$

$$\tilde{F}_{A,+}(y; \pi \pm 0^+) = \mp \tilde{F}_{B,-}(y; \pi), \quad (5.59)$$

$$\tilde{F}_{B,+}(y; \pi \pm 0^+) = \mp \tilde{F}_{A,-}(y; \pi) \quad (5.60)$$



(see Appendix G for details). Although it might have been expected that the Wannier functions in the ML gauge change discontinuously as function of  $\gamma$  when an edge state touches the band edge, it is quite remarkable that the Wannier functions in the AF gauge stay continuous but show a non-exponential decay exactly at the touching point although the gap is still present. Up to our best knowledge this has not been reported before.

In summary, we find an exponential decay with a preexponential power law for  $|y| \gg 1$  (similar power laws for the preexponential functions have been obtained in Ref. [41] for a variety of other localized single-particle wave functions) whereas, for  $|y| \lesssim 1$ , we obtain the scaling

$$|y| \lesssim 1 \Rightarrow F_{A/B,\tau}(y; \gamma), \tilde{F}_{A/B,\tau}(y; \gamma) \sim O(1). \quad (5.61)$$

This scaling together with the exponential behavior at large  $y$  is the essential reason why the moments scale as

$$C_{r\tau}^{(v)}, \tilde{C}_{r\tau}^{(v)} \sim (\xi^{(v)})^{r-1}, \quad (5.62)$$

leading via (4.49) and (4.54) to the universal scaling  $a\xi$  (with  $\xi \sim \xi^{(\alpha)}, \xi^{(\alpha-1)}$ ) for the quadratic spread of the lattice Wannier functions. This is in contrast to the visual impression of the lattice Wannier functions leading to the incorrect scaling  $a^2$ , as we will discuss in all detail in Sec. VE.

### C. Moments of Dirac Wannier functions

The moments in the ML gauge are given by

$$\begin{aligned} \tilde{C}_{r\tau}^{(v)} &= \sum_{p=\pm} \int \frac{dq}{2\pi} (\tilde{\chi}_{q\tau p}^{(v)})^* (i\partial_q)^r \tilde{\chi}_{q\tau p}^{(v)} \\ &= \frac{i^r (2\xi^{(v)})^{r-1}}{4\pi} \sum_{p=\pm} \int d\bar{q} \sqrt{1 + \frac{p\bar{q}}{\bar{\epsilon}_q}} \partial_{\bar{q}}^r \sqrt{1 + \frac{p\bar{q}}{\bar{\epsilon}_q}}. \end{aligned} \quad (5.63)$$

Noticing that only even moments ( $r = 2l$ ) are nonzero, we rewrite this expression in the symmetrized form

$$\tilde{C}_{2l,\tau}^{(v)} = (\xi^{(v)})^{2l-1} \tilde{c}_{2l}, \quad (5.64)$$

$$\tilde{c}_{2l} = \frac{2^{2l-2}}{2\pi} \sum_{p=\pm} \int d\bar{q} \left[ \partial_{\bar{q}}^l \sqrt{1 + \frac{p\bar{q}}{\bar{\epsilon}_q}} \right]^2 > 0. \quad (5.65)$$

In particular, we find that the moments are independent of  $\tau$  and depend on  $\nu$  only via the decay length  $\xi^{(v)}$ .

The scaling behavior of  $\tilde{C}_{2l,\tau}^{(v)}$  is given by (5.64), while Eq. (5.65) provides positive dimensionless coefficients. Making the integration variable change  $\bar{q} = \tan \theta$ , we conveniently represent them in the form

$$\tilde{c}_{2l} = \frac{1}{4\pi} \int_{-\pi}^{\pi} d\theta \left[ \cos \theta \left( \frac{\partial}{\partial \theta} 2 \cos^2 \theta \right)^{l-1} \sin \frac{\theta}{2} \right]^2. \quad (5.66)$$

In particular, the second moment is recovered from

$$\tilde{C}_{2\tau}^{(v)} = \frac{\xi^{(v)}}{4\pi} \int_{-\pi}^{\pi} d\theta \cos^2 \theta \sin^2 \frac{\theta}{2} = \frac{\xi^{(v)}}{8}. \quad (5.67)$$

At large  $l$  the sequence (5.66) is well approximated by  $\frac{(2l-2)!}{e}$  (see Fig. 4).

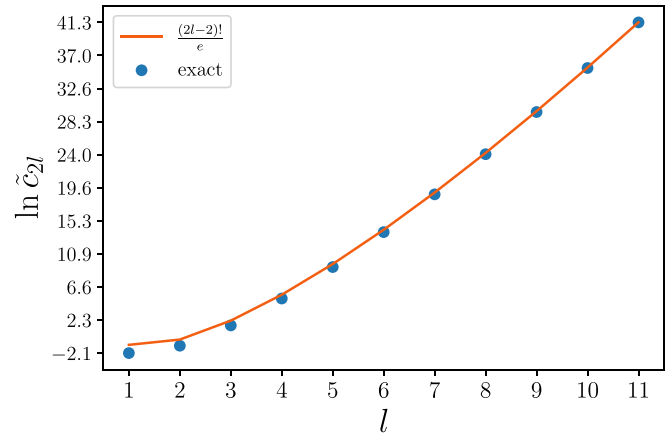


FIG. 4. Logarithmic plot of first 10 coefficients  $\tilde{c}_{2l}$  and of their asymptotic value  $\frac{(2l-2)!}{e}$ . As one can notice, the sequence is very well approximated by its asymptotic even for moderate values of  $l$ .

In the AF gauge, the moments (4.38) are more complicated and can be expressed as

$$\begin{aligned} C_{r\tau}^{(v)} &= \sum_{p=\pm} \int \frac{dq}{2\pi} (\chi_{q\tau p}^{(v)})^* (i\partial_q)^r \chi_{q\tau p}^{(v)} \\ &= i^r \sum_{p=\pm} \int \frac{dq}{2\pi} (\tilde{\chi}_{q\tau p}^{(v)})^* e^{i\phi_{q\tau}^{(v)}} \partial_q^r [e^{-i\phi_{q\tau}^{(v)}} \tilde{\chi}_{q\tau p}^{(v)}], \end{aligned} \quad (5.68)$$

such that

$$C_{r\tau}^{(v)} = (\xi^{(v)})^{r-1} c_{r\tau}^{(v)}, \quad (5.69)$$

$$\begin{aligned} c_{r\tau}^{(v)} &= i^r 2^{r-1} \sum_{p=\pm} \int \frac{d\bar{q}}{4\pi} \sqrt{1 + \frac{p\bar{q}}{\bar{\epsilon}_q}} e^{i\phi_{q\tau}^{(v)}} \\ &\quad \times \sum_{s=0}^r \frac{r!}{s!(r-s)!} \left[ \partial_{\bar{q}}^{r-s} e^{-i\phi_{q\tau}^{(v)}} \partial_{\bar{q}}^s \sqrt{1 + \frac{p\bar{q}}{\bar{\epsilon}_q}} \right]. \end{aligned} \quad (5.70)$$

In contrast to the ML gauge, the dimensionless coefficients  $c_{r\tau}^{(v)}$  depend on  $\tau$  and  $\nu$  in the AF gauge via the phase factors involving  $\phi_{q\tau}^{(v)}$ .

In particular, we find

$$\begin{aligned} C_{1\tau}^{(v)} &= \int \frac{dq}{2\pi} \partial_q \phi_{q\tau}^{(v)} \\ &= \begin{cases} \frac{\gamma^{(v)}}{2\pi}, & \tau = + \\ -\frac{\gamma^{(v)}}{2\pi} + \frac{1}{2} S_{\gamma^{(v)}}, & \tau = - \end{cases} \end{aligned} \quad (5.71)$$

which agrees with (4.39) and (3.31), and

$$\begin{aligned} C_{2\tau}^{(v)} &= \tilde{C}_{2\tau}^{(v)} + \int \frac{dq}{2\pi} [\partial_q \phi_{q\tau}^{(v)}]^2 \\ &= \frac{\xi^{(v)}}{8} + \frac{\xi^{(v)}}{4} \left[ \tan^2 \gamma^{(v)} + \frac{4 \cot 2\gamma^{(v)} \Theta(-\tau)}{\cos \gamma^{(v)}} S_{\gamma^{(v)}} \right. \\ &\quad \left. + \frac{2\tau(2\gamma^{(v)} \cot 2\gamma^{(v)} - 1)}{\pi \cos \gamma^{(v)}} \right]. \end{aligned} \quad (5.72)$$

One finds that  $C_{2\tau}^{(v)} > \tilde{C}_{2\tau}^{(v)} = \xi^{(v)}/8$  as expected and a divergence when the first moment (5.71) jumps (for  $\tau = +$  at  $\gamma^{(v)} = \pm\pi$ , and for  $\tau = -$  at  $\gamma^{(v)} = 0$ ), compare with the discussion after (5.10).

**D. Scaling properties of lattice Wannier functions**

In this section we reveal the scaling properties of the lattice Wannier functions as given by (4.44), (4.45), (4.52), and (4.53) via the Dirac Wannier functions. Inserting the decompositions (5.22) and (5.23) of the Dirac Wannier functions in real and imaginary parts, and using

$$\cos\left(\frac{\nu\pi}{Z}m\right) = (-1)^{\nu(n-1)} \cos\left(\pi \frac{\nu j}{Z}\right), \quad \sin\left(\frac{\nu\pi}{Z}m\right) = (-1)^{\nu(n-1)} \sin\left(\pi \frac{\nu j}{Z}\right), \tag{5.73}$$

where  $m = Z(n - 1) + j$ , we obtain

$$\frac{1}{\sqrt{Z}}(-1)^{\alpha(n-1)}m w_{u,\alpha}(n, j) = F_{-}\left(\frac{ma}{\xi^{(\alpha)}}; \frac{\alpha j}{Z}, \gamma^{(\alpha)}\right), \tag{5.74}$$

$$\frac{1}{\sqrt{Z}}(-1)^{(\alpha-1)(n-1)}m w_{d,\alpha}(n, j) = F_{+}\left(\frac{ma}{\xi^{(\alpha-1)}}; \frac{(\alpha-1)j}{Z}, \gamma^{(\alpha-1)}\right), \tag{5.75}$$

$$\frac{1}{\sqrt{Z}}(-1)^{\alpha(n-1)}\tilde{m}_{\alpha} \tilde{w}_{u,\alpha}(n, j) = \tilde{F}_{-}\left(\frac{\tilde{m}_{\alpha}a}{\xi^{(\alpha)}}; \frac{\alpha j}{Z}, \gamma^{(\alpha)}\right), \tag{5.76}$$

$$\frac{1}{\sqrt{Z}}(-1)^{(\alpha-1)(n-1)}\tilde{m}_{\alpha} \tilde{w}_{d,\alpha}(n, j) = \tilde{F}_{+}\left(\frac{\tilde{m}_{\alpha}a}{\xi^{(\alpha-1)}}; \frac{(\alpha-1)j}{Z}, \gamma^{(\alpha-1)}\right), \tag{5.77}$$

where  $\alpha = 1, \dots, Z - 1$  for  $w_{u,\alpha}$  and  $\tilde{w}_{u,\alpha}$ , and  $\alpha = 2, \dots, Z$  for  $w_{d,\alpha}$  and  $\tilde{w}_{d,\alpha}$ . The scaling functions  $F_{\tau}$  and  $\tilde{F}_{\tau}$  are defined in the following way in terms of the fundamental scaling functions  $F_{A/B,\tau}$  and  $\tilde{F}_{A/B,\tau}$  introduced in Sec. VB:

$$F_{\tau}(y; s, \gamma) = F_{B,\tau}(y; \gamma) \cos(\pi s) - F_{A,\tau}(y; \gamma) \sin(\pi s), \quad \tilde{F}_{\tau}(y; s, \gamma) = \tilde{F}_{B,\tau}(y; \gamma) \cos(\pi s) - \tilde{F}_{A,\tau}(y; \gamma) \sin(\pi s). \tag{5.78}$$

As a result, for given value  $j = 1, \dots, Z$  of the site index within a unit cell, the up and down parts of the lattice Wannier functions reveal universal scaling with a single length scale. By adding the up and down parts we get for the total Wannier function of band  $\alpha$

$$\frac{1}{\sqrt{Z}}(-1)^{\alpha(n-1)}m w_{\alpha}(n, j) = F_{-}\left(\frac{ma}{\xi^{(\alpha)}}; \frac{\alpha j}{Z}, \gamma^{(\alpha)}\right) + (-1)^{n-1}F_{+}\left(\frac{ma}{\xi^{(\alpha-1)}}; \frac{(\alpha-1)j}{Z}, \gamma^{(\alpha-1)}\right), \tag{5.79}$$

$$\frac{1}{\sqrt{Z}}(-1)^{\alpha(n-1)}\tilde{m}_{\alpha} \tilde{w}_{\alpha}(n, j) = \tilde{F}_{-}\left(\frac{\tilde{m}_{\alpha}a}{\xi^{(\alpha)}}; \frac{\alpha j}{Z}, \gamma^{(\alpha)}\right) + (-1)^{n-1}\tilde{F}_{+}\left(\frac{\tilde{m}_{\alpha}a}{\xi^{(\alpha-1)}}; \frac{(\alpha-1)j}{Z}, \gamma^{(\alpha-1)}\right), \tag{5.80}$$

where we leave out the second (first) term on the right-hand side for  $\alpha = 1$  ( $\alpha = Z$ ). For  $\alpha = 2, \dots, Z - 1$  universal scaling appears with two different length scales corresponding to the gaps at the bottom and the top of the band. Therefore, to reveal universal scaling, one has to keep the ratio of the two length scales fixed.

For the universal scaling of the moments

$$C_{u/d,r\alpha}(Ma) = \sum_{m=-M}^M (ma)^r |w_{u/d,\alpha}(ma)|^2, \quad \tilde{D}_{u/d,r\alpha}(Ma) = \sum_{m=-M}^M (\tilde{m}_{\alpha}a)^r |\tilde{w}_{u/d,\alpha}(ma)|^2, \tag{5.81}$$

we get from the above equations after neglecting the strongly oscillating terms

$$\frac{C_{u,r\alpha}(Ma)}{Za(\xi^{(\alpha)})^{r-1}} \approx \int_{\frac{-Ma}{\xi^{(\alpha)}}}^{\frac{Ma}{\xi^{(\alpha)}}} dy y^{r-2} G_{-}(y; \gamma^{(\alpha)}), \quad \frac{C_{d,r\alpha}(Ma)}{Za(\xi^{(\alpha-1)})^{r-1}} \approx \int_{\frac{-Ma}{\xi^{(\alpha-1)}}}^{\frac{Ma}{\xi^{(\alpha-1)}}} dy y^{r-2} G_{+}(y; \gamma^{(\alpha-1)}), \tag{5.82}$$

$$\frac{\tilde{D}_{u,r\alpha}(Ma)}{Za(\xi^{(\alpha)})^{r-1}} \approx \int_{\frac{-Ma}{\xi^{(\alpha)}}}^{\frac{Ma}{\xi^{(\alpha)}}} dy y^{r-2} \tilde{G}_{-}(y; \gamma^{(\alpha)}), \quad \frac{\tilde{D}_{d,r\alpha}(Ma)}{Za(\xi^{(\alpha-1)})^{r-1}} \approx \int_{\frac{-Ma}{\xi^{(\alpha-1)}}}^{\frac{Ma}{\xi^{(\alpha-1)}}} dy y^{r-2} \tilde{G}_{+}(y; \gamma^{(\alpha-1)}), \tag{5.83}$$

where

$$G_{\tau}(y; \gamma) = \frac{1}{2}\{|F_{A,\tau}(y; \gamma)|^2 + |F_{B,\tau}(y; \gamma)|^2\}, \quad \tilde{G}_{\tau}(y; \gamma) = \frac{1}{2}\{|\tilde{F}_{A,\tau}(y; \gamma)|^2 + |\tilde{F}_{B,\tau}(y; \gamma)|^2\}, \tag{5.84}$$

or explicitly in terms of the right and left moving Dirac Wannier functions

$$G_{\tau}(x/\xi^{(v)}; \gamma^{(v)}) = x^2 \sum_p |w_{\tau p}^{(v)}(x)|^2, \tag{5.85}$$

$$\tilde{G}_{\tau}(x/\xi^{(v)}; \gamma^{(v)}) = x^2 \sum_p |\tilde{w}_{\tau p}^{(v)}(x)|^2. \tag{5.86}$$

By using (5.35) and (5.37) we note that  $\tilde{G}_{\tau}$  is a symmetric function

$$\tilde{G}_{\tau}(-y; \gamma) = \tilde{G}_{\tau}(y; \gamma). \tag{5.87}$$

As a result, the right-hand sides of (5.83) are exactly zero for odd values of  $r$ . As mentioned already at the end of Sec. IV C 2, this does not mean that the lattice moments  $\tilde{D}_{u/d,r\alpha}(Ma)$  are zero for odd values of  $r$ . It only means that by dividing them by the leading order  $a\xi^{r-1}$  [with  $\xi \equiv \xi^{(\alpha)}$  for  $\tilde{D}_{u,r\alpha}(Ma)$  and  $\xi \equiv \xi^{(\alpha-1)}$  for  $\tilde{D}_{d,r\alpha}(Ma)$ ], one gets zero in the limit  $\xi \rightarrow \infty$ . This is in contrast to the even moments in the ML gauge which stay finite in this limit after divided by this order. Therefore, the odd moments in the ML gauge are negligible and are not considered in the following. In the AF gauge the function  $G_\tau(y; \gamma)$  is asymmetric due to the asymmetry of the scaling function  $F_{A,\tau}(y; \gamma)$  [see (5.32) and Fig. 3], with a significant larger part for either positive or negative values of  $y$ . Therefore, in the AF gauge, all moments are of order  $a\xi^{r-1}$  and stay finite in the limit  $\xi \rightarrow \infty$  when divided by this order.

The total moment for  $\alpha = 2, \dots, Z-1$  is obtained from the sum [neglecting strongly oscillating terms involving  $(-1)^n$ ]

$$C_{r\alpha}(Ma) \approx C_{u,r\alpha}(Ma) + C_{d,r\alpha}(Ma), \quad (5.88)$$

$$\tilde{D}_{r\alpha}(Ma) \approx \tilde{D}_{u,r\alpha}(Ma) + \tilde{D}_{d,r\alpha}(Ma), \quad (5.89)$$

where we leave out the second (first) term on the right-hand side for  $\alpha = 1$  ( $\alpha = Z$ ).

According to (4.49) and (4.54), we get for the asymptotic value  $M \rightarrow \infty$

$$C_{u,r\alpha}(Ma) \rightarrow Za C_{r-}^{(\alpha)} = Za(\xi^{(\alpha)})^{r-1} c_{r-}^{(\alpha)}, \quad (5.90)$$

$$C_{d,r\alpha}(Ma) \rightarrow Za C_{r+}^{(\alpha-1)} = Za(\xi^{(\alpha-1)})^{r-1} c_{r+}^{(\alpha-1)}, \quad (5.91)$$

$$\tilde{D}_{u,2l,\alpha}(Ma) \rightarrow Za \tilde{C}_{2l,-}^{(\alpha)} = Za(\xi^{(\alpha)})^{2l-1} \tilde{c}_{2l}, \quad (5.92)$$

$$\tilde{D}_{d,2l,\alpha}(Ma) \rightarrow Za \tilde{C}_{2l,+}^{(\alpha-1)} = Za(\xi^{(\alpha-1)})^{2l-1} \tilde{c}_{2l}, \quad (5.93)$$

where we used (5.69) and (5.64). According to (5.71) and (5.67) we get for  $r = 1, 2$

$$\frac{C_{u,1\alpha}(Ma)}{Za} \rightarrow -\frac{\gamma^{(\alpha)}}{2\pi} + \frac{1}{2} s_{\gamma^{(\alpha)}}, \quad (5.94)$$

$$\frac{C_{d,1\alpha}(Ma)}{Za} \rightarrow \frac{\gamma^{(\alpha-1)}}{2\pi}, \quad (5.95)$$

$$\frac{\tilde{D}_{u,2\alpha}(Ma)}{Za\xi^{(\alpha)}} \rightarrow \frac{1}{8}, \quad (5.96)$$

$$\frac{\tilde{D}_{d,2\alpha}(Ma)}{Za\xi^{(\alpha-1)}} \rightarrow \frac{1}{8}. \quad (5.97)$$

Using (5.88) and (5.89) together with (3.30) and (3.31), we find for the asymptotic values of the total moments

$$\frac{C_{1\alpha}(Ma)}{Za} \rightarrow \frac{\gamma_\alpha}{2\pi}, \quad (5.98)$$

$$\frac{\tilde{D}_{2\alpha}(Ma)}{Za\xi^{(\alpha)}} \rightarrow \frac{1}{8} \left( 1 + \frac{\xi^{(\alpha-1)}}{\xi^{(\alpha)}} \right). \quad (5.99)$$

In the following two subsections we will demonstrate the universal scaling of the lattice Wannier functions and their moments for two examples with  $Z = 2$  and 3. For simplicity, we assume that the dominant contribution to the gap param-

eter (2.21) is the term  $H'$  of the effective Hamiltonian (2.20). This means that all Fourier components

$$\delta\tilde{v}_v = \frac{1}{Z} \sum_{j=1}^Z e^{-i2\pi jv/Z} \delta v_j, \quad (5.100)$$

$$\delta\tilde{t}_v = \frac{1}{Z} \sum_{j=1}^Z e^{-i2\pi jv/Z} \delta t_j \quad (5.101)$$

have approximately the same order of magnitude. By convention we set

$$\delta\tilde{v}_0 = \delta\tilde{t}_0 = 0. \quad (5.102)$$

By inserting the form (2.3) of  $H'$  and using (2.18) for the eigenstates of  $H_0$ , it is then straightforward to show that

$$\Delta^{(v)} e^{i\gamma^{(v)}} = \delta\tilde{v}_v - 2e^{-i\pi v/Z} \delta\tilde{t}_v. \quad (5.103)$$

The  $2(Z-1)$  lattice parameters  $\delta v_j$  and  $\delta t_j$ , with  $j = 1, \dots, Z$  and  $\sum_j \delta v_j = \sum_j \delta t_j = 0$ , are then fixed via the  $Z-1$  complex gap parameters  $\Delta^{(v)} e^{i\gamma^{(v)}}$ , with  $v = 1, \dots, Z-1$ , by using (5.103) and the inverse Fourier transform of (5.100) and (5.101):

$$\delta v_j = \sum_{v=1}^{Z-1} e^{i2\pi jv/Z} \delta\tilde{v}_v, \quad (5.104)$$

$$\delta t_j = \sum_{v=1}^{Z-1} e^{i2\pi jv/Z} \delta\tilde{t}_v. \quad (5.105)$$

### 1. Universal scaling for $Z=2$

For  $Z = 2$  (the so-called Rice-Mele model [39,42] which, for  $\delta v_{1/2} = 0$ , turns into the Su-Schrieffer-Heeger model [43]) we show in Figs. 1(a) and 1(b) the universal scaling of the Wannier functions  $w(ma) \equiv w_1(ma) = w_{u,1}(ma)$  and  $\tilde{w}(ma) \equiv \tilde{w}_1(ma) = \tilde{w}_{u,1}(ma)$  of the lowest band  $\alpha = 1$ , together with the scaling of their first and second moment, respectively. The lattice parameters are fixed by the complex gap parameter  $\Delta e^{i\gamma} \equiv \Delta^{(1)} e^{i\gamma^{(1)}}$ . The lattice Wannier functions are calculated numerically from their definitions (4.5) and (4.7) via the Bloch states  $u_{k1}(ma)$  and the gauge phase  $\varphi_{k1}$ . The latter follows from (3.13) via the Zak-Berry connection  $A_{k1}$ . The Bloch states and the Zak-Berry connection are calculated analytically in Appendix H. The Zak-Berry phase  $\gamma_1$  of the lowest band can either be calculated from its definition (3.9) via numerical integration over the Zak-Berry connection or from the approximate formulas (3.30) and (3.31) as

$$\gamma_1 \approx \gamma_-^{(1)} = -\gamma^{(1)} + \pi s_{\gamma^{(1)}}. \quad (5.106)$$

As shown in Fig. 5 the two ways to calculate the Zak-Berry phase coincide quite well in the low-energy regime of small gaps. Therefore, it makes no visible difference in Fig. 1(b) which choice is taken to calculate the shift variable  $\tilde{m}_1 = m - \gamma_1/\pi$ .

With  $m = 2(n-1) + j$  and  $j = 1, 2$ , we find from (5.74) and (5.78) for odd and even values of  $m$  (i.e., for  $j = 1, 2$ )

$$\frac{1}{2} |m w(ma)|^2 \approx \begin{cases} |F_{A,-}(y; \gamma)|^2 & \text{for } m \text{ odd,} \\ |F_{B,-}(y; \gamma)|^2 & \text{for } m \text{ even,} \end{cases} \quad (5.107)$$

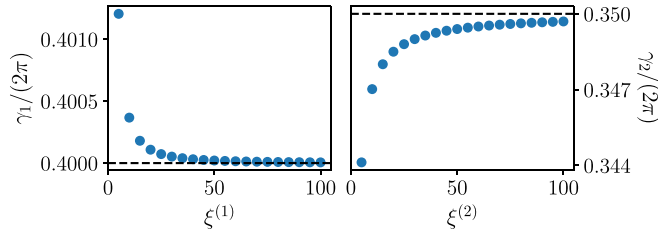


FIG. 5. The Zak-Berry phase  $\gamma_\alpha/(2\pi)$  calculated from (3.9) as function of  $\xi^{(\alpha)}$  for  $\alpha = 1$ ,  $Z = 2$  (left panel) and for  $\alpha = 2$ ,  $Z = 3$  (right panel). The parameters are the same as in Figs. 1 and 10, respectively. For all  $\xi^{(\alpha)}$  used in these figures we find that the deviation from the approximation  $\gamma_1/(2\pi) \approx 0.4$  [left panel, see (5.106)] and  $\gamma_2 \approx \frac{7}{20} = 0.35$  [right panel, see (5.115)] is less than  $\approx 0.02\%$ .

$$\frac{1}{2} |\tilde{m} \tilde{w}(ma)|^2 \approx \begin{cases} |\tilde{F}_{A,-}(\tilde{y}; \gamma)|^2 & \text{for } m \text{ odd,} \\ |\tilde{F}_{B,-}(\tilde{y}; \gamma)|^2 & \text{for } m \text{ even,} \end{cases} \quad (5.108)$$

where we used the abbreviations

$$y \equiv ma/\xi, \quad \tilde{y} \equiv \tilde{m}a/\xi, \quad (5.109)$$

with  $\tilde{m} \equiv \tilde{m}_1 = m - \gamma_1/\pi$ . As a consequence, the four fundamental scaling functions  $F_{A/B,-}(y; \gamma)$  and  $\tilde{F}_{A/B,-}(y; \gamma)$  show up naturally in the scaling behavior of the lower band (in AF or ML gauge, respectively) of the Rice-Mele model for odd (A) and even (B) sites. Similarly, for the upper band the corresponding scaling functions with  $\tau = +$  appear, which are related via (5.38)–(5.41) to the scaling functions with  $\tau = -$ . For larger values of  $Z > 2$ , the same scaling functions appear but in subtle combinations, as we will demonstrate in the next subsection for  $Z = 3$ .

In the main panels of Figs. 1(a) and 1(b) we reveal the universal scaling properties by plotting the left-hand side of Eqs. (5.107) and (5.108) for different values of  $\xi$  as function of  $y = ma/\xi$  or  $\tilde{y} = \tilde{m}a/\xi$ , and find that all curves fall on top of the universal functions  $|F_{A/B,-}(y; \gamma)|^2$  and  $|\tilde{F}_{A/B,-}(y; \gamma)|^2$  (for  $m$  odd or even), respectively [note that we omitted the subindex  $\tau = -$  and the superindex (1) for all quantities used in the caption of these figures]. Although (5.107) and (5.108) are expected to hold only for large  $|m| \gg 1$  and for small gaps (where  $\xi \gg a$ ), it is remarkable that the coincidence is even quite well for  $|m| \sim O(1)$  and for large gaps [where  $\xi \sim O(a)$ ]. We will come back to this point in Sec. VE, where we discuss the behavior for small scales  $m \sim O(1)$ . All the universal scalings are in accordance with the symmetries and asymptotic behaviors discussed for the scaling functions in Sec. VB (compare with Fig. 3). For even  $m$ , the function  $|m w(ma)|^2$  scales symmetrically with zero value at  $m = 0$ ; all other cases are asymmetric and have a finite value at  $m = 0$  (or  $\tilde{m} = 0$ ), in accordance with (5.43)–(5.46). For large  $|m| \gg 1$  all functions are exponentially decaying with a preexponential power-law  $\sim |ma/\xi|^r \exp(-|ma|/\xi)$ . Since  $\gamma = 0.2\pi$  and  $\tau = -$ , we get  $\tau = -\text{sgn}(\cos \gamma) \neq 0$ , leading via (5.49) and (5.50) to  $r = \pm 1$  for  $|m w(ma)|^2$  and odd  $m \geq 0$ , and to  $r = -1$  for  $|m w(ma)|^2$  and even  $m$ . For  $|\tilde{m} \tilde{w}(ma)|^2$  we get  $r = 0.5$  for both  $m$  even or odd [see (5.53) and (5.54)].

In the insets of Figs. 1(a) and 1(b) we show the universal scaling of the first moment  $C_1(Ma)/(2a)$  of  $w(ma)$  and of the

quadratic spread  $\tilde{D}_2(Ma)/(2a\xi)$  of  $\tilde{w}(ma)$  (with  $C_1 \equiv C_{11} = C_{u,11}$  and  $\tilde{D}_2 \equiv \tilde{D}_{21} = \tilde{D}_{u,21}$ ) when plotted against  $Ma/\xi$  for different  $\xi$ . As demonstrated, all discrete points fall on top of the universal curves (5.82) (with  $r = 1$ ) and (5.83) (with  $r = 2$ ), i.e.,

$$\frac{C_1(Ma)}{2a} \approx \int_{-Ma/\xi}^{Ma/\xi} dy \frac{1}{y} G_-(y; \gamma), \quad (5.110)$$

$$\frac{\tilde{D}_2(Ma)}{2a\xi} \approx \int_{-Ma/\xi}^{Ma/\xi} dy \tilde{G}_-(y; \gamma), \quad (5.111)$$

with  $\tilde{G}_-(y; \gamma) = \tilde{G}_-(-y; \gamma)$  due to (5.87). In accordance with (5.94) and (5.96) they converge smoothly to the values

$$\frac{C_1}{2a} = \lim_{M \rightarrow \infty} \frac{C_1(Ma)}{2a} = \frac{\gamma_1}{2\pi}, \quad (5.112)$$

$$\frac{\tilde{D}_2}{2a\xi} = \lim_{M \rightarrow \infty} \frac{\tilde{D}_2(Ma)}{2a\xi} = \frac{1}{8}. \quad (5.113)$$

The last result gives the universal value for the quadratic spread  $\langle \Delta x^2 \rangle = \tilde{D}_2 = 2a\xi/8$  of the Wannier function  $\tilde{w}(ma)$  in the ML gauge (for arbitrary size  $Za$  of the unit cell it turns into  $Za\xi/8$  for the lowest band).

## 2. Universal scaling for $Z=3$

For  $Z = 3$  the Bloch states and the Zak-Berry connection are calculated analytically in Appendix H. In contrast to the case  $Z = 2$ , two gap parameters  $\Delta^{(v)} e^{i\gamma^{(v)}}$ , with  $v = 1, 2$ , are needed to fix the lattice parameters. They correspond to the two gaps and contain different phases  $\gamma^{(v)}$  and different length scales  $\xi^{(v)} = v_F^{(v)}/(2\Delta^{(v)})$ . For the Zak-Berry phases of the three bands we obtain approximately from (3.30) and (3.31)

$$\gamma_1 \approx \gamma_-^{(1)} = -\gamma^{(1)} + \pi s_{\gamma^{(1)}}, \quad (5.114)$$

$$\gamma_2 \approx \gamma_+^{(1)} + \gamma_-^{(2)} = \gamma^{(1)} - \gamma^{(2)} + \pi s_{\gamma^{(2)}}, \quad (5.115)$$

$$\gamma_3 \approx \gamma_+^{(2)} = \gamma^{(2)}, \quad (5.116)$$

where, as shown in Fig. 5, it makes a negligible difference for small gaps whether one takes these approximate formulas to calculate the shift variables  $\tilde{m}_\alpha = m - \frac{3}{2\pi} \gamma_\alpha$  or the precise definition (3.9) by integrating over the lattice Zak-Berry connection.

According to (5.74) and (5.76), the scaling of the Wannier functions of the first band involves the scaling functions  $F_-(ma/\xi^{(1)}; j/3; \gamma^{(1)})$  and  $\tilde{F}_-(\tilde{m}_1 a/\xi^{(1)}; j/3; \gamma^{(1)})$ , with  $j = 1, 2, 3$ . Similar to the first band for  $Z = 2$ , only a single length scale  $\xi^{(1)}$  appears, but the scaling is different since, for  $j \neq 3$ , linear combinations of the scaling functions  $F_{A,-}(y; \gamma^{(1)})$  and  $F_{B,-}(y; \gamma^{(1)})$  [or  $\tilde{F}_{A,-}(y; \gamma^{(1)})$  and  $\tilde{F}_{B,-}(y; \gamma^{(1)})$ ] will occur [see Eqs. (5.78) and (5.121)–(5.123)]. However, aside from this no other change of the scaling appears.

Of particular interest is the scaling of the second band  $\alpha = 2$  where two different length scales occur for the up and down

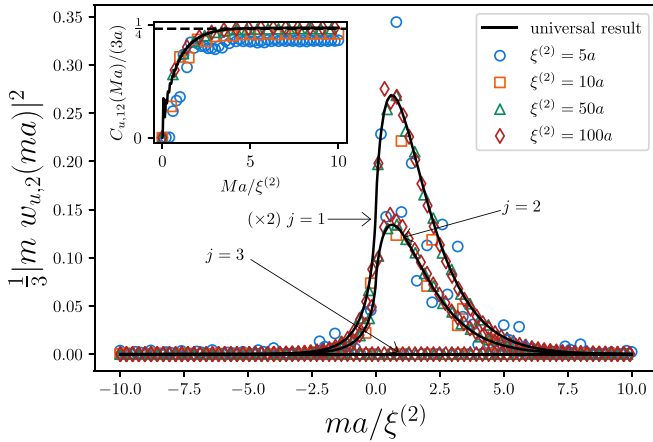


FIG. 6. Scaling of  $\frac{1}{3}|m w_{u,2}(n, j)|^2$  for  $\alpha = 2$ ,  $Z = 3$ ,  $\gamma^{(1)} = 0.2\pi$ ,  $\gamma^{(2)} = 0.5\pi$ , and  $\xi^{(1)} = 30a$ , as function of  $ma/\xi^{(2)}$  for different values of  $\xi^{(2)}$  and  $j$ . According to (5.117) we obtain the universal curve  $|F_-(ma/\xi^{(2)}; 2j/3, \gamma^{(2)})|^2$ . The inset shows the scaling of the first moment  $C_{u,1\alpha}(Ma)/(3a)$  for  $\alpha = 2$ ,  $Z = 3$  as function of  $s_2 = Ma/\xi^{(2)}$ . According to (5.127) we obtain the universal curve  $\int_{-s_2}^{s_2} dy y^{-1} G_-(y; \gamma^{(2)})$  with saturation at the correct value  $-\frac{\gamma^{(2)}}{2\pi} + \frac{1}{2}s_{2,\gamma^{(2)}} = \frac{1}{4}$ .

components. From (5.74)–(5.77) we find

$$\frac{1}{3}|m w_{u,2}(n, j)|^2 \approx \left| F_-\left(\frac{ma}{\xi^{(2)}}; \frac{2j}{3}, \gamma^{(2)}\right) \right|^2, \quad (5.117)$$

$$\frac{1}{3}|m w_{d,2}(n, j)|^2 \approx \left| F_+\left(\frac{ma}{\xi^{(1)}}; \frac{j}{3}, \gamma^{(1)}\right) \right|^2, \quad (5.118)$$

$$\frac{1}{3}|\tilde{m}_2 \tilde{w}_{u,2}(n, j)|^2 \approx \left| \tilde{F}_-\left(\frac{\tilde{m}_2 a}{\xi^{(2)}}; \frac{2j}{3}, \gamma^{(2)}\right) \right|^2, \quad (5.119)$$

$$\frac{1}{3}|\tilde{m}_2 \tilde{w}_{d,2}(n, j)|^2 \approx \left| \tilde{F}_+\left(\frac{\tilde{m}_2 a}{\xi^{(1)}}; \frac{j}{3}, \gamma^{(1)}\right) \right|^2, \quad (5.120)$$

i.e., the up and down components scale with  $\xi^{(2)}$  and  $\xi^{(1)}$ , respectively, corresponding to the gap at the top and bottom of the second band. From (5.78) we find that  $F_\tau$  involves the following combinations of the scaling functions  $F_{A,\tau}$  and  $F_{B,\tau}$ :

$$\begin{aligned} F_\tau\left(y; \frac{1}{3}, \gamma\right) &= -F_\tau\left(y; \frac{4}{3}, \gamma\right) \\ &= \frac{1}{2}\{F_{B,\tau}(y; \gamma) - \sqrt{3}F_{A,\tau}(y; \gamma)\}, \end{aligned} \quad (5.121)$$

$$F_\tau\left(y; \frac{2}{3}, \gamma\right) = -\frac{1}{2}\{F_{B,\tau}(y; \gamma) + \sqrt{3}F_{A,\tau}(y; \gamma)\}, \quad (5.122)$$

$$F_\tau(y; 1, \gamma) = -F_\tau(y; 2, \gamma) = -F_{B,\tau}(y; \gamma), \quad (5.123)$$

and similar equations for  $F_\tau \rightarrow \tilde{F}_\tau$  and  $F_{A/B,\tau} \rightarrow \tilde{F}_{A/B,\tau}$ . The scaling of the four functions (5.117)–(5.120) is demonstrated in Fig. 6 (up component in AF gauge), Fig. 7 (down component in AF gauge), Fig. 8 (up component in ML gauge), and Fig. 9 (down component in ML gauge), for all components  $j = 1, 2, 3$ , using the choice  $\gamma^{(1)} = 0.2\pi$  and  $\gamma^{(2)} = \pi/2$ . Due to the special choice  $\gamma^{(2)} = \pi/2$  it turns out that

$$\frac{1}{3}|m w_{u,2}(n, 3)|^2 \approx |F_{B,-}(y; \pi/2)|^2 = 0 \quad (5.124)$$

due to (5.117), (5.123), and (5.52), leading to the vanishing  $j = 3$  component in Fig. 6.

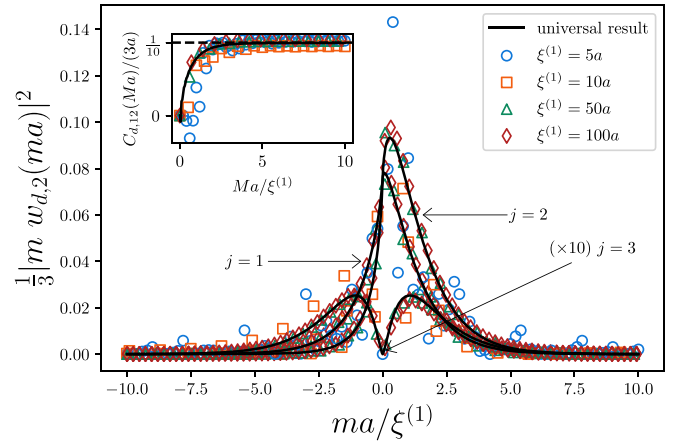


FIG. 7. Scaling of  $\frac{1}{3}|m w_{d,2}(n, j)|^2$  for  $\alpha = 2$ ,  $Z = 3$ ,  $\gamma^{(1)} = 0.2\pi$ ,  $\gamma^{(2)} = 0.5\pi$ , and  $\xi^{(2)} = 30a$ , as function of  $ma/\xi^{(1)}$  for different values of  $\xi^{(1)}$  and  $j$ . According to (5.118) we obtain the universal curve  $|F_+(ma/\xi^{(1)}; j/3, \gamma^{(1)})|^2$ . The inset shows the scaling of the first moment  $C_{d,1\alpha}(Ma)/(3a)$  for  $\alpha = 2$ ,  $Z = 3$  as function of  $s_1 = Ma/\xi^{(1)}$ . According to (5.128) we obtain the universal curve  $\int_{-s_1}^{s_1} dy y^{-1} G_+(y; \gamma^{(1)})$  with saturation at the correct value  $\frac{\gamma^{(1)}}{2\pi} = \frac{1}{10}$ .

In Figs. 10 and 11 we show the scaling of the total Wannier functions  $|m w_2(ma)|^2$  and  $|\tilde{m}_2 \tilde{w}_2(ma)|^2$  of the second band, respectively. Since two different length scales appear, the ratio  $\xi^{(1)}/\xi^{(2)}$  has to be kept fixed and universality is demonstrated for different choices of  $\xi^{(2)}$ . Furthermore, since the scaling of the up and down components contains a different sign factor  $(-1)^{n-1}$  [see Eqs. (5.74)–(5.77)], the total Wannier function is the sum or difference of the scaling functions of the up

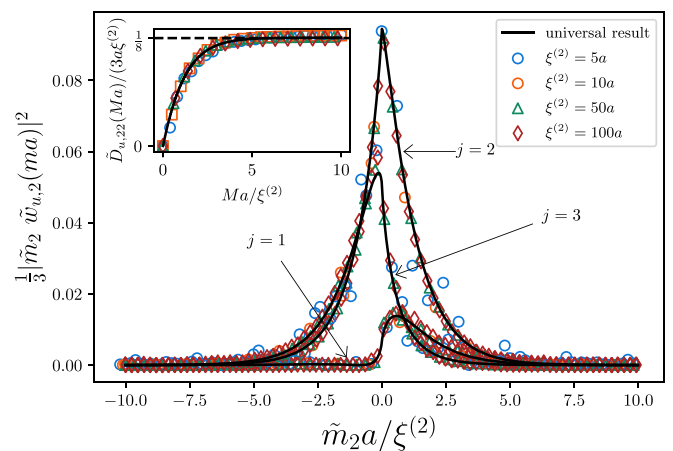


FIG. 8. Scaling of  $\frac{1}{3}|\tilde{m}_2 \tilde{w}_{u,2}(n, j)|^2$  for  $\alpha = 2$ ,  $Z = 3$ ,  $\gamma^{(1)} = 0.2\pi$ ,  $\gamma^{(2)} = 0.5\pi$ , and  $\xi^{(1)} = 30a$ , as function of  $\tilde{m}_2 a/\xi^{(2)}$  for different values of  $\xi^{(2)}$  and  $j$ . According to (5.119) we obtain the universal curve  $|\tilde{F}_-(\tilde{m}_2 a/\xi^{(2)}; 2j/3, \gamma^{(2)})|^2$ . The inset shows the scaling of the second moment  $\tilde{D}_{u,2\alpha}(Ma)/(3a\xi^{(2)})$  for  $\alpha = 2$ ,  $Z = 3$  as function of  $s_2 = Ma/\xi^{(2)}$ . According to (5.129) we obtain the universal curve  $\int_{-s_2}^{s_2} dy y \tilde{G}_-(y; \gamma^{(2)})$  with saturation at the correct value  $\frac{1}{8}$ .

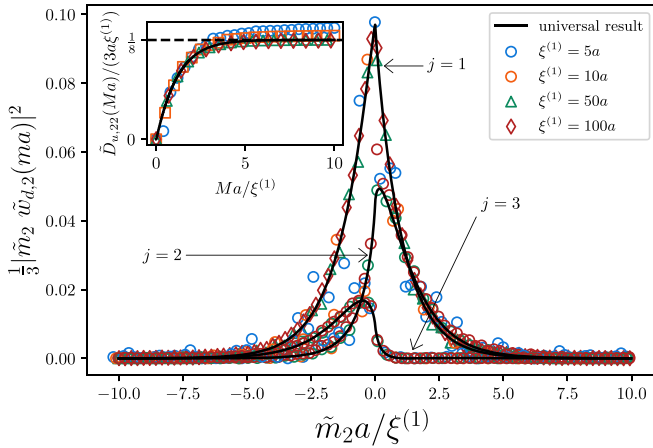


FIG. 9. Scaling of  $\frac{1}{3}|\tilde{m}_2 \tilde{w}_{d,\alpha}(n, j)|^2$  for  $\alpha = 2$ ,  $Z = 3$ ,  $\gamma^{(1)} = 0.2\pi$ ,  $\gamma^{(2)} = 0.5\pi$ , and  $\xi^{(2)} = 30a$ , as function of  $\tilde{m}_2 a/\xi^{(1)}$  for different values of  $\xi^{(1)}$  and  $j$ . According to (5.120) we obtain the universal curve  $|\tilde{F}_+(\tilde{m}_2 a/\xi^{(1)}; j/3, \gamma^{(1)})|^2$ . The inset shows the scaling of the second moment  $\tilde{D}_{u,2\alpha}(Ma)/(Za \xi^{(\alpha-1)})$  for  $\alpha = 2$ ,  $Z = 3$  as function of  $s_1 = Ma/\xi^{(1)}$ . According to (5.130) we obtain the universal curve  $\int_{-s_1}^{s_1} dy \tilde{G}_+(y; \gamma^{(1)})$  with saturation at the correct value  $\frac{1}{8}$ .

and down components depending on the parity of  $n$  [see Eqs. (5.79)–(5.80)]. For  $Z = 3$  and  $\alpha = 2$  one obtains

$$\frac{1}{3}|m w_2(n, j)|^2 = \left| F_-\left(\frac{ma}{\xi^{(2)}}; \frac{2j}{3}, \gamma^{(2)}\right) + (-1)^{n-1} F_+\left(\frac{ma}{\xi^{(1)}}; \frac{j}{3}, \gamma^{(1)}\right) \right|^2, \quad (5.125)$$

$$\frac{1}{3}|\tilde{m}_2 \tilde{w}_2(n, j)|^2 = \left| \tilde{F}_-\left(\frac{\tilde{m}_2 a}{\xi^{(2)}}; \frac{2j}{3}, \gamma^{(2)}\right) + (-1)^{n-1} \tilde{F}_+\left(\frac{\tilde{m}_2 a}{\xi^{(1)}}; \frac{j}{3}, \gamma^{(1)}\right) \right|^2, \quad (5.126)$$

leading to different scaling behavior for  $n$  even or odd.

In the insets of Figs. 6–9 we show the scaling of the first moment (for AF gauge) and the second moment (for ML gauge) of the up and down components of the second band. According to (5.82) and (5.83), they follow the following scaling laws with asymptotic behavior according to (5.94)–(5.97):

$$\begin{aligned} \frac{C_{u,12}(Ma)}{Za} &\approx \int_{-\frac{Ma}{\xi^{(2)}}}^{\frac{Ma}{\xi^{(2)}}} dy y^{-1} G_-(y; \gamma^{(2)}) \\ &\rightarrow -\frac{\gamma^{(2)}}{2\pi} + \frac{1}{2} s_{\gamma^{(2)}}, \end{aligned} \quad (5.127)$$

$$\frac{C_{d,12}(Ma)}{Za} \approx \int_{-\frac{Ma}{\xi^{(1)}}}^{\frac{Ma}{\xi^{(1)}}} dy y^{-1} G_+(y; \gamma^{(1)}) \rightarrow \frac{\gamma^{(1)}}{2\pi}, \quad (5.128)$$

$$\frac{\tilde{D}_{u,22}(Ma)}{Za \xi^{(2)}} \approx \int_{-\frac{Ma}{\xi^{(2)}}}^{\frac{Ma}{\xi^{(2)}}} dy \tilde{G}_-(y; \gamma^{(2)}) \rightarrow \frac{1}{8}, \quad (5.129)$$

$$\frac{\tilde{D}_{d,22}(Ma)}{Za \xi^{(1)}} \approx \int_{-\frac{Ma}{\xi^{(1)}}}^{\frac{Ma}{\xi^{(1)}}} dy \tilde{G}_+(y; \gamma^{(1)}) \rightarrow \frac{1}{8}. \quad (5.130)$$

The scaling of the total moments of the second band in the AF and ML gauges are shown in the insets of Figs. 10 and 11,

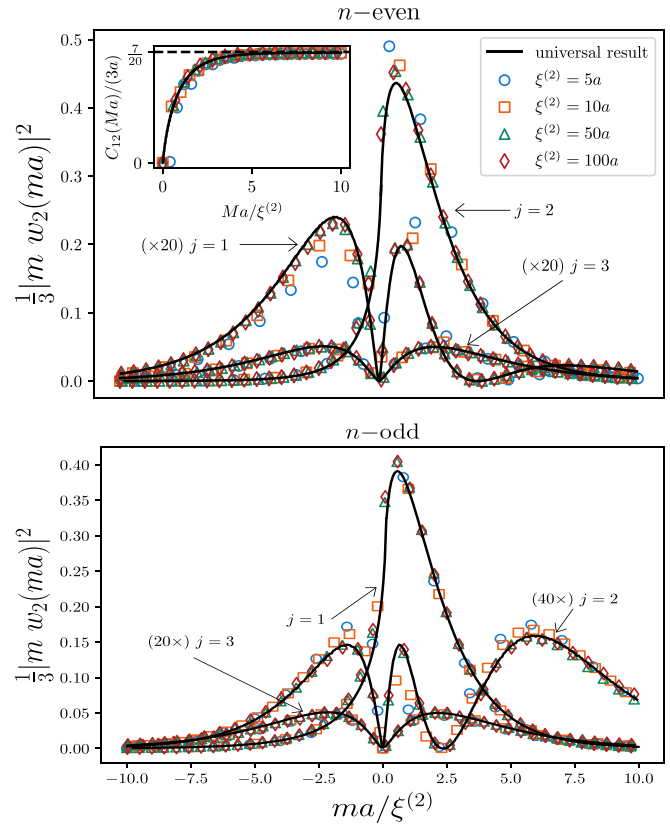


FIG. 10. Scaling of  $\frac{1}{3}|m w_\alpha(n, j)|^2$  for  $\alpha = 2$ ,  $Z = 3$ ,  $\gamma^{(1)} = 0.2\pi$ ,  $\gamma^{(2)} = 0.5\pi$ , and  $\xi^{(1)} = 2\xi^{(2)}$ , as function of  $ma/\xi^{(2)}$  for different values of  $\xi^{(2)}$  and  $j$ , with  $n$  even (odd) for upper (lower) panel. According to (5.125) we obtain the universal curve  $|F_-(\frac{ma}{\xi^{(2)}}; \frac{2j}{3}, \gamma^{(2)}) \pm F_+(\frac{ma}{\xi^{(1)}}; \frac{j}{3}, \gamma^{(1)})|^2$  for  $n$  even or odd. The inset shows the scaling of the first moment  $C_{1\alpha}(Ma)/(Za)$  for  $\alpha = 2$ ,  $Z = 3$  as function of  $s = Ma/\xi^{(2)}$ . According to (5.131), we obtain the universal curve  $\int_{-s}^s dy y^{-1} \tilde{G}_-(y; \gamma^{(2)}) + \int_{-s/2}^{s/2} dy y^{-1} \tilde{G}_+(y; \gamma^{(1)})$  with saturation at the correct value  $\frac{7}{2\pi} = \frac{7}{20}$ , where  $\gamma_2$  has been evaluated from (5.115).

respectively. According to (5.88) and (5.89) they follow from the sum of the up and down components as

$$\begin{aligned} \frac{C_{12}(Ma)}{Za} &\approx \int_{-s_2}^{s_2} dy y^{-1} G_-(y; \gamma^{(2)}) \\ &+ \int_{-s_2/r_{12}}^{s_2/r_{12}} dy y^{-1} G_+(y; \gamma^{(1)}) \rightarrow \frac{\gamma_2}{2\pi}, \end{aligned} \quad (5.131)$$

$$\begin{aligned} \frac{\tilde{D}_{22}(Ma)}{Za \xi^{(2)}} &\approx \int_{-s_2}^{s_2} dy \tilde{G}_-(y; \gamma^{(2)}) \\ &+ r_{12} \int_{-s_2/r_{12}}^{s_2/r_{12}} dy \tilde{G}_+(y; \gamma^{(1)}) \rightarrow \frac{1}{8}(1 + r_{12}), \end{aligned} \quad (5.132)$$

where we defined  $s_2 = Ma/\xi^{(2)}$  and  $r_{12} = \xi^{(1)}/\xi^{(2)}$ . As a result, the ratio  $r_{12}$  of the two length scales has to be kept fixed to see universal scaling by varying  $\xi^{(2)}$ .

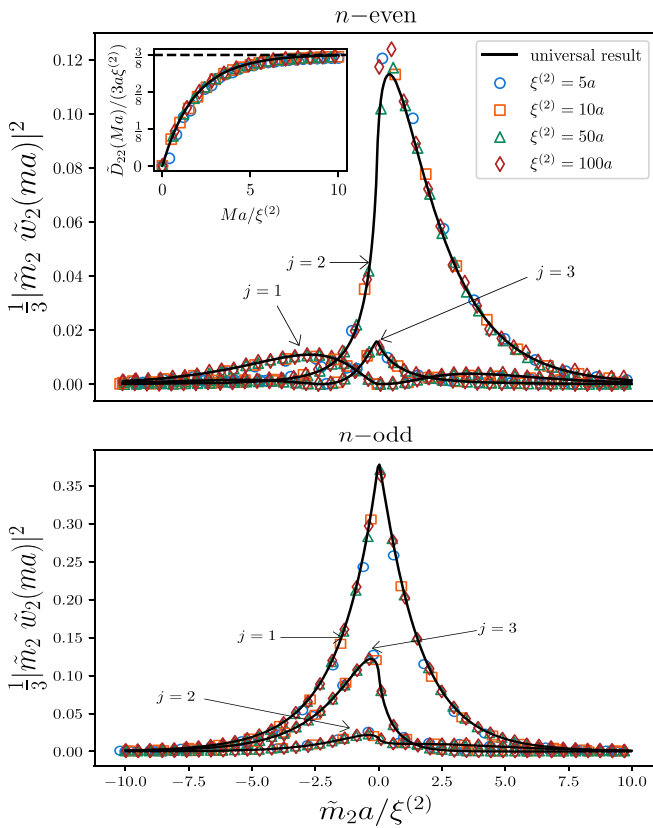


FIG. 11. Scaling of  $\frac{1}{3} |\tilde{m}_\alpha \tilde{w}_\alpha(n, j)|^2$  for  $\alpha = 2$ ,  $Z = 3$ ,  $\gamma^{(1)} = 0.2\pi$ ,  $\gamma^{(2)} = 0.5\pi$ , and  $\xi^{(1)} = 2\xi^{(2)}$ , as function of  $\tilde{m}_2 a / \xi^{(2)}$  for different values of  $\xi^{(2)}$  and  $j$ , with  $n$  even (odd) for upper (lower) panel. According to (5.126) we obtain the universal curve  $|\tilde{F}_-(\frac{\tilde{m}_2 a}{\xi^{(2)}}; \frac{2j}{3}, \gamma^{(2)}) \pm \tilde{F}_+(\frac{\tilde{m}_2 a}{\xi^{(1)}}; \frac{j}{3}, \gamma^{(1)})|^2$  for  $n$  even or odd. The inset shows the scaling of the second moment  $\tilde{D}_{2\alpha}(Ma)/(Za \xi^{(\alpha)})$  for  $\alpha = 2$ ,  $Z = 3$  as function of  $s = Ma/\xi^{(2)}$ . According to (5.132), we obtain the universal curve  $\int_{-s}^s dy \tilde{G}_-(y; \gamma^{(2)}) + 2 \int_{-s/2}^{s/2} dy \tilde{G}_-(y; \gamma^{(1)})$  with saturation at the correct value  $\frac{1}{8}(1 + \xi^{(1)}/\xi^{(2)}) = \frac{3}{8}$ .

### E. Properties of lattice Wannier functions on different scales

In this section we exhibit and compare the important properties of lattice Wannier functions on different scales, always taking the case of small gaps where a clear separation of

length scales is present  $\xi \gg a$  (here,  $\xi$  denotes some typical order of all  $\xi^{(\nu)}$ ,  $\nu = 1, \dots, Z - 1$ ). We distinguish three different regimes, called (F) (for free or gapless case), (S) (for scaling region where the length scale  $\xi$  appears), and (E) (for exponentially decaying region):

$$\text{Regime (F) : } |ma| \ll \xi, \quad (5.133)$$

$$\text{Regime (S) : } |ma| \lesssim \xi, \quad (5.134)$$

$$\text{Regime (E) : } |ma| \gg \xi. \quad (5.135)$$

Regime (E) is the regime where all Wannier functions decay exponentially with some universal preexponential power law [see Eqs. (5.47)–(5.54)]. Therefore, in this regime the Wannier functions have a negligible contribution to the moments and the normalization. Regime (S) is the most important issue of our work where the universal scaling on length scale  $\xi$  is visible. As shown in the previous Sec. VD (see also the insets of Figs. 6–11) all moments  $C_{r\alpha}(Ma)$  (with  $r \geq 1$ ) and  $\tilde{D}_{r\alpha}(Ma)$  (with  $r \geq 2$  even) approach their asymptotic values on the length scale  $\xi$ . Finally, the region (F) is the regime of small length scales where the presence of the gap does not play any role. As a consequence, the Wannier functions are identical to the zero gap limit in this regime (although the phases  $\gamma^{(\nu)}$  from the gap parameter occur in the ML gauge due to the special choice of this gauge). As we will see in the following, the regime (F) is only important for the correct normalization of the Wannier function but is of no significance for the scaling and provides a misleading visual impression of the Wannier function.

In regime (F) of small scales  $|y| \sim |ma|/\xi \ll 1$ , we can use the  $y = 0$  limit of the universal scaling functions, as given by Eqs. (5.43)–(5.46). Using  $|e^{i\gamma} - 1| = 2s_\gamma \sin(\gamma/2)$  and  $|e^{i\gamma} + 1| = 2 \cos(\gamma/2)$  we find for the scaling functions (5.78) the free result at zero gap

$$F_\tau(y; s, \gamma) \approx -\frac{\tau}{\pi} \sin(\pi s), \quad (5.136)$$

$$\tilde{F}_+(y; s, \gamma) \approx -\frac{1}{\pi} \sin(\pi s + \gamma/2), \quad (5.137)$$

$$\tilde{F}_-(y; s, \gamma) \approx -\frac{1}{\pi} s_\gamma \cos(\pi s + \gamma/2). \quad (5.138)$$

Inserting these scaling functions in (5.79) and (5.80) we get for the lattice Wannier functions on small scales  $|ma| \ll \xi$ :

$$\frac{1}{\sqrt{Z}} m w_\alpha(ma) \approx \frac{1}{\pi} \left\{ \sin\left(\frac{\pi\alpha}{Z} m\right) - \sin\left(\frac{\pi(\alpha-1)}{Z} m\right) \right\}, \quad (5.139)$$

$$\frac{1}{\sqrt{Z}} \tilde{m}_\alpha \tilde{w}_\alpha(ma) \approx -\frac{1}{\pi} \left\{ s_{\gamma^{(\alpha)}} \cos\left(\frac{\pi\alpha}{Z} m + \frac{1}{2}\gamma^{(\alpha)}\right) + \sin\left(\frac{\pi(\alpha-1)}{Z} m + \frac{1}{2}\gamma^{(\alpha-1)}\right) \right\}. \quad (5.140)$$

It is a straightforward exercise to see that these are indeed the exact results for the Wannier functions in the absence of a gap. Only due to the special choice of the ML gauge, the phases of the gap parameters enter in (5.140). In Fig. 12 we show the square  $|w_\alpha(ma)|^2$  and  $|\tilde{w}_\alpha(ma)|^2$  of the Wannier functions in the AF and ML gauges as function of  $m$  for the special case  $Z = 3$  and  $\alpha = 2$ . As can be seen, the numerical lattice

result agrees perfectly with the analytical low-energy result (5.79) and (5.80) for large  $\xi \gg a$ , as well as with the gapless result (5.139) and (5.140) in the small-scale regime  $|ma| \ll \xi$ . The scaling region (S) is not visible at all in this figure since the Wannier functions are of order  $1/m^2 \sim (a/\xi)^2 \ll 1$  in this regime. Therefore, the visible impression of the square of the Wannier functions is approximately the result in the absence

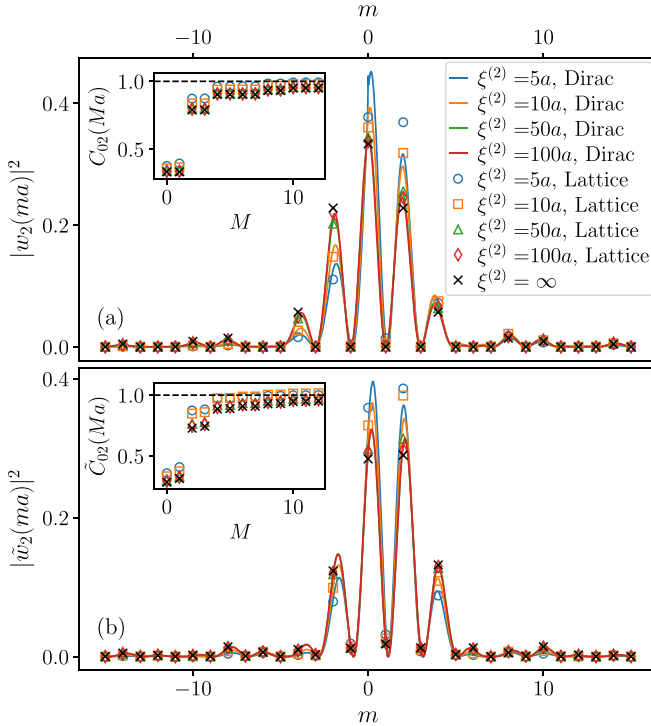


FIG. 12. Comparison between  $|w_\alpha(ma)|^2$  (upper panel) and  $|\tilde{w}_\alpha(ma)|^2$  (lower panel) as function of  $m$  for  $Z = 3$  and  $\alpha = 2$  with  $\xi^{(1)} = 2\xi^{(2)}$ ,  $\gamma^{(1)} = 0.2\pi$ , and  $\gamma^{(2)} = 0.5\pi$  with their Dirac counterpart for a variety of  $\xi^{(2)}$ . The results for large  $\xi^{(2)} \gg a$  are in perfect agreement with the analytical results (5.139) and (5.140) at zero gap. In the two insets we show the scaling of the normalization  $C_{0\alpha}(Ma) = \sum_{m=-M}^M |w_\alpha(ma)|^2$  and  $\tilde{C}_{0\alpha}(Ma) = \sum_{m=-M}^M |\tilde{w}_\alpha(ma)|^2$  for  $\alpha = 2$ . As can be seen, the normalization reaches unity already on very small scales  $M \sim Z$  for large  $\xi^{(2)} \gg a$ .

of the gap. As shown in the left insets of Fig. 12 the region (F) covers almost completely the correct normalization of the Wannier functions, such that the scaling of the normalizations

$$C_{0\alpha}(Ma) = \sum_{m=-M}^M |w_\alpha(ma)|^2, \quad (5.141)$$

$$\tilde{C}_{0\alpha}(Ma) = \sum_{m=-M}^M |\tilde{w}_\alpha(ma)|^2 \quad (5.142)$$

approaches unity already on scales  $M \sim O(Z)$ . In contrast, the region (S) contributes only the negligible order  $\sim (\xi/a)(a/\xi)^2 \sim a/\xi \ll 1$  to the normalization.

We note that even the value at  $m = 0$  is perfectly reproduced by the free results (5.139) and (5.140) and gives an important contribution to the correct normalization. We obtain [note that the right-hand side of (5.139) has to be expanded up to linear order in  $m$  to get the correct result for  $m = 0$ ]

$$w_\alpha(0) = \frac{1}{\sqrt{Z}}, \quad (5.143)$$

$$\tilde{w}_\alpha(0) = \frac{2}{\gamma_\alpha \sqrt{Z}} \left\{ s_{\gamma^{(\alpha)}} \cos\left(\frac{\gamma^{(\alpha)}}{2}\right) + \sin\left(\frac{\gamma^{(\alpha-1)}}{2}\right) \right\}. \quad (5.144)$$

With  $\gamma_\alpha = \gamma^{(\alpha-1)} - \gamma^{(\alpha)} + \pi s_{\gamma^{(\alpha)}}$  one finds a perfect agreement with the results in Fig. 12 for  $\xi^{(\alpha-1)}, \xi^{(\alpha)} \gg a$ .

In contrast to the normalization, all the interesting scaling behavior on the length scale  $\xi$  shows up in the regime (S) and are only visible when multiplying the Wannier function with  $m$  in the AF gauge or  $\tilde{m}_\alpha$  in the ML gauge. In a nutshell, one can express the subtle dependence of the Wannier functions on the two length scales  $a$  and  $\xi$  roughly as follows (in AF or ML gauge for any band). Replacing  $ma$  by the continuous variable  $x$  and rescaling the Wannier function via

$$x \equiv ma, \quad w(x) \equiv \frac{1}{\sqrt{a}} w(ma), \quad (5.145)$$

the qualitative form can be stated as follows (omitting strongly oscillating terms on scale  $Za$  which contribute a negligible amount to the moments):

$$|w(x)|^2 \sim \delta_a(x) f(x/\xi) e^{-|x|/\xi}, \quad (5.146)$$

where  $\delta_a(x)$  is a Lorentzian delta function on scale  $a$ , and  $f(y)$  is a dimensionless function of order  $O(1)$ :

$$\delta_a(x) = \frac{1}{\pi} \frac{a}{x^2 + a^2}, \quad f(y) \sim O(1). \quad (5.147)$$

The most important fact is that the delta function covers the complete normalization on scales  $x \sim a$ ,

$$\sum_m |w(ma)|^2 \sim \int dx |w(x)|^2 \sim f(0) \sim O(1), \quad (5.148)$$

whereas, due to the Lorentzian form of the delta function, the scaling of all moments is determined from the region  $|x| \lesssim \xi$  as

$$\begin{aligned} C_r &\sim \sum_m (ma)^r |w(ma)|^2 \sim \int dx x^r |w(x)|^r \\ &\sim a \int dx x^{r-2} f(x/\xi) e^{-|x|/\xi} \\ &\sim a \xi^{r-1} \int_{|y| \lesssim 1} dy y^{r-2} f(y) \sim a \xi^{r-1} \end{aligned} \quad (5.149)$$

for all  $r \geq 1$ . Note that the region  $|x| \sim a$  contributes only the negligible amount  $a^r \ll a \xi^{r-1}$  to the moment for  $r \geq 2$ . Strictly speaking, this argument is only rigorous for even values of  $r$  since in this case all terms of the integrand are positive. For odd values of  $r$  it can happen that the prefactor in front of the leading term is zero. This happens for the scaling of the odd moments  $\tilde{D}_{2l+1,\alpha}(Ma)$  and  $\tilde{C}_{2l+1,\alpha}(Ma)$  in the ML gauge, which do not show universality but are of no importance since their asymptotic value is of negligible order  $\sim a^2 \xi^{2l-1}$  for  $l > 0$  [see the detailed discussion after Eq. (5.87)]. In contrast, in the AF gauge the odd moments scale according to the estimate (5.149) for all  $r \geq 1$  and their order of magnitude is fully determined by the scaling region (S) (even for  $r = 1$ ). They can be written as

$$C_{2l+1} \sim a \int_{0 < y \lesssim 1} dy y^{2l-1} \{f(y) - f(-y)\} \sim a \xi^{2l}, \quad (5.150)$$

such that convergence is guaranteed at  $y = 0$  even for  $l = 0$ . Importantly, in the AF gauge, the function  $f(y) \approx f(-y)$  is



nearly symmetric for  $|y| \ll 1$  [see (5.139)]. Therefore, the region of small scales does not contribute and the odd moments show universal scaling from the regime (S) due to a significant asymmetry of the scaling function  $f(y)$  for  $|y| \sim O(1)$  [see the discussion after Eq. (5.87)]. For example, it is quite remarkable that the first moment in the AF gauge is fully determined by the nearly invisible asymmetry of the Wannier function on large scales  $|ma| \sim \xi \gg a$ , whereas the visible impression of a symmetric shape on small scales  $|ma| \sim O(a)$  would predict an incorrect vanishing first moment.

The scaling behavior of the moments is demonstrated numerically in the insets of Figs. 1(a), 1(b), 10, and 11, where it is shown that the first moments  $C_{1\alpha}(Ma)$  in the AF gauge and the second moments  $\tilde{D}_{2\alpha}(Ma)$  in the ML gauge scale indeed smoothly to their universal asymptotic values on the scale  $Ma \sim \xi$  and obtain a negligible contribution from the small-scale regime (F). Similar numerical evidence can be shown for all the other higher moments  $C_{r\alpha}(Ma)$  with  $r \geq 1$  and  $\tilde{D}_{2l,\alpha}(Ma)$  with  $l \geq 1$ .

## VI. NON-ABELIAN WANNIER FUNCTIONS

In this final section we analyze another class of lattice Wannier functions which arise from a special non-Abelian gauge transformation which mixes the Bloch states from all bands  $\alpha = 1, \dots, \nu$  below a given one labeled by  $\nu$ . In particular, the non-Abelian gauge of maximally localized Wannier functions has been proposed in the literature and the connection of their spread to the polarization fluctuations has been put forward. Here we analyze the universal scaling properties of the non-Abelian lattice Wannier functions and find the interesting result that they scale up to a surprisingly high precision according to the Dirac Wannier function of the lower band of the Dirac model corresponding to gap  $\nu$  or, equivalently, similar to the upper part of the Wannier function of band  $\nu$  as shown in Fig. 8. In Sec. VIA we summarize the general definition of the non-Abelian gauge of maximal localization and show our central result of its universal scaling. The technical part of the construction of this non-Abelian gauge is outlined in Sec. VIB, supplemented by three Appendices I, J, and K. We show this construction in the main part of this work since it provides a very efficient way to analyze non-Abelian gauges analytically via the Wilson propagator which, up to our best knowledge, has not been reported before.

### A. Non-Abelian lattice gauge and summary of results

So far we have discussed Abelian gauge transformations for a fixed band index  $\alpha$  by allowing for a phase factor  $e^{i\varphi_{k\alpha}}$  to transform the Bloch state as

$$|u_{k\alpha}\rangle \rightarrow |\tilde{u}_{k\alpha}\rangle = e^{i\varphi_{k\alpha}} |u_{k\alpha}\rangle, \quad (6.1)$$

where  $\varphi_{k\alpha} = \varphi_{k+\frac{2\pi}{Za},\alpha}$  is a periodic phase variable. In multi-band systems, more general non-Abelian gauge transformations have been discussed (see Ref. [10] or reviews in Refs. [3,9]). These gauge transformations mix the Bloch states of all band indices  $\alpha = 1, \dots, \nu$  up to a certain value  $\nu$  (defining the valence band if a chemical potential is present in gap

$\nu$ ) via a  $k$ -dependent unitary transformation as

$$|\hat{u}_{k\alpha}^{(\nu)}\rangle = \sum_{\alpha'=1}^{\nu} |u_{k\alpha'}\rangle (\hat{U}_k^{(\nu)})_{\alpha'\alpha}. \quad (6.2)$$

Here,  $\hat{U}_k^{(\nu)}$  is a unitary  $\nu \times \nu$  matrix which is periodic under  $k \rightarrow k + \frac{2\pi}{Za}$ :

$$\hat{U}_k^{(\nu)} (\hat{U}_k^{(\nu)})^\dagger = 1, \quad \hat{U}_{k+\frac{2\pi}{Za}}^{(\nu)} = \hat{U}_k^{(\nu)}. \quad (6.3)$$

From the normalization (2.13) of the Bloch states and the unitarity of  $\hat{U}_k^{(\nu)}$  we note the properties

$$\langle \hat{u}_{k\alpha}^{(\nu)} | \hat{u}_{k\alpha'}^{(\nu)} \rangle = \delta_{\alpha\alpha'}, \quad (6.4)$$

$$\sum_{\alpha=1}^{\nu} |\hat{u}_{k\alpha}^{(\nu)}\rangle \langle \hat{u}_{k\alpha}^{(\nu)}| = \sum_{\alpha=1}^{\nu} |u_{k\alpha}\rangle \langle u_{k\alpha}|. \quad (6.5)$$

Although the Bloch states  $|\hat{u}_{k\alpha}^{(\nu)}\rangle$  are no longer eigenstates of the Bloch Hamiltonian  $h_k$  as defined in (2.10), the corresponding non-Abelian Wannier functions

$$\hat{w}_\alpha^{(\nu)}(ma) = \frac{Za}{2\pi} \int_{-\frac{\pi}{Za}}^{\frac{\pi}{Za}} dk \hat{u}_{k\alpha}^{(\nu)}(ja) e^{ikma} \quad (6.6)$$

have interesting properties [10] which we summarize in the following.

Defining the Zak-Berry connection and the geometric tensor in the non-Abelian case analog to (3.1) and (3.3) as

$$(\hat{A}_k^{(\nu)})_{\alpha\beta} = (\hat{A}_k^{(\nu)})_{\beta\alpha}^* = i \langle \hat{u}_{k\alpha}^{(\nu)} | \partial_k \hat{u}_{k\beta}^{(\nu)} \rangle, \quad (6.7)$$

$$(\hat{Q}_k^{(\nu)})_{\alpha\beta} = (\hat{Q}_k^{(\nu)})_{\beta\alpha} \quad (6.8)$$

$$= \langle \partial_k \hat{u}_{k\alpha}^{(\nu)} | \partial_k \hat{u}_{k\beta}^{(\nu)} \rangle \delta_{\alpha\beta} - |\langle \hat{u}_{k\alpha}^{(\nu)} | \partial_k \hat{u}_{k\beta}^{(\nu)} \rangle|^2, \quad (6.9)$$

and introducing the following definition for the non-Abelian Zak-Berry phase matrix

$$\hat{\gamma}_{\alpha\beta}^{(\nu)} = \int_{-\frac{\pi}{Za}}^{\frac{\pi}{Za}} dk (\hat{A}_k^{(\nu)})_{\alpha\beta}, \quad (6.10)$$

together with  $\hat{\gamma}_\alpha^{(\nu)} = \hat{\gamma}_{\alpha\alpha}^{(\nu)}$  for the diagonal components, one finds the following useful transformation properties:

$$\hat{A}_k^{(\nu)} = (\hat{U}_k^{(\nu)})^\dagger A_k^{(\nu)} \hat{U}_k^{(\nu)} + i (\hat{U}_k^{(\nu)})^\dagger \partial_k \hat{U}_k^{(\nu)}, \quad (6.11)$$

$$\sum_{\alpha=1}^{\nu} \hat{\gamma}_\alpha^{(\nu)} = \sum_{\alpha=1}^{\nu} \gamma_\alpha - 2\pi \text{wn}[\det \hat{U}_k^{(\nu)}], \quad (6.12)$$

$$\sum_{\alpha,\beta=1}^{\nu} (\hat{Q}_k^{(\nu)})_{\alpha\beta} = \sum_{\alpha,\beta=1}^{\nu} (Q_k)_{\alpha\beta}, \quad (6.13)$$

where  $\text{wn}[\det \hat{U}_k^{(\nu)}]$  denotes the winding number of the determinant of  $\hat{U}_k^{(\nu)}$ . Here,  $A_k^{(\nu)}$  is a  $\nu \times \nu$  matrix with matrix elements  $(A_k^{(\nu)})_{\alpha\beta} = (A_k)_{\alpha\beta}$  given by the Zak-Berry connection (3.1) from all band indices  $\alpha, \beta = 1, \dots, \nu$ .

Furthermore, defining the moments of the Wannier functions in the non-Abelian case analog to (4.16) as

$$\hat{C}_{r\alpha}^{(\nu)} = \langle x^r \rangle_\alpha^{(\nu)} = \sum (ma)^r |\hat{w}_\alpha^{(\nu)}(ma)|^2 \quad (6.14)$$

$$= \frac{Za}{2\pi} \int_{-\pi/Za}^{\pi/Za} dk \langle \hat{u}_{k\alpha}^{(\nu)} | (i\partial_k)^r | \hat{u}_{k\alpha}^{(\nu)} \rangle, \quad (6.15)$$

and introducing the following definitions for the sum over the positions and the quadratic spreads

$$\langle x \rangle^{(\nu)} = \sum_{\alpha=1}^{\nu} \hat{C}_{1\alpha}^{(\nu)}, \quad (6.16)$$

$$\langle \Delta x^2 \rangle^{(\nu)} = \sum_{\alpha=1}^{\nu} (\hat{C}_{2\alpha}^{(\nu)} - (\hat{C}_{1\alpha}^{(\nu)})^2), \quad (6.17)$$

we find from (6.15) and the above definitions after some straightforward algebra

$$\frac{\langle x \rangle^{(\nu)}}{Za} = \sum_{\alpha=1}^{\nu} \frac{\hat{\gamma}_{\alpha}^{(\nu)}}{2\pi}, \quad \hat{C}_{1\alpha}^{(\nu)} = \frac{Za}{2\pi} \hat{\gamma}_{\alpha}^{(\nu)}, \quad (6.18)$$

$$\begin{aligned} \langle \Delta x^2 \rangle^{(\nu)} = & \frac{Za}{2\pi} \sum_{\alpha,\beta=1}^{\nu} \int_{-\frac{\pi}{2a}}^{\frac{\pi}{2a}} dk \left\{ (\hat{Q}_k^{(\nu)})_{\alpha\beta} \right. \\ & + \left[ (\hat{A}_k^{(\nu)})_{\alpha\beta} - \frac{Za}{2\pi} \hat{\gamma}_{\alpha\beta}^{(\nu)} \right]^2 \\ & \left. + (1 - \delta_{\alpha\beta}) \left( \frac{Za}{2\pi} \hat{\gamma}_{\alpha\beta}^{(\nu)} \right)^2 \right\}. \end{aligned} \quad (6.19)$$

Since the first term on the right-hand side of (6.19) is gauge invariant according to (6.13), and the last two terms are positive, the quadratic spread  $\langle \Delta x^2 \rangle^{(\nu)}$  is minimal in the non-Abelian gauge of maximal localization (NA-ML), defined by a  $k$ -independent and band-diagonal Zak-Berry connection

$$(\hat{A}_k^{(\nu)})_{\alpha\beta} = \frac{Za}{2\pi} \hat{\gamma}_{\alpha\beta}^{(\nu)} = \frac{Za}{2\pi} \hat{\gamma}_{\alpha}^{(\nu)} \delta_{\alpha\beta}. \quad (6.20)$$

As a consequence, using the surface fluctuation theorem (3.51), we find that the minimal quadratic spread in the NA-ML gauge is related to the boundary charge fluctuations as

$$l_p (\Delta Q_B^{(v)})^2 = \frac{1}{Za} \langle \Delta x^2 \rangle_{\min}^{(\nu)} = \sum_{\alpha=1}^{\nu} \hat{D}_{2\alpha}^{(\nu)}, \quad (6.21)$$

where we defined the moments

$$\hat{D}_{r\alpha}^{(\nu)} = \sum_m (\hat{m}_{\alpha}^{(\nu)} a)^r |\hat{w}_{\alpha}^{(\nu)}(ma)|^2 \quad (6.22)$$

relative to the shift by the first moment, i.e., with

$$\hat{m}_{\alpha}^{(\nu)} = m - \hat{C}_{1\alpha}^{(\nu)} = m - \frac{Za}{2\pi} \hat{\gamma}_{\alpha}^{(\nu)}. \quad (6.23)$$

Most importantly, the surface fluctuation theorem (6.21) formulated in terms of the non-Abelian quadratic spread shows that the NA-ML gauge is unique in the sense that the boundary charge fluctuations can be written as the sum over the second moments of the Wannier functions of the individual bands. Within the Abelian ML gauge this is *not* possible, as already discussed after (3.57). For the boundary charge itself it seems to be similar at first sight to the Abelian AF and ML gauges, where it is also related to the sum over the Zak-Berry phases of the individual bands. However, as we have seen in Sec. III C 1, a subtle cancellation procedure happens such that the sum over the Zak-Berry phases of the individual bands  $\alpha = 1, \dots, \nu$  is related to the phase  $\gamma^{(\nu)}$  of the gap parameter of gap  $\nu$  [see the central equation (3.50)]. As we will see in the following, this is not needed in the non-Abelian gauge where

all Zak-Berry phases  $\hat{\gamma}_{\alpha}^{(\nu)}$  are equally spaced and related to  $\gamma^{(\nu)}$  in the following universal way:

$$\begin{aligned} \hat{\gamma}_{\alpha}^{(\nu)} = & -\frac{1}{\nu} \gamma^{(\nu)} \\ & + \frac{\pi}{\nu} \begin{cases} \nu - 2\alpha + 1 & \text{for } \nu \text{ even,} \\ \nu - 2\alpha + 1 + s_{\gamma^{(\nu)}} & \text{for } \nu \text{ odd} \end{cases} \end{aligned} \quad (6.24)$$

such that  $-\pi < \hat{\gamma}_{\alpha}^{(\nu)} < \pi$ .

Consistent with the way the fluctuations are distributed among the bands in the non-Abelian case, we will find that all Wannier functions  $\hat{w}_{\alpha}^{(\nu)}$  in the NA-ML gauge scale in the same way according to the upper component  $\tilde{w}_{u,\nu}$  of the Wannier function in the Abelian ML gauge for band  $\nu$ . The central result shown in the next Sec. VI B is the precise relation

$$\hat{w}_{\alpha}^{(\nu)}(ma) \approx \frac{1}{\sqrt{\nu}} e^{i\theta_{\alpha}^{(\nu)}} a \sqrt{Z} \sum_p \tilde{w}_{-p}^{(\nu)}(\hat{m}_{\alpha}^{(\nu)} a) e^{ipk_F^{(\nu)} ma}, \quad (6.25)$$

where  $e^{i\theta_{\alpha}^{(\nu)}}$  is some unimportant phase factor. Up to this phase factor, we find in comparison to (4.52) the central result that *all* Wannier functions  $\hat{w}_{\alpha}^{(\nu)}(ma)$  in the NA-ML gauge scale precisely as  $\frac{1}{\sqrt{\nu}} \tilde{w}_{u,\nu}(ma)$  in the Abelian ML gauge, provided one replaces  $\gamma_{\nu} \rightarrow \hat{\gamma}_{\alpha}^{(\nu)}$  in the shift variable such that  $\tilde{m}_{\nu} \rightarrow \hat{m}_{\alpha}^{(\nu)}$ . Using the scaling property (5.76) of the upper component of the Wannier function in the ML gauge, this gives the following scaling for all non-Abelian Wannier functions:

$$\frac{1}{Z} |\hat{m}_{\alpha}^{(\nu)} \hat{w}_{\alpha}^{(\nu)}(n, j)|^2 \approx \frac{1}{\nu} \left| \tilde{F}_{-} \left( \frac{\hat{m}_{\alpha}^{(\nu)} a}{\xi^{(\nu)}}; \frac{\nu j}{Z}, \gamma^{(\nu)} \right) \right|^2. \quad (6.26)$$

As a consequence, defining the following scaling functions for the non-Abelian moments

$$\hat{D}_{r\alpha}^{(\nu)}(Ma) = \sum_{m=-M}^M (\hat{m}_{\alpha}^{(\nu)} a)^r |\hat{w}_{\alpha}^{(\nu)}(ma)|^2, \quad (6.27)$$

we find from (5.83) and (5.96) the universal scaling

$$\frac{\hat{D}_{r\alpha}^{(\nu)}(Ma)}{Za(\xi^{(\nu)})^{r-1}} \approx \frac{1}{\nu} \int_{-\frac{Ma}{\xi^{(\nu)}}}^{\frac{Ma}{\xi^{(\nu)}}} dy y^{r-2} \tilde{G}_{-}(y; \gamma^{(\nu)}), \quad (6.28)$$

with the following asymptotic value for the second moment:

$$\frac{\hat{D}_{2\alpha}^{(\nu)}(Ma)}{Za\xi^{(\nu)}} \rightarrow \frac{1}{8\nu}. \quad (6.29)$$

The universal relation (6.25) of all non-Abelian Wannier functions  $\alpha = 1, \dots, \nu$  to the Dirac Wannier function of the lower band of the Dirac theory corresponding to gap  $\nu$  is one of the most important results of this work. It shows that all non-Abelian Wannier functions scale independent of the band index in the same way and depend only on the low-energy properties of the model, provided that the condition of small gaps is fulfilled. Up to our best knowledge, the non-Abelian Wannier functions have not been studied analytically in the literature and their universal scaling behavior for all generalized AAH models in terms of a single length scale  $\xi^{(\nu)}$  has not been reported before.

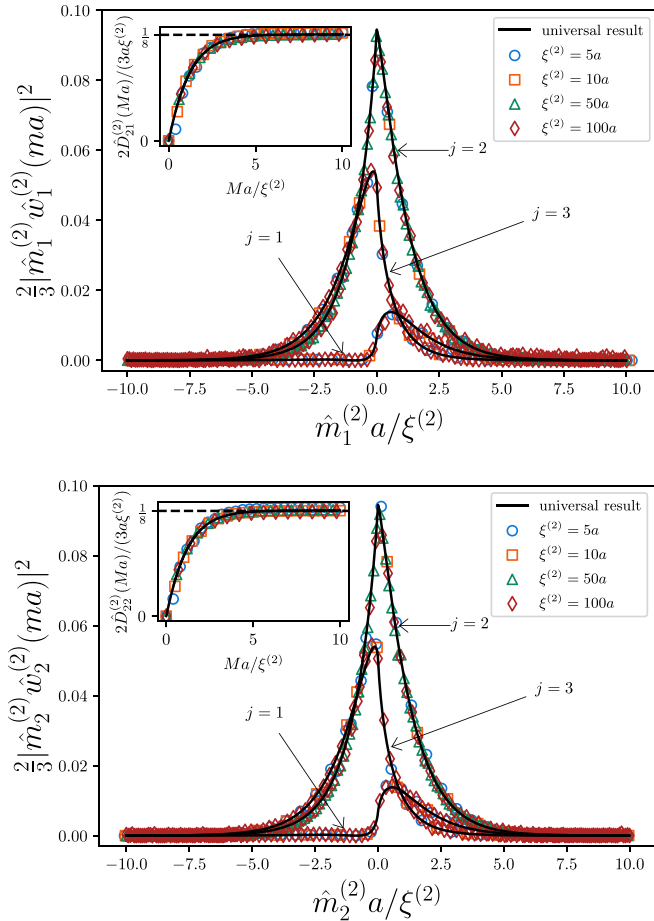


FIG. 13. Scaling of  $\frac{\nu}{Z} |\hat{m}_\alpha^{(v)} \hat{w}_\alpha^{(v)}(n, j)|^2$  for  $\alpha = 1, 2$  (upper and lower panels),  $\nu = 2$ ,  $Z = 3$ ,  $\gamma^{(1)} = 0.2\pi$ ,  $\gamma^{(2)} = 0.5\pi$ , and  $\xi^{(1)} = 30a$ , as function of  $\hat{m}_\alpha^{(v)} a / \xi^{(v)}$  for different values of  $\xi^{(2)}$  and  $j = 1, 2, 3$ . According to (6.26) we obtain the universal curve  $|\hat{F}_-(\hat{m}_\alpha^{(2)} a / \xi^{(2)}; 2j/3, \gamma^{(2)})|^2$ , i.e., the same scaling for both  $\alpha = 1, 2$  as for  $\frac{1}{3} |\hat{m}_2 \hat{w}_{u,2}(ma)|^2$  as function of  $\hat{m}_2 a / \xi^{(2)}$  (see Fig. 8). The two insets show the scaling of the corresponding second moments  $\nu \hat{D}_{2\alpha}^{(v)}(Ma) / (Za \xi^{(v)})^2$  as function of  $s_2 = Ma / \xi^{(2)}$ . According to (6.28) and (6.29) we obtain the universal curve  $\int_{-s_2}^{s_2} dy \tilde{G}_-(y; \gamma^{(2)})$  with saturation at the correct value  $\frac{1}{8}$ . The scaling is independent of  $\alpha = 1, 2$  and identical to the scaling of the second moment  $\hat{D}_{u,22}(Ma)$  as shown in the inset of Fig. 8.

For the special case  $Z = 3$  and  $\nu = 2$  we demonstrate in Fig. 13 the universal scaling (6.26) of the Wannier functions  $\hat{w}_\alpha^{(v)}(ma)$  in NA-ML gauge and the corresponding scaling (6.28) of the second moments  $\hat{D}_{2\alpha}^{(v)}(Ma)$ . Indeed, we find that the scaling is independent of  $\alpha = 1, 2$  and follows the scaling of the upper part  $\hat{w}_{u,\nu}(ma)$  of the Wannier function and its second moment in ML gauge for band  $\alpha = \nu$  (compare with Fig. 8).

We note that also the winding number of the determinant  $\det \hat{U}_k^{(v)}$  contains interesting information about the total number of edge states  $N_E^{(v)}$  present in all gaps  $\nu' = 1, \dots, \nu$ . According to (6.12), the winding number controls the change of the sum over all Zak-Berry phases from the bands  $\alpha = 1, \dots, \nu$  in the gauge transformation. On the other hand, we

can calculate the sum over the Zak-Berry phases in the AF and NA-ML gauge separately by using (3.30), (3.31), and (6.24):

$$\sum_{\alpha=1}^{\nu} \gamma_\alpha = -\gamma^{(v)} + \pi \sum_{\nu'=1}^{\nu} s_{\gamma^{(\nu')}} \quad (6.30)$$

$$\sum_{\alpha=1}^{\nu} \hat{\gamma}_\alpha^{(v)} = -\gamma^{(v)} + \pi \delta_{\nu, \text{odd}} s_{\gamma^{(v)}} \quad (6.31)$$

Since an edge state is present in gap  $\nu'$  if  $s_{\gamma^{(\nu')}} > 0$ , we get in addition for the total number of edge states the expression

$$N_E^{(v)} = \frac{1}{2} \sum_{\nu'=1}^{\nu} (1 + s_{\gamma^{(\nu')}}) = \frac{1}{2} \left( \nu + \sum_{\nu'=1}^{\nu} s_{\gamma^{(\nu')}} \right) \quad (6.32)$$

Comparing these equations we find the following relation between the number of edge states and the winding number of  $\det \hat{U}_k^{(v)}$ ,

$$N_E^{(v)} = \text{wn}[\det \hat{U}_k^{(v)}] + \frac{1}{2} (\nu + \delta_{\nu, \text{odd}} s_{\gamma^{(v)}}) \quad (6.33)$$

which is obviously an integer number for both  $\nu$  even or odd. This relation can be viewed as a bulk-boundary correspondence relating the total number of edge states up to a certain gap  $\nu$  to a winding number defined in terms of the bulk states. In contrast to other topological invariants defined in one-dimensional systems, this relation holds independent of any symmetry constraints and includes all edge states, independent of whether they are at zero energy or not.

Using (3.46) and (3.49), we finally find for the total boundary charge without or with including the edge states

$$\sum_{\alpha=1}^{\nu} Q_{B,\alpha} = -\frac{1}{2\pi} \sum_{\alpha=1}^{\nu} \gamma_\alpha - \frac{1}{2} \nu + \frac{\nu}{2Z} \quad (6.34)$$

$$Q_B^{(v)} = -\frac{1}{2\pi} \sum_{\alpha=1}^{\nu} \hat{\gamma}_\alpha^{(v)} + \frac{1}{2} \delta_{\nu, \text{odd}} s_{\gamma^{(v)}} + \frac{\nu}{2Z} \quad (6.35)$$

The last result states the surface charge theorem in terms of the non-Abelian Zak-Berry phases  $\hat{\gamma}_\alpha^{(v)}$  or the first moments  $\hat{C}_{1\alpha}^{(v)} = \frac{Za}{2\pi} \hat{\gamma}_\alpha^{(v)}$  of the non-Abelian Wannier functions.

## B. Explicit construction of the NA-ML gauge

We proceed in this section with the explicit construction of the NA-ML gauge and the analytical proof of the central identity (6.25) relating all non-Abelian Wannier functions  $\alpha = 1, \dots, \nu$  to the upper part of the Wannier function in the Abelian ML gauge for the highest band  $\nu$ . The NA-ML gauge can be constructed explicitly via the unitary transformation

$$\hat{U}_k^{(v)} = U^{(v)}(k, k_0) V^{(v)} e^{-ik \frac{Za}{2\pi} \hat{\gamma}^{(v)}}, \quad (6.36)$$

where

$$U^{(v)}(k_1, k_2) = P e^{i \int_{k_2}^{k_1} dk A_k^{(v)}} \quad (6.37)$$

denotes the Wilson propagator along the path from  $k_2 \rightarrow k_1$  (here,  $P$  denotes the  $k$ -ordering operator analog to the time-ordering operator),  $k_0$  is an arbitrary reference point, and  $V^{(v)}$  is the unitary transformation which diagonalizes the Wilson

loop operator

$$e^{i\hat{\gamma}^{(\nu)}} = (V^{(\nu)})^\dagger U_L^{(\nu)} V^{(\nu)}, \quad (6.38)$$

$$U_L^{(\nu)} \equiv U^{(\nu)}\left(k_0 + \frac{2\pi}{Za}, k_0\right). \quad (6.39)$$

Indeed, using the differential equation

$$i\partial_k \hat{U}_k^{(\nu)} = -A_k^{(\nu)} \hat{U}_k^{(\nu)} + \hat{U}_k^{(\nu)} \frac{Za}{2\pi} \gamma^{(\nu)}, \quad (6.40)$$

and multiplying it from the left with  $(U_k^{(\nu)})^\dagger$ , one obtains directly the condition (6.20) defining the NA-ML gauge. Furthermore, using the group property of the Wilson propagator together with its periodicity (following from the periodicity of the Zak-Berry connection  $A_k^{(\nu)} = A_{k+\frac{2\pi}{Za}}^{(\nu)}$ )

$$U^{(\nu)}(k_1, k_2) = U^{(\nu)}(k_1, k_3) U^{(\nu)}(k_3, k_2), \quad (6.41)$$

$$U^{(\nu)}(k_1, k_2) = U^{(\nu)}\left(k_1 + \frac{2\pi}{Za}, k_2 + \frac{2\pi}{Za}\right), \quad (6.42)$$

we find periodicity of the non-Abelian gauge transformation

$$\begin{aligned} \hat{U}_{k+\frac{2\pi}{Za}}^{(\nu)} &= U^{(\nu)}\left(k + \frac{2\pi}{Za}, k_0 + \frac{2\pi}{Za}\right) \hat{U}_L^{(\nu)} V^{(\nu)} e^{-i\hat{\gamma}^{(\nu)}} e^{-ik\frac{Za}{2\pi}\hat{\gamma}^{(\nu)}} \\ &= U^{(\nu)}(k, k_0) V^{(\nu)} e^{-ik\frac{Za}{2\pi}\hat{\gamma}^{(\nu)}} = \hat{U}_k^{(\nu)}. \end{aligned} \quad (6.43)$$

Here we used in the first step the group property (6.41), and in the second step the periodicity (6.42) together with the definition (6.38) of the transformation  $V^{(\nu)}$ .

We note that the non-Abelian Zak-Berry phases  $\hat{\gamma}_\alpha^{(\nu)}$  are independent of the choice  $k_0$  of the reference point since the Wilson loop operator  $(U_L^{(\nu)})'$  for another reference point  $k'_0$  is related to  $U_L^{(\nu)}$  via a unitary transformation

$$(U_L^{(\nu)})' = U^{(\nu)}(k_0, k'_0) U_L^{(\nu)} U^{(\nu)}(k_0, k'_0)^\dagger, \quad (6.44)$$

where we used the group property (6.41) and the periodicity (6.42). This means that the unitary operator  $(V^{(\nu)})'$  with respect to the reference point  $k'_0$  is related to  $V^{(\nu)}$  by

$$(V^{(\nu)})' = U^{(\nu)}(k'_0, k_0) V^{(\nu)}. \quad (6.45)$$

As a consequence, up to trivial phase factors to define the unitary transformation  $V^{(\nu)}$ , we find that the unitary transformation  $\hat{U}_k^{(\nu)}$  is unique and independent of  $k_0$ .

For convenience, we will choose in the following  $k_0 = -\pi/(2Za)$  as reference point and discuss the two intervals (I)  $k \in [-\pi/(2Za), \pi/(2Za)]$  and (II)  $k \in [\pi/(2Za), 3\pi/(2Za)]$  separately since they are related to two different Dirac theories (see Fig. 14). Using (2.8), (2.30), and (2.31), we get for the relation between the Bloch and Dirac states in this convention

$$\begin{aligned} \text{(I): } \nu \text{ even, } k = q \\ u_{k\alpha}(ma) &= \frac{1}{\sqrt{Z}} \sum_{p=\pm} \chi_{q\tau p}^{(\nu)} e^{ipk_F^{(\nu)} ma}, \end{aligned} \quad (6.46)$$

$$\begin{aligned} \text{(II): } \nu \text{ odd, } k = \frac{\pi}{Za} + q \\ u_{k\alpha}(ma) &= e^{-i\frac{\pi}{2}m} \frac{1}{\sqrt{Z}} \sum_{p=\pm} \chi_{q\tau p}^{(\nu)} e^{ipk_F^{(\nu)} ma}, \end{aligned} \quad (6.47)$$

where in both cases  $|q| < \frac{\pi}{2Za}$ . By convention we have added two trivial Dirac theories, one for  $\alpha = 1$  in subinterval (I) with

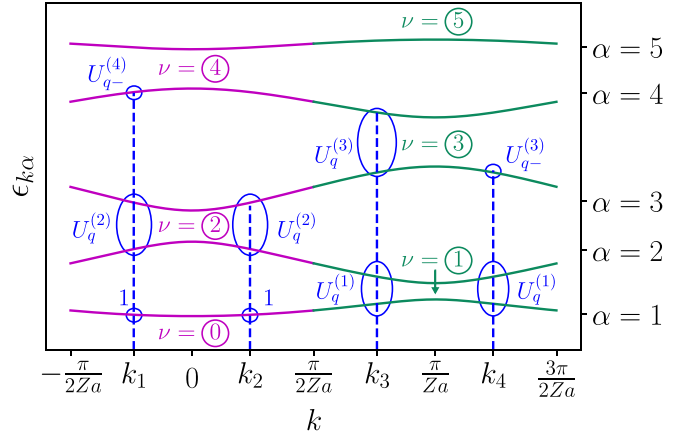


FIG. 14. Sketch of the typical band structure for the interval  $-\frac{\pi}{Za} < k < \frac{3\pi}{2Za}$  for  $Z = 5$ . The two subintervals (I) with  $-\frac{\pi}{Za} < k < \frac{\pi}{2Za}$  and (II) with  $\frac{\pi}{Za} < k < \frac{3\pi}{2Za}$  are indicated in violet and green color, respectively. The lowest (highest) bands in subinterval (I) [(II)] (corresponding to  $\nu = 0$  and  $Z$ , respectively) are characterized by  $k$ -independent Bloch states  $u_{k\alpha}$ . Four different  $k$  values are shown, indicating the block structures (6.61) and (6.62) of the lattice Wilson propagator for  $\nu$  even or odd in the two subintervals. Here,  $k_1$  corresponds to  $U^{(\nu)}(k_1, -\frac{\pi}{2Za})$  for  $\nu = 4$  even in subinterval (I) [see the first equation of (6.61)],  $k_2$  to  $U^{(\nu)}(k_2, -\frac{\pi}{2Za})$  for  $\nu = 3$  odd in subinterval (I) [see the first equation of (6.62)],  $k_3$  to  $U^{(\nu)}(k_3, \frac{\pi}{2Za})$  for  $\nu = 4$  even in subinterval (II) [see the second equation of (6.61)], and  $k_4$  to  $U^{(\nu)}(k_4, \frac{\pi}{2Za})$  for  $\nu = 3$  odd in subinterval (II) [see the second equation of (6.62)].

$\nu = 0$  and  $\tau = +$ , and another for  $\alpha = Z$  in subinterval (II) with  $\nu = Z$  and  $\tau = -$ . In these two cases the Bloch states are given by the gapless ones and are independent of  $k$ , such that the Zak-Berry connection is zero.

The non-Abelian Wannier functions are then split as

$$\hat{w}_\alpha^{(\nu)}(ma) = \hat{w}_{I,\alpha}^{(\nu)}(ma) + \hat{w}_{II,\alpha}^{(\nu)}(ma), \quad (6.48)$$

$$\hat{w}_{I,\alpha}^{(\nu)}(ma) = \frac{Za}{2\pi} \int_{-\frac{\pi}{2Za}}^{\frac{\pi}{2Za}} dk \hat{u}_{k\alpha}^{(\nu)}(ma) e^{ikma}, \quad (6.49)$$

$$\hat{w}_{II,\alpha}^{(\nu)}(ma) = \frac{Za}{2\pi} \int_{\frac{\pi}{2Za}}^{\frac{3\pi}{2Za}} dk \hat{u}_{k\alpha}^{(\nu)}(ma) e^{ikma}, \quad (6.50)$$

where the transformed Bloch states follow from (6.2) and (6.36) as

$$\begin{aligned} \hat{u}_{k\alpha}^{(\nu)}(ma) e^{ikma} \\ = e^{ik\hat{m}_\alpha^{(\nu)}a} \sum_{\alpha'=1}^{\nu} u_{k\alpha'}^{(\nu)}(ma) \left[ U^{(\nu)}\left(k, -\frac{\pi}{2Za}\right) V^{(\nu)} \right]_{\alpha'\alpha}. \end{aligned} \quad (6.51)$$

It turns out to be useful to define the propagators when crossing the two subintervals as

$$U_I^{(\nu)} = U^{(\nu)}\left(\frac{\pi}{2Za}, -\frac{\pi}{2Za}\right), \quad (6.52)$$

$$U_{II}^{(\nu)} = U^{(\nu)}\left(\frac{3\pi}{2Za}, \frac{\pi}{2Za}\right), \quad (6.53)$$

and use

$$U^{(v)}\left(k, -\frac{\pi}{2Za}\right) = U^{(v)}\left(k, \frac{\pi}{2Za}\right) U_I^{(v)} \quad (6.54)$$

for the representation of the propagator within subinterval (II). The Wilson loop operator defining the transformation  $V^{(v)}$  via (6.38) can be expressed as

$$U_L^{(v)} = U_{II}^{(v)} U_I^{(v)}. \quad (6.55)$$

As outlined in detail in Appendix I one can extend the integration limits in (6.49) and (6.50) to  $\pm\infty$  by using the  $k$ -independent values of the Bloch states and the propagator at the edges of the two subintervals. This means that the propagators for the two subintervals are extended as

$$(I) : U^{(v)}\left(k, -\frac{\pi}{2Za}\right) \rightarrow \begin{cases} U_I^{(v)} & \text{for } k > \frac{\pi}{2Za}, \\ 1 & \text{for } k < -\frac{\pi}{2Za}, \end{cases} \quad (6.56)$$

$$(II) : U^{(v)}\left(k, \frac{\pi}{2Za}\right) \rightarrow \begin{cases} U_{II}^{(v)} & \text{for } k > \frac{3\pi}{2Za}, \\ 1 & \text{for } k < \frac{\pi}{2Za}. \end{cases} \quad (6.57)$$

The additional contributions are shown to cancel each other when taking the sum (6.48) up to trivial delta functions which are subtracted as usual both in lattice and Dirac theories.

$v$  even :

$$U^{(v)}\left(k, -\frac{\pi}{2Za}\right) = \begin{pmatrix} U_{q^-}^{(v)} & & & & \\ & U_q^{(v-2)} & & & \\ & & U_q^{(v-4)} & & \\ & & & \ddots & \\ & & & & U_q^{(2)} \\ & & & & & 1 \end{pmatrix}, \quad U^{(v)}\left(k, \frac{\pi}{2Za}\right) = \begin{pmatrix} U_q^{(v-1)} & & & & \\ & U_q^{(v-3)} & & & \\ & & \ddots & & \\ & & & \ddots & \\ & & & & U_q^{(1)} \\ & & & & & 1 \end{pmatrix}, \quad (6.61)$$

$v$  odd :

$$U^{(v)}\left(k, -\frac{\pi}{2Za}\right) = \begin{pmatrix} U_q^{(v-1)} & & & & \\ & U_q^{(v-3)} & & & \\ & & \ddots & & \\ & & & \ddots & \\ & & & & U_q^{(2)} \\ & & & & & 1 \end{pmatrix}, \quad U^{(v)}\left(k, \frac{\pi}{2Za}\right) = \begin{pmatrix} U_{q^-}^{(v)} & & & & \\ & U_q^{(v-2)} & & & \\ & & U_q^{(v-4)} & & \\ & & & \ddots & \\ & & & & U_q^{(1)} \\ & & & & & 1 \end{pmatrix}, \quad (6.62)$$

where  $U_{q^-}^{(v)}$  is the Abelian Dirac propagator for the upper part of the last band  $v$ , which can be expressed with the help of (5.10) and (5.9) as

$$U_{q^-}^{(v)} = e^{i \int_{-\infty}^q dq' A_{q'}^{(v)}} = e^{i\phi_{q^-}^{(v)}} e^{-i\gamma^{(v)}}. \quad (6.63)$$

The  $2 \times 2$  matrix  $U_q^{(v)}$  of the Wilson propagator in Dirac theory is calculated in Appendix J. Most importantly, it is shown that this propagator transforms the Dirac states as

$$\hat{\chi}_{q\tau p}^{(v)} \equiv \sum_{\tau'} \chi_{q\tau' p}^{(v)} (U_q^{(v)})_{\tau'\tau} = \delta_{p,-\tau}, \quad (6.64)$$

i.e., leads to a  $q$ -independent contribution. This has the effect that, after inserting (6.46), (6.47), (6.51), (6.54), (6.61), and (6.62) in (6.49) and (6.50) (with  $k = q$  and  $k = \frac{\pi}{2Za} + q$ , respectively, and the  $q$  integrals extending to  $\pm\infty$ ), all prop-

agators  $U_q^{(v)}$  give rise to integrands where the  $q$  dependence of the Dirac states disappears and only a purely oscillating exponential remains in the integrand. This leads to unphysical delta function contributions which can be disregarded (see Appendix I). The only remaining terms are those from the Abelian propagators  $U_{q^-}^{(v)}$ , leading to the following result for the two subintervals:

$$(I) : U^{(v)}\left(k, -\frac{\pi}{2Za}\right)_{\alpha_1\alpha_2} = \delta_{v_1 v_2} (U_q^{(v_1)})_{\tau_1\tau_2}, \quad (6.58)$$

$$(II) : U^{(v)}\left(k, \frac{\pi}{2Za}\right)_{\alpha_1\alpha_2} = \delta_{v_1 v_2} (U_q^{(v_1)})_{\tau_1\tau_2}, \quad (6.59)$$

where  $U_q^{(v)}$  denotes the Wilson propagator in Dirac theory with respect to reference point  $q = -\infty$ :

$$U_q^{(v)} = P e^{i \int_{-\infty}^q dq' A_{q'}^{(v)}}. \quad (6.60)$$

Here,  $A_q^{(v)}$  denotes the  $2 \times 2$  matrix of the Dirac Zak-Berry connection (3.18). According to Fig. 14 we then get the following block structure of the two propagators for  $v$  even and odd in the corresponding subintervals [where we use  $k = q$  for subinterval (I) and  $k = \frac{\pi}{2Za} + q$  for subinterval (II)]

agators  $U_q^{(v)}$  give rise to integrands where the  $q$  dependence of the Dirac states disappears and only a purely oscillating exponential remains in the integrand. This leads to unphysical delta function contributions which can be disregarded (see Appendix I). The only remaining terms are those from the Abelian propagators  $U_{q^-}^{(v)}$ , leading to the following result for the two subintervals:

$$\hat{w}_{I,\alpha}^{(v)}(ma) \approx \delta_{v,\text{even}} \frac{Za}{2\pi} \int dq e^{iq\hat{m}_\alpha^{(v)} a} u_{qv}(ma) U_{q^-}^{(v)} (V^{(v)})_{v\alpha}, \quad (6.65)$$

$$\hat{w}_{II,\alpha}^{(v)}(ma) \approx \delta_{v,\text{odd}} \frac{Za}{2\pi} \int dq e^{i(\frac{\pi}{2Za} + q)\hat{m}_\alpha^{(v)} a} \times u_{\frac{\pi}{2Za} + q, v}(ma) U_{q^-}^{(v)} (U_I^{(v)} V^{(v)})_{v\alpha}. \quad (6.66)$$

Inserting (6.46), (6.47), (6.63), and using the definition (4.34) of the Dirac Wannier function in ML gauge, we find

$$\hat{w}_{I,\alpha}^{(\nu)}(ma) \approx \delta_{\nu,\text{even}} e^{-i\gamma^{(\nu)}} (V^{(\nu)})_{\nu\alpha} \times a\sqrt{Z} \sum_p \tilde{w}_{-,p}^{(\nu)}(\hat{m}_\alpha^{(\nu)}) e^{ipk_F^{(\nu)}ma}, \quad (6.67)$$

$$\hat{w}_{II,\alpha}^{(\nu)}(ma) \approx \delta_{\nu,\text{odd}} e^{-i\gamma^{(\nu)}} e^{-i\frac{1}{2}\hat{\gamma}_\alpha^{(\nu)}} (U_I^{(\nu)} V^{(\nu)})_{\nu\alpha} \times a\sqrt{Z} \sum_p \tilde{w}_{-,p}^{(\nu)}(\hat{m}_\alpha^{(\nu)}) e^{ipk_F^{(\nu)}ma}. \quad (6.68)$$

In Appendix K we analyze the unitary transformations  $V^{(\nu)}$  and  $U_I^{(\nu)} V^{(\nu)}$ , with the result that all matrix elements consist of phase factors multiplied with  $1/\sqrt{\nu}$  (from the normalization). As a result, we find the central result (6.25), relating all Wannier functions in NA-ML gauge to the Dirac Wannier function corresponding to the upper part of the highest band in ML gauge. In addition, we will analyze in Appendix K the spectrum of the Wilson loop operator (6.55) and find indeed that all non-Abelian Zak-Berry phases are equidistantly distributed according to (6.24).

## VII. SUMMARY AND OUTLOOK

The universal scaling properties of Wannier functions found in this work are fundamental in several respects. Of particular interest is the high-precision value of the universal scaling even for rather large gaps, as it appears most prominently in the non-Abelian gauge of maximal localization (NA-ML) (see Fig. 13). The NA-ML gauge has been put forward in the literature [3,9,10] since in this gauge the total quadratic spread probes the fluctuations of the bulk polarization which, according to the surface fluctuation theorem [18], is related to the fluctuations of a directly measurable observable, the boundary charge  $Q_B^{(\nu)}$ . Importantly,  $Q_B^{(\nu)}$  is the *total* boundary charge corresponding to a chemical potential located in gap  $\nu$ , in contrast to the boundary charge  $Q_{B,\alpha}$  of a single band  $\alpha$  which is not measurable. The observable  $Q_B^{(\nu)}$  probes essentially low-energy properties of the system since the low-lying states are occupied and contribute a fixed amount to the boundary charge which can not fluctuate. This characteristic feature of the boundary charge is directly linked to our result that the Wannier functions in the NA-ML gauge are probing low-energy properties and reveal universal scaling to high precision. This motivates further investigations of the NA-ML gauge in other multichannel or higher-dimensional systems to find similar universal scaling assuming the small gap limit.

Our proposal of how to define Wannier functions for field-theoretical Dirac models opens up pathways to describe universal scaling properties of Wannier functions for a whole class of multiband lattice models very efficiently via a simplified continuum model with a low number of bands only. We have exemplified this here for a two-band Dirac model in 1+1 dimensions, covering the whole class of one-dimensional tight-binding models with nearest-neighbor hopping, one orbital per site, and generic periodic onsite potential and hopping modulation. These are special models in the sense that the gap opens up at a single point in quasimomentum space [in our case either at  $k = 0$  or at  $k =$

$\pm\pi/(Za)$ ]. This gives rise to a field-theoretical low-energy model consisting of one pair of a slowly varying right and left movers. In other more general multichannel cases or models with longer-ranged hopping, several anticrossings can appear in quasimomentum space, each of them characterized by the presence of several channels. The same will apply to higher-dimensional systems. However, also for these more general cases, the boundary charge and the Wannier functions in the NA-ML gauge are expected to probe only the low-energy properties of the system provided that all gap openings close to the Fermi level remain small. Therefore, based on our proposal of how to define Wannier functions within continuum field theories, it will be of high interest to study the Wannier functions of more general field-theoretical models, consisting of multichannel Dirac models or several Dirac models which are coupled to each other (each of them corresponding to one anticrossing point).

It is a quite remarkable result of our work that field-theoretical models also allow for the definition of the Zak-Berry connection and the geometric tensor, although these are quantities which are used in lattice models to characterize global topological properties of the whole band structure in a nonlocal way. In contrast, a field theory can only characterize the physics close to the Fermi level where the gap opens, i.e., it probes essentially local aspects of the band structure. However, this does not mean that a field theory is per se not capable of characterizing topological properties of lattice models. If the gaps at the bottom and the top of a given band are both small, it is possible to set up two independent field theories around the two gaps and supplement them by appropriate asymptotic conditions to connect them in a unique way. This allows to obtain full control over the whole Zak-Berry connection in the lattice theory via the Zak-Berry connection defined in the field-theoretical models. For the lattice models studied in this work the band structure in the middle of the bands is characterized by free plane waves, leading naturally to the choice of asymptotically free plane waves in the corresponding field-theoretical models. For more general multichannel models, as well as in the presence of spin-orbit interaction and Zeeman fields, generalizations have to be studied for small gaps by expanding around a different reference system [called  $H_0$  in our work, see Eq. (2.2)]. This reference system will again prescribe certain asymptotic conditions for the field theory far away from the gap, which can be used to set up unique correlations between gaps opening up at different energies. Therefore, we expect that our way to describe universal properties of Wannier functions and the boundary charge via low-energy models will be also of interest for the description of global topological aspects in more general models.

## ACKNOWLEDGMENTS

We thank S. Blügel, E. Koch, and V. Meden for fruitful discussions. This work was supported by the Deutsche Forschungsgemeinschaft via RTG 1995, the Swiss National Science Foundation (SNSF) and NCCR QSIT, and by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy–Cluster of Excellence Matter and Light for Quantum Computing

(ML4Q) EXC 2004/1-390534769. We acknowledge support from the Max Planck New York City Center for Non-Equilibrium Quantum Phenomena. Simulations were performed with computing resources granted by RWTH Aachen University. Funding was received from the European Union's Horizon 2020 research and innovation program (ERC Starting Grant, Grant Agreement No. 757725).

K.P. and M.P. contributed equally to this work.

### APPENDIX A: GAPLESS CASE, NORMALIZATION, AND COMPLETENESS RELATION

Here we show that the normalization and completeness relations (2.13) and (2.14) are consistent with the corresponding ones in the low-energy region, given by (2.35) and (2.36). To achieve this we first discuss the gapless case and show that (2.29) reproduces the correct result for the eigenstates of  $H_0$ .

For the gapless case  $\delta v_m = \delta t_m = 0$  the Hamiltonian  $h_k$  given by (2.10) has the form

$$h_k = -t \sum_{j=1}^Z (|j+1\rangle\langle j|e^{-ika} + |j\rangle\langle j+1|e^{ika}). \quad (\text{A1})$$

One obtains straightforwardly for the dispersion and the eigenfunctions

$$\epsilon_k = -2t \cos\left(ka + \frac{2\pi}{Z}n_{k\alpha}\right), \quad (\text{A2})$$

$$u_{k\alpha}(ja) = \frac{1}{\sqrt{Z}} e^{i\frac{2\pi}{Z}n_{k\alpha}j}, \quad (\text{A3})$$

where  $n_{k\alpha}$  are  $Z$  integers in some interval of size  $Z$ . To achieve the ordering  $\epsilon_k^{(1)} < \epsilon_k^{(2)} < \dots < \epsilon_k^{(Z)}$  in the reduced zone scheme  $-\frac{\pi}{Za} < k < \frac{\pi}{Za}$ , we choose  $n_{k1} = 0$ ,  $n_{k2} = -\text{sgn}(k)$ ,  $n_{k3} = \text{sgn}(k)$ ,  $n_{k4} = -2\text{sgn}(k)$ ,  $n_{k5} = 2\text{sgn}(k)$ , etc., which gives

$$n_{k\alpha} = \frac{1}{2}\text{sgn}(k) \begin{cases} -\alpha & \text{for } \alpha \text{ even,} \\ \alpha - 1 & \text{for } \alpha \text{ odd} \end{cases} \quad (\text{A4})$$

leading to

$$u_{k\alpha}(ja) = \frac{1}{\sqrt{Z}} \begin{cases} e^{-i\text{sgn}(k)\frac{\pi}{2}\alpha j} & \text{for } \alpha \text{ even,} \\ e^{i\text{sgn}(k)\frac{\pi}{2}(\alpha-1)j} & \text{for } \alpha \text{ odd.} \end{cases} \quad (\text{A5})$$

According to (2.30) and (2.31) it is straightforward to show that this corresponds precisely to the choice  $\chi_{q\tau}^{(v)} = \delta_{p,\text{sgn}(q\tau)}$  in Dirac theory, as stated in (2.29).

To check the normalization (2.13) we prove (2.40) rigorously for any integers  $r, s \geq 0$  via the relations (2.30) and (2.31). Except for the lower half of the lowest band (i.e., for  $|k| < \frac{\pi}{2Za}$  and  $\alpha = 1$ ) or the upper half of the highest band (i.e., for  $|k| > \frac{\pi}{2Za}$  and  $\alpha = Z$ ), the gaps with even  $\nu$  describe the region  $|k| < \frac{\pi}{2Za}$ , whereas the gaps with odd  $\nu$  correspond to  $|k| > \frac{\pi}{2Za}$ . Therefore, in almost all cases, the parity of  $\nu$  and  $\nu'$  must be the same for given  $k$ . In this case we get

$$\begin{aligned} & \langle \partial_k^r u_{k\alpha} | \partial_k^s u_{k\alpha'} \rangle \\ &= \sum_{p,p'=\pm} \partial_q^r (\chi_{q\tau p}^{(v)})^* \partial_q^s \chi_{q\tau' p'}^{(v')} \frac{1}{Z} \sum_{j=1}^Z e^{-i\pi(p\nu - p'\nu')j/Z} \end{aligned}$$

$$\begin{aligned} &= \sum_{p,p'=\pm} \partial_q^r (\chi_{q\tau p}^{(v)})^* \partial_q^s \chi_{q\tau' p'}^{(v')} \delta_{p\nu, p'\nu'} \\ &= \delta_{\nu\nu'} \langle \partial_q^r \chi_{q\tau}^{(v)} | \partial_q^s \chi_{q\tau'}^{(v')} \rangle, \end{aligned} \quad (\text{A6})$$

where we used  $\delta_{p\nu, p'\nu'} = \delta_{pp'} \delta_{\nu\nu'}$  in the last step (since  $\nu, \nu' > 0$ ) and the fact that  $p\nu - p'\nu'$  is always an even number (since  $\nu$  and  $\nu'$  have the same parity). The special case  $\nu = 1$  and  $\nu'$  even is only possible for  $\tau = -$ ,  $p = \text{sgn}(k) = -\text{sgn}(q)$ , and  $|k| < \frac{\pi}{2Za}$ . It leads to

$$\begin{aligned} & \langle \partial_k^r u_{k1} | \partial_k^s u_{k\alpha'} \rangle \\ &= \sum_{p,p'=\pm} \partial_q^r (\chi_{q,-,p}^{(1)})^* \partial_q^s \chi_{q\tau' p'}^{(v')} \frac{1}{Z} \sum_{j=1}^Z e^{-i\pi(p+\text{sgn}(q)-p'\nu')j/Z}. \end{aligned}$$

This expression must be zero since, by using  $p + \text{sgn}(q) = 0$  and  $\nu'$  even, we conclude that  $p'\nu'$  must be zero which is not possible. Finally, the special case  $\nu = Z - 1$  and  $\nu'$  having a different parity than  $Z - 1$  can be treated in a similar way and leads also to a contradiction.

To prove the completeness relationship (2.14) for given  $k$  we note that, for each given gap  $\nu$ , one has to sum over both band indices  $\tau = \pm$  of Dirac theory, except for  $\nu = 1$  or  $\nu = Z - 1$  when  $|k| < \frac{\pi}{2Za}$  or  $|k| > \frac{\pi}{2Za}$ , respectively. In the latter case only the Dirac bands  $\tau = -$  (for  $\nu = 1$  and  $|k| < \frac{\pi}{2Za}$ ) or  $\tau = +$  (for  $\nu = Z - 1$  and  $|k| > \frac{\pi}{2Za}$ ) have to be taken, respectively. This means that only the eigenstates far away from the gap  $\nu = 1$  or  $\nu = Z - 1$  are involved for the lowest or highest lattice band, where we can take approximately the free solution (A5). The same happens for the contribution from the pairs  $\tau = \pm$  since one finds from (2.30) and (2.31) that the expression  $\sum_{\tau} u_{k\alpha}(ja)u_{k\alpha'}(j'a)^*$  involves always the combination

$$\sum_{\tau} \chi_{q\tau p}^{(v)} (\chi_{q\tau p'}^{(v)})^* = \delta_{pp'}, \quad (\text{A7})$$

where we used the completeness relationship (2.36) of the Dirac states. Since the same result arises for the free Dirac states  $\chi_{q\tau p}^{(v)} = \delta_{p,\text{sgn}(q\tau)}$ , we find that the completeness relationship in the presence of a gap is not changed compared to the one for the free Bloch states (A5).

### APPENDIX B: CANCELLATION OF ZAK-BERRY PHASES

Here we show on quite general grounds that the sum of the Zak-Berry phases in the AF gauge of the two bands of the Dirac model is quantized in odd units of  $\pi$ . We first need that one can write (with  $\chi_{q\tau} \equiv |\chi_{q\tau}\rangle$  defining column vectors in  $p = \pm$ )

$$\begin{aligned} \sum_{\tau} \gamma_{\tau}^{(v)} &= \sum_{\tau} \int dq A_{q\tau}^{(v)} = \sum_{\tau} \int dq (\chi_{q\tau}^{(v)})^{\dagger} i \partial_q \chi_{q\tau}^{(v)} \\ &= \int dq \text{Tr}(M_q^{(v)})^{\dagger} i \partial_q M_q^{(v)} = i \int \text{Tr}(M_q^{(v)})^{\dagger} dM_q^{(v)} \\ &= i \int d \ln(\det M_q^{(v)}), \end{aligned} \quad (\text{B1})$$

where we defined the unitary matrix

$$M_q^{(v)} = \begin{pmatrix} \chi_{q+}^{(v)} & \chi_{q-}^{(v)} \end{pmatrix}. \quad (\text{B2})$$

This means that (up to multiples of  $2\pi$ ) the sum of the Zak-Berry phases of the two Dirac bands can be written as minus the phase difference of the determinant of  $M_q^{(v)}$  between the asymptotic values at  $q = \infty$  and  $-\infty$ :

$$\sum_{\tau} \gamma_{\tau}^{(v)} = - \sum_{r=1,2} (\varphi_{q=\infty,r}^{(v)} - \varphi_{q=-\infty,r}^{(v)}) \pmod{2\pi}, \quad (\text{B3})$$

where  $e^{i\varphi_{qr}^{(v)}}$ , with  $r = 1, 2$ , denote the two eigenvalues of the unitary matrix  $M_q^{(v)}$ . In the AF gauge the asymptotic values of  $\chi_{q\tau p}^{(v)}$  are given by  $\delta_{p,\text{sgn}(q\tau)}$  according to (2.29), leading to

$$\lim_{q \rightarrow \infty} M_q^{(v)} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad (\text{B4})$$

$$\lim_{q \rightarrow -\infty} M_q^{(v)} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \quad (\text{B5})$$

This gives

$$\sum_r \varphi_{q=\infty,r}^{(v)} = 0, \quad \sum_r \varphi_{q=-\infty,r}^{(v)} = \pi, \quad (\text{B6})$$

and we get from (B3) the final result

$$\sum_{\tau} \gamma_{\tau}^{(v)} = \pi \pmod{2\pi}, \quad (\text{B7})$$

consistent with (3.50).

### APPENDIX C: SPLITTING OF WANNIER FUNCTION

Here we show how to split the lattice Wannier function for band  $\alpha = 2, \dots, Z-1$  into the contributions corresponding to the upper and lower halves of the band and how to extend the momentum integration to infinity [see Eqs. (4.8), (4.9), (4.12), and (4.13)]. Since the eigenfunctions of the lattice and the Dirac model are related via (2.22), (2.23), (2.24), (2.30), and (2.31), we can perform this analysis directly within the Dirac model by calculating the contributions from the region  $|q| > \Lambda_{\alpha}$  of the Dirac Wannier functions occurring on the right-hand sides of (4.44), (4.45), (4.52), and (4.53). Denoting the contribution of the various Wannier functions from this region by  $\delta w_{u,d,\alpha}(ma)$  we get

$$\delta w_{u,\alpha}(ma) = a\sqrt{Z} \sum_p \delta w_{-,p}^{(\alpha)}(ma) e^{ipk_F^{(\alpha)} ma}, \quad (\text{C1})$$

$$\delta w_{d,\alpha}(ma) = a\sqrt{Z} \sum_p \delta w_{+,p}^{(\alpha-1)}(ma) e^{ipk_F^{(\alpha-1)} ma}, \quad (\text{C2})$$

$$\delta \tilde{w}_u^{(\alpha)}(ma) = a\sqrt{Z} \sum_p e^{-i\phi_{0,-}^{(\alpha)}} \delta \tilde{w}_{-,p}^{(\alpha)}(\tilde{m}_{\alpha} a) e^{ipk_F^{(\alpha)} ma}, \quad (\text{C3})$$

$$\delta \tilde{w}_d^{(\alpha)}(ma) = a\sqrt{Z} \sum_p e^{-i\phi_{0,+}^{(\alpha-1)}} \delta \tilde{w}_{+,p}^{(\alpha-1)}(\tilde{m}_{\alpha} a) e^{ipk_F^{(\alpha-1)} ma}, \quad (\text{C4})$$

where

$$\delta w_{\tau p}^{(v)}(x) = \int_{|q| > \Lambda_{\alpha}} \frac{dq}{2\pi} \chi_{q\tau p}^{(v)} e^{iqx} e^{-\eta|q|}, \quad (\text{C5})$$

$$\delta \tilde{w}_{\tau p}^{(v)}(x) = \int_{|q| > \Lambda_{\alpha}} \frac{dq}{2\pi} \tilde{\chi}_{q\tau p}^{(v)} e^{iqx} e^{-\eta|q|}, \quad (\text{C6})$$

with  $\tilde{\chi}_{q\tau p}^{(v)} = e^{i\phi_{qr}^{(v)}} \chi_{q\tau p}^{(v)}$ . We have introduced a convenient regularization of the integrals via the convergence factor  $e^{-\eta|q|}$ .

For  $|q| > \Lambda_{\alpha}$  we can take the free form  $\chi_{q\tau p}^{(v)} \approx \delta_{\text{sgn}(q),p\tau}$  [cf. (2.29)], and get with (3.36), (3.26), and (3.33) the following value for the phase  $\phi_{q\tau}^{(v)}$  in the asymptotic high-momentum region

$$\begin{aligned} \phi_{q\tau}^{(v)} - \phi_{0,\tau}^{(v)} &= \frac{1}{2} \int_{-q}^q dq' A_{q'\tau}^{(v)} \\ &\approx \frac{1}{2} \text{sgn}(q) \int dq' A_{q'\tau}^{(v)} = \frac{1}{2} \text{sgn}(q) \gamma_{\tau}^{(v)}. \end{aligned} \quad (\text{C7})$$

Using these relations we obtain for (C5) and (C6)

$$\delta w_{\tau p}^{(v)}(x) = \frac{i}{2\pi} \frac{e^{ip\tau \Lambda_{\alpha} x}}{\tau px + i\eta}, \quad (\text{C8})$$

$$e^{-i\phi_{0,\tau}^{(v)}} \delta \tilde{w}_{\tau p}^{(v)}(x) = \frac{i}{2\pi} \frac{e^{ip\tau(\Lambda_{\alpha} x + \frac{1}{2}\gamma_{\tau}^{(v)})}}{\tau px + i\eta}. \quad (\text{C9})$$

Leaving out the contribution proportional to the delta function  $\delta(x)$  (this is the regularization we use in Sec. V to define the Dirac Wannier function), we obtain after inserting these expressions into Eqs. (C1)–(C4)

$$\delta w_{u,\alpha}(ma) = \frac{\sqrt{Z}}{\pi m} \sin \{ (k_F^{(\alpha)} - \Lambda_{\alpha}) ma \}, \quad (\text{C10})$$

$$\delta w_{d,\alpha}(ma) = -\frac{\sqrt{Z}}{\pi m} \sin \{ (k_F^{(\alpha-1)} + \Lambda_{\alpha}) ma \}, \quad (\text{C11})$$

$$\begin{aligned} \delta \tilde{w}_u^{(\alpha)}(ma) &= \frac{\sqrt{Z}}{\pi \tilde{m}_{\alpha}} \sin \left\{ k_F^{(\alpha)} ma - \Lambda_{\alpha} \tilde{m}_{\alpha} a - \frac{1}{2} \gamma_{-}^{(\alpha)} \right\}, \end{aligned} \quad (\text{C12})$$

$$\begin{aligned} \delta \tilde{w}_d^{(\alpha)}(ma) &= -\frac{\sqrt{Z}}{\pi \tilde{m}_{\alpha}} \sin \left\{ k_F^{(\alpha-1)} ma + \Lambda_{\alpha} \tilde{m}_{\alpha} a + \frac{1}{2} \gamma_{+}^{(\alpha-1)} \right\}. \end{aligned} \quad (\text{C13})$$

For  $\alpha = 2, \dots, Z-1$  we can insert  $\Lambda_{\alpha} = \frac{\pi}{2Za}$  and  $\gamma_{\alpha} = \gamma_{+}^{(\alpha-1)} + \gamma_{-}^{(\alpha)}$ , and get with  $k_F^{(\alpha)} = \alpha \frac{\pi}{Za}$  and  $\tilde{m}_{\alpha} = m - \frac{Z}{2\pi} \gamma_{\alpha}$  the final result

$$\delta w_{u,\alpha}(ma) = -\delta w_{d,\alpha}(ma) = \delta w_{\alpha}(ma), \quad (\text{C14})$$

$$\delta \tilde{w}_u^{(\alpha)}(ma) = -\delta \tilde{w}_d^{(\alpha)}(ma) = \delta \tilde{w}_{\alpha}(ma), \quad (\text{C15})$$

where  $\delta w_{\alpha}(ma)$  and  $\delta \tilde{w}_{\alpha}(ma)$  are given by (4.12) and (4.13), respectively.

In contrast, for  $\alpha = 1, Z$ , we can use  $\Lambda_{1/Z} = \frac{\pi}{Za}$ ,  $\gamma_1 = \gamma_{-}^{(1)}$ , and  $\gamma_Z = \gamma_{+}^{(Z-1)}$  with the result

$$\delta w_{u,1}(ma) = \delta w_{d,Z}(ma) = 0, \quad (\text{C16})$$

$$\delta \tilde{w}_{u,1}(ma) = \delta \tilde{w}_{d,Z}(ma) = 0, \quad (\text{C17})$$

in accordance with (4.15).

### APPENDIX D: REAL MOMENTUM INTEGRAL REPRESENTATIONS OF DIRAC WANNIER FUNCTIONS AND SCALING FUNCTIONS

In this Appendix we present the representation of the Dirac Wannier functions  $w_{p\tau}^{(v)}(x)$  and  $\tilde{w}_{p\tau}^{(v)}(x)$  together with the scaling functions  $F_{A/B,\tau}(y; \gamma)$  and  $\tilde{F}_{A/B,\tau}(y; \gamma)$  via convergent momentum integrals on the real axis. Starting from



(5.14)–(5.17), we first subtract the large  $|\bar{q}| \gg 1$  behavior of the integrands to get convergent integrals and to explicitly exhibit the part which remains for  $\bar{\eta} \rightarrow 0$ . Using (2.29) and (C7) together with (C8) and (C9) (with  $\Lambda_\alpha = 0$ ) we find

$$w_{\tau p}^{(v)}(x) = \frac{i}{2\pi} \frac{1}{p\tau x + i\eta} + \frac{1}{4\pi \xi^{(v)}} \int d\bar{q} e^{i\bar{q}x} [\chi_{\tau p}^{(v)}(\bar{q}) - \delta_{\text{sgn}\bar{q}, p\tau}], \quad (\text{D1})$$

$$\tilde{w}_{\tau p}^{(v)}(x) = \frac{i}{2\pi} \frac{e^{i\phi_{0,\tau}^{(v)}} e^{i\frac{1}{2}p\tau\gamma_\tau^{(v)}}}{p\tau x + i\eta} + \frac{1}{4\pi \xi^{(v)}} \int d\bar{q} e^{i\bar{q}x} [\tilde{\chi}_{\tau p}^{(v)}(\bar{q}) - e^{i\phi_{0,\tau}^{(v)}} e^{i\frac{1}{2}p\tau\gamma_\tau^{(v)}} \delta_{\text{sgn}\bar{q}, p\tau}]. \quad (\text{D2})$$

The first term on the right-hand side of the Dirac Wannier functions contains an unphysical contribution  $\sim \delta(x)$  which arises from large momenta and has to be subtracted in the regularization procedure. In the scaling functions this part does anyhow not appear since one multiplies with  $x$ . In contrast, the principal part  $\sim \frac{1}{x}$  of the first term is an important contribution to the scaling functions. It should be contrasted to the contributions (C8) and (C9) which arise from extending the momentum integration to infinity and cancel in the lattice Wannier function of a certain band when adding the

contributions from the upper and lower parts of the band to the Wannier function (see Appendix C).

Using (5.11) and (3.31) we note the identities

$$e^{i\phi_{0,\tau}^{(v)}} e^{i\frac{1}{2}p\tau\gamma_\tau^{(v)}} = \begin{cases} e^{i\gamma^{(v)}} & \text{for } p = +, \\ \tau & \text{for } p = -, \end{cases} \quad (\text{D3})$$

$$e^{i\frac{1}{2}p\tau\gamma_\tau^{(v)}} = e^{i\frac{1}{2}p\gamma^{(v)}} \begin{cases} 1 & \text{for } \tau = +, \\ -ips_{\gamma^{(v)}} & \text{for } \tau = -. \end{cases} \quad (\text{D4})$$

Multiplying (D1) and (D2) with  $x$ , we obtain with (D3) in the limit  $\eta \rightarrow 0$  the useful representations

$$x w_{\tau p}^{(v)}(x) = i \frac{p\tau}{2\pi} + \frac{y}{2} \int \frac{d\bar{q}}{2\pi} e^{i\bar{q}y} [\chi_{\tau p}^{(v)}(\bar{q}) - \delta_{\text{sgn}\bar{q}, p\tau}], \quad (\text{D5})$$

$$e^{-i\phi_{0,\tau}^{(v)}} x \tilde{w}_{\tau p}^{(v)}(x) = e^{i\frac{1}{2}p\tau\gamma_\tau^{(v)}} \left\{ i \frac{p\tau}{2\pi} + \frac{y}{2} \int \frac{d\bar{q}}{2\pi} \times e^{i\bar{q}y} \left[ \frac{\sqrt{\bar{\epsilon}_q + p\tau\bar{q}}}{\sqrt{2\bar{\epsilon}_q}} - \delta_{\text{sgn}\bar{q}, p\tau} \right] \right\}, \quad (\text{D6})$$

where we inserted the form (5.17) for  $\tilde{\chi}_{\tau p}^{(v)}(\bar{q})$ , and used the definition  $y = x/\xi^{(v)} = 2\bar{x}$ . Inserting (5.16) and using (D4) one can calculate the scaling functions from (5.24)–(5.29) and finds after some straightforward manipulations the compact forms

$$F_{A,\tau}(y; \gamma) = \frac{\tau}{\pi} + \tau y \int_0^\infty \frac{d\bar{q}}{2\pi} \left[ \frac{\bar{q} \sin \frac{\bar{q}y}{2} + \sin \gamma \cos \frac{\bar{q}y}{2}}{\sqrt{\bar{\epsilon}_q}(\bar{\epsilon}_q + \tau \cos \gamma)} - \sin \frac{\bar{q}y}{2} \right], \quad (\text{D7})$$

$$F_{B,\tau}(y; \gamma) = y \int_0^\infty \frac{d\bar{q}}{2\pi} \cos \frac{\bar{q}y}{2} \left[ \sqrt{1 + \frac{\tau \cos \gamma}{\bar{\epsilon}_q}} - 1 \right], \quad (\text{D8})$$

$$\tilde{F}_{A,-}(y; \gamma) = -\frac{s_\gamma}{\pi} \left\{ \sin \frac{\gamma}{2} + \frac{y}{2\sqrt{2}} \int_0^\infty d\bar{q} \left[ \left( \sqrt{1 + \frac{\bar{q}}{\bar{\epsilon}_q}} - \sqrt{2} \right) \cos \frac{\bar{q}y - \gamma}{2} + \sqrt{1 - \frac{\bar{q}}{\bar{\epsilon}_q}} \cos \frac{\bar{q}y + \gamma}{2} \right] \right\}, \quad (\text{D9})$$

$$\tilde{F}_{B,-}(y; \gamma) = -\frac{s_\gamma}{\pi} \left\{ \cos \frac{\gamma}{2} + \frac{y}{2\sqrt{2}} \int_0^\infty d\bar{q} \left[ \left( \sqrt{1 + \frac{\bar{q}}{\bar{\epsilon}_q}} - \sqrt{2} \right) \sin \frac{\bar{q}y - \gamma}{2} - \sqrt{1 - \frac{\bar{q}}{\bar{\epsilon}_q}} \sin \frac{\bar{q}y + \gamma}{2} \right] \right\}, \quad (\text{D10})$$

$$\tilde{F}_{A,+}(y; \gamma) = \frac{1}{\pi} \cos \frac{\gamma}{2} + \frac{y}{2\pi\sqrt{2}} \int_0^\infty d\bar{q} \left[ \left( \sqrt{1 + \frac{\bar{q}}{\bar{\epsilon}_q}} - \sqrt{2} \right) \sin \frac{\bar{q}y + \gamma}{2} - \sqrt{1 - \frac{\bar{q}}{\bar{\epsilon}_q}} \sin \frac{\bar{q}y - \gamma}{2} \right], \quad (\text{D11})$$

$$\tilde{F}_{B,+}(y; \gamma) = -\frac{1}{\pi} \sin \frac{\gamma}{2} + \frac{y}{2\pi\sqrt{2}} \int_0^\infty d\bar{q} \left[ \left( \sqrt{1 + \frac{\bar{q}}{\bar{\epsilon}_q}} - \sqrt{2} \right) \cos \frac{\bar{q}y + \gamma}{2} + \sqrt{1 - \frac{\bar{q}}{\bar{\epsilon}_q}} \cos \frac{\bar{q}y - \gamma}{2} \right], \quad (\text{D12})$$

with  $\bar{\epsilon}_q = \sqrt{1 + \bar{q}^2}$  and  $s_\gamma = \text{sgn} \gamma$ . These formulas can be straightforwardly evaluated numerically and have been used to produce the results shown in Fig. 3.

#### APPENDIX E: INTEGRAL REPRESENTATION ON THE IMAGINARY MOMENTUM AXIS AND ASYMPTOTICS OF WANNIER FUNCTIONS IN DIRAC THEORY

Here we show how to obtain the integral representations (5.18) and (5.19) for the Wannier functions  $w_{\tau p}^{(v)}(x)$  and  $\tilde{w}_{\tau p}^{(v)}(x)$  in the universal regime  $|x| \gg a$  of Dirac theory, together with their asymptotics at  $|x| \ll \xi^{(v)}$  and  $|x| \gg \xi^{(v)}$ , as given by (E9), (E10), (E14)–(E18), and (E20).

To obtain (5.18) and (5.19), we use (5.14) and (5.15), and close the integration contour in the upper or lower half of the complex plane for  $s_x = \text{sgn}(x) = \pm$ , respectively. Choosing the branch cuts of the various square roots and of  $e^{-\bar{\eta}|\bar{q}|}$  on the imaginary axis, we can close the integration contour around the branch cut, and obtain with  $\bar{q} = is_x\kappa$  and  $0 < \kappa < \infty$  the form

$$w_{\tau p}^{(v)}(x) = \frac{is_x}{4\pi \xi^{(v)}} \int_0^\infty d\kappa e^{-|\bar{x}|\kappa} \left\{ e^{-is_x\bar{\eta}\kappa} \chi_{\tau p}^{(v)}(is_x\kappa + 0^+) - e^{is_x\bar{\eta}\kappa} \chi_{\tau p}^{(v)}(is_x\kappa - 0^+) \right\}, \quad (\text{E1})$$

and an analog equation for  $\tilde{w}_{\tau p}^{(v)}(x)$  with  $\chi_{\tau p}^{(v)} \rightarrow \tilde{\chi}_{\tau p}^{(v)}$ . Due to the exponentially decaying factor  $e^{-|\bar{x}|\kappa}$ , these integrals are

convergent in the limit  $\bar{\eta} \rightarrow 0$ , and we obtain the results (5.18) and (5.19) in this limit.

To evaluate the integrands  $\delta\chi_{\tau p}^{(v)}(is_x\kappa)$  and  $\delta\tilde{\chi}_{\tau p}^{(v)}(is_x\kappa)$  of (5.18) and (5.19), we note the following values of the various square roots for  $\bar{q} = is_x\kappa \pm 0^+$ :

$$\bar{\epsilon}_{\bar{q}} = |1 - \kappa^2|^{\frac{1}{2}} \begin{cases} 1 \pm is_x 0^+ & \text{for } \kappa < 1, \\ \pm is_x & \text{for } \kappa > 1, \end{cases} \quad (\text{E2})$$

$$\sqrt{\bar{\epsilon}_{\bar{q}}} = |1 - \kappa^2|^{\frac{1}{4}} \begin{cases} 1 & \text{for } \kappa < 1, \\ e^{\pm is_x \frac{\pi}{4}} & \text{for } \kappa > 1, \end{cases} \quad (\text{E3})$$

$$\sqrt{\bar{\epsilon}_{\bar{q}} + p\tau\bar{q}} = \begin{cases} \sqrt{|1 - \kappa^2|^{\frac{1}{2}} + ip\tau s_x\kappa} & \text{for } \kappa < 1, \\ \sqrt{\pm is_x |1 - \kappa^2|^{\frac{1}{2}} + ip\tau s_x\kappa} & \text{for } \kappa > 1. \end{cases} \quad (\text{E4})$$

Furthermore, for  $\tau \cos \gamma^{(v)} > 0$ , we get

$$\begin{aligned} & \sqrt{\bar{\epsilon}_{\bar{q}} + \tau \cos \gamma^{(v)}} \\ &= \begin{cases} \sqrt{|1 - \kappa^2|^{\frac{1}{2}} + |\cos \gamma^{(v)}|} & \text{for } \kappa < 1, \\ \sqrt{\pm is_x |1 - \kappa^2|^{\frac{1}{2}} + |\cos \gamma^{(v)}|} & \text{for } \kappa > 1, \end{cases} \end{aligned} \quad (\text{E5})$$

and for  $\tau \cos \gamma^{(v)} < 0$

$$\begin{aligned} & \sqrt{\bar{\epsilon}_{\bar{q}} + \tau \cos \gamma^{(v)}} \\ &= \begin{cases} \sqrt{|1 - \kappa^2|^{\frac{1}{2}} - |\cos \gamma^{(v)}|} & \text{for } \kappa < |\sin \gamma^{(v)}|, \\ \pm is_x \sqrt{-|1 - \kappa^2|^{\frac{1}{2}} + |\cos \gamma^{(v)}|} & \text{for } |\sin \gamma^{(v)}| < \kappa < 1, \\ \sqrt{\pm is_x |1 - \kappa^2|^{\frac{1}{2}} - |\cos \gamma^{(v)}|} & \text{for } \kappa > 1. \end{cases} \end{aligned} \quad (\text{E6})$$

For  $|\bar{x}| \ll 1$ , the integrals are dominated by the regime  $\kappa \gg 1$ , where we get

$$\delta\chi_{\tau p}^{(v)}(is_x\kappa) \rightarrow p\tau, \quad (\text{E7})$$

$$\delta\tilde{\chi}_{\tau p}^{(v)}(is_x\kappa) \rightarrow p\tau \begin{cases} e^{i\gamma^{(v)}} & \text{for } p = +, \\ \tau & \text{for } p = -. \end{cases} \quad (\text{E8})$$

Inserting these forms in (5.18) and (5.19), we find straightforwardly the following result for small  $|x| \ll \xi^{(v)}$ :

$$w_{\tau p}^{(v)}(x) \rightarrow \frac{ip\tau}{2\pi x}, \quad (\text{E9})$$

$$\tilde{w}_{\tau p}^{(v)}(x) \rightarrow \frac{ip\tau}{2\pi x} \begin{cases} e^{i\gamma^{(v)}} & \text{for } p = +, \\ \tau & \text{for } p = -. \end{cases} \quad (\text{E10})$$

For  $|\bar{x}| \sim O(1)$ , the integrals are of  $O(1)$ , leading to the universal scaling of the Wannier functions in the regime  $|x| \sim \xi^{(v)}$ ,

$$|x w_{\tau p}^{(v)}(x)| \sim |x \tilde{w}_{\tau p}^{(v)}(x)| \sim O(1). \quad (\text{E11})$$

Due to the results (E9) and (E10), this scaling holds also for  $|x| \ll \xi^{(v)}$ .

For  $|\bar{x}| \gg 1$ , the integrals are dominated by the regime around that branching point of the various square roots which has the smallest imaginary part. We start with  $w_{\tau p}^{(v)}(x)$ . For  $\tau \cos \gamma^{(v)} > 0$ , there is only a single branching point at  $\kappa = 1$ , and close to this branching point we get for small  $0 < \kappa' = \kappa - 1 \ll 1$

$$\delta\chi_{\tau p}^{(v)}(is_x(1 + \kappa')) \approx \frac{\tau(ip s_x + e^{ip\gamma^{(v)}})}{is_x 2^{\frac{3}{4}} \sqrt{|\cos \gamma^{(v)}|}} (\kappa')^{-\frac{1}{4}}. \quad (\text{E12})$$

Inserting this form in (5.18) and using

$$\int_0^\infty d\kappa' e^{-|\bar{x}|\kappa'} (\kappa')^r = \frac{\Gamma(1+r)}{|\bar{x}|^{1+r}}, \quad (\text{E13})$$

we find the following result for  $|x| \gg \xi^{(v)}$ :

$$\tau = \text{sgn}(\cos \gamma^{(v)}):$$

$$w_{\tau p}^{(v)}(x) \rightarrow \frac{\tau(ip s_x + e^{ip\gamma^{(v)}})\Gamma(\frac{3}{4})}{4\pi \sqrt{|\cos \gamma^{(v)}|} \xi^{(v)}} \left(\frac{\xi^{(v)}}{|x|}\right)^{\frac{3}{4}} e^{-\frac{|x|}{2\xi^{(v)}}}. \quad (\text{E14})$$

In contrast, for  $\tau \cos \gamma^{(v)} < 0$ , the branching point with smallest imaginary part is located at  $\kappa = |\sin \gamma^{(v)}|$ . Excluding the case  $\sin \gamma^{(v)} = 0$  (which is treated in detail in Appendix G) we obtain close to this branching point for small  $0 < \kappa' = \kappa - |\sin \gamma^{(v)}| \ll 1$

$$\begin{aligned} \delta\chi_{\tau p}^{(v)}(is_x(|\sin \gamma^{(v)}| + \kappa')) &\approx \frac{ip\tau |\sin \gamma^{(v)}| [s_x + s_{\gamma^{(v)}}] + (-|\tan \gamma^{(v)}| + ip\tau s_x)\kappa'}{is_x \sqrt{|\sin \gamma^{(v)}|} \sqrt{\kappa'}} \\ &\approx \begin{cases} 2p\tau \sqrt{|\sin \gamma^{(v)}|} (\kappa')^{-\frac{1}{2}} & \text{for } s_x = s_{\gamma^{(v)}}, \\ \frac{-|\tan \gamma^{(v)}| + ip\tau s_x}{is_x \sqrt{|\sin \gamma^{(v)}|}} (\kappa')^{\frac{1}{2}} & \text{for } s_x = -s_{\gamma^{(v)}}. \end{cases} \end{aligned} \quad (\text{E15})$$

Inserting this form in (5.18) and using (E13), we find for  $|x| \gg \xi^{(v)}$  the asymptotic behavior

$$\tau = -\text{sgn}(\cos \gamma^{(v)}):$$

$$w_{\tau p}^{(v)}(x) \rightarrow \frac{ip\tau s_x - |\tan \gamma^{(v)}|}{2\sqrt{2\pi} \sqrt{|\sin \gamma^{(v)}|} \xi^{(v)}} \left(\frac{\xi^{(v)}}{|x|}\right)^{\frac{3}{2}} e^{-|\sin \gamma^{(v)}| \frac{|x|}{2\xi^{(v)}}} + \frac{1 + s_x s_{\gamma^{(v)}}}{2} \frac{ip\tau s_x \sqrt{|\sin \gamma^{(v)}|}}{\sqrt{2\pi} \xi^{(v)}} \left(\frac{\xi^{(v)}}{|x|}\right)^{\frac{1}{2}} e^{-|\sin \gamma^{(v)}| \frac{|x|}{2\xi^{(v)}}}. \quad (\text{E16})$$

For the special case  $\cos \gamma^{(v)} = 0$ , the two branching points fall together and start at  $\kappa = 1$ . Close to this branching point we get for small  $0 < \kappa' = \kappa - 1 \ll 1$

$$\delta\chi_{\tau p}^{(v)}(is_x(1 + \kappa')) \approx p\tau \frac{1 + s_x s_{\gamma^{(v)}} + \kappa'}{\sqrt{2\kappa'}}. \quad (\text{E17})$$

Inserting this form in (5.18) we find the following result for  $|x| \gg \xi^{(v)}$ :

$$\cos \gamma^{(v)} = 0 :$$

$$w_{\tau p}^{(v)}(x) \rightarrow \frac{i p \tau s_x}{4 \sqrt{\pi} \xi^{(v)}} e^{-\frac{|x|}{2 \xi^{(v)}}} \times \left\{ (1 + s_x s_{\gamma^{(v)}}) \left( \frac{\xi^{(v)}}{|x|} \right)^{\frac{1}{2}} + \left( \frac{\xi^{(v)}}{|x|} \right)^{\frac{3}{2}} \right\}. \quad (E18)$$

Finally, for the Wannier function  $\tilde{w}_{\tau p}^{(v)}(x)$ , we obtain a single branching point at  $\kappa = 1$ , and get close to this branching point for small  $0 < \kappa' = \kappa - 1 \ll 1$ :

$$\delta \tilde{\chi}_{\tau p}^{(v)}(i s_x (1 + \kappa')) \approx \frac{e^{i p \tau s_x \frac{\pi}{4}}}{i s_x 2^{\frac{1}{4}}} (\kappa')^{-\frac{1}{4}} \begin{cases} e^{i \gamma^{(v)}} & \text{for } p = +, \\ \tau & \text{for } p = -. \end{cases} \quad (E19)$$

Inserting this form in (5.18) and using (E13), we find the following asymptotic behavior for  $|x| \gg \xi^{(v)}$  in all cases:

$$\tilde{w}_{\tau p}^{(v)}(x) \rightarrow \frac{e^{i p \tau s_x \frac{\pi}{4}} \Gamma\left(\frac{3}{4}\right)}{2^{\frac{3}{2}} \pi \xi^{(v)}} \left( \frac{\xi^{(v)}}{|x|} \right)^{\frac{3}{4}} e^{-|x|/(2 \xi^{(v)})} \times \begin{cases} e^{i \gamma^{(v)}} & \text{for } p = +, \\ \tau & \text{for } p = -. \end{cases} \quad (E20)$$

**APPENDIX F: PROPERTIES OF DIRAC WANNIER FUNCTIONS**

From (5.14)–(5.17) one can straightforwardly derive the following useful properties for  $w_{\tau p}^{(v)}(x)$  and  $\tilde{w}_{\tau p}^{(v)}(x)$ :

$$w_{\tau, -p}^{(v)}(x) = w_{\tau p}^{(v)}(x)^*, \quad (F1)$$

$$w_{\tau p}^{(v)}(x) = w_{\tau p}^{(v)}(-x)^* |_{\gamma^{(v)} \rightarrow -\gamma^{(v)}}, \quad (F2)$$

$$\sum_p w_{\tau p}^{(v)}(-x) = \sum_p w_{\tau p}^{(v)}(x), \quad (F3)$$

$$\sum_p p w_{\tau p}^{(v)}(-x) = - \sum_p p w_{\tau p}^{(v)}(x) |_{\gamma^{(v)} \rightarrow -\gamma^{(v)}}, \quad (F4)$$

$$w_{-\tau, p}^{(v)}(x) = w_{\tau p}^{(v)}(x)^* |_{\gamma^{(v)} \rightarrow -\gamma^{(v)} + \pi s_{\gamma^{(v)}}}, \quad (F5)$$

$$\tilde{w}_{\tau, -p}^{(v)}(x) = \tau e^{i \gamma^{(v)}} \tilde{w}_{\tau p}^{(v)}(x)^*, \quad (F6)$$

$$\tilde{w}_{\tau p}^{(v)}(x) = \tilde{w}_{\tau p}^{(v)}(-x)^* |_{\gamma^{(v)} \rightarrow -\gamma^{(v)}}, \quad (F7)$$

$$= -p \tilde{w}_{\tau p}^{(v)}(-x)^* |_{\gamma^{(v)} \rightarrow -\gamma^{(v)} + \pi s_{\gamma^{(v)}}}, \quad (F8)$$

$$\tilde{w}_{\tau p}^{(v)}(-x) = \tau e^{i p \gamma^{(v)}} \tilde{w}_{\tau, -p}^{(v)}(x), \quad (F9)$$

$$\tilde{w}_{-\tau, p}^{(v)}(x) = p \tilde{w}_{\tau p}^{(v)}(-x). \quad (F10)$$

Correspondingly, we find for the real functions  $w_{A/B, \tau}^{(v)}(x)$  and  $\tilde{w}_{A/B, \tau}^{(v)}(x)$ , as defined by (5.24)–(5.27), the properties

$$w_{A, \tau}^{(v)}(-x) = -w_{A, \tau}^{(v)}(x) |_{\gamma^{(v)} \rightarrow -\gamma^{(v)}}, \quad (F11)$$

$$w_{B, \tau}^{(v)}(-x) = w_{B, \tau}^{(v)}(x) = w_{B, \tau}^{(v)}(x) |_{\gamma^{(v)} \rightarrow -\gamma^{(v)}}, \quad (F12)$$

$$w_{A, -\tau}^{(v)}(x) = -w_{A, \tau}^{(v)}(x) |_{\gamma^{(v)} \rightarrow -\gamma^{(v)} + \pi s_{\gamma^{(v)}}}, \quad (F13)$$

$$w_{B, -\tau}^{(v)}(x) = w_{B, \tau}^{(v)}(x) |_{\gamma^{(v)} \rightarrow -\gamma^{(v)} + \pi s_{\gamma^{(v)}}}, \quad (F14)$$

$$\tilde{w}_{A, \tau}^{(v)}(-x) = -\tilde{w}_{A, \tau}^{(v)}(x) |_{\gamma^{(v)} \rightarrow -\gamma^{(v)}} \quad (F15)$$

$$= \tau s_{\gamma^{(v)}} \tilde{w}_{B, \tau}^{(v)}(x) |_{\gamma^{(v)} \rightarrow -\gamma^{(v)} + \pi s_{\gamma^{(v)}}} \quad (F16)$$

$$= \tau [-\cos \gamma^{(v)} \tilde{w}_{A, \tau}^{(v)}(x) + \sin \gamma^{(v)} \tilde{w}_{B, \tau}^{(v)}(x)], \quad (F17)$$

$$\tilde{w}_{B, \tau}^{(v)}(-x) = w_{B, \tau}^{(v)}(x) |_{\gamma^{(v)} \rightarrow -\gamma^{(v)}} \quad (F18)$$

$$= \tau s_{\gamma^{(v)}} \tilde{w}_{A, \tau}^{(v)}(x) |_{\gamma^{(v)} \rightarrow -\gamma^{(v)} + \pi s_{\gamma^{(v)}}} \quad (F19)$$

$$= \tau [\sin \gamma^{(v)} \tilde{w}_{A, \tau}^{(v)}(x) + \cos \gamma^{(v)} \tilde{w}_{B, \tau}^{(v)}(x)], \quad (F20)$$

$$\tilde{w}_{A, -\tau}^{(v)}(x) = -\tau s_{\gamma^{(v)}} \tilde{w}_{B, \tau}^{(v)}(-x), \quad (F21)$$

$$\tilde{w}_{B, -\tau}^{(v)}(x) = \tau s_{\gamma^{(v)}} \tilde{w}_{A, \tau}^{(v)}(-x). \quad (F22)$$

**APPENDIX G: SCALING FUNCTIONS FOR  $\sin \gamma = 0$**

In this Appendix we discuss the properties of the scaling functions  $F_{A/B, \tau}(y; \gamma)$  and  $\tilde{F}_{A/B, \tau}(y; \gamma)$  for the special case  $\sin \gamma = 0$ . Based on the relations (5.32)–(5.41) for  $\gamma = O^\pm$  and  $\gamma = \pi \pm O^+ = -\pi \pm 0^+$ , one finds that the scaling functions  $F_{A/B, \tau}(y; \gamma)$  are continuous in  $\gamma$  at  $\gamma = 0, \pi$  and have the following properties:

$$F_{A, \tau}(-y; 0) = F_{A, \tau}(y; 0), \quad (G1)$$

$$F_{A, \tau}(-y; \pi) = F_{A, \tau}(y; \pi), \quad (G2)$$

$$F_{B, \tau}(-y; 0) = -F_{B, \tau}(y; 0), \quad (G3)$$

$$F_{B, \tau}(-y; \pi) = F_{B, \tau}(y; \pi), \quad (G4)$$

$$F_{A, -\tau}(y; 0) = -F_{A, \tau}(y; \pi), \quad (G5)$$

$$F_{B, -\tau}(y; 0) = F_{B, \tau}(y; \pi). \quad (G6)$$

For the scaling functions  $\tilde{F}_{A/B, \tau}(y; \gamma)$  it turns out that they are continuous at  $\gamma = 0, \pi$  for  $\tau \cos \gamma > 0$  (i.e., for  $\gamma = \pi, \tau = -$  or for  $\gamma = 0, \tau = +$ ) but discontinuous for  $\tau \cos \gamma < 0$  (i.e., for  $\gamma = 0, \tau = -$  or for  $\gamma = \pi, \tau = +$ ), with the properties

$$\tilde{F}_{A, +}(-y; 0) = \tilde{F}_{A, +}(y; 0), \quad (G7)$$

$$\tilde{F}_{A, -}(-y; \pi) = \tilde{F}_{A, -}(y; \pi), \quad (G8)$$

$$\tilde{F}_{B, +}(-y; 0) = -\tilde{F}_{B, +}(y; 0), \quad (G9)$$

$$\tilde{F}_{B, -}(-y; \pi) = -\tilde{F}_{B, -}(y; \pi), \quad (G10)$$

$$\tilde{F}_{A, -}(y; 0^\pm) = \mp \tilde{F}_{B, +}(y; 0), \quad (G11)$$

$$\tilde{F}_{A, +}(y; \pi \pm 0^+) = \mp \tilde{F}_{B, -}(y; \pi), \quad (G12)$$

$$\tilde{F}_{B, -}(y; 0^\pm) = \mp \tilde{F}_{A, +}(y; 0), \quad (G13)$$

$$\tilde{F}_{B, +}(y; \pi \pm 0^+) = \mp \tilde{F}_{A, -}(y; \pi). \quad (G14)$$

For  $|y| \ll 1$  and  $|y| \gg 1$ , the values can be obtained straightforwardly from (5.43)–(5.54). A special case are the scaling functions  $F_{A/B,-}(y; 0)$  and  $F_{A/B,+}(y; \pi)$  which are not exponentially decaying since the branch cuts in the complex plane start at the origin. For  $\gamma = 0$  and  $\tau = -$  we obtain from (5.16)

$$\chi_{-,p}^{(v)}(\bar{q}) = \frac{\bar{\epsilon}_{\bar{q}} - p\bar{q} - 1}{2\sqrt{\bar{\epsilon}_{\bar{q}}}\sqrt{\bar{\epsilon}_{\bar{q}} - 1}}. \quad (\text{G15})$$

Expanding around  $\bar{q} = 0$  with  $\bar{q} = is_y\kappa \pm \eta$ , we obtain

$$\sum_p p \chi_{-,p}^{(v)}(is_y\kappa \pm \eta) \approx \mp\sqrt{2}, \quad (\text{G16})$$

$$\sum_p \chi_{-,p}^{(v)}(is_y\kappa \pm \eta) \approx \pm \frac{is_y\kappa}{\sqrt{2}}, \quad (\text{G17})$$

leading to

$$\sum_p p \delta \chi_{-,p}^{(v)}(is_y\kappa) \approx -2\sqrt{2}, \quad (\text{G18})$$

$$\sum_p \delta \chi_{-,p}^{(v)}(is_y\kappa) = i\sqrt{2} s_y\kappa. \quad (\text{G19})$$

Inserting in (5.18) we obtain for the asymptotic behavior

$$w_{A,-}^{(v)}(x) = -i \sum_p p w_{-,p}^{(v)}(x) \rightarrow -\frac{\sqrt{2}}{\pi x}, \quad (\text{G20})$$

$$w_{B,-}^{(v)}(x) = \sum_p w_{-,p}^{(v)}(x) \rightarrow -\frac{\sqrt{2}}{\pi \xi^{(v)}} \left(\frac{\xi^{(v)}}{x}\right)^2, \quad (\text{G21})$$

which gives for the scaling functions the results (5.55) and (5.56):

$$F_{A,-}(y; 0) = -F_{A,+}(y; \pi) \rightarrow -\frac{\sqrt{2}}{\pi}, \quad (\text{G22})$$

$$F_{B,-}(y; 0) = F_{B,+}(y; \pi) \rightarrow -\frac{\sqrt{2}}{\pi} \frac{1}{y}. \quad (\text{G23})$$

### APPENDIX H: BLOCH STATES FOR $Z = 2, 3$

The states  $u_{k\alpha}(ja) = \langle j|u_{k\alpha} \rangle$  follow from diagonalizing the Hamiltonian  $h_k$  given by (2.10). For the particular cases  $Z = 2$  and 3 we obtain the matrices

$$h_k = \begin{pmatrix} \delta v_1 & -t_1 e^{ika} - t_2 e^{-ika} \\ -t_1 e^{-ika} - t_2 e^{ika} & \delta v_2 \end{pmatrix}, \quad (\text{H1})$$

with  $\delta v_1 = -\delta v_2$  and  $\frac{1}{2}(t_1 + t_2) = t$ , and

$$h_k = \begin{pmatrix} \delta v_1 & -t_1 e^{ika} & -t_3 e^{-ika} \\ -t_1 e^{-ika} & \delta v_2 & -t_2 e^{ika} \\ -t_3 e^{ika} & -t_2 e^{-ika} & \delta v_3 \end{pmatrix}, \quad (\text{H2})$$

with  $\delta v_1 + \delta v_2 + \delta v_3 = 0$  and  $\frac{1}{3}(t_1 + t_2 + t_3) = t$ .

It is convenient to write

$$|u_{k\alpha}\rangle = \frac{(-1)^{\alpha+1}}{\sqrt{N_{k\alpha}}} \underline{x}_{k\alpha}, \quad N_{k\alpha} = \sum_{j=1}^Z |\underline{x}_{k\alpha}|^2, \quad (\text{H3})$$

where  $\underline{x}_{k\alpha}$  denotes a  $Z$ -dimensional vector.

For  $Z = 2$  this gives the explicit formulas

$$\underline{x}_{k\alpha} = \begin{pmatrix} t_1 e^{ika} + t_2 e^{-ika} \\ \delta v - \epsilon_{k\alpha} \end{pmatrix}, \quad (\text{H4})$$

$$N_{k\alpha} = 2\epsilon_{k\alpha}(\epsilon_{k\alpha} - \delta v), \quad (\text{H5})$$

$$\epsilon_{k,1/2} = \mp\sqrt{\Delta^2 + 4t_1 t_2 \cos^2(ka)}, \quad (\text{H6})$$

$$\Delta = \sqrt{\delta v^2 + 4\delta t^2}, \quad (\text{H7})$$

where  $\delta v = \delta v_1 = -\delta v_2$ , and  $t_{1/2} = t \pm \delta t$ . Using (5.103), (5.100), (5.101), and  $v_F^{(1)} = 2ta$ , we obtain for the gap parameter and the correlation length  $\xi^{(1)} = v_F^{(1)}/(2\Delta^{(1)})$

$$\Delta e^{i\gamma} \equiv \Delta^{(1)} e^{i\gamma^{(1)}} = -\delta v - 2i\delta t, \quad (\text{H8})$$

$$\xi \equiv \xi^{(1)} = \frac{at}{\Delta}. \quad (\text{H9})$$

For  $Z = 3$  we get the following explicit formulas:

$$\underline{x}_{k\alpha} = \begin{pmatrix} \delta \bar{v}_2 t_3 e^{-ika} + t_1 t_2 e^{2ika} \\ \delta \bar{v}_1 t_2 e^{ika} + t_1 t_3 e^{-2ika} \\ \delta \bar{v}_1 \delta \bar{v}_2 - t_1^2 \end{pmatrix}, \quad (\text{H10})$$

$$N_{k\alpha} = (t_1^2 - \delta \bar{v}_1 \delta \bar{v}_2)(t_1^2 + t_2^2 + t_3^2 - \delta \bar{v}_1 \delta \bar{v}_2 - \delta \bar{v}_1 \delta \bar{v}_3 - \delta \bar{v}_2 \delta \bar{v}_3), \quad (\text{H11})$$

$$2t_1 t_2 t_3 \cos(3ka) = \delta \bar{v}_1 \delta \bar{v}_2 \delta \bar{v}_3 - \delta \bar{v}_1 t_2^2 - \delta \bar{v}_2 t_3^2 - \delta \bar{v}_3 t_1^2, \quad (\text{H12})$$

where we defined

$$\delta \bar{v}_j = \delta v_j - \epsilon_{k\alpha}. \quad (\text{H13})$$

We do not indicate the dependence of  $\delta \bar{v}_j \equiv \delta \bar{v}_{jk\alpha}$  on  $k$  and  $\alpha$  since it is clear which band index has to be taken for the Abelian case.

The two gap parameters  $\Delta^{(v)} e^{i\gamma^{(v)}}$ , with  $v = 1, 2$ , follow via (5.103) from the Fourier components  $\delta \bar{v}_j$  and  $\delta \tilde{t}_j$ . The latter can be determined from (5.100) and (5.101) as

$$\begin{aligned} \delta \bar{v}_1 &= \delta \tilde{t}_2^* \\ &= -\frac{1}{2}(\delta v_1 + \delta v_2) - \frac{i}{2\sqrt{3}}(\delta v_1 - \delta v_2), \end{aligned} \quad (\text{H14})$$

$$\begin{aligned} \delta \tilde{t}_1 &= \delta \tilde{t}_2^* \\ &= -\frac{1}{2}(\delta t_1 + \delta t_2) - \frac{i}{2\sqrt{3}}(\delta t_1 - \delta t_2). \end{aligned} \quad (\text{H15})$$

Inserting this result in (5.103) we find after some algebra

$$\begin{aligned} \Delta^{(1)} e^{i\gamma^{(1)}} &= \delta t_1 - \frac{1}{2}(\delta v_1 + \delta v_2) \\ &\quad - \frac{i}{2\sqrt{3}}(2\delta t_1 + 4\delta t_2 + \delta v_1 - \delta v_2), \end{aligned} \quad (\text{H16})$$

$$\begin{aligned} \Delta^{(2)} e^{i\gamma^{(2)}} &= -\delta t_1 - \frac{1}{2}(\delta v_1 + \delta v_2) \\ &\quad - \frac{i}{2\sqrt{3}}(2\delta t_1 + 4\delta t_2 - \delta v_1 + \delta v_2). \end{aligned} \quad (\text{H17})$$

The correlation lengths for  $\nu = 1, 2$  follow from  $\xi^{(\nu)} = \frac{v_F^{(\nu)}}{2\Delta^{(\nu)}}$ . With  $v_F^{(\nu)} = 2ta \sin(k_F^{(\nu)} a)$  and  $k_F^{(\nu)} = \frac{\nu\pi}{3a}$  we obtain

$$\xi^{(1/2)} = \frac{\sqrt{3}}{2} \frac{at}{\Delta^{(1/2)}}. \quad (\text{H18})$$

We note that we have included a sign factor  $(-1)^{\alpha+1}$  to  $|u_{k\alpha}\rangle$ , such that, according to Ref. [15], our convention that  $u_{k\alpha}(Za) > 0$  is positive [see Eq. (5.6)] is fulfilled.

The Zak-Berry connection can be calculated most elegantly from

$$A_{k\alpha} = \frac{1}{N_{k\alpha}} \text{Re}(x_{k\alpha})^\dagger i\partial_k x_{k\alpha}. \quad (\text{H19})$$

Inserting the forms for  $x_{k\alpha}$  from (H4) and (H10), one finds for  $Z = 2$

$$A_{k\alpha} = \frac{a(t_2^2 - t_1^2)}{N_{k\alpha}}, \quad (\text{H20})$$

and for  $Z = 3$

$$A_{k\alpha} = \frac{a}{N_{k\alpha}} \{2t_1^2(t_3^2 - t_2^2) - \delta\bar{v}_1^2 t_2^2 + \delta\bar{v}_2^2 t_3^2 + (\delta\bar{v}_1 - \delta\bar{v}_2)t_1 t_2 t_3 \cos(3ka)\}. \quad (\text{H21})$$

For the non-Abelian case we also need the nondiagonal elements of the Zak-Berry connection which can be calculated from

$$\alpha \neq \beta: \quad (A_k)_{\alpha\beta} = \frac{(-1)^{(\alpha+\beta)}}{\sqrt{N_{k\alpha}N_{k\beta}}} \langle x_{k\alpha} | i\partial_k x_{k\beta} \rangle. \quad (\text{H22})$$

For the special case  $Z = 3$  we obtain the expression

$$\begin{aligned} \langle x_{k1} | i\partial_k x_{k2} \rangle &= i(\delta\bar{v}_{11}\delta\bar{v}_{21} - t_1^2)(\epsilon_{k1} + 2\epsilon_{k2}) \frac{d\epsilon_{k2}}{dk} \\ &+ a(\delta\bar{v}_{21}t_3 e^{ika} + t_1 t_2 e^{-2ika}) \\ &\times (\delta\bar{v}_{22}t_3 e^{-ika} - 2t_1 t_2 e^{2ika}) \\ &+ a(\delta\bar{v}_{11}t_2 e^{-ika} + t_1 t_3 e^{2ika}) \\ &\times (-\delta\bar{v}_{12}t_2 e^{ika} + 2t_1 t_3 e^{-2ika}), \end{aligned} \quad (\text{H23})$$

where we used the notation  $\delta\bar{v}_{j\alpha} = v_j - \epsilon_{k\alpha}$ .

For the derivative of the dispersion we use the results obtained in Ref. [15], which give for  $Z = 2$

$$\begin{aligned} \frac{d\epsilon_{k\alpha}}{dk} &= 4at_1 t_2 \frac{\delta\bar{v}_1}{N_{k\alpha}} \sin(2ka) \\ &= -\frac{2at_1 t_2 \sin(2ka)}{\epsilon_{k\alpha}}, \end{aligned} \quad (\text{H24})$$

and for  $Z = 3$

$$\begin{aligned} \frac{d\epsilon_{k\alpha}}{dk} &= 6at_1 t_2 t_3 \frac{\delta\bar{v}_1 \delta\bar{v}_2 - t_1^2}{N_{k\alpha}} \sin(3ka) \\ &= -\frac{6at_1 t_2 t_3 \sin(3ka)}{t_1^2 + t_2^2 + t_3^2 - \delta\bar{v}_1 \delta\bar{v}_2 - \delta\bar{v}_1 \delta\bar{v}_3 - \delta\bar{v}_2 \delta\bar{v}_3}. \end{aligned} \quad (\text{H25})$$

## APPENDIX I: SPLITTING OF NON-ABELIAN WANNIER FUNCTIONS

In this Appendix we show that extending the momentum integrations in the representations (6.49) and (6.50) of the Wannier functions in NA-ML gauge for the two subintervals gives only rise to delta function contributions in the total sum (6.48) which can be disregarded. After inserting (6.51) we replace the propagators beyond the subintervals via (6.56), (6.57), and (6.54) as

$$\text{(I)}: U^{(\nu)}\left(k, -\frac{\pi}{2Za}\right) \rightarrow \begin{cases} U_1^{(\nu)} & \text{for } k > \frac{\pi}{2Za}, \\ 1 & \text{for } k < -\frac{\pi}{2Za}, \end{cases} \quad (\text{I1})$$

$$\text{(II)}: U^{(\nu)}\left(k, -\frac{\pi}{2Za}\right) \rightarrow \begin{cases} U_L^{(\nu)} & \text{for } k > \frac{3\pi}{2Za}, \\ U_1^{(\nu)} & \text{for } k < \frac{\pi}{2Za} \end{cases} \quad (\text{I2})$$

and the Bloch states are replaced with the help of (A5) and (2.8) by the  $k$ -independent values

$$\text{(I)}: u_{k\alpha'}(ma) \rightarrow \begin{cases} f_{\alpha'}(m) & \text{for } k > \frac{\pi}{2Za}, \\ f_{\alpha'}(m)^* & \text{for } k < -\frac{\pi}{2Za}, \end{cases} \quad (\text{I3})$$

$$\text{(II)}: u_{k\alpha'}(ma) \rightarrow \begin{cases} f_{\alpha'}(m)^* e^{-i\frac{2\pi}{Z}m} & \text{for } k > \frac{3\pi}{2Za}, \\ f_{\alpha'}(m) & \text{for } k < \frac{\pi}{2Za} \end{cases} \quad (\text{I4})$$

where we defined

$$f_{\alpha'}(m) = \frac{1}{\sqrt{Z}} \begin{cases} e^{-i\frac{\pi}{Z}\alpha'm} & \text{for } \alpha' \text{ even,} \\ e^{i\frac{\pi}{Z}(\alpha'-1)m} & \text{for } \alpha' \text{ odd.} \end{cases} \quad (\text{I5})$$

Obviously, the two contributions  $k > \frac{\pi}{2Za}$  from the extension of subinterval (I) and  $k < -\frac{\pi}{2Za}$  from the extension of subinterval (II) lead to the same integrand, such that the sum contains  $\int dk e^{ik\hat{m}_\alpha^{(\nu)}a} = 2\pi\delta(\hat{m}_\alpha^{(\nu)}a)$ . These are unphysical delta function contributions which we always disregard. Similar delta function contributions are generated from the high-momentum integrals in Dirac theory which cancel these terms.

For the two remaining contributions we first shift  $k \rightarrow k + \frac{2\pi}{Za}$  for the contribution  $k > \frac{3\pi}{2Za}$  from the extension of subinterval (II) to get the same starting point  $-\frac{2\pi}{Za}$  for the  $k$  integration as the end point for the contribution  $k < -\frac{\pi}{2Za}$  from the extension of subinterval (I). This shift gives an additional  $e^{i\frac{2\pi}{Z}\hat{m}_\alpha^{(\nu)}a}$  factor from the exponential  $e^{ik\hat{m}_\alpha^{(\nu)}a}$  in the integrand. The integrands for the two contributions are then again the same since

$$e^{i\frac{2\pi}{Z}\hat{m}_\alpha^{(\nu)}a} e^{-i\frac{2\pi}{Z}m} = e^{-i\hat{v}_\alpha^{(\nu)}}, \quad (\text{I6})$$

$$(U_L^{(\nu)} V^{(\nu)})_{\alpha'\alpha} = (V^{(\nu)})_{\alpha'\alpha} e^{i\hat{v}_\alpha^{(\nu)}}, \quad (\text{I7})$$

such that only  $(V^{(\nu)})_{\alpha'\alpha}$  remains for the product of (I6) and (I7). Here, we have used the definition  $\hat{m}_\alpha^{(\nu)} = m - \frac{Z}{2\pi}\hat{v}_\alpha^{(\nu)}$  for the derivation of (I6), and the definition (6.38) of the transformation  $V^{(\nu)}$  for the derivation of (I7). As a result, we get again a delta function from the two remaining contributions which we can disregard.

## APPENDIX J: WILSON PROPAGATOR IN DIRAC THEORY

In this Appendix we calculate the Wilson propagator (6.60) in Dirac theory and prove the central property (6.64). We first

transform the propagator via the Abelian ML gauge to

$$U_q^{(v)} = \Lambda_q^{(v)} \tilde{U}_q^{(v)} (\Lambda_{-\infty}^{(v)})^\dagger, \quad (\text{J1})$$

where

$$\tilde{U}_q^{(v)} = P e^{i \int_{-\infty}^q d q' \tilde{A}_{q'}^{(v)}}, \quad (\text{J2})$$

$$\tilde{A}_q^{(v)} = (\Lambda_q^{(v)})^\dagger A_q^{(v)} \Lambda_q^{(v)} + i (\Lambda_q^{(v)})^\dagger \partial_q \Lambda_q^{(v)}, \quad (\text{J3})$$

$$(\tilde{A}_q^{(v)})_{\tau\tau'} = i \langle \tilde{\chi}_{q\tau}^{(v)} | \partial_q \tilde{\chi}_{q\tau'}^{(v)} \rangle, \quad (\text{J4})$$

$$\Lambda_q^{(v)} = \begin{pmatrix} e^{i\phi_{q\tau}^{(v)}} & 0 \\ 0 & e^{i\phi_{q-}^{(v)}} \end{pmatrix}. \quad (\text{J5})$$

Using  $\tilde{A}_{q\tau}^{(v)} = (\tilde{A}_q^{(v)})_{\tau\tau} = 0$  [which defines the Abelian ML gauge via (3.34)], we need in addition the nondiagonal matrix elements. They follow from the form (5.1) of  $|\tilde{\chi}_{q\tau}^{(v)}\rangle$  and (J4) in dimensionless units as

$$\tilde{A}(\bar{q}) \equiv \frac{1}{2\xi^{(v)}} \tilde{A}_q^{(v)} = \frac{1}{2} \frac{1}{1 + \bar{q}^2} \sigma_y, \quad (\text{J6})$$

where  $\bar{q} = 2\xi^{(v)}q$ . With

$$\tilde{U}(\bar{q}) \equiv \tilde{U}_q^{(v)} = P e^{i \int_{-\infty}^{\bar{q}} d\bar{q}' \tilde{A}(\bar{q}')}, \quad (\text{J7})$$

we find the differential equation

$$\frac{d}{d\bar{q}} \tilde{U}(\bar{q}) = i \tilde{A}(\bar{q}) \tilde{U}(\bar{q}), \quad (\text{J8})$$

which has the solution

$$\tilde{U}(\bar{q}) = e^{i \frac{1}{2} \sigma_y \arctan \bar{q}} e^{i \sigma_y \frac{\pi}{4}}, \quad (\text{J9})$$

$$\nu \text{ even : } U_I^{(v)} = \begin{pmatrix} -e^{-i\gamma^{(v)}} & & & & \\ & \sigma_x & & & \\ & & \sigma_x & & \\ & & & \ddots & \\ \sigma_x & & & & 1 \end{pmatrix}, \quad U_{II}^{(v)} = \begin{pmatrix} \sigma_x & & & & \\ & \sigma_x & & & \\ & & \ddots & & \\ & & & \ddots & \\ & & & & \sigma_x \end{pmatrix}, \quad (\text{K1})$$

$$\nu \text{ odd : } U_I^{(v)} = \begin{pmatrix} \sigma_x & & & & \\ & \sigma_x & & & \\ & & \ddots & & \\ & & & \ddots & \\ & & & & \sigma_x \\ & & & & & 1 \end{pmatrix}, \quad U_{II}^{(v)} = \begin{pmatrix} -e^{-i\gamma^{(v)}} & & & & \\ & \sigma_x & & & \\ & & \sigma_x & & \\ & & & \ddots & \\ & & & & \sigma_x \end{pmatrix}, \quad (\text{K2})$$

where we used

$$\lim_{q \rightarrow \infty} U_{q-}^{(v)} = -e^{-i\gamma^{(v)}}, \quad (\text{K3})$$

$$\lim_{q \rightarrow \infty} U_q^{(v)} = \sigma_x, \quad (\text{K4})$$

which follows from (J1), (J5), (J10), and (J13), together with (5.8) and (5.9). Thereby, only  $\sigma_x$  are  $2 \times 2$  submatrices, whereas  $-e^{-i\gamma^{(v)}}$  and 1 denote  $1 \times 1$  numbers.

From (K1) and (K2) one can see directly that the Wilson loop operator  $U_L^{(v)} = U_{II}^{(v)} U_I^{(v)}$  is a unitary matrix which contains only one nonzero phase factor in each row and column. Therefore, all eigenvectors consist only of phase factors with a normalization factor  $1/\sqrt{\nu}$ . We conclude that the matrix

such that the boundary condition  $\tilde{U}(-\infty) = 1$  is fulfilled. A straightforward analysis leads to the explicit result

$$\tilde{U}(\bar{q}) = \frac{1}{\sqrt{2\bar{\epsilon}_{\bar{q}}}} \begin{pmatrix} \sqrt{\bar{\epsilon}_{\bar{q}} - \bar{q}} & \sqrt{\bar{\epsilon}_{\bar{q}} + \bar{q}} \\ -\sqrt{\bar{\epsilon}_{\bar{q}} + \bar{q}} & \sqrt{\bar{\epsilon}_{\bar{q}} - \bar{q}} \end{pmatrix}, \quad (\text{J10})$$

with  $\bar{\epsilon}_{\bar{q}} = \sqrt{1 + \bar{q}^2}$ .

To prove (6.64), we use (J1) and find

$$\hat{\chi}^{(v)}(\bar{q}) = \tilde{\chi}^{(v)}(\bar{q}) \tilde{U}(\bar{q}) (\Lambda_{-\infty}^{(v)})^\dagger, \quad (\text{J11})$$

where we defined the  $2 \times 2$  matrices  $[\hat{\chi}^{(v)}(\bar{q})]_{p\tau} \equiv \hat{\chi}_{q\tau p}^{(v)}$  and  $[\tilde{\chi}^{(v)}(\bar{q})]_{p\tau} \equiv \tilde{\chi}_{q\tau p}^{(v)}$ . Using (5.1), (J5), (5.8), and (5.8), we find

$$\tilde{\chi}^{(v)}(\bar{q}) = \frac{1}{\sqrt{2\bar{\epsilon}_{\bar{q}}}} \begin{pmatrix} \sqrt{\bar{\epsilon}_{\bar{q}} + \bar{q}} e^{i\gamma^{(v)}} & \sqrt{\bar{\epsilon}_{\bar{q}} - \bar{q}} e^{i\gamma^{(v)}} \\ \sqrt{\bar{\epsilon}_{\bar{q}} - \bar{q}} & -\sqrt{\bar{\epsilon}_{\bar{q}} + \bar{q}} \end{pmatrix}, \quad (\text{J12})$$

$$\Lambda_{-\infty}^{(v)} = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\gamma^{(v)}} \end{pmatrix}. \quad (\text{J13})$$

Multiplying (J12) with the product of (J10) and (J13), we get

$$\hat{\chi}^{(v)}(\bar{q}) = \sigma_x, \quad (\text{J14})$$

which proves (6.64).

### APPENDIX K: SPECTRUM OF WILSON LOOP OPERATOR

The Wilson loop operator defining the unitary transformation  $V^{(v)}$  can be calculated from (6.55) as the product  $U_L^{(v)} = U_{II}^{(v)} U_I^{(v)}$  of the propagators for the two subintervals. The propagators  $U_I^{(v)}$  and  $U_{II}^{(v)}$  follow from the limit  $q \rightarrow \infty$  of (6.61) and (6.62) as

elements  $(V^{(v)})_{\nu\alpha}$  and  $(U_I^{(v)} V^{(v)})_{\nu\alpha}$  occurring in (6.67) and (6.68) are given by  $1/\sqrt{\nu}$  times a phase factor.

Furthermore, by using the special structure of the Wilson loop matrix  $U_L^{(v)}$ , one can easily find the eigenvalues  $\lambda$ . A straightforward analysis gives

$$\det(U_L^{(v)} - \lambda \mathbb{1}) = (-\lambda)^\nu + (-1)^\nu e^{-i\gamma^{(v)}} = 0, \quad (\text{K5})$$

such that for both  $\nu$  even or odd we get

$$\lambda^\nu = -e^{-i\gamma^{(v)}}. \quad (\text{K6})$$

This equation has  $\nu$  solutions for  $\lambda = e^{i\gamma_\alpha^{(v)}}$ , with  $\gamma_\alpha^{(v)}$  given by (6.24).

- [1] G. H. Wannier, The structure of electronic excitation levels in insulating crystals, *Phys. Rev.* **52**, 191 (1937).
- [2] W. Kohn, Analytic properties of bloch waves and wannier functions, *Phys. Rev.* **115**, 809 (1959).
- [3] N. Marzari, A. A. Mostofi, J. R. Yates, I. Souza, and D. Vanderbilt, Maximally localized wannier functions: Theory and applications, *Rev. Mod. Phys.* **84**, 1419 (2012).
- [4] B. Bradlyn, L. Elcoro, J. Cano, M. G. Vergniory, Z. Wang, C. Felser, M. I. Aroyo, and B. A. Bernevig, Topological quantum chemistry, *Nature (London)* **547**, 298 (2017).
- [5] H. C. Po, A. Vishwanath, and H. Watanabe, Symmetry-based indicators of band topology in the 230 space groups, *Nat. Commun.* **8**, 50 (2017).
- [6] H. C. Po, H. Watanabe, and A. Vishwanath, Fragile Topology and Wannier Obstructions, *Phys. Rev. Lett.* **121**, 126402 (2018).
- [7] R. Resta, Macroscopic polarization in crystalline dielectrics: the geometric phase approach, *Rev. Mod. Phys.* **66**, 899 (1994).
- [8] C. Sgierovello, M. Peressi, and R. Resta, Electron localization in the insulating state: Application to crystalline semiconductors, *Phys. Rev. B* **64**, 115202 (2001).
- [9] D. Vanderbilt, *Berry Phases in Electronic Structure Theory: Electric Polarization, Orbital Magnetization and Topological Insulators* (Cambridge University Press, Cambridge, 2018).
- [10] N. Marzari and D. Vanderbilt, Maximally localized generalized wannier functions for composite energy bands, *Phys. Rev. B* **56**, 12847 (1997).
- [11] I. Souza, T. Wilkens, and R. M. Martin, Polarization and localization in insulators: Generating function approach, *Phys. Rev. B* **62**, 1666 (2000).
- [12] J.-H. Park, G. Yang, J. Klinovaja, P. Stano, and D. Loss, Fractional boundary charges in quantum dot arrays with density modulation, *Phys. Rev. B* **94**, 075416 (2016).
- [13] M. Thakurathi, J. Klinovaja, and D. Loss, From fractional boundary charges to quantized hall conductance, *Phys. Rev. B* **98**, 245404 (2018).
- [14] M. Pletyukhov, D. M. Kennes, J. Klinovaja, D. Loss, and H. Schoeller, Topological invariants to characterize universality of boundary charge in one-dimensional insulators beyond symmetry constraints, *Phys. Rev. B* **101**, 161106(R) (2020).
- [15] M. Pletyukhov, D. M. Kennes, J. Klinovaja, D. Loss, and H. Schoeller, Surface charge theorem and topological constraints for edge states: Analytical study of one-dimensional nearest-neighbor tight-binding models, *Phys. Rev. B* **101**, 165304 (2020).
- [16] Y.-T. Lin, D. M. Kennes, M. Pletyukhov, C. S. Weber, H. Schoeller, and V. Meden, Interacting rice-mele model: Bulk and boundaries, *Phys. Rev. B* **102**, 085122 (2020).
- [17] M. Pletyukhov, D. M. Kennes, K. Piasotski, J. Klinovaja, D. Loss, and H. Schoeller, Rational boundary charge in one-dimensional systems with interaction and disorder, *Phys. Rev. Res.* **2**, 033345 (2020).
- [18] C. S. Weber, K. Piasotski, M. Pletyukhov, J. Klinovaja, D. Loss, H. Schoeller, and D. M. Kennes, Universality of Boundary Charge Fluctuations, *Phys. Rev. Lett.* **126**, 016803 (2021).
- [19] D. Vanderbilt and R. D. King-Smith, Electric polarization as a bulk quantity and its relation to surface charge, *Phys. Rev. B* **48**, 4442 (1993).
- [20] P. G. Harper, Single band motion of conduction electrons in a uniform magnetic field, *Proc. Phys. Soc., London, Sect. A* **68**, 874 (1955).
- [21] S. Aubry and G. André, Analyticity breaking and anderson localization in incommensurate lattices, *Ann. Isr. Phys. Soc.* **3**, 133 (1980).
- [22] S. Ganesan, K. Sun, and S. Das Sarma, Topological Zero-Energy Modes in Gapless Commensurate Aubry-André-Harper Models, *Phys. Rev. Lett.* **110**, 180403 (2013).
- [23] Y. Lahini, R. Pugatch, F. Pozzi, M. Sorel, R. Morandotti, N. Davidson, and Y. Silberberg, Observation of a Localization Transition in Quasiperiodic Photonic Lattices, *Phys. Rev. Lett.* **103**, 013901 (2009).
- [24] Y. E. Kraus, Y. Lahini, Z. Ringel, M. Verbin, and O. Zeitler, Topological States and Adiabatic Pumping in Quasicrystals, *Phys. Rev. Lett.* **109**, 106402 (2012).
- [25] W. DeGottardi, D. Sen, and S. Vishveshwara, Majorana Fermions in Superconducting 1d Systems having Periodic, Quasiperiodic, and Disordered Potentials, *Phys. Rev. Lett.* **110**, 146404 (2013).
- [26] M. Schreiber, S. S. Hodgman, P. Bordia, H. P. Lüschen, M. H. Fischer, R. Vosk, E. Altman, U. Schneider, and I. Bloch, Observation of many-body localization of interacting fermions in a quasirandom optical lattice, *Science* **349**, 842 (2015).
- [27] S. Gangadharaiah, L. Trifunovic, and D. Loss, Localized End States in Density Modulated Quantum Wires and Rings, *Phys. Rev. Lett.* **108**, 136803 (2012).
- [28] M. Atala, M. Aidelsburger, J. T. Barreiro, D. Abanin, T. Kitagawa, E. Demler, and I. Bloch, Direct measurement of the zak phase in topological bloch bands, *Nat. Phys.* **9**, 795 (2013).
- [29] E. J. Meier, F. A. An, and B. Gadway, Observation of the topological soliton state in the Su-Schrieffer-Heeger model, *Nat. Commun.* **7**, 13986 (2016).
- [30] M. J. Yoo, T. A. Fulton, H. F. Hess, R. L. Willett, L. N. Dunkleberger, R. J. Chichester, L. N. Pfeiffer, and K. W. West, Scanning single-electron transistor microscopy: Imaging individual charges, *Science* **276**, 579 (1997).
- [31] S. H. Tessmer, P. I. Glicofridis, R. C. Ashoori, L. S. Levitov, and M. R. Melloch, Subsurface charge accumulation imaging of a quantum hall liquid, *Nature (London)* **392**, 51 (1998).
- [32] G. Finkelstein, P. I. Glicofridis, R. C. Ashoori, and M. Shayegan, Topographic mapping of the quantum hall liquid using a few-electron bubble, *Science* **289**, 90 (2000).
- [33] G. Ben-Shach, A. Haim, I. Appelbaum, Y. Oreg, A. Yacoby, and B. I. Halperin, Detecting majorana modes in one-dimensional wires by charge sensing, *Phys. Rev. B* **91**, 045403 (2015).
- [34] M. Xiao, G. Ma, Z. Yang, P. Sheng, Z. Q. Zhang, and C. T. Chan, Geometric phase and band inversion in periodic acoustic systems, *Nat. Phys.* **11**, 240 (2015).
- [35] Y. Efroni, S. Ilani, and E. Berg, Topological Transitions and Fractional Charges Induced by Strain and a Magnetic Field in Carbon Nanotubes, *Phys. Rev. Lett.* **119**, 147704 (2017).
- [36] S.-R. E. Yang, Soliton fractional charges in graphene nanoribbon and polyacetylene: Similarities and differences, *Nanomaterials* **9**, 885 (2019).
- [37] J. Klinovaja, P. Stano, and D. Loss, Transition from Fractional to Majorana Fermions in Rashba Nanowires, *Phys. Rev. Lett.* **109**, 236801 (2012).

- [38] D. Rainis, A. Saha, J. Klinovaja, L. Trifunovic, and D. Loss, Transport Signatures of Fractional Fermions in Rashba Nanowires, *Phys. Rev. Lett.* **112**, 196803 (2014).
- [39] M. J. Rice and E. J. Mele, Elementary Excitations of a Linearly Conjugated Diatomic Polymer, *Phys. Rev. Lett.* **49**, 1455 (1982).
- [40] R. Resta and S. Sorella, Electron Localization in the Insulating State, *Phys. Rev. Lett.* **82**, 370 (1999).
- [41] L. He and D. Vanderbilt, Exponential Decay Properties of Wannier Functions and Related Quantities, *Phys. Rev. Lett.* **86**, 5341 (2001).
- [42] A. J. Heeger, Nobel lecture: Semiconducting and metallic polymers: The fourth generation of polymeric materials, *Rev. Mod. Phys.* **73**, 681 (2001).
- [43] W. P. Su, J. R. Schrieffer, and A. J. Heeger, Solitons in Polyacetylene, *Phys. Rev. Lett.* **42**, 1698 (1979).