

**Constructing optimal Wannier functions via potential theory: isolated
single band for matrix models**

Hanwen Zhang
Research Report YALEU/DCS/TR-1573
February 28, 2025

We present a rapidly convergent scheme for computing globally optimal Wannier functions of isolated single bands for matrix models in two dimensions. The scheme proceeds first by constructing provably exponentially localized Wannier functions directly from parallel transport (with simple analytically computable corrections) when topological obstruction is absent. We prove that the corresponding Wannier functions are real when the matrix model possesses time-reversal symmetry. When a band has nonzero Berry curvature, the resulting Wannier function is not optimal, but it is transformed into the global optimum by a single gauge transformation that eliminates the divergence of the Berry connection. Complete analysis of the construction is presented, paving the way for further improvements and generalization. The performance of the scheme is illustrated with several numerical examples.

**Constructing optimal Wannier functions via potential
theory: isolated single band for matrix models**

Hanwen Zhang [†]
February 28, 2025

[†] Dept. of Applied and Computational Mathematics, Yale University, New Haven, CT 06511

Contents

1	Introduction	3
2	Problem statement	5
3	Mathematical and physical preliminaries	6
3.1	Notations	6
3.2	Periodic lattices in $\mathbb{R}^2$	6
3.3	Tight-binding models	7
3.4	Time-reversal symmetry	9
3.5	Wannier functions	10
3.6	Systems of coordinates	12
3.7	Poisson's equation on tori	13
3.8	Vector fields on tori	13
3.9	Analyticity of eigenvalues and eigenvectors	15
3.10	Ordinary differential equations	18
3.11	Pfaffian systems	19
4	Numerical preliminaries	19
4.1	Trapezoidal rule for periodic functions	20
4.2	Discrete Fourier transform in two dimensions	20
4.3	Fourth-order Runge-Kutta method	21
5	Analytic apparatus	22
5.1	Parallel transport equation	23
5.2	Berry connection, curvature and integrability	24
5.3	Gauge transformation	26
5.4	First Chern number	27
5.5	Gauge choice and Wannier localization	29
5.5.1	Variance formulas for Wannier functions	29
5.5.2	Optimal Berry connections	30
6	Construction of optimal Wannier functions	33
6.1	Stage 1: constructing an assignment on a line	35
6.2	Stage 2: constructing an assignment on the torus	37
6.3	Stage 3: eliminating the divergence of the Berry connection	40
6.4	Reality of Wannier functions	42

7	Numerical procedure	44
7.1	Discretizing the torus D^*	45
7.2	Step 1: computing an assignment on the line γ_0	46
7.3	Step 2: computing parallel transport on lines γ_{κ_1}	47
7.4	Step 3: determining the topological obstruction	48
7.5	Step 4: computing the Berry connection	49
7.6	Step 5: eliminating the divergence of the Berry connection	51
7.7	Step 6: computing the Wannier function	51
7.8	Detailed description of the algorithms	52
8	Numerical results	55
8.1	Example 1: A 3×3 matrix model	56
8.2	Example 2: Haldane model (topologically trivial case)	66
8.3	Example 3: Haldane model (topologically non-trivial case)	73
9	Conclusion and generalization	76
10	Appendix	80
10.1	Alternative approach: direct computation of optimal gauge	80
10.1.1	Stage 1: computing the Berry curvature	80
10.1.2	Stage 2: computing the divergence-free component	80
10.1.3	Stage 3: compute the harmonic component	82
10.1.4	Reality of Wannier functions	85
10.2	Derivation of formulas in Lemma 5.5	87

1 Introduction

A standard task in solid state physics and quantum chemistry is the computation of localized molecular orbitals known as Wannier functions [17]. This paper is a follow up to [12], extending its approach from isolated single bands of Schrödinger operators in one dimension to matrix models in two dimensions. We present a rapidly convergent scheme for computing globally optimal Wannier functions based on solving the parallel transport equation and the Helmholtz-Hodge decomposition of Berry connections (gauge fields). First, we construct provably exponentially localized Wannier functions by *solely* solving the parallel transport equation (with simple corrections) to assign eigenvectors smoothly when no topological obstruction is present. Although the resulting Wannier functions are not optimally localized when the Berry curvature is nonzero, it is *automatically* real (given a simple choice of the initial condition)

when the model has time-reversal symmetry. Next, globally optimal Wannier functions are obtained via a single gauge transformation that eliminates the divergence of Berry connections. Such gauge transformations are obtained by solving the Poisson’s equation on tori, which can be done in an efficient and accurate manner. (An alternative approach that utilizes the gauge invariance of the Berry curvature can be found in Remark 6.1 and Appendix 10.1.) In addition, when topological obstruction is encountered, the resulting assignments are still analytic (but it is discontinuous when viewed as a periodic function); the assignments are amenable to efficient interpolation schemes based on, for example, Chebyshev nodes, instead of equispaced ones in the Fourier basis in this paper.

The construction of Wannier functions can be viewed as a problem of assigning eigenvectors of some analytic family of matrices/operators to be as smooth as possible, so that coefficients of the Fourier expansion of the eigenvectors decay exponentially. For computational purposes, it is a common practice to view this problem as a nonlinear and nonconvex optimization problem. As a result, the focus of most work has been obtaining robust initial guesses. For example, high quality initial guesses are obtained by applying the interpolative decomposition to the so-called density matrices [10, 9]. Parallel-transport-based approaches have been applied for assigning eigenvectors [6], but it results in continuous assignments, corresponding to slowly decaying Wannier functions. Then, the initial guesses are optimized by minimizing the spread of the Fourier coefficients (Marzari–Vanderbilt functional) via gradient descent [18, 17], where the gradients are often computed by finite differences. From this optimization point of view, the first step of the scheme in this paper can be viewed as obtaining nearly optimal initial assignment of eigenvectors (corresponding to exponentially localized Wannier functions) by only carrying out parallel transport. Moreover, the minimization procedures can be drastically accelerated by making use of the potential theory of the physical quantities involved; for the single band case in this paper, it is an unconstrained quadratic problem whose global optimum is achieved in a single Newton step, where the Hessian inversion is essentially free.

It should be observed that, although we solve the parallel transport equation explicitly in this paper to obtain the band structure and Wannier functions simultaneously, the computation of the two can be easily decoupled by using the scheme for parallel transport in Remark 7.2 at the cost of lower accuracy. Such an approach produces the results of lower accuracy with almost no additional computational cost once the band structure is obtained. We refer the reader to Section 8.1 for details.

The extension of the schemes in [12] and this paper to isolated multi-bands has been worked out and are in preparation for publication. We refer the readers to Section 9 for the generalization to Schrödinger operators and higher dimensions.

The structure of this paper is as follows. Section 2 contains the description of the problem to be solved. Section 2 contains the physical and mathematical preliminaries, followed by Section 4 consisting of the numerical tools used in this paper. In Section 5, we introduce the

analytic apparatus for constructing Wannier functions. Section 5 contains the procedures for the construction of optimal Wannier functions, followed by the corresponding numerical procedures in Section 7. Section 8 contains several examples that illustrate the performance of the schemes in this paper.

2 Problem statement

Let D^* be a two-dimensional (deformed) torus

$$D^* = \{\kappa_1 \mathbf{b}_1 + \kappa_2 \mathbf{b}_2 : \kappa_1, \kappa_2 \in [-1/2, 1/2]\}$$

parameterized by $(\kappa_1, \kappa_2) \in [-\frac{1}{2}, \frac{1}{2}] \times [-\frac{1}{2}, \frac{1}{2}]$, where $\mathbf{b}_1$ and $\mathbf{b}_2$ are basis vectors in $\mathbb{R}^2$ that span a Bravais lattice. Suppose that $H : D^* \rightarrow \mathbb{C}^{n \times n}$ is a family of Hermitian matrices that satisfies the following periodicity condition

$$H(\mathbf{k} + m_1 \mathbf{b}_1 + m_2 \mathbf{b}_2) = H(\mathbf{k}), \quad \mathbf{k} \in D^*, m_1, m_2 \in \mathbb{Z}. \quad (1)$$

Suppose further that the elements in H are analytic functions in D^* . We consider an eigenvalue $E(\mathbf{k})$ and its corresponding normalized eigenvector $\mathbf{u}(\mathbf{k})$ for some $\mathbf{k} \in D^*$ defined by the formula

$$H(\mathbf{k})\mathbf{u}(\mathbf{k}) = E(\mathbf{k})\mathbf{u}(\mathbf{k}), \quad (2)$$

and we assume that the eigenvalue $E(\mathbf{k})$ never becomes degenerate for any $\mathbf{k} \in D^*$. Since we may multiply the vector $\mathbf{u}(\mathbf{k})$ by a $\mathbf{k}$ dependent phase factor $e^{-i\varphi(\mathbf{k})}$ (with a real $\varphi(\mathbf{k})$) without affecting the definition in (2), the eigenvector $\mathbf{u}(\mathbf{k})$ is only unique up to a phase factor. The phase factors are often referred to as the “choice of gauge” in physics literature. In this paper, we will use the terms phase factors and choice of gauge (or simply gauge) interchangeably.

In this paper, given $H : D^* \rightarrow \mathbb{C}^{n \times n}$ in (1) above, we wish to construct an assignment of $\mathbf{u}$ in D^* such that $\mathbf{u}$ is analytic in D^* and satisfies the same periodicity in (1) (so that the Fourier coefficients of elements in $\mathbf{u}$ decay exponentially asymptotically). Furthermore, we also require the assignment of $\mathbf{u}$ to be optimal over all possible choice of φ in the sense that the Fourier coefficients of $\mathbf{u}$ have a minimum spread defined by some variance function. Such a problem can be viewed as a matrix version of the construction of Wannier functions for the Schrödinger operator case. (The physical meaning of the quantities considered above will be further clarified in Section 3.3 and the variance function is defined in (38) in Section 3.5.)

3 Mathematical and physical preliminaries

3.1 Notations

We let δ_{ij} denote the Kronecker delta function, defined by

$$\delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}. \quad (3)$$

Suppose $\mathbf{v}$ is a vector in $\mathbb{C}^N$. We denote its norm by $\|\mathbf{v}\|$, which is given by the standard inner product as $\|\mathbf{v}\|^2 = \mathbf{v}^* \mathbf{v}$. For vectors $\mathbf{a}, \mathbf{b}$ in $\mathbb{R}^N$, we also use the dot product notation $\mathbf{a} \cdot \mathbf{b}$ to represent their standard inner product $\mathbf{a}^* \mathbf{b}$.

Unless otherwise specified, we choose the principal branch of the complex log function, i.e. $\log(z) = \log(|z|) + i \arg(z)$, where $-\pi < \arg(z) \leq \pi$.

3.2 Periodic lattices in $\mathbb{R}^2$

In this section, we introduce standard definitions for a periodic lattice in $\mathbb{R}^2$.

Suppose the canonical basis vectors in $\mathbb{R}^2$ are given by

$$\mathbf{e}_x = (1, 0), \quad \mathbf{e}_y = (0, 1). \quad (4)$$

We define a periodic lattice in two dimension by two linearly independent vectors $\mathbf{a}_1$ and $\mathbf{a}_2$ in $\mathbb{R}^2$:

$$\Lambda = \{\mathbf{R} = m_1 \mathbf{a}_1 + m_2 \mathbf{a}_2 : m_1, m_2 \in \mathbb{Z}\}. \quad (5)$$

The vectors $\mathbf{a}_1$ and $\mathbf{a}_2$ are referred to as the real space primitive lattice vectors. Given such vectors, we also denote the reciprocal space primitive lattice vectors by vectors $\mathbf{b}_1$ and $\mathbf{b}_2$ in $\mathbb{R}^2$, such that their inner product satisfies

$$\mathbf{a}_i \cdot \mathbf{b}_j = 2\pi \delta_{ij}, \quad i, j = 1, 2. \quad (6)$$

Similarly, the reciprocal lattice is defined as

$$\Lambda^* = \{\mathbf{G} = m_1 \mathbf{b}_1 + m_2 \mathbf{b}_2 : m_1, m_2 \in \mathbb{Z}\}. \quad (7)$$

Due to the relation in (7), for any vector $\mathbf{R} \in \Lambda$ and $\mathbf{G} \in \Lambda^*$, we have

$$\mathbf{G} \cdot \mathbf{R} = 2\pi l \quad (8)$$

for some integer l .

We define the primitive unit cell in the real space by the formula

$$D = \{r_1 \mathbf{a}_1 + r_2 \mathbf{a}_2 : r_1, r_2 \in [-1/2, 1/2]\}, \quad (9)$$

and we denote its area $|D|$ by V_{puc} . We also define the primitive unit cell in the reciprocal space by the formula

$$D^* = \{\kappa_1 \mathbf{b}_1 + \kappa_2 \mathbf{b}_2 : \kappa_1, \kappa_2 \in [-1/2, 1/2]\}. \quad (10)$$

For any $\mathbf{k}$ in D^* , we denotes its x, y component by k_x, k_y given by the formula

$$\mathbf{k} = k_x \mathbf{e}_x + k_y \mathbf{e}_y. \quad (11)$$

We define a torus T

$$T = \left[-\frac{1}{2}, \frac{1}{2}\right] \times \left[-\frac{1}{2}, \frac{1}{2}\right], \quad (12)$$

which provides a natural domain for parameterizing $\mathbf{k} \in D^*$ by the formula

$$\mathbf{k} = \mathbf{k}(\kappa_1, \kappa_2) = \kappa_1 \mathbf{b}_1 + \kappa_2 \mathbf{b}_2, \quad (\kappa_1, \kappa_2) \in T. \quad (13)$$

The components of $\mathbf{k}$ are given by

$$k_x = k_x(\kappa_1, \kappa_2), \quad k_y = k_y(\kappa_1, \kappa_2), \quad (\kappa_1, \kappa_2) \in T. \quad (14)$$

Remark 3.1. *The region D^* defined in (10) is not necessarily the Wigner-Seitz cell of the lattice Λ^* that defines the first Brillouin zone (BZ) in physics literature. Due to periodicity, the domain D^* serves the same purpose as the BZ, except for not capturing the full symmetry of the lattice.*

3.3 Tight-binding models

In this section, we introduce the so-called tight-binding model to provide the physical context for the eigenvalue problem in (2). They are standard theory in solid state physics and can be found, for example, in [16].

Given a Schrödinger operator with a periodic potential on the lattice Λ defined in (5), a common strategy in band structure modeling is to approximate the operator by a set of n Bloch-like functions $\chi_{\mathbf{k},i} : \mathbb{R}^2 \rightarrow \mathbb{C}$ defined by the formula

$$\chi_{\mathbf{k},i}(\mathbf{r}) = \sum_{\mathbf{R} \in \Lambda} e^{-i\mathbf{k} \cdot \mathbf{R}} \phi_i(\mathbf{r} + \mathbf{R}), \quad \mathbf{k} \in D^*, \mathbf{r} \in \mathbb{R}^2, \quad i = 1, 2, \dots, n, \quad (15)$$

where $\phi_i : \mathbb{R}^2 \rightarrow \mathbb{R}$ represents some pre-specified atomic orbitals (localized near $\mathbf{R} = \mathbf{0}$). The eigenfunctions of the j -th eigenvalue of the Schrödinger operator can thus be approximated by some linear combination of $\chi_{\mathbf{k},i}$ in (15) given by the formula

$$\psi_{\mathbf{k}}^{(j)}(\mathbf{r}) = \sum_{i=1}^n u_i^{(j)}(\mathbf{k}) \chi_{\mathbf{k},i}(\mathbf{r}), \quad (16)$$

where $u_i^{(j)}(\mathbf{k})$ for $i = 1, 2, \dots, n$ are the coefficients to be found.

By approximating the Schrödinger operator in the subspace spanned by $\chi_{\mathbf{k},i}$, the original eigenvalue problem of the operator is reduced into that of a family of n by n matrices $H(\mathbf{k})$ for $\mathbf{k} \in D^*$. The eigenvectors of the matrices give the coefficients $u_i^{(j)}(\mathbf{k})$ in (15), thus obtaining the approximating eigenfunctions $\psi_{\mathbf{k}}^{(j)}$. The family of matrices $H(\mathbf{k})$ is referred to as the tight-binding Hamiltonian. In general, various assumptions are made about the properties of the inner product among functions ϕ_i in (15); parameters in $H(\mathbf{k})$ are often determined by least-square procedures against experimental data or first-principle calculations. We refer the reader to [16, 21] for details.

We form a family of vector $\mathbf{u}_n : D^* \rightarrow \mathbb{C}^n$ containing all coefficients in (16) in the form

$$\mathbf{u}_j(\mathbf{k}) = (u_1^{(j)}(\mathbf{k}), u_2^{(j)}(\mathbf{k}), \dots, u_n^{(j)}(\mathbf{k})), \quad (17)$$

so that the eigenvalue problem for the tight-binding Hamiltonian becomes

$$H(\mathbf{k})\mathbf{u}_j(\mathbf{k}) = E_j(\mathbf{k})\mathbf{u}_j(\mathbf{k}), \quad (18)$$

with the normalization condition

$$\mathbf{u}_j^*(\mathbf{k})\mathbf{u}_j(\mathbf{k}) = 1. \quad (19)$$

The eigenvalue index j is also referred to as the band index, and the family of eigenvalues $E_j(\mathbf{k})$ for $\mathbf{k} \in D^*$ is the energy levels of the j -th band. Combining (18) with (16), the family of vectors $\mathbf{u}_j(\mathbf{k})$ for $\mathbf{k} \in D^*$ defines the Bloch-like functions for the j -th band. The eigenvalue problem defined in (2) can be viewed as the one in (18) for a particular band with the band index j dropped.

We assume that the elements in the matrix H are analytic in D^* and H are periodic in the following form

$$H(\mathbf{k} + \mathbf{G}) = H(\mathbf{k}), \quad \mathbf{k} \in D^*, \mathbf{G} \in \Lambda^*, \quad (20)$$

which is identical to the condition in (1). We refer to periodic functions in the form of (20) as being Λ^* -periodic. Thus the eigenvalue $E_j(\mathbf{k})$ is also Λ^* -periodic:

$$E_j(\mathbf{k} + \mathbf{G}) = E_j(\mathbf{k}), \quad \mathbf{k} \in D^*, \mathbf{G} \in \Lambda^*. \quad (21)$$

Throughout this paper, we assume the family of eigenvalues $E_j(\mathbf{k})$ of interests never becomes degenerate. The eigenvector $\mathbf{u}_j(\mathbf{k})$ is chosen to be periodic copies of that in D^* so that we have

$$\mathbf{u}_j(\mathbf{k} + \mathbf{G}) = \mathbf{u}_j(\mathbf{k}), \quad \mathbf{k} \in D^*, \mathbf{G} \in \Lambda^*. \quad (22)$$

The projector $P_j(\mathbf{k}) = \mathbf{u}_j(\mathbf{k})\mathbf{u}_j^*(\mathbf{k})$ has the same periodicity

$$P_j(\mathbf{k} + \mathbf{G}) = P_j(\mathbf{k}), \quad \mathbf{k} \in D^*, \mathbf{G} \in \Lambda^*. \quad (23)$$

We observe that, although the eigenvector $\mathbf{u}_j$ is only determined up to a phase factor, the projector P_j is independent of the such choices, thus is uniquely defined in D^* .

Remark 3.2. *In this paper, most functions of interests both analytic and Λ^* -periodic in D^* , such as H in (20). However, there are cases where it is still convenient to view D^* not as a deformed torus but as a subset of $\mathbb{R}^2$, so that a function can be analytic in D^* without being Λ^* -periodic. For this reason, we always make it explicit when a function is both analytic and Λ^* -periodic to indicate that D^* is viewed as a torus, not as a subset of $\mathbb{R}^2$.*

Due to the periodicity in (22), each element $u_i^{(j)}$ in $\mathbf{u}_n$ can be represented by its Fourier series of the form

$$u_j^{(n)}(\mathbf{k}) = \sum_{\mathbf{R} \in \Lambda} u_{j,\mathbf{R}}^{(n)} \cdot e^{i\mathbf{R} \cdot \mathbf{k}}. \quad (24)$$

The orthogonality for two different lattice vectors $\mathbf{R}, \mathbf{R}' \in \Lambda$

$$\int_{D^*} d\mathbf{k} e^{i(\mathbf{R}-\mathbf{R}') \cdot \mathbf{k}} = \frac{(2\pi)^2}{V_{\text{puc}}} \delta_{\mathbf{R},\mathbf{R}'} \quad (25)$$

allows us to compute the Fourier coefficients by the formula

$$u_{j,\mathbf{R}}^{(n)} = \frac{V_{\text{puc}}}{(2\pi)^2} \int_{D^*} d\mathbf{k} e^{-i\mathbf{R} \cdot \mathbf{k}} \cdot u_j^{(n)}(\mathbf{k}), \quad (26)$$

where V_{puc} is the volume of the primitive unit cell D in (9).

3.4 Time-reversal symmetry

We refer to H in (18) having time-reversal symmetry if it satisfies

$$\overline{H}(\mathbf{k}) = H(-\mathbf{k}), \quad \mathbf{k} \in D^*. \quad (27)$$

Assuming H has time-reversal symmetry, it is straightforward to show that the eigenvalue E_j and projector P_j have the same symmetry

$$\overline{P}_j(\mathbf{k}) = P_j(-\mathbf{k}), \quad \mathbf{k} \in D^*, \quad (28)$$

$$E_j(\mathbf{k}) = E_j(-\mathbf{k}), \quad \mathbf{k} \in D^*. \quad (29)$$

3.5 Wannier functions

In this section, we define Wannier functions for the matrix models and their variance that measures the spread. It should be observed that the variance definition for the matrix case in this paper is different from the standard one in the Schrödinger operator [18]; the purpose of the new definition is to make the matrix case almost identical to the operator one in terms of minimizing the variance.

Analogous to the Wannier functions in the Schrödinger operator case [18, 26], we define the Wannier function for the j -th band centered at some $\mathbf{R} \in \Lambda$ by the formula

$$W_{\mathbf{R}}^{(j)}(\mathbf{r}) = \frac{(2\pi)^2}{V_{\text{puc}}} \int_{D^*} d\mathbf{k} e^{-i\mathbf{k} \cdot \mathbf{R}} \psi_{\mathbf{k}}^{(j)}(\mathbf{r}), \quad (30)$$

where $\psi_{\mathbf{k}}^{(j)}(\mathbf{r})$ is given in (16) and the integration domain is D^* in (10).

Due to the Bloch-like nature of (15), the Wannier functions $W_{\mathbf{R}}^{(j)}(\mathbf{r})$ centered at $\mathbf{R}$ are copies of those at $\mathbf{R} = 0$. More explicitly, we have

$$W_{\mathbf{R}}^{(j)}(\mathbf{r}) = W_{\mathbf{0}}^{(j)}(\mathbf{r} - \mathbf{R}). \quad (31)$$

Hence, we only need to consider those centered around $\mathbf{R} = \mathbf{0}$:

$$W_{\mathbf{0}}^{(j)}(\mathbf{r}) = \frac{V_{\text{puc}}}{(2\pi)^2} \int_{D^*} d\mathbf{k} \psi_{\mathbf{k}}^{(j)}(\mathbf{r}). \quad (32)$$

By the definition of $\psi_{\mathbf{k}}^{(j)}$ in (15), (16) and the Fourier series formulas (24), (26), we rewrite $W_{\mathbf{0}}^{(j)}$ in the following form

$$W_{\mathbf{0}}^{(j)}(\mathbf{r}) = \sum_{i=1}^n \sum_{\mathbf{R} \in \Lambda} u_{i,\mathbf{R}}^{(j)} \cdot \phi_i(\mathbf{r} + \mathbf{R}), \quad (33)$$

where $u_{i,\mathbf{R}}^{(j)}$ is the Fourier coefficients of $u_i^{(j)}(\mathbf{k})$ given by the formula

$$u_{i,\mathbf{R}}^{(j)}(\mathbf{k}) = \sum_{\mathbf{R} \in \Lambda} u_{i,\mathbf{R}}^{(j)} \cdot e^{i\mathbf{R} \cdot \mathbf{k}}. \quad (34)$$

Due to (33), the Wannier function defined in (32) is well-localized near $\mathbf{R} = \mathbf{0}$, provided $|u_{i,\mathbf{R}}^{(j)}|$ decays quickly as $\|\mathbf{R}\|$ gets large. This can be achieved if the eigenvector $\mathbf{u}$ is chosen as analytic and Λ^* -periodic functions in D^* , so that $|u_{i,\mathbf{R}}|$ decays exponentially for large $\|\mathbf{R}\|$.

For the rest of this paper, we drop the band index label j since we only deal with a single band. Hence the eigenvalue equation for the eigenvalue E and eigenvector $\mathbf{u}$ of interests for $\mathbf{k} \in D^*$ becomes

$$H(\mathbf{k})\mathbf{u}(\mathbf{k}) = E(\mathbf{k})\mathbf{u}(\mathbf{k}). \quad (35)$$

Thus the Wannier function is given by the formula

$$W_0(\mathbf{r}) = \frac{V_{\text{puc}}}{(2\pi)^2} \int_{D^*} d\mathbf{k} \psi_{\mathbf{k}}(\mathbf{r}) = \sum_{i=1}^n \sum_{\mathbf{R} \in \Lambda} u_{i,\mathbf{R}} \cdot \phi_i(\mathbf{r} + \mathbf{R}), \quad (36)$$

and the Fourier series of the component u_i in the eigenvector $\mathbf{u}$ is given by

$$u_i(\mathbf{k}) = \sum_{\mathbf{R} \in \Lambda} u_{i,\mathbf{R}} \cdot e^{i\mathbf{R} \cdot \mathbf{k}}. \quad (37)$$

We quantify the localization of $W_0(\mathbf{r})$ by the following variance function

$$\langle \|\mathbf{R}\|^2 \rangle - \|\langle \mathbf{R} \rangle\|^2. \quad (38)$$

where the first and second moment functions are defined by the formulas

$$\langle \mathbf{R} \rangle := \sum_{\mathbf{R} \in \Lambda} |u_{i,\mathbf{R}}|^2 (-\mathbf{R}), \quad \langle \|\mathbf{R}\|^2 \rangle := \sum_{\mathbf{R} \in \Lambda} |u_{i,\mathbf{R}}|^2 \|\mathbf{R}\|^2. \quad (39)$$

The first moment $\langle \mathbf{R} \rangle$ is commonly referred to the Wannier center, where the minus sign is introduced for reasons as follows.

We observe that the moment functions in (39) are different from the conventional definitions

$$\langle \mathbf{r} \rangle = \int_{\mathbb{R}^2} d\mathbf{r} |W_0(\mathbf{r})|^2 \mathbf{r}, \quad \langle r^2 \rangle = \int_{\mathbb{R}^2} d\mathbf{r} |W_0(\mathbf{r})|^2 r^2. \quad (40)$$

The justification for choosing (39) over (40) is as follows. First, since the tight-binding model is already a physical approximation of the original problem, the choice of (39) does not affect much the physical relevance of the final results. Second, and more importantly, since this paper is a stepping stone for the Schrödinger operator case, it is desirable for all definitions to mimic the Wannier problem for the operator case mathematically.

Suppose that $\mathbf{u}$ in (17) is chosen to be analytic and Λ^* -periodic (see (22)) in D^* . Then the Fourier coefficients in (39) decay exponentially and the moment functions (39) are well-defined. Furthermore, by differentiating (34) with respect to $\mathbf{k}$, we obtain the following expressions of (39) in terms of the eigenvector $\mathbf{u}$

$$\langle \mathbf{R} \rangle = i \frac{V_{\text{puc}}}{(2\pi)^2} \int_{D^*} d\mathbf{k} \mathbf{u}^*(\mathbf{k}) \nabla_{\mathbf{k}} \mathbf{u}(\mathbf{k}), \quad (41)$$

$$\langle \|\mathbf{R}\|^2 \rangle = \frac{V_{\text{puc}}}{(2\pi)^2} \int_{D^*} d\mathbf{k} \|\nabla_{\mathbf{k}} \mathbf{u}(\mathbf{k})\|^2. \quad (42)$$

We observe that (41) and (42) resemble the formulas of (40) in terms of Bloch functions for the operator case [18] with only the eigenfunction $u_{\mathbf{k}}$ replaced by the eigenvector $\mathbf{u}(\mathbf{k})$.

3.6 Systems of coordinates

This section contains basic formulas for changing systems of coordinates between variables k_x, k_y and κ_1, κ_2 for parameterizing D^* in (10) defined in (13).

In this paper, we mainly deal with Λ^* -periodic functions in D^* of the form

$$f(\mathbf{k}) = f(\mathbf{k} + \mathbf{G}), \quad \mathbf{k} \in D^*, \mathbf{G} \in \Lambda^*. \quad (43)$$

By the parameterization in (13), the Λ^* -periodicity of f is equivalent to

$$f(\mathbf{k}(\kappa_1 + m, \kappa_2 + n)) = f(\mathbf{k}(\kappa_1, \kappa_2)), \quad (\kappa_1, \kappa_2) \in T, m, n \in \mathbb{Z}. \quad (44)$$

For derivatives, by the orthogonality in (6), we have

$$\frac{\partial f}{\partial k_x} = \frac{1}{2\pi} \mathbf{a}_1 \cdot \mathbf{e}_x \frac{\partial f}{\partial \kappa_1} + \frac{1}{2\pi} \mathbf{a}_2 \cdot \mathbf{e}_x \frac{\partial f}{\partial \kappa_2}, \quad \frac{\partial f}{\partial k_y} = \frac{1}{2\pi} \mathbf{a}_1 \cdot \mathbf{e}_y \frac{\partial f}{\partial \kappa_1} + \frac{1}{2\pi} \mathbf{a}_2 \cdot \mathbf{e}_y \frac{\partial f}{\partial \kappa_2}, \quad (45)$$

where $\mathbf{a}_1$ and $\mathbf{a}_2$ are the real-space primitive lattice vectors in (9). By (13), we have similarly

$$\frac{\partial f}{\partial \kappa_1} = \mathbf{b}_1 \cdot \mathbf{e}_x \frac{\partial f}{\partial k_x} + \mathbf{b}_1 \cdot \mathbf{e}_y \frac{\partial f}{\partial k_y}, \quad \frac{\partial f}{\partial \kappa_2} = \mathbf{b}_2 \cdot \mathbf{e}_x \frac{\partial f}{\partial k_x} + \mathbf{b}_2 \cdot \mathbf{e}_y \frac{\partial f}{\partial k_y}. \quad (46)$$

Accordingly, for integrals over D^* , we have

$$\int_{D^*} d\mathbf{k} f(\mathbf{k}) = \int_{D^*} dk_x dk_y f(\mathbf{k}) = \frac{(2\pi)^2}{V_{\text{puc}}} \int_T d\kappa_1 d\kappa_2 f(\mathbf{k}(\kappa_1, \kappa_2)), \quad (47)$$

where the torus T is given by (12). By applying (44) and (47), we have

$$\int_{D^*} d\mathbf{k} \frac{\partial f(\mathbf{k})}{\partial \kappa_1} = \frac{(2\pi)^2}{V_{\text{puc}}} \int_{-\frac{1}{2}}^{\frac{1}{2}} d\kappa_2 [f(\mathbf{k}(1/2, \kappa_2)) - f(\mathbf{k}(-1/2, \kappa_2))] = 0, \quad (48)$$

$$\int_{D^*} d\mathbf{k} \frac{\partial f(\mathbf{k})}{\partial \kappa_2} = \frac{(2\pi)^2}{V_{\text{puc}}} \int_{-\frac{1}{2}}^{\frac{1}{2}} d\kappa_1 [f(\mathbf{k}(\kappa_1, 1/2)) - f(\mathbf{k}(\kappa_1, -1/2))] = 0. \quad (49)$$

Combing (45) with (48) and (49), we thus have

$$\int_{D^*} d\mathbf{k} \frac{\partial f(\mathbf{k})}{\partial k_x} = 0, \quad \int_{D^*} d\mathbf{k} \frac{\partial f(\mathbf{k})}{\partial k_y} = 0. \quad (50)$$

3.7 Poisson's equation on tori

This section contains basic results for Poisson's equation defined in the torus D^* defined in (10). Consider the following Poisson's equation

$$\frac{\partial^2 f}{\partial k_x^2} + \frac{\partial^2 f}{\partial k_y^2} = -g, \quad f \text{ is } \Lambda^*\text{-periodic in } D^*, \quad (51)$$

where $g : D^* \rightarrow \mathbb{R}$ is assumed to be an analytic and Λ^* -periodic function, whose Fourier series is of the form

$$g(\mathbf{k}) = \sum_{\mathbf{R} \in \Lambda} g_{\mathbf{R}} \cdot e^{i\mathbf{R} \cdot \mathbf{k}}, \quad \mathbf{k} \in D^*. \quad (52)$$

The analyticity of g implies $|g_{\mathbf{R}}|$ decays exponentially for large $\|\mathbf{R}\|$. Suppose that its zeroth Fourier coefficient g_0 is zero; equivalently, this condition is given by formula

$$\int_{D^*} d\mathbf{k} g(\mathbf{k}) = 0. \quad (53)$$

Due to the condition in (53), it is easy to verify that

$$f(\mathbf{k}) = f_0 + \sum_{\substack{\mathbf{R} \in \Lambda \\ \mathbf{R} \neq \mathbf{0}}} \frac{g_{\mathbf{R}}}{\|\mathbf{R}\|^2} \cdot e^{i\mathbf{R} \cdot \mathbf{k}}, \quad \mathbf{k} \in D^*, \quad (54)$$

solves (51), where f_0 can be any real constant and may be chosen to be zero, i.e. $f_0 = 0$. Obviously, the solution $f : D^* \rightarrow \mathbb{R}$ in (54) is analytic since its Fourier coefficients also decay exponentially. Thus we have the following observation.

Theorem 3.3. *There exists a unique solution (up to a constant) to Poisson's equation in (51) if and only if*

$$\int_{D^*} d\mathbf{k} g(\mathbf{k}) = 0. \quad (55)$$

Moreover, the solution is analytic in D^* and given by (54).

3.8 Vector fields on tori

This section contains the Helmholtz-Hodge decomposition [19] of vector fields defined on the torus D^* in (10). In this special case, the decomposition of such vector fields is easily obtained from their Fourier series.

Consider a vector field $\mathbf{f} : D^* \rightarrow \mathbb{R}^2$ given by the formula

$$\mathbf{f}(\mathbf{k}) = (f_x(\mathbf{k}), f_y(\mathbf{k})). \quad (56)$$

We assume that $\mathbf{f}$ is an analytic and Λ^* -periodic function (see (20)) in D^* . Thus we expand $\mathbf{f}$ in its Fourier series

$$\mathbf{f}(\mathbf{k}) = \sum_{\mathbf{R} \in \Lambda} \mathbf{f}_{\mathbf{R}} \cdot e^{i\mathbf{R} \cdot \mathbf{k}}, \quad (57)$$

where the constant vector $\mathbf{f}_{\mathbf{R}} \in \mathbb{R}^2$ defined by

$$\mathbf{f}_{\mathbf{R}} = (f_{x,\mathbf{R}}, f_{y,\mathbf{R}}) \quad (58)$$

contains the Fourier coefficients of $f_{x,\mathbf{R}}$ and $f_{y,\mathbf{R}}$ of f_x and f_y at the lattice point $\mathbf{R}$. We decompose $\mathbf{f}_{\mathbf{R}}$ into

$$\mathbf{f}_{\mathbf{R}} = \mathbf{R} \frac{\mathbf{R} \cdot \mathbf{f}_{\mathbf{R}}}{\|\mathbf{R}\|^2} + \left(\mathbf{f}_{\mathbf{R}} - \mathbf{R} \frac{\mathbf{R} \cdot \mathbf{f}_{\mathbf{R}}}{\|\mathbf{R}\|^2} \right), \quad \text{for any } \mathbf{R} \neq \mathbf{0}, \quad (59)$$

where the first term is parallel to $\mathbf{R}$ and the second is orthogonal to $\mathbf{R}$. By components in $\mathbf{R} = (R_x, R_y)$, the second term in (59) can be written as

$$\mathbf{f}_{\mathbf{R}} - \mathbf{R} \frac{\mathbf{R} \cdot \mathbf{f}_{\mathbf{R}}}{\|\mathbf{R}\|^2} = (R_y, -R_x) \frac{R_y f_{x,\mathbf{R}} - R_x f_{y,\mathbf{R}}}{\|\mathbf{R}\|^2}. \quad (60)$$

Based on the decomposition of Fourier series in (59) and (60), we define a potential $\psi : D^* \rightarrow \mathbb{R}$ by the formula

$$\psi(\mathbf{k}) = - \sum_{\substack{\mathbf{R} \in \Lambda \\ \mathbf{R} \neq \mathbf{0}}} e^{i\mathbf{R} \cdot \mathbf{k}} \cdot \frac{i\mathbf{R} \cdot \mathbf{f}_{\mathbf{R}}}{\|\mathbf{R}\|^2}, \quad \mathbf{k} \in D^*, \quad (61)$$

for the curl-free component, and a potential $F : D^* \rightarrow \mathbb{R}$ by the formula

$$F(\mathbf{k}) = \sum_{\substack{\mathbf{R} \in \Lambda \\ \mathbf{R} \neq \mathbf{0}}} e^{i\mathbf{R} \cdot \mathbf{k}} \cdot \frac{iR_x f_{y,\mathbf{R}} - iR_y f_{x,\mathbf{R}}}{\|\mathbf{R}\|^2}, \quad \mathbf{k} \in D^*, \quad (62)$$

for the divergence-free component. Obviously, both ψ and F are analytic and Λ^* -periodic. We substitute (59) and (60) into (57) and utilize $\frac{\partial}{\partial k_x} e^{i\mathbf{R} \cdot \mathbf{k}} = iR_x e^{i\mathbf{R} \cdot \mathbf{k}}$, $\frac{\partial}{\partial k_y} e^{i\mathbf{R} \cdot \mathbf{k}} = iR_y e^{i\mathbf{R} \cdot \mathbf{k}}$ to arrive at the Helmholtz-Hodge decomposition of the vector field $\mathbf{f}$, given by the formula

$$\mathbf{f} = - \left(\frac{\partial \psi}{\partial k_x}, \frac{\partial \psi}{\partial k_y} \right) + \left(\frac{\partial F}{\partial k_y}, -\frac{\partial F}{\partial k_x} \right) + (h_x, h_y), \quad (63)$$

where the harmonic component (h_x, h_y) is given by the zeroth Fourier coefficients of $\mathbf{f}$ by the formula

$$(h_x, h_y) = (f_{x,\mathbf{0}}, f_{y,\mathbf{0}}), \quad (64)$$

which is obviously both curl-free and divergence-free. It is easily verified that (63) is identical to (57) due to (61), (62) and (64). We summarize this decomposition in the following observation.

Theorem 3.4. *Let $\mathbf{f} : D^* \rightarrow \mathbb{R}^2$ be an analytic and Λ^* -periodic vector field. Then there is a unique decomposition of $\mathbf{f}$ into a curl-free component due to the potential ψ , a divergence-free component due to the potential F and a harmonic component (h_x, h_y) in the form of (63).*

Furthermore, according to Theorem 3.3, (61) and (62) imply that ψ and F satisfy the following Poisson's equations

$$\frac{\partial^2 \psi}{\partial k_x^2} + \frac{\partial^2 \psi}{\partial k_y^2} = - \left(\frac{\partial f_x}{\partial k_x} + \frac{\partial f_y}{\partial k_y} \right), \quad (65)$$

$$\frac{\partial^2 F}{\partial k_x^2} + \frac{\partial^2 F}{\partial k_y^2} = - \left(\frac{\partial f_y}{\partial k_x} - \frac{\partial f_x}{\partial k_y} \right), \quad (66)$$

subject to the condition that ψ and F are both Λ^* -periodic in D^* . We observe that solving (65) and (66) only determines ψ and F up to a constant by Theorem 3.3. It does not pose any ambiguity since we are mostly interested in their derivatives as in (63).

3.9 Analyticity of eigenvalues and eigenvectors

In this section, we first introduce standard smoothness results for eigenvalues/vectors of an analytic family of matrices. The results in this section can be viewed as the higher-dimensional extension of the one-dimensional results [12], which is a summery of those in [15]. Moreover, we also include, in this section, formulas for the directional derivative of eigenvalues and eigenvectors; they are basically identical to those in [12] and only included here for completeness.

In this section, we consider a family of eigenvalues λ and its corresponding eigenvectors $\mathbf{v} \in \mathbb{C}^n$ of an analytic family of $n \times n$ Hermitian matrices $M : R \subseteq \mathbb{R}^2 \rightarrow \mathbb{C}^{n \times n}$, defined by the formula

$$M(x_1, x_2)\mathbf{v}(x_1, x_2) = \lambda(x_1, x_2)\mathbf{v}(x_1, x_2), \quad (x_1, x_2) \in R, \quad (67)$$

where the elements in M are assumed to be analytic both in R . Furthermore, similar to the one-dimensional case in [12], we assume that, for any $(z_1, z_2) \in \mathbb{C}^2$ in a complex neighborhood of $(x_1, x_2) \in R$, we have the extension of A into the complex plane $\mathbb{C}^2$ satisfying the condition

$$M(z_1, z_2)^* = M(\bar{z}_1, \bar{z}_2). \quad (68)$$

It should be observed that the generality of the one-dimensional results is lost (see Remark 2.6.3 in [15]); it is essential that we assume the eigenvalue family $\lambda(x_1, x_2)$ is *never degenerate* since we only consider a single eigenvalue in this section.

By the definition of the projector P via the formula [15]

$$P(x_1, x_2) = -\frac{1}{2\pi i} \oint_{\mathcal{C}} (M(x_1, x_2) - \xi)^{-1} d\xi, \quad (x_1, x_2) \in R, \quad (69)$$

where the integration is over a contour $\mathcal{C} \subset \mathbb{C}$ enclosing only the eigenvalue $\lambda(x_1, x_2)$ for any $(x_1, x_2) \in R$, it is obvious that the projector P is analytic in R . (Such a contour $\mathcal{C}$ is possible since we assume λ is never degenerate.) The eigenvalue family $\lambda(x_1, x_2)$ is also analytic in R by the formula

$$\lambda(x_1, x_2) = \text{Tr}(P(x_1, x_2)M(x_1, x_2)), \quad (x_1, x_2) \in R. \quad (70)$$

We summarize this observation in the following.

Theorem 3.5. *Suppose that λ and $\mathbf{v}$ are the family of eigenvalues and eigenvectors of the matrices A defined in (67) in $R \subseteq \mathbb{R}^2$, Suppose further that the eigenvalues λ are never degenerate in R . Then both the eigenvalues λ and the eigenprojectors P are analytic in R .*

We observe that the projector, which can also be expressed as $P = \mathbf{v}\mathbf{v}^*$, is analytic even when $\mathbf{v}$ is not, since P is independent of the phase choice of $\mathbf{v}$. In Section 2.4.2 of [15], Kato also gives a construction of an analytic family of $\mathbf{v}$ by solving an ODE. We define an analytic curve $\gamma : [a, b] \rightarrow \mathbb{R}^2$, starting at some point $\mathbf{x}_0$:

$$\gamma(s) = (x_1(s), x_2(s)) \in R, \quad \gamma(a) = \mathbf{x}_0, \quad s \in [a, b]. \quad (71)$$

Applying this construction to the eigenvalue problem (67) on γ yields the following result.

Theorem 3.6. *Let γ be the curve given in (71), A be the family of matrix given in (67) with eigenvalue family λ , and $\mathbf{v}_0$ be the eigenvector of a non-degenerate $\lambda(\mathbf{x}_0)$ at $\gamma(a) = \mathbf{x}_0$. Consider the following initial value problem*

$$\frac{d}{ds}\mathbf{v}(\gamma(s)) = Q(\gamma(s))\mathbf{v}(\gamma(s)), \quad \mathbf{v}(\mathbf{x}_0) = \mathbf{v}_0, \quad (72)$$

where the matrix Q is given by the commutator

$$Q(\gamma(s)) = \left[\frac{dP(\gamma(s))}{ds}, P(\gamma(s)) \right] = \frac{dP(\gamma(s))}{ds}P(\gamma(s)) - P(\gamma(s))\frac{dP(\gamma(s))}{ds}. \quad (73)$$

The solution $\mathbf{v}$ on the curve γ satisfies

$$M(\gamma(s))\mathbf{v}(\gamma(s)) = \lambda(\gamma(s))\mathbf{v}(\gamma(s)), \quad s \in [a, b] \quad (74)$$

and is analytic in $[a, b]$.

By differentiating $P^2 = P$, we obtain

$$P\frac{dP}{ds} + \frac{dP}{ds}P = \frac{dP}{ds}. \quad (75)$$

By multiplying (75) with P on the left, we obtain

$$P \frac{dP}{ds} P = 0. \quad (76)$$

Since $\mathbf{v}$ stays an eigenvector along the curve γ in Theorem 3.6, we have $P \frac{dP}{ds} \mathbf{v} = P \frac{dP}{ds} P \mathbf{v} = 0$ by (76) and $\frac{dP}{ds} P \mathbf{v} = \frac{dP}{ds} \mathbf{v}$. We use these two relations to simplify the ODE in (72) to the following

$$\frac{d}{ds} \mathbf{v}(\gamma(s)) = \frac{dP(\gamma(s))}{ds} \mathbf{v}(\gamma(s)). \quad (77)$$

Next, we apply the well-know perturbation formulas for projectors to (77) (see Section 2.5.4 in [15]) to derive the derivative of eigenvectors in terms of eigenvectors. We summarize this fact together with the derivative of eigenvalues in the following theorem.

Theorem 3.7. *Let γ be the curve given in (71) and M be the family of matrix given in (67) with a non-degenerate eigenvalue λ and its corresponding eigenvector $\mathbf{v}$ in R . We have the following formulas:*

$$\frac{d\lambda(\gamma(s))}{ds} = \mathbf{v}^*(\gamma(s)) \frac{dM(\gamma(s))}{ds} \mathbf{v}(\gamma(s)), \quad (78)$$

$$\frac{d\mathbf{v}(\gamma(s))}{ds} = -(M(\gamma(s)) - \lambda(\gamma(s)))^\dagger \frac{dM(\gamma(s))}{ds} \mathbf{v}(\gamma(s)), \quad (79)$$

where $(M - \lambda)^\dagger$ is the pseudoinverse of the matrix $A - \lambda$ ignoring the eigensubspace of λ .

By the definition of pseudoinverse, the range of $(A - \lambda)^\dagger$ is orthogonal to that of P , so we have the following formula

$$(M - \lambda)^\dagger P = P(M - \lambda)^\dagger = 0. \quad (80)$$

We observe that (72), (77) and (79) are mathematically equivalent. In this paper, we will use (77) for analytic purposes and (79) for numerically ones. Moreover, similar to Remark 2.10 in [12], we can easily replace the pseudoinverse in (79) with a true inverse for a similar matrix, which is more attractive numerically. Since we only deal with small matrices in this paper, this is not significant.

Remark 3.8. *The formula (79) can be written by the spectral decomposition of A in the following form*

$$\frac{d\mathbf{v}}{ds} = \sum_{\substack{j=1 \\ \lambda_j \neq \lambda}}^N \mathbf{v}_j \frac{\mathbf{v}_j^* \frac{dM}{ds} \mathbf{v}}{\lambda - \lambda_j}, \quad (81)$$

where λ_j and $\mathbf{v}_j$ are the rest of the eigenvalues and eigenvectors of M . This is equivalent to computing $\frac{d\mathbf{v}}{ds}$ by the singular value decomposition of the matrix $M - \lambda$.

3.10 Ordinary differential equations

In this section, we introduce a smoothness result for solutions to systems of linear ODEs with analytic coefficients. It is a straightforward generalization of continuity of solutions on parameters for ODEs with continuous coefficients, which can be found, for example, in [22].

Consider $R = [t_0, t_1] \times [-\frac{a}{2}, \frac{a}{2}]$ for some $t_1 > t_0$ and $a > 0$, and a family of $n \times n$ matrices $M : R \rightarrow \mathbb{C}^{n \times n}$. We assume that M is analytic in R and a periodic function of μ with period a :

$$M(t, \mu) = M(t, \mu + ma), \quad (t, \mu) \in R, m \in \mathbb{Z}. \quad (82)$$

Consider the following ODE system

$$\frac{d}{dt} \mathbf{v}(t, \mu) = M(t, \mu) \mathbf{v}(t, \mu), \quad (t, \mu) \in R, \quad (83)$$

with the initial condition

$$\mathbf{v}(t_0, \mu) = \mathbf{v}^0(\mu), \quad \mu \in \left[-\frac{a}{2}, \frac{a}{2}\right], \quad (84)$$

where we also assume $\mathbf{v}_0$ is analytic and has period a

$$\mathbf{v}^0(\mu) = \mathbf{v}^0(\mu + na), \quad \mu \in \left[-\frac{a}{2}, \frac{a}{2}\right], \quad n \in \mathbb{Z}. \quad (85)$$

The variable t parameterizes the ODE solution $\mathbf{v}$ and the other variable μ can be viewed as a numerical parameter.

By applying Picard iteration to (83), the solution $\mathbf{v}$ is given by the limit of the following uniformly convergent sequence of functions

$$\mathbf{v}_0(\mu), \quad \mathbf{v}_0(\mu) + \int_{t_0}^t ds M(s, \mu) \mathbf{v}_0(\mu), \quad \dots \quad (86)$$

Obviously, the form of the functions shows that the solution has the same periodicity as M and $\mathbf{v}^0$. Moreover, since the limit of a sequence of uniformly convergent analytic functions is analytic, we conclude that the solution $\mathbf{v}$ is also analytic in the region R . If we further assume that M is Hermitian, i.e. $M^* = M$ in R , and that the transpose of the initial condition $(\mathbf{v}^0)^*$ is also analytic in $[-\frac{1}{2}, \frac{1}{2}]$, the same argument applied to the transpose of (83) shows that the transpose of the solution $\mathbf{v}^*$ is also analytic in R .

Lemma 3.9. *The solution $\mathbf{v}$ to (83) subject to the initial condition (84) is analytic in R and period of the form*

$$\mathbf{v}(t, \mu) = \mathbf{v}(t, \mu + ma), \quad (t, \mu) \in R, m \in \mathbb{Z}. \quad (87)$$

Moreover, if we assume $M = M^*$ in R and $(\mathbf{v}^0)^*$ is analytic, the transpose of the solution $\mathbf{v}^*$ is also analytic in R and has the same periodicity as $\mathbf{v}$.

3.11 Pfaffian systems

In this section, we introduce the integrability condition for systems of first order partial differential equations in $\mathbb{R}^2$.

Let R be a subset of $\mathbb{R}^2$ and $M_1, M_2 : R \rightarrow \mathbb{C}^{n \times n}$ be two families of matrices containing twice continuously differentiable coefficients. Consider a system of first order partial differential equations given by the formulas

$$\frac{\partial}{\partial x_1} \mathbf{v}(x_1, x_2) = M_1(x_1, x_2) \mathbf{v}(x_1, x_2), \quad \frac{\partial}{\partial x_2} \mathbf{v}(x_1, x_2) = M_2(x_1, x_2) \mathbf{v}(x_1, x_2), \quad (88)$$

with the initial condition

$$\mathbf{v}(x_1^0, x_2^0) = \mathbf{v}^0 \quad (89)$$

for some $(x_1^0, x_2^0) \in R$ and a constant vector $\mathbf{v}^0 \in \mathbb{C}^n$. The system (88) is often referred to as a Pfaffian system [7].

We observe that each equation in (88) together with the initial condition (89) defines an initial value problem in both directions, so the system (88) is overdetermined and may not define $\mathbf{v}$ as a function in R . The following theorem states that the system (88) becomes integrable, i.e. it admits a solution $\mathbf{v}$ in R , if and only if the second order derivatives are commutative and R is simply connected. (It is analogous to the classical Poincaré lemma for vector fields.) For the particular choice of derivatives with respect to x_1, x_2 , the commutativity condition is given by the formula

$$\frac{\partial}{\partial x_1} \frac{\partial}{\partial x_2} \mathbf{v} = \frac{\partial}{\partial x_2} \frac{\partial}{\partial x_1} \mathbf{v}. \quad (90)$$

The solution is also unique by the uniqueness theorem for initial value problems. Thus we have the following observation, which is a special case of Theorem 6.20 in [7] stated in a slightly different form.

Theorem 3.10. *Suppose that R is a simply connected domain in $\mathbb{R}^2$. Then the Pfaffian system (88) with the initial condition (89) has a unique solution $\mathbf{v}$ in R if and only if the second order derivatives of $\mathbf{v}$ are commutative in R .*

4 Numerical preliminaries

This section introduces the numerical tools used in this paper.

4.1 Trapezoidal rule for periodic functions

We introduce the trapezoidal rule for approximating integrals of periodic functions.

Let $f : [0, L] \rightarrow \mathbb{C}$ be an analytic and periodic function of the form

$$f(x) = f(x + Lm), \quad x \in [0, L], \quad m \in \mathbb{Z}. \quad (91)$$

Let N be a positive integer and $h = \frac{L}{N}$. We denote an N -point trapezoidal rule approximation of the integral of f over $[0, L]$ by $I_N(f)$, defined by the formula

$$I_N(f) = h \sum_{j=0}^{N-1} f(jh). \quad (92)$$

The following fact is a slightly different version of Theorem 3.2 in [25], stating that the approximation I_N converges exponentially for f as the number of points N increases.

Theorem 4.1. *Suppose that $f : [0, L] \rightarrow \mathbb{C}$ is analytic and periodic in the form of (91). Then for any positive integer N , there exist positive real numbers C and a such that*

$$\left| I_N(f) - \int_0^L f(x) dx \right| < C e^{-aN}. \quad (93)$$

We observe that Theorem 4.1 is easily generalized to higher dimensions by recursively applying (92) to each dimension.

4.2 Discrete Fourier transform in two dimensions

Suppose that $g : D^* \rightarrow \mathbb{C}$ is analytic and Λ^* -periodic. We denote its Fourier coefficient by $g_{\mathbf{R}}$ at the lattice point $\mathbf{R} = m_1 \mathbf{a}_1 + m_2 \mathbf{a}_2 \in \Lambda$ for some integer m_1, m_2 and $g_{\mathbf{R}}$ is given by the formula (see (25))

$$g_{\mathbf{R}} = \frac{(2\pi)^2}{V_{\text{puc}}} \int_{D^*} d\mathbf{k} e^{-i\mathbf{k} \cdot \mathbf{R}} g(\mathbf{k}) \quad (94)$$

$$= \int_T d\kappa_1 d\kappa_2 e^{-2\pi i m_1 \kappa_1} e^{-2\pi i m_2 \kappa_2} g(\mathbf{k}(\kappa_1, \kappa_2)), \quad (95)$$

where (47) and (6) are used for obtaining the second equality. For simplicity, let N be a positive even integer and $h = \frac{1}{N}$. Suppose that m_1, m_2 are restricted to $-N/2, -(N/2-1), \dots, N/2-1$. Applying the trapezoidal rule approximation (92) to both variables κ_1 and κ_2 in (95) yields the approximation $\hat{g}_{m_1 m_2}$ of the Fourier coefficient $g_{\mathbf{R}}$ via a discrete Fourier transform

$$\hat{g}_{m_1 m_2} = \frac{1}{N^2} \sum_{j_1=-N/2}^{N/2-1} \sum_{j_2=-N/2}^{N/2-1} e^{-\frac{2\pi i}{N} m_1 j_1} e^{-\frac{2\pi i}{N} m_2 j_2} g(\mathbf{k}(j_1 h, j_2 h)). \quad (96)$$

We observe that the approximation $\hat{g}_{mn}$ converges to $g_{\mathbf{R}}$ exponentially as N increases by the two-dimensional version of Theorem (4.1). We denote the discrete Fourier transform on the right of (96) by $\mathcal{F}_N$, turning (96) into

$$\hat{g}_{m_1 m_2} = (\mathcal{F}_N(g))_{m_1 m_2} . \quad (97)$$

It is straightforward to verify that the inverse $\mathcal{F}_N^{-1}$ is given by the formula

$$g(\mathbf{k}(j_1 h, j_2 h)) = (\mathcal{F}_N^{-1}(\hat{g}))_{j_1 j_2} = \sum_{m_1=-N/2}^{N/2-1} \sum_{m_2=-N/2}^{N/2-1} e^{\frac{2\pi i}{N} m_1 j_1} e^{\frac{2\pi i}{N} m_2 j_2} \hat{g}_{m_1 m_2} , \quad (98)$$

for $j_1, j_2 = -N/2, -(N/2-1), \dots, N/2-1$. Both $\mathcal{F}_N$ and $\mathcal{F}_N^{-1}$ can be applied via the Fast Fourier transform (FFT) in $O(N^2 \log N)$ operations. The details can be found, for example, in [4].

Since derivatives $\frac{\partial}{\partial \kappa_1}$ and $\frac{\partial}{\partial \kappa_2}$ on the basis $e^{2\pi i m_1 \kappa_1} e^{2\pi i m_2 \kappa_2}$ produce $2\pi i m_1$ and $2\pi i m_2$ respectively, we compute an approximate of $\frac{\partial}{\partial \kappa_1} g$ and $\frac{\partial}{\partial \kappa_2} g$ by the formulas

$$\begin{aligned} \frac{\partial}{\partial \kappa_1} g(\mathbf{k}(j_1 h, j_2 h)) &\approx \sum_{m_1=-N/2}^{N/2-1} \sum_{m_2=-N/2}^{N/2-1} e^{\frac{2\pi i}{N} m_1 j_1} e^{\frac{2\pi i}{N} m_2 j_2} (2\pi i m_1 \hat{g}_{m_1 m_2}) , \\ \frac{\partial}{\partial \kappa_2} g(\mathbf{k}(j_1 h, j_2 h)) &\approx \sum_{m_1=-N/2}^{N/2-1} \sum_{m_2=-N/2}^{N/2-1} e^{\frac{2\pi i}{N} m_1 j_1} e^{\frac{2\pi i}{N} m_2 j_2} (2\pi i m_2 \hat{g}_{m_1 m_2}) , \end{aligned} \quad (99)$$

for $j_1, j_2 = -N/2, -(N/2-1), \dots, N/2-1$. It is convenient to introduce the notation $\mathcal{D}_1$ and $\mathcal{D}_2$ for the component-wise multiplication by $2\pi i m_1$ and $2\pi i m_2$ in (99) respectively, so that (99) is compactly written as

$$\frac{\partial}{\partial \kappa_1} g \approx \mathcal{F}_N^{-1} \mathcal{D}_1 \mathcal{F}_N(g) , \quad \frac{\partial}{\partial \kappa_2} g \approx \mathcal{F}_N^{-1} \mathcal{D}_2 \mathcal{F}_N(g) . \quad (100)$$

Since the approximation $\hat{g}$ in (96) converges exponentially, so are those in (99).

4.3 Fourth-order Runge-Kutta method

A standard fourth-order Runge-Kutta method (see [8], for example) solves the following initial value problem

$$y'(t) = f(t, y) , \quad y(t_0) = y_0 ,$$

defined for $t \in [t_0, t + L]$ via the formulas

$$\begin{aligned} t_{i+1} &= t_i + h \\ k_1 &= hf(t_i, y_i), \quad k_2 = hf(t_i + \frac{1}{2}h, y_i + \frac{1}{2}k_1), \\ k_3 &= hf(t_i + \frac{1}{2}h, y_i + \frac{1}{2}k_2), \quad k_4 = hf(t_i + h, y_i + k_3), \\ y(t_{i+1}) &= y(t_i) + \frac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4), \end{aligned}$$

with $i = 0, 1, \dots, N$ and $h = \frac{L}{N}$. For a smooth f , the global truncation error is $O(h^4)$. More explicitly, the approximation $y(t, h)$ at a fixed point t has an expansion of the form

$$y(t, h) = y(t) + c_4(t)h^4 + c_5(t)h^5 + O(h^6). \quad (101)$$

Applying the Richardson extrapolation via the formulas

$$\begin{aligned} \Delta_1(t) &= \frac{1}{15}(16y(t, h/2) - y(t, h)), \quad \Delta_2(t) = \frac{1}{15}(16y(t, h/4) - y(t, h/2)), \\ \Delta_3(t) &= \frac{1}{31}(32\Delta_2(t) - \Delta_1(t)), \end{aligned} \quad (102)$$

yields an approximation $\Delta_3(t)$ with $O(h^6)$ global truncation error.

5 Analytic apparatus

This section introduces the analytic apparatus for constructing optimal Wannier functions in Section 6. In Section 5.1, we define the parallel transport equation for assigning eigenvectors. In Section 5.2, we introduce the Berry connection and curvature. Moreover, we relate the integrability of Pfaffian systems arising from parallel transport to the Berry curvature. Section 5.3 introduces gauge transformations. Section 5.4 introduces the first Chern number as the topological obstruction to exponentially localized Wannier functions. Section 5.5 contains the formulas of the variance of Wannier functions and the optimal gauge choice.

Consider the eigenvalue problem defined by some analytic family of matrix $H : D^* \rightarrow \mathbb{C}^{n \times n}$ introduced in Section 3.3 by the formula

$$H(\mathbf{k})\mathbf{u}(\mathbf{k}) = E(\mathbf{k})\mathbf{u}(\mathbf{k}), \quad \mathbf{k} \in D^*, \quad (103)$$

where $\mathbf{u}$ and E are the family of eigenvectors and eigenvalues. We assume that E is never degenerate. The family of matrices H is Λ^* -periodic and analytic in D^* (see (20)) and is assumed to satisfy the condition (68). By applying Theorem 3.5 to (103), we conclude the both the eigenvalue E and the projection $P = \mathbf{u}\mathbf{u}^*$ are Λ^* -periodic and analytic in D^* . More explicitly, the projector P satisfies

$$P(\mathbf{k}) = P(\mathbf{k} + \mathbf{G}), \quad \mathbf{k} \in D^*, \mathbf{G} \in \Lambda^*. \quad (104)$$

5.1 Parallel transport equation

In this section, we apply results in Section 3.9 to derive the parallel transport equation to assign eigenvector $\mathbf{u}$ in (103) along curves in D^* .

Suppose that $\gamma : [0, 1] \rightarrow D^*$ is an analytic curve with $\gamma(0) = \mathbf{k}^0$. From Kato's construction in Theorem 3.6, we can assign eigenvectors analytically as a function of $\mathbf{k}$ along $\gamma \subset D^*$ by solving the following ODE

$$\frac{d\mathbf{u}(\gamma(s))}{ds} = \frac{dP(\gamma(s))}{ds} \mathbf{u}(\gamma(s)), \quad (105)$$

subject to the initial condition

$$\mathbf{u}(\mathbf{k}^0) = \mathbf{u}^0, \quad (106)$$

where $\mathbf{u}_0$ is the eigenvector of $H(\mathbf{k}_0)$. We observe that

$$\mathbf{u}^* \frac{dP}{ds} \mathbf{u} = \mathbf{u}^* P \frac{dP}{ds} P \mathbf{u} \quad (107)$$

since $\mathbf{u}$ is an eigenvector by Theorem 3.6. By (76), we always have

$$\mathbf{u}^*(\gamma(s)) \frac{d\mathbf{u}(\gamma(s))}{ds} = 0 \quad (108)$$

along γ . Hence the change of $\mathbf{u}$ is always orthogonal to $\mathbf{u}$ and (105) is referred to as the parallel-transport equation. Moreover, the condition (108) implies that the normalization is unchanged:

$$\|\mathbf{u}(\mathbf{k})\| = 1, \quad \text{for any } \mathbf{k} \text{ in } \gamma. \quad (109)$$

The apparent difficulty in extending such solutions along curves (families of one dimensional assignments) to be a globally smooth two-dimensional assignment of $\mathbf{u}$ is as follows. Consider two curves passing through some $\mathbf{k} \in D^*$ with the tangent vector in two different directions. If we choose the directions to be in the $\mathbf{e}_x$ and $\mathbf{e}_y$ (see (13)), we obtain a Pfaffian system (see Section 3.11) given by the formulas

$$\frac{\partial \mathbf{u}(\mathbf{k})}{\partial k_x} = \frac{\partial P(\mathbf{k})}{\partial k_x} \mathbf{u}(\mathbf{k}), \quad \frac{\partial \mathbf{u}(\mathbf{k})}{\partial k_y} = \frac{\partial P(\mathbf{k})}{\partial k_y} \mathbf{u}(\mathbf{k}). \quad (110)$$

Equivalently, if the directions are given by $\mathbf{b}_1$ and $\mathbf{b}_2$, we obtain a similar system

$$\frac{\partial \mathbf{u}(\mathbf{k})}{\partial \kappa_1} = \frac{\partial P(\mathbf{k})}{\partial \kappa_1} \mathbf{u}(\mathbf{k}), \quad \frac{\partial \mathbf{u}(\mathbf{k})}{\partial \kappa_2} = \frac{\partial P(\mathbf{k})}{\partial \kappa_2} \mathbf{u}(\mathbf{k}), \quad (111)$$

where the parameterization $\mathbf{k} = \mathbf{k}(\kappa_1, \kappa_2)$ is given in (13). (We will use (110) primarily for analytic purposes and (111) for numerical ones.) For any solution $\mathbf{u}$ obtained from solving (105)

in any subset of D^* , the solution $\mathbf{u}$ must satisfy both equations in (110). However, the system (110) is overdetermined and does not always meet the integrability condition in Theorem 3.10.

The non-integrability of the Pfaffian system (110) can also be understood by solving (105) with the initial condition (106) around a closed loop γ_c (that starts and ends at the same point $\mathbf{k}_0$). The solution at the end point is also an eigenvector of $H(\mathbf{k}_0)$ of the same band (as we assume the eigenvalue is non-degenerate), but it could differ from the initial $\mathbf{u}_0$ by a path-dependent phase factor $e^{i\varphi_{\gamma_c}}$:

$$\mathbf{u}^0 \rightarrow e^{i\varphi_{\gamma_c}} \mathbf{u}^0. \quad (112)$$

This is analogous to the case in classical differential geometry, where parallel transporting a vector on a surface around a closed loop results in a change in the angle between the initial and final vector when the Gaussian curvature is nonzero [1].

5.2 Berry connection, curvature and integrability

This section introduces the so-called Berry connection and Berry curvature related to the system (110). These quantities play important role in defining the assignment of $\mathbf{u}$ in D^* and the localization of its corresponding Wannier functions. The approach here is based on standard ideas in differential geometry but is somewhat unconventional. The standard definition of Berry connections and curvatures in physics literature are reproduced.

Let U be a subset of D^* that contains the curve γ in (105). Suppose that we define a continuously differentiable vector field $\mathbf{A} : U \rightarrow \mathbb{R}^2$ by the formula

$$\mathbf{A}(\mathbf{k}) = (A_x(\mathbf{k}), A_y(\mathbf{k})). \quad (113)$$

We modify (110) by adding an extra term, so that we obtain a new Pfaffian system, given by the formulas

$$\frac{\partial \tilde{\mathbf{u}}}{\partial k_x} = \frac{\partial P}{\partial k_x} \tilde{\mathbf{u}} - iA_x \tilde{\mathbf{u}}, \quad \frac{\partial \tilde{\mathbf{u}}}{\partial k_y} = \frac{\partial P}{\partial k_y} \tilde{\mathbf{u}} - iA_y \tilde{\mathbf{u}}. \quad (114)$$

When (110) and (114) are solved on the same curve γ subject to the same initial condition (106), we observe that the solution $\mathbf{u}(\mathbf{k})$ to (110) only differs from the one $\tilde{\mathbf{u}}(\mathbf{k})$ to (114) by a phase factor $e^{i\varphi_\gamma(\mathbf{k})}$ at any $\mathbf{k}$ in γ , where φ_γ is given by the line integral of $\mathbf{A}$ from $\mathbf{k}^0$ to $\mathbf{k}$ along γ :

$$\varphi_\gamma(\mathbf{k}) = \int_{\mathbf{k}^0}^{\mathbf{k}} \mathbf{A} \cdot d\mathbf{l}. \quad (115)$$

The quantities $\mathbf{A}$ and the phase φ_γ are referred to as the Berry connection and the Berry phase [2, 19]. Due to (108), A_x, A_y are also given by the formulas

$$A_x = i\tilde{\mathbf{u}}^* \frac{\partial \tilde{\mathbf{u}}}{\partial k_x}, \quad A_y = i\tilde{\mathbf{u}}^* \frac{\partial \tilde{\mathbf{u}}}{\partial k_y}, \quad (116)$$

which are the definition in physics literature [2, 19]. Similarly, the system (110) becomes

$$\frac{\partial \tilde{\mathbf{u}}}{\partial \kappa_1} = \frac{\partial P}{\partial \kappa_1} \tilde{\mathbf{u}} - \mathbf{i} A_1 \tilde{\mathbf{u}}, \quad \frac{\partial \tilde{\mathbf{u}}}{\partial \kappa_2} = \frac{\partial P}{\partial \kappa_2} \tilde{\mathbf{u}} - \mathbf{i} A_2 \tilde{\mathbf{u}}, \quad (117)$$

where A_1 and A_2 are respectively the components of $\mathbf{A}$ in (113) in the $\mathbf{b}_1$ and $\mathbf{b}_2$ direction

$$A_1 = \frac{\mathbf{b}_1}{\|\mathbf{b}_1\|} \cdot \mathbf{A}, \quad A_2 = \frac{\mathbf{b}_2}{\|\mathbf{b}_2\|} \cdot \mathbf{A}. \quad (118)$$

Next, we show that the new system (114) can be chosen to be integrable locally by choosing A_x and A_y appropriately. Consider a point $\mathbf{k}$ in a simply-connected subset $U \subset D^*$. Theorem 3.10 states that the Pfaffian system (114) is integrable (so it defines a function $\tilde{\mathbf{u}}$ for in U) if and only if the mixed derivatives are commutative:

$$\frac{\partial}{\partial k_y} \frac{\partial}{\partial k_x} \tilde{\mathbf{u}}(\mathbf{k}) = \frac{\partial}{\partial k_x} \frac{\partial}{\partial k_y} \tilde{\mathbf{u}}(\mathbf{k}), \quad \mathbf{k} \in U. \quad (119)$$

By differentiating the two equations in (114) followed simple manipulations, the condition (119) is equivalent to

$$\left(\frac{\partial A_y}{\partial k_x} - \frac{\partial A_x}{\partial k_y} \right) \tilde{\mathbf{u}} = \mathbf{i} \left[\frac{\partial P}{\partial k_x}, \frac{\partial P}{\partial k_y} \right] \tilde{\mathbf{u}}, \quad (120)$$

where $\left[\frac{\partial P}{\partial k_x}, \frac{\partial P}{\partial k_y} \right]$ is the commutator between $\frac{\partial P}{\partial k_x}$ and $\frac{\partial P}{\partial k_y}$ given by the formula

$$\left[\frac{\partial P}{\partial k_x}, \frac{\partial P}{\partial k_y} \right] = \frac{\partial P}{\partial k_x} \frac{\partial P}{\partial k_y} - \frac{\partial P}{\partial k_y} \frac{\partial P}{\partial k_x}. \quad (121)$$

We observe that $\frac{\partial P}{\partial k_x} \frac{\partial P}{\partial k_y} \tilde{\mathbf{u}}$ is in the span of $\tilde{\mathbf{u}}$ since

$$P \frac{\partial P}{\partial k_x} \frac{\partial P}{\partial k_y} \tilde{\mathbf{u}} = \left(-\frac{\partial P}{\partial k_x} P + \frac{\partial P}{\partial k_x} \right) \frac{\partial P}{\partial k_y} \tilde{\mathbf{u}} = \frac{\partial P}{\partial k_x} \frac{\partial P}{\partial k_y} \tilde{\mathbf{u}}, \quad (122)$$

where the first equality is obtained by applying (75) and the second equality comes from (76). Similar argument can be applied to show that $\frac{\partial P}{\partial k_y} \frac{\partial P}{\partial k_x} \tilde{\mathbf{u}}$ is also in the span of $\tilde{\mathbf{u}}$. As a result, by projecting the condition (120) onto $\tilde{\mathbf{u}}$, we obtain the condition equivalent to (120) given by the formula

$$\frac{\partial A_y}{\partial k_x} - \frac{\partial A_x}{\partial k_y} = \mathbf{i} \tilde{\mathbf{u}}^* \left[\frac{\partial P}{\partial k_x}, \frac{\partial P}{\partial k_y} \right] \tilde{\mathbf{u}}. \quad (123)$$

By the cyclic property of trace and $P = \tilde{\mathbf{u}} \tilde{\mathbf{u}}^*$, we obtain

$$\frac{\partial A_y}{\partial k_x} - \frac{\partial A_x}{\partial k_y} = \mathbf{i} \text{Tr} \left(P \left[\frac{\partial P}{\partial k_x}, \frac{\partial P}{\partial k_y} \right] \right). \quad (124)$$

We observe that the right-hand side of (124) is a real quantity and is referred to as the Berry curvature of the band [5]. We denote the Berry curvature by Ω_{xy} and it is given by the formula

$$\Omega_{xy} = \text{i Tr} \left(P \left[\frac{\partial P}{\partial k_x}, \frac{\partial P}{\partial k_y} \right] \right). \quad (125)$$

The quantity on the left of (124) is the typical definition of the Berry curvature in physics literature. As long as the Berry connection (A_x, A_y) is chosen to satisfy the condition (124) in U , the Pfaffian system (114) gives a well-defined assignment of $\tilde{\mathbf{u}}$ in U . We summarize this observation in the following.

Theorem 5.1. *Suppose U is a simply-connected subset of $\mathbb{R}^2$ and the system (114) satisfies an initial condition of the form (106) at some $\mathbf{k}_0 \in U$. Suppose further that the components of the Berry connection $\mathbf{A}$ in (114) are chosen to satisfy the integrability condition (124) in U . Then the Pfaffian system (114) defines a unique solution $\tilde{\mathbf{u}}$ in U by Theorem 3.10.*

It should be observed that Theorem 5.1 only ensures local integrability in a simply-connected region in D^* . As we will see in Section 6, extending it to a global one in D^* may encounter topological obstruction since D^* as a torus is not simply-connected. Moreover, even when there is no such obstruction, the condition (124) is not sufficient to ensure a Λ^* -periodic assignment of $\tilde{\mathbf{u}}$ in D^* since (124) only specifies the divergence-free part (see Theorem 3.4) of the Berry connection.

5.3 Gauge transformation

Suppose that $\tilde{\mathbf{u}}$ satisfies (114) defined by some Berry connection $\mathbf{A}$ satisfying the condition (124) in a simply-connected domain $U \subseteq D^*$. Given some initial condition, by Theorem 5.1, we have a well-defined assignment of $\tilde{\mathbf{u}}$ in U . Thus, we can multiply $\tilde{\mathbf{u}}$ by a phase factor $e^{-\text{i}\varphi}$ with a continuously differentiable function $\varphi : U \rightarrow \mathbb{R}$. We define the gauge transformation of $\tilde{\mathbf{u}}$ by the formula

$$\tilde{\tilde{\mathbf{u}}}(\mathbf{k}) = e^{-\text{i}\varphi(\mathbf{k})} \tilde{\mathbf{u}}(\mathbf{k}), \quad \mathbf{k} \in U. \quad (126)$$

We observe that the new $\tilde{\tilde{\mathbf{u}}}$ satisfies the same system of differential equations as (114) but with a new Berry connection $\tilde{\mathbf{A}} = (\tilde{A}_x, \tilde{A}_y)$:

$$\frac{\partial \tilde{\tilde{\mathbf{u}}}}{\partial k_x} = \frac{\partial P}{\partial k_x} \tilde{\tilde{\mathbf{u}}} - \text{i} \tilde{A}_x \tilde{\tilde{\mathbf{u}}}, \quad \frac{\partial \tilde{\tilde{\mathbf{u}}}}{\partial k_y} = \frac{\partial P}{\partial k_y} \tilde{\tilde{\mathbf{u}}} - \text{i} \tilde{A}_y \tilde{\tilde{\mathbf{u}}}. \quad (127)$$

where the components of $\tilde{\mathbf{A}}$ are given by the formulas

$$\tilde{A}_x = A_x + \frac{\partial \varphi}{\partial k_x}, \quad \tilde{A}_y = A_y + \frac{\partial \varphi}{\partial k_y}. \quad (128)$$

We observe that gauge transformations do not affect the divergence-free part of $\mathbf{A}$. Namely, we have

$$\frac{\partial \tilde{A}_y}{\partial k_x} - \frac{\partial \tilde{A}_x}{\partial k_y} = \frac{\partial A_y}{\partial k_x} - \frac{\partial A_x}{\partial k_y} = \Omega_{xy}, \quad (129)$$

where the Berry curvature Ω_{xy} is defined in (125).

Remark 5.2. *The Berry curvature Ω_{xy} given by (125) only depends on the projector, thus is independent of the gauge choice of the eigenvector $\tilde{\mathbf{u}}$ in (126). As a result, in physics literature, the Berry curvature is often computed via the following formula [23]*

$$\Omega_{xy} = \frac{\partial A_y}{\partial k_x} - \frac{\partial A_x}{\partial k_y} = i \frac{\partial \mathbf{u}^*}{\partial k_x} \frac{\partial \mathbf{u}}{\partial k_y} - i \frac{\partial \mathbf{u}^*}{\partial k_y} \frac{\partial \mathbf{u}}{\partial k_x}, \quad (130)$$

where the derivatives $\frac{\partial \mathbf{u}}{\partial k_x}$ and $\frac{\partial \mathbf{u}}{\partial k_y}$ are given in (110). Then the perturbation formula (81) is applied to compute (110) via the formulas

$$\frac{\partial \mathbf{u}}{\partial k_x} = -(H - E)^\dagger \frac{\partial H}{\partial k_x} \mathbf{u}, \quad \frac{\partial \mathbf{u}}{\partial k_y} = -(H - E)^\dagger \frac{\partial H}{\partial k_y} \mathbf{u}. \quad (131)$$

5.4 First Chern number

In this section, we introduce the first Chern number associated with a non-degenerate eigenvalue and various formulas for computing it. The first Chern number is the only obstruction to the construction of exponentially localized Wannier functions [5].

Consider the eigenvalue problem in (103). Analogous to the Gauss-Bonnet formula in classical differential geometry, the first Chern number a band is defined by the integral of the Berry curvature (see (125)) over D^* [23, 19]

$$C_1 = \frac{1}{2\pi} \int_{D^*} d\mathbf{k} \Omega_{xy}(\mathbf{k}) = \frac{1}{2\pi} \int_{D^*} dk_x dk_y \Omega_{xy}(\mathbf{k}), \quad (132)$$

and C_1 only takes integer values, i.e. $C_1 \in \mathbb{Z}$. Since Ω_{xy} is independent of gauge transformations (see (129)), so is the first Chern number C_1 . Thus C_1 is an intrinsic property of the matrix H .

The following theorem states that the first Chern number is the topological obstruction to exponentially localized Wannier functions. It is first observed in [24] and later formalized in [20, 5]. This fact will emerge in the construction of Wannier functions in Section 6.

Theorem 5.3. *There exists an analytic and Λ^* -periodic assignment of $\mathbf{u}$ in D^* (so that the corresponding Wannier function is exponentially localized) if and only if the first Chern number is zero, i.e. $C_1 = 0$.*

The integral in the definition (132) is parameterized by variables k_x, k_y . When the κ_1, κ_2 parameterization in (13) is used, the formula in (124) is modified accordingly as

$$\frac{\partial A_2}{\partial \kappa_1} - \frac{\partial A_1}{\partial \kappa_2} = \Omega_{12}, \quad (133)$$

where the Berry curvature Ω_{12} in the κ_1, κ_2 parameterization is given by the formula

$$\Omega_{12} = i \operatorname{Tr} \left(P \left[\frac{\partial P}{\partial \kappa_1}, \frac{\partial P}{\partial \kappa_2} \right] \right). \quad (134)$$

By applying the change of variables formulas (45-47), we have the following identity for the two parameterizations:

$$C_1 = \frac{1}{2\pi} \int_{D^*} dk_x dk_y \Omega_{xy}(\mathbf{k}) = \frac{1}{2\pi} \int_T d\kappa_1 d\kappa_2 \Omega_{12}(\mathbf{k}(\kappa_1, \kappa_2)). \quad (135)$$

By integrating (133) over T and applying (135), we obtain the formula

$$\int_T d\kappa_1 d\kappa_2 \left(\frac{\partial A_2}{\partial \kappa_1} - \frac{\partial A_1}{\partial \kappa_2} \right) = 2\pi C_1. \quad (136)$$

Applying Green's theorem to the left-hand side gives the following formula for computing C_1 by an line integral over the boundary of T :

$$\begin{aligned} 2\pi C_1 = & \int_{-\frac{1}{2}}^{\frac{1}{2}} d\kappa_1 (A_1(\mathbf{k}(\kappa_1, -1/2)) - A_1(\mathbf{k}(\kappa_1, 1/2))) \\ & + \int_{-\frac{1}{2}}^{\frac{1}{2}} d\kappa_2 (A_2(\mathbf{k}(1/2, \kappa_2)) - A_2(\mathbf{k}(-1/2, \kappa_2))). \end{aligned} \quad (137)$$

It should be observed that (137) does not require A_1, A_2 to be periodic in T ; only continuity in their derivatives in T is needed in order to apply the Green's theorem to (136). In fact, periodicity of A_1, A_2 will imply $C_1 = 0$ automatically by (137).

Remark 5.4. When H has time-reversal symmetry (see (27)), by the reality of the Berry curvature Ω_{xy} and the symmetry of the projector in (28), we have

$$\Omega_{xy}(-\mathbf{k}) = -\Omega_{xy}(\mathbf{k}), \quad \mathbf{k} \in D^*. \quad (138)$$

In the κ_1, κ_2 parameterization, this symmetry becomes

$$\Omega_{12}(\mathbf{k}(-\kappa_1, -\kappa_2)) = -\Omega_{12}(\mathbf{k}(\kappa_1, \kappa_2)), \quad (\kappa_1, \kappa_2) \in T. \quad (139)$$

Thus C_1 is automatically zero by (132). As a result, time-reversal symmetry ensures that no topological obstruction will be encountered.

5.5 Gauge choice and Wannier localization

In this section, we introduce formulas for the moment functions (41) and (42) in terms of the Berry connection $\mathbf{A}$ introduced in Section 5.2 and the optimal choice of the Berry connection in terms of the variance of the Wannier functions. We show that optimal Wannier functions correspond to divergence-free Berry connections and are unique up to lattice vector translations. Results of this form are known and can be found in [3, 18]. The discussion in this section assumes that we have found a vector field $\mathbf{A}$ that is Λ^* -periodic and analytic in D^* . Furthermore, we also assume that the system (114) by the vector field $\mathbf{A}$ produces a Λ^* -periodic and analytic assignment $\tilde{\mathbf{u}}$ in D^* .

5.5.1 Variance formulas for Wannier functions

We apply Theorem 3.4 to the vector field $\mathbf{A}$ to yield the formula

$$\mathbf{A} = (A_x, A_y) = - \left(\frac{\partial \psi}{\partial k_x}, \frac{\partial \psi}{\partial k_y} \right) + \left(\frac{\partial F}{\partial k_y}, -\frac{\partial F}{\partial k_x} \right) + (h_x, h_y), \quad (140)$$

where ψ and F are Λ^* -periodic potentials in D^* that generate the curl-free and divergence-free component respectively, and h_x, h_y are constants that define the harmonic component. Obviously, by Theorem 5.1, it is necessary that $\mathbf{A}$ satisfies (124) in D^* . This is equivalent to the following condition

$$\frac{\partial^2 F}{\partial k_x^2} + \frac{\partial^2 F}{\partial k_y^2} = -\Omega_{xy}, \quad F \text{ is } \Lambda^*\text{-periodic in } D^*. \quad (141)$$

In other words, the divergence-free component of $\mathbf{A}$ is determined by the Berry curvature completely.

In the following lemma, we express the moment functions for the Wannier function determined by $\tilde{\mathbf{u}}$ in terms of quantities in (140). The formulas are a simple consequence of substituting (114) into (41) and (42) and using

$$\tilde{\mathbf{u}}^* \frac{\partial P}{\partial k_x} \tilde{\mathbf{u}} = 0, \quad \tilde{\mathbf{u}}^* \frac{\partial P}{\partial k_y} \tilde{\mathbf{u}} = 0, \quad (142)$$

which are consequences of (108). The details can be found in Appendix 10.2.

Lemma 5.5. *Suppose that the Berry connection $\mathbf{A}$ is defined in (140) and its corresponding system (110) defines a Λ^* -periodic and analytic assignment of $\tilde{\mathbf{u}}$ in D^* . We have the following formula for the moment functions:*

$$\langle \mathbf{R} \rangle = (h_x, h_y), \quad (143)$$

$$\langle \|\mathbf{R}\|^2 \rangle = \frac{V_{\text{puc}}}{(2\pi)^2} \int_{D^*} d\mathbf{k} \left(\left\| \frac{\partial P(\mathbf{k})}{\partial k_x} \tilde{\mathbf{u}}(\mathbf{k}) \right\|^2 + \left\| \frac{\partial P(\mathbf{k})}{\partial k_y} \tilde{\mathbf{u}}(\mathbf{k}) \right\|^2 + \|\mathbf{A}(\mathbf{k})\|^2 \right), \quad (144)$$

where

$$\begin{aligned} \frac{V_{\text{puc}}}{(2\pi)^2} \int_{D^*} d\mathbf{k} \|\mathbf{A}(\mathbf{k})\|^2 &= \frac{V_{\text{puc}}}{(2\pi)^2} \int_{D^*} d\mathbf{k} \left[\left(\frac{\partial \psi}{\partial k_x} \right)^2 + \left(\frac{\partial \psi}{\partial k_y} \right)^2 \right. \\ &\quad \left. + \left(\frac{\partial F}{\partial k_x} \right)^2 + \left(\frac{\partial F}{\partial k_y} \right)^2 \right] + h_x^2 + h_y^2. \end{aligned} \quad (145)$$

Thus the variance is given by the formula

$$\begin{aligned} \langle \|\mathbf{R}\|^2 \rangle - \|\langle \mathbf{R} \rangle\|^2 &= \frac{V_{\text{puc}}}{(2\pi)^2} \int_{D^*} d\mathbf{k} \left[\left\| \frac{\partial P(\mathbf{k})}{\partial k_x} \tilde{\mathbf{u}}(\mathbf{k}) \right\|^2 + \left\| \frac{\partial P(\mathbf{k})}{\partial k_y} \tilde{\mathbf{u}}(\mathbf{k}) \right\|^2 \right. \\ &\quad \left. + \left(\frac{\partial \psi}{\partial k_x} \right)^2 + \left(\frac{\partial \psi}{\partial k_y} \right)^2 + \left(\frac{\partial F}{\partial k_x} \right)^2 + \left(\frac{\partial F}{\partial k_y} \right)^2 \right]. \end{aligned} \quad (146)$$

Remark 5.6. We observe that Wannier center $\langle \mathbf{R} \rangle$ is determined by the harmonic component in (140), which is analogous to the Zak phase in the one-dimensional case [27, 12]. Thus h_x and h_y are proportional to the Zak phase in the $\mathbf{e}_x$ and $\mathbf{e}_y$ direction respectively.

5.5.2 Optimal Berry connections

The derivatives of F is determined by the Berry curvature completely (see (141)). Thus the only variable quantities are those related to ψ . We show next that a gauge transformation can be applied to make ψ vanish, thus obtaining an optimal Wannier function whose variance only contains gauge-independent quantities.

In order to minimize the variance of the Wannier function defined by $\tilde{\mathbf{u}}$ in Lemma 5.5, we apply the following gauge transformation

$$\tilde{\tilde{\mathbf{u}}}(\mathbf{k}) = e^{-i\varphi(\mathbf{k})} \tilde{\mathbf{u}}(\mathbf{k}), \quad \mathbf{k} \in D^*, \quad (147)$$

where $\varphi : D^* \rightarrow \mathbb{R}$ is analytic (not necessarily Λ^* -periodic) in D^* . In the following, we first give the constraint on φ such that the transformed $\tilde{\tilde{\mathbf{u}}}$ stays analytic.

The Berry connection $\mathbf{A}$ in Lemma 5.5 transforms into a new $\tilde{\tilde{\mathbf{A}}}$ according to (128). Since $\mathbf{A}$ is Λ^* -periodic, for the new $\tilde{\tilde{\mathbf{A}}}$ to remain Λ^* -periodic, both $\frac{\partial \varphi}{\partial k_x}$ and $\frac{\partial \varphi}{\partial k_y}$ must be Λ^* -periodic. Suppose their Fourier series are given by the formulas

$$\frac{\partial \varphi(\mathbf{k})}{\partial k_x} = \varphi_{x,0} + \sum_{\substack{\mathbf{R} \in \Lambda \\ \mathbf{R} \neq 0}} \varphi_{x,\mathbf{R}} \cdot e^{i\mathbf{R} \cdot \mathbf{k}}, \quad \frac{\partial \varphi(\mathbf{k})}{\partial k_y} = \varphi_{y,0} + \sum_{\substack{\mathbf{R} \in \Lambda \\ \mathbf{R} \neq 0}} \varphi_{y,\mathbf{R}} \cdot e^{i\mathbf{R} \cdot \mathbf{k}}, \quad \mathbf{k} \in D^*, \quad (148)$$

where $\varphi_{x,\mathbf{R}}$ and $\varphi_{y,\mathbf{R}}$ are the Fourier coefficients for the derivatives and the zeroth ones are singled out. Integrating (148) shows that φ must be of the form

$$\varphi(\mathbf{k}) = \mathbf{c}_0 \cdot \mathbf{k} + f(\mathbf{k}), \quad \mathbf{k} \in D^*, \quad (149)$$

where $\mathbf{c}_0 = (\varphi_{x,0}, \varphi_{y,0})$ contains the zeroth Fourier coefficients in (148) and f is some Λ^* -periodic function. Furthermore, we have assumed that the assignment of $\tilde{\mathbf{u}}$ is analytic and Λ^* -periodic in D^* . As a result, in order for the new $e^{-i\varphi}\tilde{\mathbf{u}}$ to be analytic and Λ^* -periodic, we must impose the following condition for the term linear in $\mathbf{k}$ in (149):

$$\mathbf{c}_0 \cdot (\mathbf{k} + \mathbf{G}) = \mathbf{c}_0 \cdot \mathbf{k} + \mathbf{c}_0 \cdot \mathbf{G} = \mathbf{c}_0 \cdot \mathbf{k} + 2\pi l, \quad \mathbf{k} \in D^*, \mathbf{G} \in \Lambda^*, \quad (150)$$

for some integer l . By (5) and (8), we observe that the above requirement is equivalent to that the constant vector $\mathbf{c}_0$ is given by some lattice point $\mathbf{R}_0$ in Λ . Thus we have the following observation.

Lemma 5.7. *Let the Berry connection $\mathbf{A}$ in the Pfaffian system (114) be given by (140). Suppose that (114) defines a Λ^* -periodic and analytic assignment of $\tilde{\mathbf{u}}$ in D^* . Then if a gauge transformation in (126) defined by a function $\varphi : D^* \rightarrow \mathbb{R}$ that does not change the smoothness of $\tilde{\mathbf{u}}$, the function φ must be of the following form*

$$\varphi(\mathbf{k}) = \mathbf{R}_0 \cdot \mathbf{k} + f(\mathbf{k}), \quad (151)$$

where $\mathbf{R}_0$ is a lattice point in Λ and f is a Λ^* -periodic and analytic function in D^* .

The transformation of the Berry connection $\mathbf{A}$ in (128) shows the gauge transformation defined by φ in Lemma 5.7 will only affect the divergence (curl-free component) and the harmonic component of $\mathbf{A}$. More explicitly, the transformed $\tilde{\mathbf{A}}$ is given by

$$\tilde{\mathbf{A}} = (\tilde{A}_x, \tilde{A}_y) = \left(-\frac{\partial\psi}{\partial k_x} + \frac{\partial f}{\partial k_x}, -\frac{\partial\psi}{\partial k_y} + \frac{\partial f}{\partial k_y} \right) + \left(\frac{\partial F}{\partial k_y}, -\frac{\partial F}{\partial k_x} \right) + (h_x, h_y) + \mathbf{R}_0. \quad (152)$$

By applying Lemma 5.5 to the new $\tilde{\mathbf{A}}$, we observe that the choice of $\mathbf{R}_0$ only shifts the Wannier center by a lattice vector without affecting the variance. Moreover, the optimal choice for minimum variance is to choose $f = \psi$ so that $\tilde{\mathbf{A}}$ becomes divergence-free

$$\tilde{\mathbf{A}} = (\tilde{A}_x, \tilde{A}_y) = \left(\frac{\partial F}{\partial k_y}, -\frac{\partial F}{\partial k_x} \right) + (h_x, h_y) + \mathbf{R}_0. \quad (153)$$

As discussed below Theorem 3.4, ψ can be obtained by solving

$$\frac{\partial^2 \psi}{\partial k_x^2} + \frac{\partial^2 \psi}{\partial k_y^2} = - \left(\frac{\partial A_x}{\partial k_x} + \frac{\partial A_y}{\partial k_y} \right), \quad g \text{ is } \Lambda^*\text{-periodic in } D^*, \quad (154)$$

where the right-hand side of (154) is the negative of the divergence of $\mathbf{A}$. Solving (154) can be viewed as computing the Newton step for minimizing the quadratic objective defined by (146) with the Laplacian being the Hessian; the Laplacian is diagonal in the Fourier series representation and can be inverted with little cost via the fast Fourier transform (see Section 3.7 and 4.2).

Furthermore, we observe that, if the divergence-free Berry connection $\tilde{\mathbf{A}}$ in (153) exists, it will be unique up to a lattice vector in Λ . In other words, if two divergence-free $\mathbf{A}_1$ and $\mathbf{A}_2$ that both produce some Λ^* -periodic and analytic assignment in D^* , we must have $\mathbf{A}_1 - \mathbf{A}_2 = \mathbf{d}$ for some $\mathbf{d} \in \Lambda$. The reason is as follows. Since both $\mathbf{A}_1$ and $\mathbf{A}_2$ are divergence-free, only their harmonic component can differ so their difference is a constant vector $\mathbf{d}$. Thus $\mathbf{A}_1$ and $\mathbf{A}_2$ can be converted to each other by a gauge transformation defined by $e^{-i\mathbf{d}\cdot\mathbf{k}}$. By Lemma 5.7, the vector $\mathbf{d}$ must be a lattice vector in Λ .

Due to the uniqueness result above, we simply choose $\mathbf{R}_0 = \mathbf{0}$ in (153). Moreover, since $\mathbf{R}_0$ in (153) only shift the Wannier center by $\mathbf{R}_0$ by Lemma 5.5 and all Wannier functions centered at difference lattice points are copies of each other (see (31)), such a choice has no effect on the physical consequences of Wannier functions. We summarize the above optimal conditions and its corresponding formulas for the Wannier center and the variance in the following theorem.

Theorem 5.8. *Let $\mathbf{A}$ be the Berry connection defined in (140) and its corresponding Pfaffian system (114) defines a Λ^* -periodic and analytic assignment of $\tilde{\mathbf{u}}$ in D^* . Suppose that ψ is given by (154). Then the new $\tilde{\tilde{\mathbf{u}}}$ given by the following gauge transform*

$$\tilde{\tilde{\mathbf{u}}}(\mathbf{k}) = e^{-i\psi(\mathbf{k})}\tilde{\mathbf{u}}(\mathbf{k}), \quad \mathbf{k} \in D^*, \quad (155)$$

is the optimal assignment for minimizing the variance (146). The optimal Berry connection corresponding to $\tilde{\tilde{\mathbf{u}}}$ is given by the formula

$$\tilde{\tilde{\mathbf{A}}} = (\tilde{\tilde{A}}_x, \tilde{\tilde{A}}_y) = \left(\frac{\partial F}{\partial k_y}, -\frac{\partial F}{\partial k_x} \right) + (h_x, h_y). \quad (156)$$

The Wannier center and the variance of the optimal Wannier function corresponding to $\tilde{\tilde{\mathbf{u}}}$ are given by

$$\langle \mathbf{R} \rangle = (h_x, h_y), \quad (157)$$

$$\begin{aligned} & \langle \|\mathbf{R}\|^2 \rangle - \|\langle \mathbf{R} \rangle\|^2 \\ &= \frac{V_{\text{puc}}}{(2\pi)^2} \int_{D^*} d\mathbf{k} \left[\left\| \frac{\partial P(\mathbf{k})}{\partial k_x} \tilde{\tilde{\mathbf{u}}}(\mathbf{k}) \right\|^2 + \left\| \frac{\partial P(\mathbf{k})}{\partial k_y} \tilde{\tilde{\mathbf{u}}}(\mathbf{k}) \right\|^2 + \left(\frac{\partial F}{\partial k_x} \right)^2 + \left(\frac{\partial F}{\partial k_y} \right)^2 \right]. \end{aligned} \quad (158)$$

Furthermore, such a divergence-free Berry connection $\tilde{\tilde{\mathbf{A}}}$ is unique up to a lattice vector in Λ .

We emphasize that Theorem 5.8 assumes the existence of analytic assignment of $\tilde{\mathbf{u}}$. Such assignment will be explicitly constructed in Section 6. The uniqueness of divergence-free Berry connection implies the optimally localized Wannier function is unique (up to a lattice translation) and an apparent constant phase factor. The constant phase will also be fixed to ensure the reality of the Wannier function in the construction in Section 6 in cases when H has time-reversal symmetry (see (27)).

Remark 5.9. *As observed in [3], the optimal condition that the Berry connection is divergence-free (or satisfies the Coulomb gauge) can also be derived by applying calculus of variations to (144). Since F -related terms are determined by the Berry curvature and h_x, h_y are fixed up to a lattice vector (see Lemma 5.7), ψ is the only variable quantity. Varying ψ gives to the condition that ψ satisfies the Laplace equation:*

$$\frac{\partial^2 \psi}{\partial k_x^2} + \frac{\partial^2 \psi}{\partial k_y^2} = 0, \quad \psi \text{ is } \Lambda^* \text{-periodic in } D^*. \quad (159)$$

Obviously, the solution ψ is a constant, which also leads to (158).

Remark 5.10. *The vector $\tilde{\mathbf{u}}$ in (158) and $\tilde{\mathbf{u}}$ in (146) can be replaced by a $\mathbf{u}$ of any phase choice since the quantity is independent of the phase of the vector.*

6 Construction of optimal Wannier functions

In this section, we describe an approach to constructing globally optimal Wannier functions in the sense of Theorem 5.8 with the minimum spread defined by the variance in (38). We assume we are given a family of n by n matrix H (a tight-banding Hamiltonian introduced in Section 3.3) that is analytic and Λ^* -periodic in D^* defined by two primitive reciprocal lattice vectors $\mathbf{b}_1$ and $\mathbf{b}_2$ as in (7). (We also implicitly assume the continuation condition (67) is satisfied for H .) The corresponding real space primitive lattice vectors $\mathbf{a}_1$ and $\mathbf{a}_2$ define a lattice Λ as in (5) and satisfy (6). Moreover, it is assumed that we have picked an eigenvalue E and eigenvector $\mathbf{u}$ of interests that define an eigenvalue equation as in (103).

The approach in section is purely based on the parallel transport equation along different lines in D^* in the form of (105). Along each line, we assign eigenvectors $\mathbf{u}$ according to the approach in [12], which constructs the optimal one-dimensional Wannier function. Doing so over a family of lines in D^* defines a two-dimensional assignment of $\mathbf{u}$. The analyticity of such an assignment is purely a consequence of smoothness result of ODEs introduced in Section 3.10. After this step, a single gauge transform is performed to eliminate the divergence of the Berry connection to achieve the optimality in Theorem 5.8. This approach can be viewed as a direct extension of the method for constructing optimal single band Wannier function in [12] in dimension one. The topological construction – first Chern number $C_1 \neq 0$ introduced

in Section 5.4 – emerges automatically in the parallel transport stage, causing the construction to fail as expected. When topological obstruction is present, this approach still manages to produce an analytic assignment of $\mathbf{u}$ in D^* viewed as a subset of $\mathbb{R}^2$. (The assignment is discontinuous when D^* is viewed as a torus.)

The method can be divided into three stages. In Stage 1, the parallel transport equation is solved on γ_0 in Figure 1. A simple correction is introduced (as in the one-dimensional case [12]) that yields an analytic and periodic assignment on γ_0 . In Stage 2, the result in Stage 1 serves as the initial conditions for the parallel transport equation on the path γ_{κ_1} for $\kappa_1 \in [-\frac{1}{2}, \frac{1}{2}]$ in Figure 1, followed by applying similar corrections as in Stage 1. We show that the assignment after Stage 2 is analytic and Λ^* -periodic in D^* when the topological obstruction is not present (i.e. $C_1 = 0$). Moreover, when the matrix H has time-reversal symmetry, we show that the assignment results in a real Wannier function automatically so long as the initial condition in Stage 1 is chosen to be real. In Stage 3, the divergence of the Berry connection of the assignment is computed and eliminated by a single gauge transformation to yield the optimal assignment. The new Wannier function remains real if the result after Stage 2 is real.

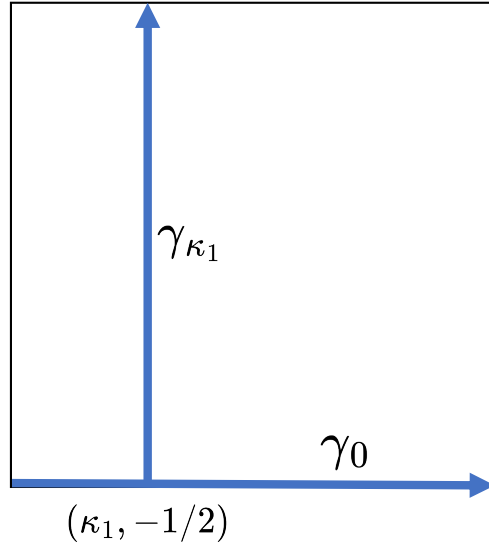


Figure 1: The path γ_0 in Stage 1 and γ_{κ_1} for $\kappa_1 \in [-\frac{1}{2}, \frac{1}{2}]$ in Stage 2 shown in T .

When describing the construction in the rest of this section, we parameterize D^* in (10) by T in (12) via (13). Hence, for any $\mathbf{k} \in D^*$, the eigenvector $\mathbf{u}$ is parameterized as

$$\mathbf{u}(\mathbf{k}) = \mathbf{u}(\mathbf{k}(\kappa_1, \kappa_2)), \quad (\kappa_1, \kappa_2) \in T. \quad (160)$$

The matrix H is also parameterized similarly by

$$H(\mathbf{k}) = H(\mathbf{k}(\kappa_1, \kappa_2)), \quad (\kappa_1, \kappa_2) \in T. \quad (161)$$

Formulas in Section 3.6 can be applied to convert various expressions of $\mathbf{k}$ into those of κ_1, κ_2 and vice versa.

Remark 6.1. *Although the approach in this section is a direct extension of the method for one-dimensional Wannier problems, it is arguably not the most natural one from a physics point of view. In Appendix 10.1, we provide an alternative construction, where the gauge-invariant vector fields in (156) are directly computed without creating the assignment of $\mathbf{u}$ first by exploiting the gauge-invariant nature of the Berry curvature in (125) (see Remark 5.2). Since the invariance no longer holds when multiple eigenvalues are considered, such an approach is not generalizable to the multi-band case. However, it highlights the roles of the gauge-invariant vector fields in constructing Wannier functions, thus it is included as a complementary viewpoint.*

6.1 Stage 1: constructing an assignment on a line

In the first stage, we start from the lower-left corner of D^* (see Figure 1) given by

$$\mathbf{k}^0 = -\frac{1}{2}\mathbf{b}_1 - \frac{1}{2}\mathbf{b}_2, \quad (162)$$

and construct an assignment of the eigenvectors on the line $\gamma_0 : [-\frac{1}{2}, \frac{1}{2}] \rightarrow D^*$ defined by

$$\gamma_0(\kappa_1) = \mathbf{k}(\kappa_1, -1/2) = \kappa_1\mathbf{b}_1 - \frac{1}{2}\mathbf{b}_2, \quad (163)$$

where the starting point $\gamma_0(-1/2) = \mathbf{k}^0$.

First, we compute the eigenvalue E^0 and eigenvector $\mathbf{v}^0$ at $\mathbf{k}^0$ given by the formula

$$H(\mathbf{k}^0)\mathbf{v}^0 = E^0\mathbf{v}^0. \quad (164)$$

Next, we assign $\mathbf{u}(\gamma_0(\kappa_1)) = \mathbf{u}(\mathbf{k}(\kappa_1, -1/2))$ for $\kappa_1 \in [-\frac{1}{2}, \frac{1}{2}]$ by solving the parallel transport equation (see (105)) given by the formula

$$\frac{\partial \mathbf{u}(\mathbf{k}(\kappa_1, -1/2))}{\partial \kappa_1} = \frac{\partial P(\mathbf{k}(\kappa_1, -1/2))}{\partial \kappa_1} \mathbf{u}(\mathbf{k}(\kappa_1, -1/2)), \quad \kappa_1 \in \left[-\frac{1}{2}, \frac{1}{2}\right], \quad (165)$$

subject to the initial condition

$$\mathbf{u}(\mathbf{k}(-1/2, -1/2)) = \mathbf{u}(\mathbf{k}^0) = \mathbf{v}^0. \quad (166)$$

By Theorem 3.6, the assignment $\mathbf{u}(\mathbf{k}(\kappa_1, -1/2))$ is analytic in $\kappa_1 \in [-\frac{1}{2}, \frac{1}{2}]$. By periodicity, we always have $H(\mathbf{k}(-1/2, -1/2)) = H(\mathbf{k}(1/2, -1/2))$, but the same condition does not hold for $\mathbf{u}$ in general. In other words, we generally have

$$\mathbf{u}(\mathbf{k}(-1/2, -1/2)) \neq \mathbf{u}(\mathbf{k}(1/2, -1/2)). \quad (167)$$

Hence, $\mathbf{u}$ is discontinuous as a periodic function in $[-\frac{1}{2}, \frac{1}{2}]$. Since both $\mathbf{u}(-1/2, -1/2)$ and $\mathbf{u}(1/2, -1/2)$ are both eigenvectors of the same (non-degenerate) eigenvalue, they can only differ by a constant phase factor, which we denote by $e^{i\varphi_1}$ for some real φ_1 :

$$\mathbf{u}(\mathbf{k}(1/2, -1/2)) = e^{i\varphi_1} \mathbf{u}(\mathbf{k}(-1/2, -1/2)) . \quad (168)$$

Obviously, φ_1 can be computed via

$$\varphi_1 = -i \log (\mathbf{u}(\mathbf{k}(-1/2, -1/2))^* \mathbf{u}(\mathbf{k}(1/2, -1/2))) . \quad (169)$$

Next, we apply a gauge transform (see (126)) to turn $\mathbf{u}$ into an analytic and periodic function in $[-\frac{1}{2}, \frac{1}{2}]$.

We apply the gauge transform given by the formula

$$\tilde{\mathbf{u}}(\mathbf{k}(\kappa_1, -1/2)) = e^{-i\varphi_1(\kappa_1 + \frac{1}{2})} \mathbf{u}(\mathbf{k}(\kappa_1, -1/2)) , \quad \kappa_1 \in \left[-\frac{1}{2}, \frac{1}{2}\right] , \quad (170)$$

so that we have

$$\tilde{\mathbf{u}}(\mathbf{k}(-1/2, -1/2)) = \tilde{\mathbf{u}}(\mathbf{k}(1/2, -1/2)) . \quad (171)$$

Thus $\tilde{\mathbf{u}}(\mathbf{k}(\kappa_1, -1/2))$ is continuous and periodic in $[-\frac{1}{2}, \frac{1}{2}]$. It is obvious that the new $\tilde{\mathbf{u}}$ satisfies the following equation

$$\frac{\partial \tilde{\mathbf{u}}(\mathbf{k}(\kappa_1, -1/2))}{\partial \kappa_1} = \frac{\partial P(\mathbf{k}(\kappa_1, -1/2))}{\partial \kappa_1} \tilde{\mathbf{u}}(\mathbf{k}(\kappa_1, -1/2)) - i\varphi_1 \tilde{\mathbf{u}}(\mathbf{k}(\kappa_1, -1/2)) , \quad (172)$$

subject to the same initial condition (166). By (171) and the periodicity of P (see (104)), the formula (172) shows that the derivative satisfies

$$\frac{\partial}{\partial \kappa_1} \tilde{\mathbf{u}}(\mathbf{k}(-1/2, -1/2)) = \frac{\partial}{\partial \kappa_1} \tilde{\mathbf{u}}(\mathbf{k}(1/2, -1/2)) . \quad (173)$$

By repetitive differentiation (172) with respect to κ_1 , the same argument shows that the derivative of $\tilde{\mathbf{u}}$ of any order $n = 0, 1, 2, \dots$ with respect to κ_1 satisfies

$$\frac{\partial^n}{\partial \kappa_1^n} \tilde{\mathbf{u}}(\mathbf{k}(-1/2, -1/2)) = \frac{\partial^n}{\partial \kappa_1^n} \tilde{\mathbf{u}}(\mathbf{k}(1/2, -1/2)) . \quad (174)$$

As a result, the constructed $\tilde{\mathbf{u}}(\mathbf{k}(\kappa_1, -1/2))$ defines an analytic and periodic function in $[-\frac{1}{2}, \frac{1}{2}]$. We summarize the result of the construction after the first stage in the following lemma.

Lemma 6.2. *Suppose that φ_1 is the constant given in (168). Then the solution $\tilde{\mathbf{u}}$ to (172) subject to the initial condition (166) is analytic and periodic on the line γ_0 defined in (163).*

By $\frac{\partial}{\partial \kappa_1} P^* = \frac{\partial}{\partial \kappa_1} P$, we can transpose (165) with its initial condition and repeat the same procedure above to show the same result for $\tilde{\mathbf{u}}^*$, the transpose of $\tilde{\mathbf{u}}$ in Lemma 6.2.

Corollary 6.3. *$\tilde{\mathbf{u}}^*$ is analytic and periodic on the line γ_0 .*

6.2 Stage 2: constructing an assignment on the torus

In the second stage, for each $\kappa_1 \in [-\frac{1}{2}, \frac{1}{2}]$, we assign the eigenvectors along a family of lines $\gamma_{\kappa_1} : [-\frac{1}{2}, \frac{1}{2}] \rightarrow D^*$ in Figure 1 defined by

$$\gamma_{\kappa_1}(\kappa_2) = \kappa_1 \mathbf{b}_1 + \kappa_2 \mathbf{b}_2, \quad (175)$$

where the starting points $\gamma_{\kappa_1}(-1/2) = \kappa_1 \mathbf{b}_1 - \frac{1}{2} \mathbf{b}_2$ for $\kappa_1 \in [-\frac{1}{2}, \frac{1}{2}]$ are on the line γ_0 in (163).

Similar to the first stage, for each $\kappa_1 \in [-\frac{1}{2}, \frac{1}{2}]$, the assignment is done by parallel transporting the eigenvector according to the equation

$$\frac{\partial \mathbf{u}(\mathbf{k}(\kappa_1, \kappa_2))}{\partial \kappa_2} = \frac{\partial P(\mathbf{k}(\kappa_1, \kappa_2))}{\partial \kappa_2} \mathbf{u}(\mathbf{k}(\kappa_1, \kappa_2)), \quad \kappa_2 \in \left[-\frac{1}{2}, \frac{1}{2}\right], \quad (176)$$

with the initial condition

$$\mathbf{u}(\mathbf{k}(\kappa_1, -1/2)) = \tilde{\mathbf{u}}(\mathbf{k}(\kappa_1, -1/2)), \quad (177)$$

where $\tilde{\mathbf{u}}$ is obtained in the first stage in Lemma 6.2. By construction, the initial condition $\tilde{\mathbf{u}}$ and its transpose $\tilde{\mathbf{u}}^*$ is analytic and periodic in $[-\frac{1}{2}, \frac{1}{2}]$. By the periodicity of $\frac{\partial}{\partial \kappa_2} P$ and $\frac{\partial}{\partial \kappa_2} P = \frac{\partial}{\partial \kappa_2} P^*$, we apply Lemma 3.9 to (176) with the initial conditions (177) to conclude the following lemma.

Lemma 6.4. *The solution $\mathbf{u}(\mathbf{k}(\kappa_1, \kappa_2))$ to (176) with the initial condition (177) and its transpose $\mathbf{u}^*(\mathbf{k}(\kappa_1, \kappa_2))$ are analytic in $\kappa_1, \kappa_2 \in [-\frac{1}{2}, \frac{1}{2}]$. Furthermore, they are periodic in κ_1 for any $\kappa_2 \in [-\frac{1}{2}, \frac{1}{2}]$:*

$$\begin{aligned} \mathbf{u}(\mathbf{k}(\kappa_1, \kappa_2)) &= \mathbf{u}(\mathbf{k}(\kappa_1 + n, \kappa_2)), \quad n \in \mathbb{Z}, \\ \mathbf{u}^*(\mathbf{k}(\kappa_1, \kappa_2)) &= \mathbf{u}^*(\mathbf{k}(\kappa_1 + n, \kappa_2)), \quad n \in \mathbb{Z}. \end{aligned} \quad (178)$$

Lemma 6.4 shows that the only task remains is to make $\mathbf{u}(\mathbf{k}(\kappa_1, \kappa_2))$ also periodic in κ_2 so that $\mathbf{u}(\mathbf{k}(\kappa_1, \kappa_2))$ is analytic and periodic in $(\kappa_1, \kappa_2) \in T$. To do so, we repeat the same procedure in the first stage along each curve γ_{κ_1} for $\kappa_1 \in [-\frac{1}{2}, \frac{1}{2}]$. First, we apply a similar argument for obtaining (168): since $\mathbf{u}(\mathbf{k}(\kappa_1, -1/2))$ and $\mathbf{u}(\mathbf{k}(\kappa_1, 1/2))$ are eigenvectors of the same eigenvalue due to periodicity, they only differ by a κ_1 -dependent phase factor $z : [-\frac{1}{2}, \frac{1}{2}] \rightarrow \mathbb{C}$, where $z(\kappa_1)$ is on the unit circle in the complex plane, i.e. $|z(\kappa_1)| = 1$. More explicitly, we have

$$\mathbf{u}(\mathbf{k}(\kappa_1, 1/2)) = z(\kappa_1) \mathbf{u}(\mathbf{k}(\kappa_1, -1/2)), \quad \kappa_1 \in \left[-\frac{1}{2}, \frac{1}{2}\right], \quad (179)$$

where $z(\kappa_1)$ can be computed via

$$z(\kappa_1) = \mathbf{u}^*(\mathbf{k}(\kappa_1, -1/2)) \mathbf{u}(\mathbf{k}(\kappa_1, 1/2)). \quad (180)$$

By Lemma 6.4 and (180), the analyticity and periodicity of $\mathbf{u}$ and $\mathbf{u}^*$ imply that z is also analytic and periodic in $[-\frac{1}{2}, \frac{1}{2}]$.

Naturally, we wish to repeat the procedure in (170) to make $\mathbf{u}(\mathbf{k}(\kappa_1, \kappa_2))$ an analytic and periodic function in κ_2 . This involves computing

$$\varphi_2(\kappa_1) = -i \log(z(\kappa_1)), \quad (181)$$

by which the following gauge transformation needs to be applied:

$$\tilde{\mathbf{u}}(\mathbf{k}(\kappa_1, \kappa_2)) = e^{-i\varphi_2(\kappa_1)(\kappa_2 + \frac{1}{2})} \mathbf{u}(\mathbf{k}(\kappa_1, \kappa_2)), \quad \kappa_1, \kappa_2 \in \left[-\frac{1}{2}, \frac{1}{2}\right]. \quad (182)$$

Using similar arguments for (171), (173) and (174), we conclude that $\tilde{\mathbf{u}}(\mathbf{k}(\kappa_1, \kappa_2))$ is analytic and periodic in $\kappa_2 \in [-\frac{1}{2}, \frac{1}{2}]$. However, since φ_2 is κ_1 -dependent, the transformed vector $\tilde{\mathbf{u}}$ is not necessarily analytic and periodic in κ_1 unless φ_2 is. Although z is analytic and periodic (see (180)), if the trajectory of z winds around the origin in the complex plane before returning to its starting point, the phase φ_2 is at best discontinuous as a periodic function in $[-\frac{1}{2}, \frac{1}{2}]$, since the log function has no continuous branch that includes the entire unit circle. To understand when this happens, we next establish the relation between the first Chern number C_1 (see (132)) and the winding number of the path of z around the origin. The relation is well-known and can be found in [13], for example.

We compute the Berry connection A_1 and A_2 on the boundary of T and use (137) to compute C_1 , where the following quantities are required:

$$A_2(\mathbf{k}(-1/2, \kappa_2)) - A_2(\mathbf{k}(1/2, \kappa_2)), \quad \kappa_2 \in \left[-\frac{1}{2}, \frac{1}{2}\right], \quad (183)$$

$$A_1(\mathbf{k}(\kappa_1, -1/2)) - A_1(\mathbf{k}(\kappa_1, 1/2)), \quad \kappa_1 \in \left[-\frac{1}{2}, \frac{1}{2}\right]. \quad (184)$$

By (116)), we compute the Berry connection $A_i = i\mathbf{u}^* \frac{\partial}{\partial \kappa_i} \mathbf{u}$ for $i = 1, 2$. Combining it with (176) and (108), we conclude that A_2 is identically zero in T , so the term in (183) is zero. Next, we show that (179) gives quantity in (184). By differentiating (179) and multiplying it by $\mathbf{u}^*(\mathbf{k}(\kappa_1, 1/2))$, we obtain

$$A_1(\mathbf{k}(\kappa_1, 1/2)) = i\mathbf{u}^*(\mathbf{k}(\kappa_1, 1/2)) \frac{\partial}{\partial \kappa_1} \mathbf{u}(\mathbf{k}(\kappa_1, 1/2)) \quad (185)$$

$$= i \frac{z'(\kappa_1)}{z(\kappa_1)} + A_1(\mathbf{k}(\kappa_1, -1/2)), \quad (186)$$

where we have used (179) and $\bar{z} = z^{-1}$ to obtain (186). Combining the fact that $A_2 = 0$ and (186), we apply (137) to obtain the following formula

$$C_1 = \frac{1}{2\pi i} \int_{-\frac{1}{2}}^{\frac{1}{2}} d\kappa_1 \frac{z'(\kappa_1)}{z(\kappa_1)} = \frac{1}{2\pi i} \int_{\Gamma} dz \frac{1}{z}, \quad (187)$$

where the second equality is obtained by a simple change of variable and Γ is the trajectory of $z(\kappa_1)$ on the unit circle as κ_1 goes from $-\frac{1}{2}$ to $\frac{1}{2}$. By the residue theorem, (187) shows that the first Chern number C_1 is the winding number of Γ around the origin as κ_1 goes from $-\frac{1}{2}$ to $\frac{1}{2}$. This fact is useful for two reasons. First, since z is analytic and periodic, (187) can be evaluated to obtain C_1 numerically in a reliable and efficient manner. Second, and more importantly, when $C_1 \neq 0$, z will at least go once around the unit circle before returning to its starting point so that φ_2 in (181) is discontinuous as a periodic function in $[-\frac{1}{2}, \frac{1}{2}]$, which is the manifestation of the topological obstruction in Theorem 5.3. Conversely, when $C_1 = 0$, z returns to the starting point without going around the origin. This means that, by choosing suitable branches if necessary, $\log$ over Γ can be chosen to be an analytic function, so that φ_2 in (181) is analytic and periodic. Thus we have the following lemma.

Lemma 6.5. *The function φ_2 in (181) is an analytic and periodic function in $[-\frac{1}{2}, \frac{1}{2}]$ if and only if the first Chern number C_1 is zero. Furthermore, φ_2 is at best a discontinuous function if C_1 is nonzero.*

We apply the gauge transform in (182) to obtain $\tilde{\mathbf{u}}$. If $C_1 = 0$, by Lemma 6.5, we conclude that $\tilde{\mathbf{u}}$ is analytic and Λ^* -periodic in D^* . Moreover, for each $\kappa_1 \in [-\frac{1}{2}, \frac{1}{2}]$, $\tilde{\mathbf{u}}$ satisfies the new differential equation, obtained from (176) by the transform (182):

$$\frac{\partial \tilde{\mathbf{u}}(\mathbf{k}(\kappa_1, \kappa_2))}{\partial \kappa_2} = \frac{\partial P(\mathbf{k}(\kappa_1, \kappa_2))}{\partial \kappa_2} \tilde{\mathbf{u}}(\mathbf{k}(\kappa_1, \kappa_2)) - i\varphi_2(\kappa_1) \tilde{\mathbf{u}}(\mathbf{k}(\kappa_1, \kappa_2)), \quad (188)$$

subject to the same initial condition (177). We summarize that results in Stage 2 in the following theorem, which is the main tool for computing exponentially localized Wannier functions.

Theorem 6.6. *Suppose that $C_1 = 0$. Then $\tilde{\mathbf{u}}$ in (182) is analytic and Λ^* -periodic in D^* . Thus the Wannier function corresponding to $\tilde{\mathbf{u}}$ is exponentially localized.*

When $C_1 \neq 0$, the assignment will be at best discontinuous and we terminate the construction. For the remainder of the construction, we assume that $C_1 = 0$ so that $\tilde{\mathbf{u}}$ is analytic and Λ^* -periodic in D^* .

It should be observed that the integrability condition (124) in Theorem 5.1 is not explicitly used in the above construction; there is no contradiction since the paths in the construction that cover D^* , namely γ_0 in (163) and γ_{κ_1} in (175), do not intersect at any point. Although the Berry connection A_2 corresponding to $\tilde{\mathbf{u}}$ is given by the formula (due to (182) and (176))

$$A_2(\mathbf{k}(\kappa_1, \kappa_2)) = \varphi_2(\kappa_1), \quad \kappa_1, \kappa_2 \in \left[-\frac{1}{2}, \frac{1}{2}\right], \quad (189)$$

information about A_1 of $\tilde{\mathbf{u}}$ is not directly available from the construction (except for the value at $\kappa_1 \in [-\frac{1}{2}, \frac{1}{2}]$ and $\kappa_2 = -1/2$ by (172), which is in term transformed based on (182)). Nonetheless, by Theorem 5.1, A_1 is automatically specified to satisfy the integrability condition (124).

Due to the lack of control over A_1 in the construction, the constructed $\tilde{\mathbf{u}}$ is not guaranteed to be optimal as stated in Theorem 5.8. In other words, although the Wannier function corresponding to $\tilde{\mathbf{u}}$ is already exponentially localized, its variance of its Fourier coefficients can still be further reduced.

Despite the fact the constructed $\tilde{\mathbf{u}}$ is not optimal, we will prove in Theorem 6.11 in Section 6.4 that, if H has time-reversal symmetry (see (27)), for any $\mathbf{k} \in D^*$, we will have

$$\widetilde{\tilde{\mathbf{u}}}(\mathbf{k}) = \tilde{\mathbf{u}}(-\mathbf{k}), \quad (190)$$

if $\mathbf{v}^0$ in the initial condition (166) is chosen to be real (which is possible again due to time-reversal symmetry). When D^* is parametrized by $(\kappa_1, \kappa_2) \in T$, this is equivalent to

$$\widetilde{\tilde{\mathbf{u}}}(\mathbf{k}(\kappa_1, \kappa_2)) = \tilde{\mathbf{u}}(\mathbf{k}(-\kappa_1, -\kappa_2)). \quad (191)$$

The condition (190) implies that the Fourier coefficients of $\tilde{\mathbf{u}}$ are real, thus the Wannier function defined by (33) is also real.

In the next stage, we turn $\tilde{\mathbf{u}}$ into the globally optimal assignment.

Remark 6.7. *The assignment constructed along γ_0 and γ_{κ_1} for each κ_1 corresponds to the so-called hybrid Wannier functions [17], lower dimensional “slices” that make up the higher dimensional ones. According to the analysis in [12] (that can be easily adapted to the matrix models here), the every assignment along those lines is globally optimal. Furthermore, the integrand $\frac{z'(\kappa_1)}{z(\kappa_1)}$ in (187) corresponds to the Wannier center of the hybrid Wannier function. It is well-known that the changes in the Wannier centers contain the topological information of the band [13].*

Remark 6.8. *Due to the global optimality of one-dimensional assignments along γ_0 and γ_{κ_1} for each κ_1 (see Remark 6.7), it is well-known that, when the Berry curvature Ω_{xy} of the band (see (125)) is zero, the global optimality of one-dimensional assignments results in that of the higher-dimensional assignment [18]. Thus, when Ω_{xy} is identically zero, the steps in Section 6.3 are not needed.*

6.3 Stage 3: eliminating the divergence of the Berry connection

Assuming $C_1 = 0$, according to Theorem 6.6 and (190), we have already obtained an analytic and Λ^* -periodic assignment $\tilde{\mathbf{u}}$. Hence, all results in Section 5.5 are applicable. By Theorem 5.8, it remains to eliminate the divergence (curl-free component) of the Berry connection corresponding to $\tilde{\mathbf{u}}$ to achieve the globally optimal assignment. In this section, we describe the steps for achieving this goal.

It should be observed that the sole purpose of Stage 3 is to modify $\tilde{\mathbf{u}}$ (by a gauge transformation) to achieve the minimum variance in terms of its Fourier coefficients. Hence, it is

optional if the goal is only to construct an exponentially localized Wannier function. (The Wannier function is also real by (190) if the matrix H has time-reversal symmetry.)

To eliminate the divergence of the Berry connection of $\tilde{\mathbf{u}}$, we first compute the Berry connection in the $\mathbf{b}_1, \mathbf{b}_2$ direction by the formulas

$$A_1 = i\tilde{\mathbf{u}}^* \frac{\partial}{\partial \kappa_1} \tilde{\mathbf{u}}, \quad A_2 = i\tilde{\mathbf{u}}^* \frac{\partial}{\partial \kappa_2} \tilde{\mathbf{u}}. \quad (192)$$

We use (45) to express the Berry connection in (192) in the $\mathbf{e}_x, \mathbf{e}_y$ basis by the formulas

$$A_x = \frac{1}{2\pi} \mathbf{a}_1 \cdot \mathbf{e}_x A_1 + \frac{1}{2\pi} \mathbf{a}_2 \cdot \mathbf{e}_x A_2, \quad A_y = \frac{1}{2\pi} \mathbf{a}_1 \cdot \mathbf{e}_y A_1 + \frac{1}{2\pi} \mathbf{a}_2 \cdot \mathbf{e}_y A_2. \quad (193)$$

Next, we compute the divergence $\frac{\partial A_x}{\partial k_x} + \frac{\partial A_y}{\partial k_y}$ and denote it by g :

$$g = \frac{\partial A_x}{\partial k_x} + \frac{\partial A_y}{\partial k_y}. \quad (194)$$

By Theorem 3.4 and (65), we compute the potential $\psi : D^* \rightarrow \mathbb{R}$ generating the divergence by solving the following Poisson's equation:

$$\frac{\partial^2 \psi}{\partial^2 k_x} + \frac{\partial^2 \psi}{\partial^2 k_y} = -g, \quad \psi \text{ is } \Lambda^*\text{-periodic in } D^*. \quad (195)$$

Since A_1 and A_2 are analytic and Λ^* -periodic in D^* by construction, so are A_x and A_y by (193). Hence, due to (50), we always have

$$\int_{D^*} g(\mathbf{k}) d\mathbf{k} = \int_{D^*} \left(\frac{\partial A_x(\mathbf{k})}{\partial k_x} + \frac{\partial A_y(\mathbf{k})}{\partial k_y} \right) d\mathbf{k} = 0. \quad (196)$$

By applying Theorem 3.3, we conclude that (259) is always solvable. Moreover, the solution ψ is analytic and Λ^* -periodic in D^* , whose Fourier series is given by the formula

$$\psi(\mathbf{k}) = \sum_{\substack{\mathbf{R} \in \Lambda \\ \mathbf{R} \neq \mathbf{0}}} \frac{g_{\mathbf{R}}}{\|\mathbf{R}\|^2} \cdot e^{i\mathbf{R} \cdot \mathbf{k}}, \quad \mathbf{k} \in D^*, \quad (197)$$

where $g_{\mathbf{R}}$ is the Fourier coefficient of g at $\mathbf{R} \in \Lambda$ given by the formula

$$g_{\mathbf{R}} = \frac{V_{\text{puc}}}{(2\pi)^2} \int_{D^*} d\mathbf{k} e^{-i\mathbf{R} \cdot \mathbf{k}} \cdot g(\mathbf{k}). \quad (198)$$

Final, we apply the following gauge transformation

$$\tilde{\tilde{\mathbf{u}}}(\mathbf{k}) = e^{-i\psi(\mathbf{k})} \tilde{\mathbf{u}}(\mathbf{k}), \quad \mathbf{k} \in D^*, \quad (199)$$

that eliminates the divergence of the Berry connection (A_x, A_y) by Theorem 5.8. Obviously, the new $\tilde{\tilde{\mathbf{u}}}$ is also analytic and Λ^* -periodic in D^* . Thus we conclude that $\tilde{\tilde{\mathbf{u}}}$ is the globally optimal assignment (up to a lattice vector in Λ), whose Wannier function's center and variance are given by (157) and (158) respectively.

Suppose that $\tilde{\mathbf{u}}$ satisfies the symmetry in (190). By combining (192), (193) with the definition of g in (194), we have

$$g(\mathbf{k}) = -g(-\mathbf{k}), \quad \mathbf{k} \in D^*. \quad (200)$$

Combining (197) and (200) shows that

$$\psi(\mathbf{k}) = -\psi(-\mathbf{k}), \quad \mathbf{k} \in D^*. \quad (201)$$

Combining (199), (201) and (190), we obtain the symmetry relation

$$\tilde{\tilde{\mathbf{u}}}(\mathbf{k}) = \tilde{\tilde{\mathbf{u}}}(-\mathbf{k}), \quad \mathbf{k} \in D^*. \quad (202)$$

As a result, the Fourier coefficients of $\tilde{\tilde{\mathbf{u}}}$ are also real, so is its corresponding Wannier function by (36).

This completes the construction of the assignment as stated in Section 2. We conclude that the Wannier function defined by (36) corresponding to the assignment $\tilde{\tilde{\mathbf{u}}}$ is exponentially localized and has the optimal variance given in (158). Furthermore, it is also real if the matrix H has time-reversal symmetry.

6.4 Realty of Wannier functions

In this section, we prove that, when the matrix H has time-reversal symmetry (see 27), the assignment $\tilde{\mathbf{u}}$ obtained in Theorem 6.6 can be chosen to satisfy (191). Consequently, its Fourier coefficients are real, so is the corresponding Wannier function (36). Such a choice can be done by choosing the initial vector $\mathbf{v}^0$ in (164) to be real. First, we show that such a choice for $\mathbf{v}^0$ is always possible.

Lemma 6.9. *Suppose the matrix H has time-reversal symmetry (27). Then $\mathbf{v}^0$ in (164) can be chosen to be a real vector.*

Proof. By periodicity and the symmetry (27) of H , we have

$$\overline{H}(\mathbf{k}(-1/2, -1/2)) = H(\mathbf{k}(1/2, 1/2)) = H(\mathbf{k}(-1/2, -1/2)). \quad (203)$$

By complex conjugating (164), the relations in (203) show that both $\mathbf{v}^0$ and $\overline{\mathbf{v}}^0$ satisfy the eigenvalue equation (164). Since the eigenvalue is non-degenerate, $\mathbf{v}^0$ and $\overline{\mathbf{v}}^0$ can only differ by a phase factor, i.e. $\overline{\mathbf{v}}^0 = e^{2i\varphi_0} \mathbf{v}^0$, for some real φ_0 . If $\mathbf{v}^0$ is not real, $\pm e^{i\varphi_0} \mathbf{v}^0$ will be real. $\square$

Next, we prove the fact that φ_2 in (181) is an even function when time-reversal symmetry is present, which will be used in proving the reality of Wannier functions.

Lemma 6.10. *Suppose that H has time-reversal symmetry. Then $\varphi_2 : [-\frac{1}{2}, \frac{1}{2}] \rightarrow \mathbb{R}$ in (181) is even, i.e. $\varphi_2(\kappa_1) = \varphi_2(-\kappa_1)$.*

Proof. We integrate (186) from $-1/2$ to κ_1 for $\kappa_1 \in [-\frac{1}{2}, \frac{1}{2}]$. By applying Green's theorem over $[-\frac{1}{2}, \kappa_1] \times [-\frac{1}{2}, \frac{1}{2}] \subseteq T$ as in Section 5.4, we obtain

$$\varphi_2(\kappa_1) = -i \log z(\kappa_1) = \int_{-\frac{1}{2}}^{\kappa_1} ds \int_{-\frac{1}{2}}^{\frac{1}{2}} d\kappa_2 \Omega_{12}(\mathbf{k}(s, \kappa_2)). \quad (204)$$

(For $\kappa_1 = 1/2$, we obtain (187).) By Remark 5.4, time-reversal symmetry implies the first Chern number $C_1 = 0$. Due to (135), $C_1 = 0$ implies that

$$\int_{-\frac{1}{2}}^{\kappa_1} ds \int_{-\frac{1}{2}}^{\frac{1}{2}} d\kappa_2 \Omega_{12}(\mathbf{k}(s, \kappa_2)) = - \int_{\kappa_1}^{\frac{1}{2}} ds \int_{-\frac{1}{2}}^{\frac{1}{2}} d\kappa_2 \Omega_{12}(\mathbf{k}(s, \kappa_2)). \quad (205)$$

We replace κ_1 with $-\kappa_1$ in (204) and apply (205) to obtain

$$\varphi_2(-\kappa_1) = - \int_{-\kappa_1}^{\frac{1}{2}} ds \int_{-\frac{1}{2}}^{\frac{1}{2}} d\kappa_2 \Omega_{12}(\mathbf{k}(s, \kappa_2)). \quad (206)$$

We apply a change of variable $s \rightarrow -s$ and $\kappa_2 \rightarrow -\kappa_2$ to (206) and use the symmetry (139) to yield the desired result:

$$\varphi_2(-\kappa_1) = \int_{-\frac{1}{2}}^{\kappa_1} ds \int_{-\frac{1}{2}}^{\frac{1}{2}} d\kappa_2 \Omega_{12}(\mathbf{k}(s, \kappa_2)) = \varphi_2(\kappa_1). \quad (207)$$

□

In the following, we prove the reality of Wannier functions. The idea is to use the time-reversal symmetry to run the construction in Section 6.1 and 6.2 backwards in time κ_1, κ_2 . It can be viewed as an extension of the proof for the reality of Wannier functions in [12] to two dimensions for matrix models.

Theorem 6.11. *Suppose the matrix H has time-reversal symmetry (27) and $\mathbf{v}^0$ is chosen to be real according to Lemma 6.9. Then the assignment $\tilde{\mathbf{u}}$ in Theorem 6.6 satisfies*

$$\tilde{\mathbf{u}}(\mathbf{k}(\kappa_1, \kappa_2)) = \tilde{\mathbf{u}}(\mathbf{k}(-\kappa_1, -\kappa_2)), \quad (\kappa_1, \kappa_2) \in T. \quad (208)$$

Hence, the Fourier coefficients of $\tilde{\mathbf{u}}$ are real, so is its corresponding Wannier function.

Proof. In Section 6.1, $\tilde{\mathbf{u}}$ is constructed by solving (172) with the initial condition (166). We define a new vector $\mathbf{w}_1(\kappa_1, -1/2) = \tilde{\mathbf{u}}(\mathbf{k}(-\kappa_1, -1/2))$. We observe that, due to (28), the periodicity of P and the chain rule, $\mathbf{v}$ satisfies the same equation (172):

$$\frac{\partial \mathbf{w}_1(\kappa_1, -1/2)}{\partial \kappa_1} = \frac{\partial P(\mathbf{k}(\kappa_1, -1/2))}{\partial \kappa_1} \mathbf{w}_1(\kappa_1, -1/2) - i\varphi_1 \mathbf{w}_1(\kappa_1, -1/2) \quad (209)$$

for $\kappa_1 \in [-\frac{1}{2}, \frac{1}{2}]$, subject to the initial condition

$$\mathbf{w}_1(-1/2, -1/2) = \tilde{\mathbf{u}}(\mathbf{k}(1/2, -1/2)) = \tilde{\mathbf{u}}(\mathbf{k}(-1/2, -1/2)) = \bar{\mathbf{v}}^0 = \mathbf{v}^0, \quad (210)$$

where the second equality is due to periodicity (see Lemma 6.2) and the third one is due to the reality of $\mathbf{v}^0$ by assumption. We observe that (209) with the initial condition (210) is identical to (172) with (166). By the uniqueness theorem of initial value problems, we conclude that $\mathbf{w}_1(\kappa_1, -1/2) = \tilde{\mathbf{u}}(\mathbf{k}(-\kappa_1, -1/2))$, which implies that

$$\tilde{\mathbf{u}}(\mathbf{k}(\kappa_1, -1/2)) = \tilde{\mathbf{u}}(\mathbf{k}(-\kappa_1, -1/2)). \quad (211)$$

In Section 6.2, for each $\kappa_1 \in [-\frac{1}{2}, \frac{1}{2}]$, $\tilde{\mathbf{u}}$ in (182) is constructed by solving (188) with the initial condition (177). Similarly, we define a new vector $\mathbf{w}_2(\kappa_1, \kappa_2) = \tilde{\mathbf{u}}(\mathbf{k}(-\kappa_1, -\kappa_2))$. Again, due to (28) and the chain rule, for each $\kappa_1 \in [-\frac{1}{2}, \frac{1}{2}]$, $\mathbf{w}_2(\kappa_1, \kappa_2)$ satisfies the same equation (188)

$$\frac{\partial \mathbf{w}_2(\kappa_1, \kappa_2)}{\partial \kappa_2} = \frac{\partial P(\mathbf{k}(\kappa_1, \kappa_2))}{\partial \kappa_2} \mathbf{w}_2(\kappa_1, \kappa_2) - i\varphi_2(\kappa_1) \mathbf{w}_2(\kappa_1, \kappa_2), \quad (212)$$

where we have used $\varphi_2(\kappa_1) = \varphi_2(-\kappa_1)$ in Lemma 6.10. The initial condition is given by the formula

$$\mathbf{w}_2(\kappa_1, -1/2) = \tilde{\mathbf{u}}(\mathbf{k}(-\kappa_1, 1/2)) = \tilde{\mathbf{u}}(\mathbf{k}(-\kappa_1, -1/2)) = \tilde{\mathbf{u}}(\mathbf{k}(\kappa_1, -1/2)), \quad (213)$$

where the second equality is due to periodicity (see Theorem 6.6) and the third one is due to (211). We observe that (212) with the initial condition (213) is identical to (188) with (177). By the uniqueness theorem of initial value problems again, we conclude that $\mathbf{w}_2(\kappa_1, \kappa_2) = \tilde{\mathbf{u}}(\mathbf{k}(\kappa_1, \kappa_2))$. Complex conjugating it yields the desirable result

$$\tilde{\mathbf{u}}(\mathbf{k}(\kappa_1, \kappa_2)) = \tilde{\mathbf{u}}(\mathbf{k}(-\kappa_1, -\kappa_2)), \quad (\kappa_1, \kappa_2) \in T. \quad (214)$$

□

7 Numerical procedure

In this section, we describe the numerical procedure for constructing the optimally localized Wannier function in Section 6. Section 7.1 and 7.2 correspond to Section 6.1, Section 7.3 and

7.4 correspond to 6.2, and Section 7.5, 7.6 and 7.7 correspond to Section 6.3. The detailed description of the algorithms is in Section 7.8.

The inputs to our method are the primitive lattice vectors $\mathbf{a}_1, \mathbf{a}_2$, which defines the lattice Λ in (5), the reciprocal primitive lattice $\mathbf{b}_1, \mathbf{b}_2$, which defines the reciprocal lattice Λ^* in (7) and the torus D^* in (10), a family of n by n matrix H , and a real number $h = \frac{1}{N}$, where N is an even integer that specifies number of points for discretizing D^* in each dimension. Furthermore, we assume the family of eigenvalues E and eigenvectors $\mathbf{u}$ of interests has been chosen.

Roughly speaking, the numerical procedure involves using a fourth-order Runger-Kutta (RK4) with Richardson extrapolation (see Section 4.3) that achieves $O(h^6)$ global truncation error for solving the parallel transport equation at N equispaced points in each dimension of D^* . Since all quantities of interests by construction are analytic and Λ^* -periodic, their Fourier coefficients are approximated (with exponential convergence) via a discrete Fourier transform introduced in Section 4.2. Subsequently, all differentiation operations for computing Berry connections (and Berry curvatures in Appendix 10.1) and solving Poisson's equations are carried out with the help of their Fourier series by inverting a diagonal matrix; almost no accuracy is lost during these steps. As a result, the accuracy of the procedure in this paper is determined by the convergence of the ODE solver, which is $O(h^6)$ in this case, for sufficiently large N . Moreover, the computation time is also dominated by solving the parallel transport equation. It should be observed that we choose RK4 in this paper for simplicity. The choice is by no means optimal and can be easily replaced by higher-order methods for better convergence, such as spectral deferred schemes [11], or multi-step methods for fewer function evaluations.

It also should be observed that, although we solve the parallel transport equation explicitly in this paper to obtain the band structure and the Wannier function simultaneously, the two parts can be decoupled according to Remark 7.2 after band structures have been obtained. The results produced by this approach are of lower accuracy but it can be done with very little cost. We refer the reader to Section 8.1 for details.

7.1 Discretizing the torus D^*

We parameterize D^* by T in the form of (13) and define $(N+1)^2$ equispaced points $\{(\kappa_1^{(j_1)}, \kappa_2^{(j_2)})\}$ in T by the formula

$$\kappa_1^{(j_1)} = j_1 h, \quad \kappa_2^{(j_2)} = j_2 h, \quad j_1, j_2 = -N/2, -(N/2 - 1), \dots, N/2, \quad (215)$$

According to (13), these points correspond to

$$\mathbf{k}^{(j_1, j_2)} = \kappa_1^{(j_1)} \mathbf{b}_1 + \kappa_2^{(j_2)} \mathbf{b}_2 \quad (216)$$

in D^* . Given any function f defined in D^* , we use the notation $f^{(j_1, j_2)}$ to represent the computed value of f at $\mathbf{k}^{(j_1, j_2)}$ so that we have

$$f^{(j_1, j_2)} \approx f(\mathbf{k}^{(j_1, j_2)}). \quad (217)$$

7.2 Step 1: computing an assignment on the line γ_0

We compute parallel transporting eigenvectors described in Section 6.1 on the line γ_0 in (163). First, we obtain an initial condition by finding the eigenvalue E^0 and eigenvector $\mathbf{v}^0$

$$H(\mathbf{k}(-1/2, -1/2))\mathbf{v}^0 = E^0\mathbf{v}^0. \quad (218)$$

This is done by the standard QR algorithm in $O(n^3)$ operations. When H has time-reversal symmetry, we choose $\mathbf{v}^0$ to be a real vector (see Lemma 6.9) for the reality of the Wannier function by Theorem 6.11. If it is not the case, any choice of $\mathbf{v}^0$ is accepted.

Instead of using (165), we solve its equivalent version stated in (78) and (79). Namely, we solve the following system of ODEs

$$\begin{aligned} \frac{\partial}{\partial \kappa_1} E(\mathbf{k}(\kappa_1, -1/2)) &= \mathbf{u}^*(\mathbf{k}(\kappa_1, -1/2)) \frac{\partial H}{\partial \kappa_1} \mathbf{u}(\mathbf{k}(\kappa_1, -1/2)), \\ \frac{\partial}{\partial \kappa_1} \mathbf{u}(\mathbf{k}(\kappa_1, -1/2)) &= -(H - E)^\dagger \frac{\partial H}{\partial \kappa_1} \mathbf{u}(\mathbf{k}(\kappa_1, -1/2)), \end{aligned} \quad (219)$$

subject to the initial condition computed in (218)

$$E(\mathbf{k}(-1/2, -1/2)) = E^0, \quad \mathbf{u}(\mathbf{k}(-1/2, -1/2)) = \mathbf{v}^0. \quad (220)$$

We solve the system above by RK4 for $\kappa_1 \in [-\frac{1}{2}, \frac{1}{2}]$ with $h = 1/N$, so that we obtain the eigenvalues and eigenvectors

$$E^{(j_1, -N/2)}, \quad \mathbf{u}^{(j_1, -N/2)}, \quad \text{for } j_1 = -N/2, \dots, N/2, j_2 = -N/2. \quad (221)$$

Richardson extrapolation is done by repeating the above steps with h replaced by $h/2$ and $h/4$, followed by taking their differences as in (102).

At each step, the pseudoinverse $(H - E)^\dagger$ is computed by a singular value decomposition of $H - E$ (see Algorithm 1A), discarding the component corresponding to the smallest singular value, and applying the inverse of the decomposition directly to the right-hand side. The details can be found, for example, in [8] and this approach is related to the form in Remark 3.8. At every step, each pseudoinverse computation requires $O(n^3)$ operations, so computing (221) requires $O(n^3 N)$ operations.

Next, we compute the phase φ_1 in (168) by (169), which is given by

$$\varphi_1 = -i \log \left(\mathbf{u}^{*(-N/2, -N/2)} \mathbf{u}^{(N/2, -N/2)} \right). \quad (222)$$

Next, we apply the gauge transform defined by (170) and we have

$$\tilde{\mathbf{u}}^{(j_1, j_2)} = e^{-i\varphi_1(j_1 h + \frac{1}{2})} \mathbf{u}^{(j_1, j_2)} \quad (223)$$

for $j_1 = -N/2, \dots, N/2 - 1$ and $j_2 = -N/2$. We observe that $j_1 = N/2$ is removed since $\tilde{\mathbf{u}}$ is periodic by Lemma 6.2.

7.3 Step 2: computing parallel transport on lines γ_{κ_1}

We compute the parallel transport described in Section 6.1 on the line γ_{κ_1} in (175) for $\kappa_1 \in [-\frac{1}{2}, \frac{1}{2}]$. For each $\kappa_1^{(j_1)} = j_1 h$ with $j_1 = -N/2, \dots, -N/2 - 1$, the parallel transport equations are given by

$$\begin{aligned} \frac{\partial}{\partial \kappa_2} E(\mathbf{k}(\kappa_1^{(j_1)}, \kappa_2)) &= \mathbf{u}^*(\mathbf{k}(\kappa_1^{(j_1)}, \kappa_2)) \frac{\partial H}{\partial \kappa_2} \mathbf{u}(\mathbf{k}(\kappa_1^{(j_1)}, \kappa_2)), \\ \frac{\partial}{\partial \kappa_2} \mathbf{u}(\mathbf{k}(\kappa_1^{(j_1)}, \kappa_2)) &= -(H - E)^\dagger \frac{\partial H}{\partial \kappa_2} \mathbf{u}(\mathbf{k}(\kappa_1^{(j_1)}, \kappa_2)), \end{aligned} \quad (224)$$

subject to the initial condition

$$E(\mathbf{k}(\kappa_1^{(j_1)}, -1/2)) = E^{(j_1, -N/2)}, \quad \mathbf{u}(\mathbf{k}(\kappa_1^{(j_1)}, -1/2)) = \tilde{\mathbf{u}}^{(j_1, -N/2)}, \quad (225)$$

which are outputs from Section 7.2. Similar to solving (221), given each $\kappa_1^{(j_1)}$, the ODE system is solved by RK4 for $\kappa_2 \in [-\frac{1}{2}, \frac{1}{2}]$ with $h = 1/N$. Thus we obtained the computed eigenvalues and eigenvectors

$$E^{(j_1, j_2)}, \quad \mathbf{u}^{(j_1, j_2)}, \quad \text{for } j_1 = -N/2, \dots, N/2 - 1, j_2 = -N/2, \dots, N/2. \quad (226)$$

Again, Richardson extrapolation is done by repeating the above steps with h replaced by $h/2$ and $h/4$, followed by taking their differences as in (102). For each $\kappa_1^{(j_1)}$, the cost of ODE solves is the same as the one in computing (221) and it is repeated N times for different j_1 . Thus the total cost of computing (226) is $O(n^3 N^2)$ operations. This is the dominant cost of the entire construction.

Remark 7.1. We observe that the dependence on the eigenvalue in (219) and (224) can be removed by computing the eigenvalue from the computed eigenvector $\mathbf{u}$ via the Rayleigh quotient

$$E = \frac{\mathbf{u}^* H \mathbf{u}}{\mathbf{u}^* \mathbf{u}}. \quad (227)$$

This also squares error of the computed eigenvector, doubling the convergence rate for E .

7.4 Step 3: determining the topological obstruction

In order to apply (187) to compute the first Chern number C_1 , we determine the value $z^{(j_1)}$ of z in (180) at $\kappa_1 = \kappa_1^{(j_1)}$ with $j_1 = -N/2, \dots, -N/2 - 1$ by the formula

$$z^{(j_1)} = \mathbf{u}^{*(j_1, -N/2)} \mathbf{u}^{(j_1, N/2)}, \quad (228)$$

where the quantities on right are outputs from Section 7.3. Lemma 6.4 implies that z is analytic and periodic in κ_1 . This permits efficient and accurate evaluation of $z'(\kappa_1^{(j_1)})$ by the Fourier series of z . We first compute the Fourier coefficients $\hat{z}_m$ with $m = -N/2, \dots, N/2 - 1$ via a discrete Fourier transform (see Section 4.2 for the two-dimensional version)

$$\hat{z}_m = \frac{1}{N} \sum_{j_1=-N/2}^{N/2-1} e^{-\frac{2\pi i}{N} m j_1} z^{(j_1)}. \quad (229)$$

We differentiate z in the Fourier basis and apply the inverse of the discrete Fourier transform, obtaining an approximation $z'^{(j_1)}$ of z' at $\kappa_1 = \kappa_1^{(j_1)}$ with $j_1 = -N/2, \dots, -N/2 - 1$ by the formula

$$z'^{(j_1)} = \sum_{m=-N/2}^{N/2-1} e^{\frac{2\pi i}{N} m j_1} (2\pi i m \hat{z}_m). \quad (230)$$

Both (229) and (230) are computed via the FFT in $O(N \log N)$ operations.

Then we apply the trapezoidal rule (92) to (187) and round the real part of the sum to the nearest integer to obtain C_1 . More explicitly, we have

$$C_1 = \left\lfloor \operatorname{Re} \left\{ \frac{h}{2\pi i} \sum_{j_1=-N/2}^{N/2-1} \frac{z'^{(j_1)}}{z^{(j_1)}} \right\} \right\rfloor, \quad (231)$$

where $\lfloor \cdot \rfloor$ denotes the function for rounding to the nearest integer. By Theorem 4.1, all approximations above converge exponentially, thus highly accurate for sufficiently large N . Hence, the rounding operation in (231) is guaranteed to produce the correct integer value.

If C_1 is computed and found to be non-zero, we terminate the construction and return $\mathbf{u}^{(j_1, j_2)}$ for $j_1 = -N/2, \dots, N/2 - 1$ and $j_2 = -N/2, \dots, N/2$ computed in Section 7.3. By Lemma 6.4, $\mathbf{u}(\mathbf{k}(\kappa_1, \kappa_2))$ is analytic in both κ_1 and κ_2 . It is also periodic in κ_1 but not in κ_2 . This is the best assignment one can hope for in the presence of the obstruction $C_1 \neq 0$ by Lemma 6.5.

If C_1 is found to be zero, we continue the construction and compute φ_2 by (181). Thus we determine the value $\varphi_2^{(j_1)}$ of φ_2 at $\kappa_1 = \kappa_1^{(j_1)}$ with $j_1 = -N/2, \dots, -N/2 - 1$ by the formula

$$\varphi_2^{(j_1)} = -i \log(z^{(j_1)}), \quad j_1 = -N/2, \dots, N/2 - 1. \quad (232)$$

By Lemma 6.5, the branch for evaluating the log function is chosen so that φ_2 is periodic and analytic. (This may not coincide with the principal branch.) Then we carry out the gauge transform in (182) to obtain the new eigenvectors

$$\tilde{\mathbf{u}}^{(j_1, j_2)} = e^{-i\varphi_2^{(j_1)}(j_2 h + \frac{1}{2})} \mathbf{u}^{(j_1, j_2)} \quad (233)$$

for $j_1 = -N/2, \dots, N/2 - 1$ and $j_2 = -N/2, \dots, N/2 - 1$. The point $j_2 = N/2$ is removed since $\tilde{\mathbf{u}}$ is periodic (and analytic) in κ_2 by Theorem 6.6.

If H has time-reversal symmetry, Theorem 6.11 shows that $\tilde{\mathbf{u}}^{(j_1, j_2)}$ has the approximate symmetry

$$\overline{\tilde{\mathbf{u}}}^{(j_1, j_2)} \approx \tilde{\mathbf{u}}^{(-j_1, -j_2)} \quad (234)$$

for $j_1 = -N/2, \dots, N/2 - 1$ and $j_2 = -N/2, \dots, N/2 - 1$.

Remark 7.2. If two vectors $\mathbf{v}_1 = \mathbf{u}(\mathbf{k})$ and $\mathbf{v}_2(\mathbf{k} + h\mathbf{n})$ for small h in the direction given by $\mathbf{n}$ are obtained by direct eigensolves, the phases of $\mathbf{v}_1$ and $\mathbf{v}_2$ are independent. In [18] and [26], it is pointed out that the (approximate) parallel transported vector $\tilde{\mathbf{v}}_2$ from $\mathbf{v}_1$ given by $\tilde{\mathbf{v}}_2 = e^{-i\beta} \mathbf{v}_2$, where β is the phase given by $\beta = -\text{Im} \log(\mathbf{v}_1^* \mathbf{v}_2)$. Provided the eigenvectors at each grid point (up to any phase) have been obtained, this approach can be used to obtain $\tilde{\mathbf{u}}^{(j_1, j_2)}$ in (233) directly. However, it results in low accurate assignment; we refer the reader to Example 1 in Section 8.1 for more details.

Remark 7.3. It should be observed that the Wannier function corresponding to $\tilde{\mathbf{u}}$ is already exponentially localized. Furthermore, it is also real if H has time-reversal symmetry by (234). The following Section 7.5 and 6.3 only compute the optimal Wannier function by reducing the variance of the Fourier coefficients of $\tilde{\mathbf{u}}$. Hence, they can be skipped for Section 7.7 directly if the optimality of the Wannier function is not needed with all $\tilde{\mathbf{u}}$ replace by $\mathbf{u}$ in (233).

7.5 Step 4: computing the Berry connection

Having computed $\tilde{\mathbf{u}}^{(j_1, j_2)}$ for $j_1, j_2 = -N/2, \dots, N/2 - 1$, we first compute the derivatives

$$\frac{\partial}{\partial \kappa_1} \tilde{\mathbf{u}}^{(j_1, j_2)}, \quad \frac{\partial}{\partial \kappa_2} \tilde{\mathbf{u}}^{(j_1, j_2)} \quad (235)$$

by the Fourier series of $\tilde{\mathbf{u}}$ in order to compute the Berry connection in (192). We denote the components of $\tilde{\mathbf{u}}$ by $\tilde{u}_i$ for $i = 1, 2, \dots, n$. We compute the derivatives above by first computing the approximate Fourier coefficients by a discrete Fourier transform, then differentiating its Fourier series with respect of κ_1 and κ_2 , followed by the inverse transform. More explicitly, by notations in Section 4.2, for each component $i = 1, \dots, n$, we have

$$\frac{\partial}{\partial \kappa_1} \tilde{u}_i^{(j_1, j_2)} = (\mathcal{F}_N^{-1} \mathcal{D}_1 \mathcal{F}_N(\tilde{u}_i))_{j_1 j_2}, \quad \frac{\partial}{\partial \kappa_2} \tilde{u}_i^{(j_1, j_2)} = (\mathcal{F}_N^{-1} \mathcal{D}_2 \mathcal{F}_N(\tilde{u}_i))_{j_1 j_2}, \quad (236)$$

for $j_1, j_2 = -N/2, \dots, N/2 - 1$, where $\mathcal{D}_1$ and $\mathcal{D}_2$ are defined below (99). This involves in total $2n$ FFTs.

By applying (192), we compute the components of the Berry connection parallel to $\mathbf{b}_1$ and $\mathbf{b}_2$ via

$$A_1^{(j_1, j_2)} = \sum_{i=1}^n \bar{u}_i^{(j_1, j_2)} \frac{\partial}{\partial \kappa_1} \tilde{u}_i^{(j_1, j_2)}, \quad A_2^{(j_1, j_2)} = \sum_{i=1}^n \bar{u}_i^{(j_1, j_2)} \frac{\partial}{\partial \kappa_2} \tilde{u}_i^{(j_1, j_2)}, \quad (237)$$

for $j_1, j_2 = -N/2, \dots, N/2 - 1$. By (193), the x, y components of the Berry connection are given by

$$\begin{aligned} A_x^{(j_1, j_2)} &= \frac{1}{2\pi} \mathbf{a}_1 \cdot \mathbf{e}_x A_1^{(j_1, j_2)} + \frac{1}{2\pi} \mathbf{a}_2 \cdot \mathbf{e}_x A_2^{(j_1, j_2)}, \\ A_y^{(j_1, j_2)} &= \frac{1}{2\pi} \mathbf{a}_1 \cdot \mathbf{e}_y A_1^{(j_1, j_2)} + \frac{1}{2\pi} \mathbf{a}_2 \cdot \mathbf{e}_y A_2^{(j_1, j_2)}, \end{aligned} \quad (238)$$

for $j_1, j_2 = -N/2, \dots, N/2 - 1$.

Remark 7.4. We observe that the moment functions given by (41) and (42) for $\tilde{\mathbf{u}}$ can be computed by first computing

$$\begin{aligned} \frac{\partial}{\partial k_x} \tilde{u}_i^{(j_1, j_2)} &= \frac{1}{2\pi} \mathbf{a}_1 \cdot \mathbf{e}_x \frac{\partial}{\partial \kappa_1} \tilde{u}_i^{(j_1, j_2)} + \frac{1}{2\pi} \mathbf{a}_2 \cdot \mathbf{e}_x \frac{\partial}{\partial \kappa_2} \tilde{u}_i^{(j_1, j_2)}, \\ \frac{\partial}{\partial k_y} \tilde{u}_i^{(j_1, j_2)} &= \frac{1}{2\pi} \mathbf{a}_1 \cdot \mathbf{e}_y \frac{\partial}{\partial \kappa_1} \tilde{u}_i^{(j_1, j_2)} + \frac{1}{2\pi} \mathbf{a}_2 \cdot \mathbf{e}_y \frac{\partial}{\partial \kappa_2} \tilde{u}_i^{(j_1, j_2)}, \end{aligned}$$

by results computed in (236) and (45). Then, for the x, y component of (236), we have

$$\begin{aligned} \langle \mathbf{R}_x \rangle &\approx i h^2 \sum_{j_1, j_2 = -N/2}^{N/2-1} \sum_{i=1}^n \bar{u}_i^{(j_1, j_2)} \frac{\partial}{\partial k_x} \tilde{u}_i^{(j_1, j_2)}, \\ \langle \mathbf{R}_y \rangle &\approx h^2 \sum_{j_1, j_2 = -N/2}^{N/2-1} \sum_{i=1}^n \bar{u}_i^{(j_1, j_2)} \frac{\partial}{\partial k_y} \tilde{u}_i^{(j_1, j_2)}, \end{aligned} \quad (239)$$

and

$$\langle \|\mathbf{R}\|^2 \rangle \approx h^2 \sum_{j_1, j_2 = -N/2}^{N/2-1} \sum_{i=1}^n \left(\left| \frac{\partial}{\partial k_x} \tilde{u}_i^{(j_1, j_2)} \right|^2 + \left| \frac{\partial}{\partial k_y} \tilde{u}_i^{(j_1, j_2)} \right|^2 \right). \quad (240)$$

By Theorem 4.1, all trapezoidal rule approximations above converge exponentially as N increases and the accuracy is only limited by how accurate $\tilde{\mathbf{u}}$ is computed.

7.6 Step 5: eliminating the divergence of the Berry connection

Based on the computed Berry connection in (238), we solve the Poisson's equation in (259) by finding the divergence g in (194). Similar to steps in Section 7.5, all derivatives are done in the Fourier domain, where the equation (259) is diagonal, thus easily solved. First, we compute the Fourier coefficients $\hat{A}_x, \hat{A}_y$ of A_x, A_y in (238) by

$$\hat{A}_x = \mathcal{F}_N(A_x), \quad \hat{A}_y = \mathcal{F}_N(A_y). \quad (241)$$

By the change of variable formulas in (45), we compute the Fourier coefficients $\hat{g}$ of g in (194) by the formula

$$\hat{g} = \left(\frac{1}{2\pi} \mathbf{a}_1 \cdot \mathbf{e}_x \mathcal{D}_1 + \frac{1}{2\pi} \mathbf{a}_2 \cdot \mathbf{e}_x \mathcal{D}_2 \right) \hat{A}_x + \left(\frac{1}{2\pi} \mathbf{a}_1 \cdot \mathbf{e}_y \mathcal{D}_1 + \frac{1}{2\pi} \mathbf{a}_2 \cdot \mathbf{e}_y \mathcal{D}_2 \right) \hat{A}_y, \quad (242)$$

where $\mathcal{D}_1$ and $\mathcal{D}_2$ are defined below (99), encoding the actions of $\frac{\partial}{\partial \kappa_1}$ and $\frac{\partial}{\partial \kappa_2}$ respectively. Having computed $\hat{g}$ in (242), we solve the Poisson's equation by (197) to obtain the potential ψ for the divergence of the Berry connection. We first compute the Fourier coefficients $\hat{\psi}$ of ψ by the formula

$$\hat{\psi}_{m_1 m_2} = \begin{cases} 0, & m_1 = m_2 = 0, \\ \frac{\hat{g}_{m_1 m_2}}{\|m_1 \mathbf{a}_1 + m_2 \mathbf{a}_2\|^2}, & m_1 \neq 0 \text{ or } m_2 \neq 0, \end{cases} \quad (243)$$

for $m_1, m_2 = -N/2, \dots, N/2 - 1$. We apply the inverse discrete Fourier transform to $\hat{\psi}$ to obtain

$$\psi^{(j_1, j_2)} = \left(\mathcal{F}_N^{-1}(\hat{\psi}) \right)_{j_1 j_2} \quad (244)$$

for $j_1, j_2 = -N/2, \dots, N/2 - 1$. Thus solving for $\psi^{(j_1, j_2)}$ takes in total 3 FFTs.

Finally, we apply the gauge transform in (199) that eliminates the divergence of the Berry connection of $\tilde{\mathbf{u}}$ to obtain

$$\tilde{\tilde{\mathbf{u}}}^{(j_1, j_2)} = e^{-i\psi^{(j_1, j_2)}} \tilde{\mathbf{u}}^{(j_1, j_2)} \quad (245)$$

for $j_1, j_2 = -N/2, \dots, N/2 - 1$, which is the optimal assignment by Theorem 5.8 .

7.7 Step 6: computing the Wannier function

To compute the Wannier function corresponding to $\tilde{\tilde{\mathbf{u}}}$, we use the definition in (36). We compute the Fourier coefficients $\hat{\tilde{\tilde{u}}}_i$ of $\tilde{\tilde{u}}_i$, the i -th component of $\tilde{\tilde{\mathbf{u}}}$ for $i = 1, \dots, n$ by the formula

$$\hat{\tilde{\tilde{u}}}_i = \mathcal{F}_N(\tilde{\tilde{u}}_i). \quad (246)$$

This takes n FFTs. By (36), we have

$$W_0(\mathbf{r}) \approx \sum_{i=1}^n \sum_{m_1, m_2=-N/2}^{N/2-1} \hat{\tilde{u}}_{i, m_1 m_2} \cdot \phi_i(\mathbf{r} + \mathbf{R}_{m_1 m_2}) \quad (247)$$

where $\mathbf{R}_{m_1 m_2} = m_1 \mathbf{a}_1 + m_2 \mathbf{a}_2$.

7.8 Detailed description of the algorithms

This section contains the detailed description of the algorithms described in Section 7.

Algorithm

Initialization

1. Choose an even integer N and compute $h = 1/N$.
2. Form $\left\{(\kappa_1^{(j_1)}, \kappa_2^{(j_2)})\right\}_{j_1, j_2=-N/2}^{N/2}$ by (215) and $\left\{\mathbf{k}^{(j_1, j_2)}\right\}_{j_1, j_2=-N/2}^{N/2}$ by (216).

Step 1 [Parallel transport on γ_0]

1. Compute E^0 and $\mathbf{v}^0$ in (218) by QR and normalize $\|\mathbf{v}^0\| = 1$.
2. **If** H has time-reversal symmetry **then**
Choose $\mathbf{v}^0$ to be a real vector.
End if
3. Solve (219) by RK4 to compute $\left\{E_{(1)}^{(j_1, -N/2)}\right\}_{j_1=-N/2}^{N/2}$ and $\left\{\mathbf{u}_{(1)}^{(j_1, -N/2)}\right\}_{j_1=-N/2}^{N/2}$,
where Algorithm 1A is at each time step for evaluating $\frac{\partial}{\partial \kappa_1} E$ and $\frac{\partial}{\partial \kappa_1} \mathbf{u}$.
4. Set $h_{(2)} = h/2$ and repeat 1 to obtain $\left\{E_{(2)}^{(j_1, -N/2)}\right\}_{j_1=-N/2}^{N/2}$ and $\left\{\mathbf{u}_{(2)}^{(j_1, -N/2)}\right\}_{j_1=-N/2}^{N/2}$.
5. Set $h_{(3)} = h/4$ and repeat 1 to obtain $\left\{E_{(3)}^{(j_1, -N/2)}\right\}_{j_1=-N/2}^{N/2}$ and $\left\{\mathbf{u}_{(3)}^{(j_1, -N/2)}\right\}_{j_1=-N/2}^{N/2}$.
6. Use outputs from 3 to 5 to perform Richardson extrapolation in (102) to obtain
 $\left\{E_{(1)}^{(j_1, -N/2)}\right\}_{j_1=-N/2}^{N/2}$ and $\left\{\mathbf{u}_{(1)}^{(j_1, -N/2)}\right\}_{j_1=-N/2}^{N/2}$ with $O(h^6)$ error.
7. Compute φ_1 by (222).
8. **Do** $j_1 = -N/2, \dots, N/2 - 1$
Set $\tilde{\mathbf{u}}^{(j_1, -N/2)} = e^{-i\varphi_1(j_1 h + \frac{1}{2})} \mathbf{u}^{(j_1, -N/2)}$.

End do

Step 2 [Parallel transport on γ_{κ_1}]

1. Solve (224) by RK4 to compute $\left\{E_{(1)}^{(j_1, j_2)}\right\}_{j_1, j_2 = -N/2}^{N/2}$ and $\left\{\mathbf{u}_{(1)}^{(j_1, j_2)}\right\}_{j_1, j_2 = -N/2}^{N/2}$,
where Algorithm 1A is at each time step for evaluating $\frac{\partial}{\partial \kappa_1} E$ and $\frac{\partial}{\partial \kappa_1} \mathbf{u}$.
2. Set $h_{(2)} = h/2$ and repeat 1 to obtain $\left\{E_{(2)}^{(j_1, j_2)}\right\}_{j_1, j_2 = -N/2}^{N/2}$ and $\left\{\mathbf{u}_{(2)}^{(j_1, j_2)}\right\}_{j_1, j_2 = -N/2}^{N/2}$.
3. Set $h_{(3)} = h/4$ and repeat 1 to obtain $\left\{E_{(3)}^{(j_1, j_2)}\right\}_{j_1, j_2 = -N/2}^{N/2}$ and $\left\{\mathbf{u}_{(3)}^{(j_1, j_2)}\right\}_{j_1, j_2 = -N/2}^{N/2}$.
4. Use outputs from 1 to 3 to perform Richardson extrapolation in (102) to obtain $\left\{E^{(j_1, j_2)}\right\}_{j_1, j_2 = -N/2}^{N/2}$ and $\left\{\mathbf{u}^{(j_1, j_2)}\right\}_{j_1, j_2 = -N/2}^{N/2}$ with $O(h^6)$ error.

Step 3 [Computing the first Chern number]

1. Compute $\left\{z^{(j_1)}\right\}_{j_1 = -N/2}^{N/2-1}$ by (228).
2. Compute $\left\{\hat{z}_m\right\}_{m = -N/2}^{N/2-1}$ in (229) and $\left\{z'^{(j_1)}\right\}_{j_1 = -N/2}^{N/2-1}$ in (230) by FFT.
3. Compute C_1 by (231).
4. **If** $C_1 \neq 0$ **then**

Return $\left\{E^{(j_1, j_2)}\right\}_{j_1, j_2 = -N/2}^{N/2}$ and $\left\{\mathbf{u}^{(j_1, j_2)}\right\}_{j_1, j_2 = -N/2}^{N/2}$.

Stop[Topological obstruction encountered]

Else if $C_1 = 0$ **then**

5. Compute $\left\{\varphi_2^{(j_1)}\right\}_{j_1 = -N/2}^{N/2-1}$ by (232).

6. **Do** $j_1, j_2 = -N/2, \dots, N/2 - 1$

Set $\tilde{\mathbf{u}}^{(j_1, j_2)} = e^{-i\varphi_2^{(j_1)}(j_2 h + \frac{1}{2})} \mathbf{u}^{(j_1, j_2)}$.

End do

End if

7. **If** Optimal Wannier function is not required **then**

Go to **Step 6**.

End if

Step 4 [Computing the Berry connection]

1. **Do** $i = 1, 2, \dots, n$

Compute $\left\{ \frac{\partial}{\partial \kappa_1} \tilde{u}_i^{(j_1, j_2)} \right\}_{j_1, j_2 = -N/2}^{N/2-1}, \left\{ \frac{\partial}{\partial \kappa_2} \tilde{u}_i^{(j_1, j_2)} \right\}_{j_1, j_2 = -N/2}^{N/2-1}$ in (236) by FFT.

End do

2. Compute $\left\{ A_1^{(j_1, j_2)} \right\}_{j_1, j_2 = -N/2}^{N/2-1}, \left\{ A_2^{(j_1, j_2)} \right\}_{j_1, j_2 = -N/2}^{N/2-1}$ by (237).
3. Compute $\left\{ A_x^{(j_1, j_2)} \right\}_{j_1, j_2 = -N/2}^{N/2-1}, \left\{ A_y^{(j_1, j_2)} \right\}_{j_1, j_2 = -N/2}^{N/2-1}$ by (238).

Step 5 [Eliminating the divergence]

1. Compute $\hat{A}_x, \hat{A}_y$ in (241) by FFT.
2. Compute $\hat{g}$ by (242).
3. Compute $\hat{\psi}$ by (243).
4. Compute $\left\{ \psi^{(j_1, j_2)} \right\}_{j_1, j_2 = -N/2}^{N/2-1}$ in (244) by FFT.
5. **Do** $j_1, j_2 = -N/2, \dots, N/2 - 1$

Set $\tilde{\tilde{\mathbf{u}}}^{(j_1, j_2)} = e^{-i\psi^{(j_1, j_2)}} \tilde{\mathbf{u}}^{(j_1, j_2)}$.

End do

Step 6 [Computing the Wannier function]

(Replace all $\tilde{\tilde{\mathbf{u}}}$ below by $\tilde{\mathbf{u}}$ if from **Step 3**.)

1. **Do** $i = 1, 2, \dots, n$

Compute $\tilde{\tilde{u}}_i$ in (246) by FFT.

End do

2. Compute W_0 by (247).

Return $\left\{ E^{(j_1, j_2)} \right\}_{j_1, j_2 = -N/2}^{N/2-1}, \left\{ \tilde{\tilde{\mathbf{u}}}^{(j_1, j_2)} \right\}_{j_1, j_2 = -N/2}^{N/2-1}$ and W_0 .

Algorithm 1A

Input: $j \in \{1, 2\}$, $\kappa_1, \kappa_2 \in [0, 1]$, $\mathbf{k} = \mathbf{k}(\kappa_1, \kappa_2)$, $H(\mathbf{k})$, $\frac{\partial}{\partial \kappa_j} H(\mathbf{k}) \in \mathbb{C}^{n \times n}$, $\mathbf{u}(\mathbf{k}) \in \mathbb{C}^n$, $E(\mathbf{k}) \in \mathbb{R}$

Output: $\frac{\partial}{\partial \kappa_j} \mathbf{u}(\mathbf{k}) \in \mathbb{C}^n$, $\frac{\partial}{\partial \kappa_j} E(\mathbf{k}) \in \mathbb{R}$

1. Compute $\frac{\partial}{\partial \kappa_j} E(\mathbf{k}) = \mathbf{u}^*(\mathbf{k}) \frac{\partial H(\mathbf{k})}{\partial \kappa_j} \mathbf{u}(\mathbf{k})$.
2. Compute $\mathbf{q} = \frac{\partial H(\mathbf{k})}{\partial \kappa_j} \mathbf{u}(\mathbf{k})$.
3. Compute the singular value decomposition of $H(\mathbf{k}) - E(\mathbf{k}) = \mathbf{U} \mathbf{\Sigma} \mathbf{V}^*$ using the QR algorithm.
4. Set $\frac{\partial}{\partial \kappa_j} \mathbf{u}(\mathbf{k}) = \mathbf{V}_{1:n, 1:n-1} (\mathbf{\Sigma}_{1:n-1, 1:n-1})^{-1} (\mathbf{U}_{1:n, 1:n-1})^* \mathbf{q}$.

8 Numerical results

This section contains numerical results of the algorithm described in Section 7 applied to a 3×3 matrix on a square lattice in Section 8.1 and the Haldane model [14] (2×2 matrix on a hexagonal lattice), where the topologically trivial version ($C_1 = 0$) is in Section 8.2 and the non-trivial version ($C_1 \neq 0$) in Section 8.3.

In all examples, we increase N , the number of discretization in both dimensions (see Section 7.1), while keeping all other parameters the same. We report the accuracy of the eigenvectors computed by numerically solving the parallel transport equation in Step 1–3 in Section 7, measured by the largest difference of the projectors in terms of the Frobenius norm

$$E_{\text{evc}} = \max_{-N/2 \leq i, j \leq N/2} \left\| P_{\text{para}}^{(i, j)} - P_{\text{eig}}^{(i, j)} \right\|_{\text{F}}, \quad (248)$$

where $P_{\text{para}}^{(i, j)}$ and $P_{\text{eig}}^{(i, j)}$ are the projectors formed by eigenvectors computed by solving the ODE and a QR eigenvalue solver, respectively, at (κ_i, κ_j) . Besides, we also report the computed error in the first Chern number by the formula

$$E_{\text{Ch}} = \left| \tilde{C}_1 - C_1 \right|, \quad (249)$$

where C_1 is the exact integer value and $\tilde{C}_1$ is the computed value by (231) without rounding to an integer:

$$\tilde{C}_1 = \text{Re} \left\{ \frac{h}{2\pi i} \sum_{j_1=-N/2}^{N/2-1} \frac{z'(j_1)}{z(j_1)} \right\}. \quad (250)$$

Moreover, we report the largest value of the potential $\tilde{\psi}$ for the divergence (curl-free component) of the Berry connection of $\tilde{\mathbf{u}}^{(j_1, j_2)}$ after Step 5 (Section 7.6) as a measure on its accuracy and optimality. We denote the largest value of the potential $\tilde{\psi}$ for the curl-free component by the formula

$$E_{\text{div}} = \max_{-N/2 \leq i, j \leq N/2-1} \left| \tilde{\psi}^{(i, j)} \right|. \quad (251)$$

Furthermore, for Example 1 and 2, we report the Wannier center (143) and variance (146) (computed according to Remark 7.4) before and after eliminating the divergence, the time t_{para} for solving the parallel transport equation (Step 1–3 in Section 7), and the time t_{div} for eliminating the divergence for achieving the optimal Wannier function (Step 4–5 in Section 7).

In Example 1, we also test the parallel transport scheme in Remark 7.2; instead of solving the ODE explicitly, we replace Step 1–3 in Section 7 with the approach in Remark 7.2. We report the maximum difference E_{para} in term of the phase between the two approaches, given by the formula

$$E_{\text{para}} = \max_{-N/2 \leq i, j \leq N/2-1} \left| \text{Re} \left\{ i \log \left[(\tilde{\mathbf{u}}^{(j_1, j_2)})^* \tilde{\mathbf{u}}_{\text{twist}}^{(j_1, j_2)} \right] \right\} \right|, \quad (252)$$

where $\tilde{\mathbf{u}}^{(j_1, j_2)}$ is obtained by solving the parallel transport equation as described in Section 7 and $\tilde{\mathbf{u}}_{\text{twist}}^{(j_1, j_2)}$ is computed by first computing the eigenvalues and eigenvectors by QR, followed by the alignment scheme in Remark 7.2. For the other two examples, all results are obtained by the approach in Section 7.

All algorithms are implemented in MATLAB R2023b. All timing experiments are performed on a MacBook Pro with a M2 Max CPU.

8.1 Example 1: A 3×3 matrix model

The 3×3 matrix H in this example is taken from Example 2.6 in [16]. The model is defined on a square lattice with the real space lattice spanned by $\mathbf{a}_1 = \frac{a}{\sqrt{2}}(1, -1)$ and $\mathbf{a}_2 = \frac{a}{\sqrt{2}}(1, 1)$, and the reciprocal lattice by $\mathbf{b}_1 = \frac{\sqrt{2}\pi}{a}(1, -1)$ and $\mathbf{b}_2 = \frac{\sqrt{2}\pi}{a}(1, 1)$. We choose the constant $a = 1$ here. The elements in the matrix H are given by the formulas

$$\begin{aligned} H_{11} &= H_{22} = \epsilon_p + 2t_{pp} \left[\cos \left((k_x - k_y) a / \sqrt{2} \right) + \cos \left((k_x + k_y) a / \sqrt{2} \right) \right], \\ H_{33} &= \epsilon_d + 2t_{dd} \left[\cos \left((k_x - k_y) a / \sqrt{2} \right) + \cos \left((k_x + k_y) a / \sqrt{2} \right) \right], \\ H_{13} &= \overline{H}_{31} = -t_{pd} e^{ik_x a / \sqrt{2}} 2i \sin \left(k_x a / \sqrt{2} \right), \\ H_{23} &= \overline{H}_{32} = t_{pd} e^{ik_x a / \sqrt{2}} 2i \sin \left(k_y a / \sqrt{2} \right), \\ H_{12} &= \overline{H}_{21} = 0, \end{aligned} \quad (253)$$

where the constants are chosen as $t_{dd} = 0.1$, $t_{pd} = 2$, $t_{pp} = -0.25$, $t_{pd} = 2$, $\epsilon_d = 1$ and $\epsilon_p = -2$. The eigenvalues parameterized by κ_1 and κ_2 are shown Figure 2(a) and the top non-degenerate band is chosen for the construction. This model has time-reversal symmetry and its constructed φ_2 (see (181)) is shown in Figure 2(b), which is an even function as proved in Lemma 5.7. The first Chern number $C_1 = 0$ and there is no topological obstruction.

Table 1 shows the timings and errors of the approach in Section 7. The computation time t_{para} scales as $O(N^2)$ as discussed in Section 7.3 and the time for eliminating the divergence is

dominated by FFTs. It also shows that the error E_{evec} of computed eigenvectors by solving the parallel transport equation scales as $O(h^6)$ with $h = 1/N$. The error in the divergence E_{div} shows that, for $N \leq 200$, the sampling in D^* is not sufficient to resolve frequency content of the Berry connection. As a result, for achieving 10-digit accuracy, the optimal choice is roughly $N = 200$.

The three components of the assignment $\tilde{\mathbf{u}}$ computed by the approach in Section 7 after Stage 2 (see Section 6.2) is shown in Figure 4 – 6, together with the absolute value of their Fourier coefficients in the log scale with base 10. The Fourier coefficients decay exponentially asymptotically. Their real and imaginary part are even and odd, respectively, under the transform $\kappa_1 \rightarrow -\kappa_1$ and $\kappa_2 \rightarrow -\kappa_2$, as proved in Theorem 6.11. After eliminating the divergence of the Berry connection of $\tilde{\mathbf{u}}$ in Stage 3 in Section 6.3, the third component of $\tilde{\mathbf{u}}$ is shown in Figure 7. The other components are not shown as they are visually very similar to those of $\tilde{\mathbf{u}}$ in Figure 4 – 5. The Berry connection of $\tilde{\mathbf{u}}$ is shown in Figure 8(a), whose Helmholtz-Hodge decomposition is shown in Figure 9. After eliminating its divergence, the Berry connection of $\tilde{\mathbf{u}}$ is shown in Figure 8(b).

In comparison to Table 1, results in Table 3 and Figure 3(a) show that the parallel transport scheme in Remark 7.2 agree with the approach in Section 7 to roughly two digits of accuracy. Figure 3(b) also shows that, although the Fourier coefficients would look visually exponentially localized, they contain high frequency components, an indication that the assigned eigenvectors are inaccurate. The advantage of this approach is that it decouples the computation of eigenvectors and Wannier functions, and it can be done with very computational little cost. It could be a viable approach when low accuracy is required.

Table 2 shows that the Wannier centers and variance of the solution obtained by parallel transport described in Section 7 and the optimal one after the divergence of the Berry connection is eliminated. The Wannier center is not changed and the variance is reduced. The solution from only doing parallel transport, although not optimal, is very close to the optimal one in terms of the variance. Table 4 contains the same results with the parallel transport part done by the scheme in Remark 7.2. Although the assigned eigenvectors are inaccurate (compared with Table 1), the Wannier centers and the optimal variance are remarkably accurate.

N	t_{para} (s)	t_{div} (s)	E_{evec}	E_{Ch}	E_{div}
50	1.76	0.053	4.16e-10	6.84e-10	1.41e-3
100	6.12	0.110	4.18e-10	3.30e-16	3.38e-6
200	23.2	0.305	6.26e-12	2.21e-19	3.07e-11
400	93.5	1.10	9.93e-14	4.49e-18	7.49e-12

Table 1: Timings and errors for Example 1 by the approach in Section 7. The error E_{evec} shows six-order convergence and the error E_{div} indicates if the sampling in D^* is sufficient.

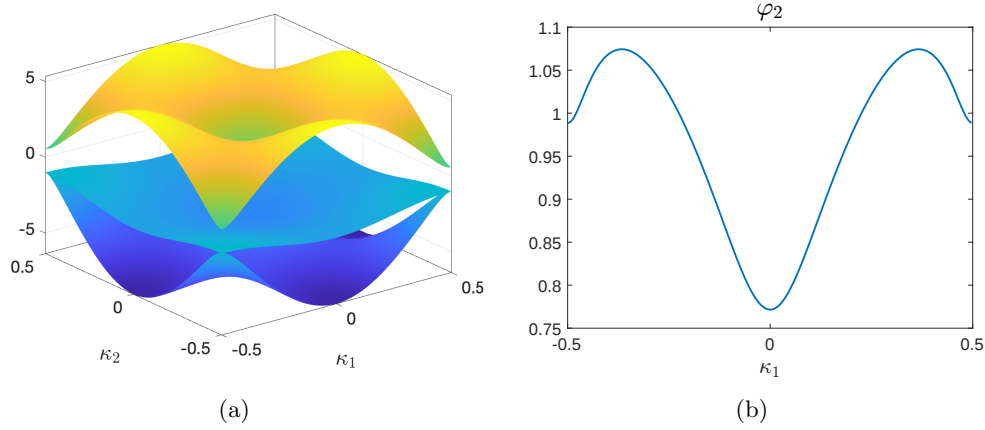


Figure 2: (a) Plot of eigenvalues of H in Example 1. The top band (non-degenerate) is picked. (b) The phase φ_2 in (181) for Example 1. It is an even function as proved in Lemma 5.7.

After Stage 2			After Stage 3 (Optimal solution)		
$\langle \mathbf{R}_x \rangle$	$\langle \mathbf{R}_y \rangle$	$\langle \ \mathbf{R}\ ^2 \rangle - \ \langle \mathbf{R} \rangle\ ^2$	$\langle \mathbf{R}_x \rangle$	$\langle \mathbf{R}_y \rangle$	$\langle \ \mathbf{R}\ ^2 \rangle - \ \langle \mathbf{R} \rangle\ ^2$
-0.217677	-2.67e-15	0.317890	-0.217677	-2.23e-15	0.313797

Table 2: The Wannier center and the variance for Example 1 computed by the approach in Section 7 before and after eliminating the divergence of the Berry connection computed for $N = 400$. The solution after Stage 2 is already close to the optimal one.

N	t_{para} (s)	t_{div} (s)	E_{para}	E_{div}
50	0.118	0.093	9.40e-2	3.55e-2
100	0.190	0.142	8.94e-2	1.38e-2
200	0.411	0.322	8.74e-2	6.86e-3
400	1.51	1.32	8.65e-2	3.44e-3
800	6.18	5.12	8.50e-2	1.72e-3

Table 3: Timings and errors for Example 1 by the parallel transport scheme in Remark 7.2. The error E_{para} barely decreases and E_{div} decreases proportional to $1/N$. However, such a parallel transport scheme is computationally cheap and decouples eigenvector and Wannier function computations.

After Stage 2 (Remark 7.2 approach)			After Stage 3 (Optimal solution)		
$\langle \mathbf{R}_x \rangle$	$\langle \mathbf{R}_y \rangle$	$\langle \ \mathbf{R}\ ^2 \rangle - \ \langle \mathbf{R} \rangle\ ^2$	$\langle \mathbf{R}_x \rangle$	$\langle \mathbf{R}_y \rangle$	$\langle \ \mathbf{R}\ ^2 \rangle - \ \langle \mathbf{R} \rangle\ ^2$
-0.217677	-3.86e-12	0.31(5124)	-0.217677	-3.91e-12	0.313797

Table 4: The Wannier center and the variance for Example 1 before and after eliminating the divergence of the Berry connection computed for $N = 400$ as in Table 2. However, Stage 1 and 2 (Step 1–3 in Section 7) are done by the parallel transport approach described in Remark 7.2. Despite low accuracy of the assigned vectors (see Table 3), a comparison with Table 2 shows that the Wannier centers and the optimal variance are remarkably accurate.

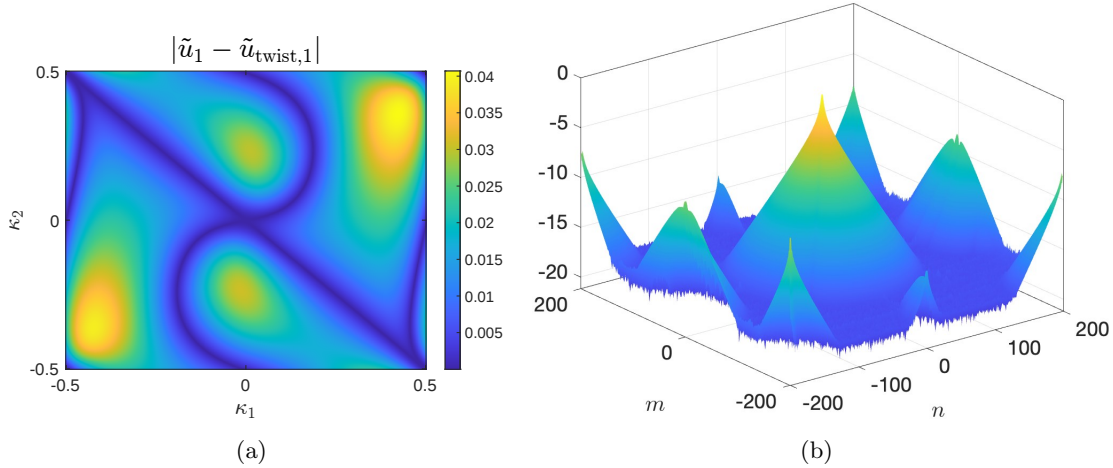


Figure 3: (a) Plot of the difference in absolute value between the first component $\tilde{u}_1$ of the assignment (also see Figure 4) obtained by the parallel transport approach in Section 7 and the first component $\tilde{u}_{\text{twist},1}$ of that obtained by the scheme in Remark 7.2 in Example 1. (b) Plot of the absolute value of the Fourier coefficients of $\tilde{u}_{\text{twist},1}$ in the log scale with base 10 for in Example 1. The high frequency components are an indication of the low accuracy of the assigned vectors.

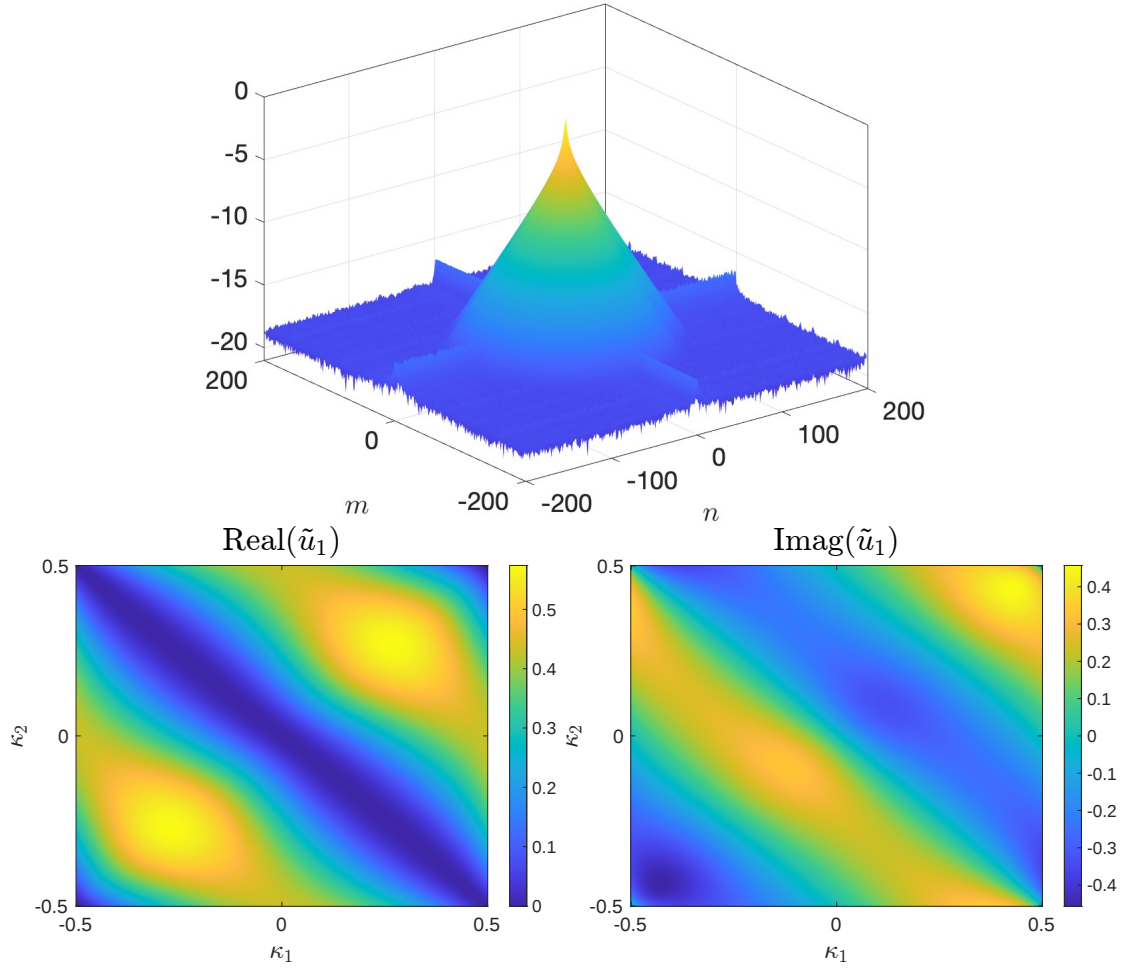


Figure 4: Plot of the real and imaginary part of the component $\tilde{u}_1$ with the absolute value of the Fourier coefficients of $\tilde{u}_1$ in the log scale with base 10 for in Example 1. Compared with Figure 3(b), the coefficients contains no high frequency components.

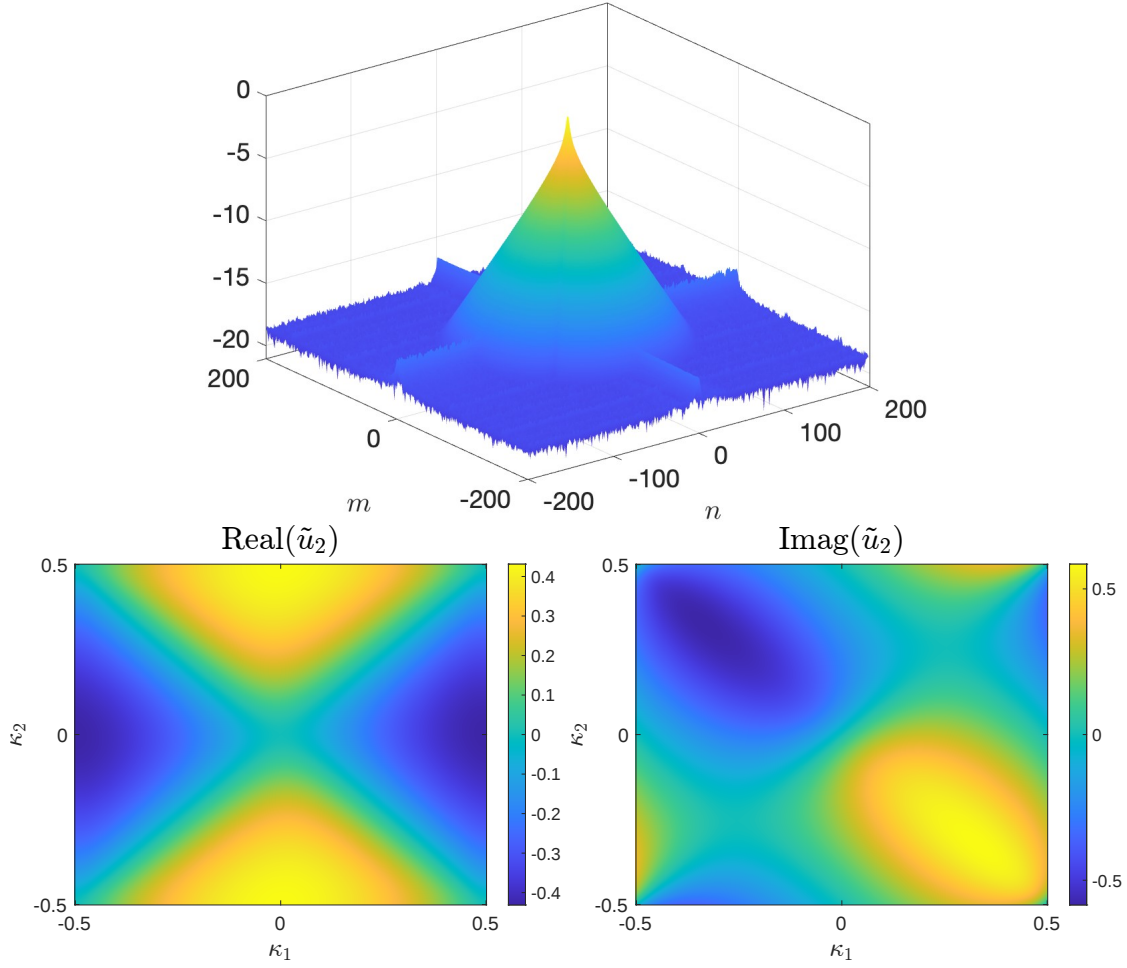


Figure 5: Plot of the real and imaginary part of the component $\tilde{u}_2$ with the absolute value of the Fourier coefficients of $\tilde{u}_2$ in the log scale with base 10 for in Example 1.

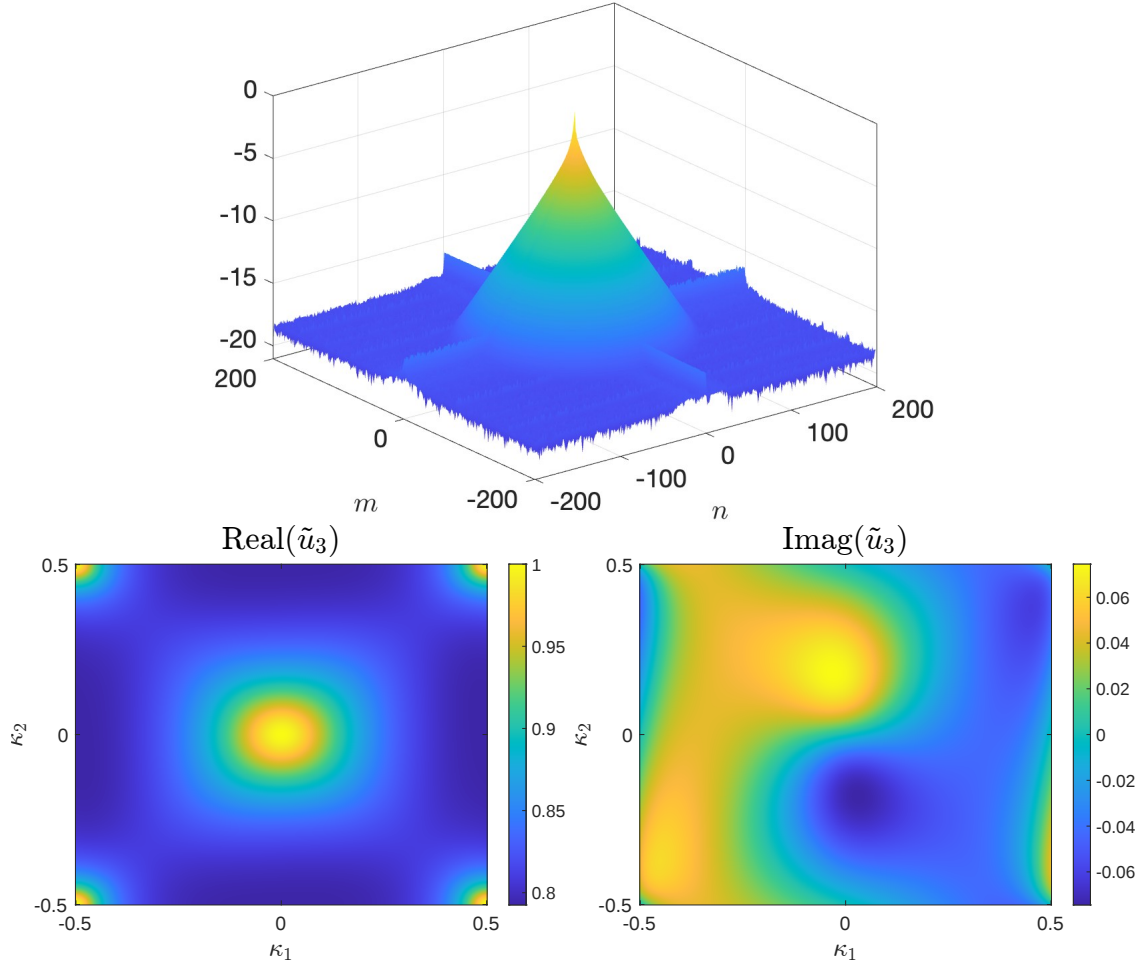


Figure 6: Plot of the real and imaginary part of the component $\tilde{u}_3$ with the absolute value of the Fourier coefficients of $\tilde{u}_3$ in the log scale with base 10 for in Example 1.

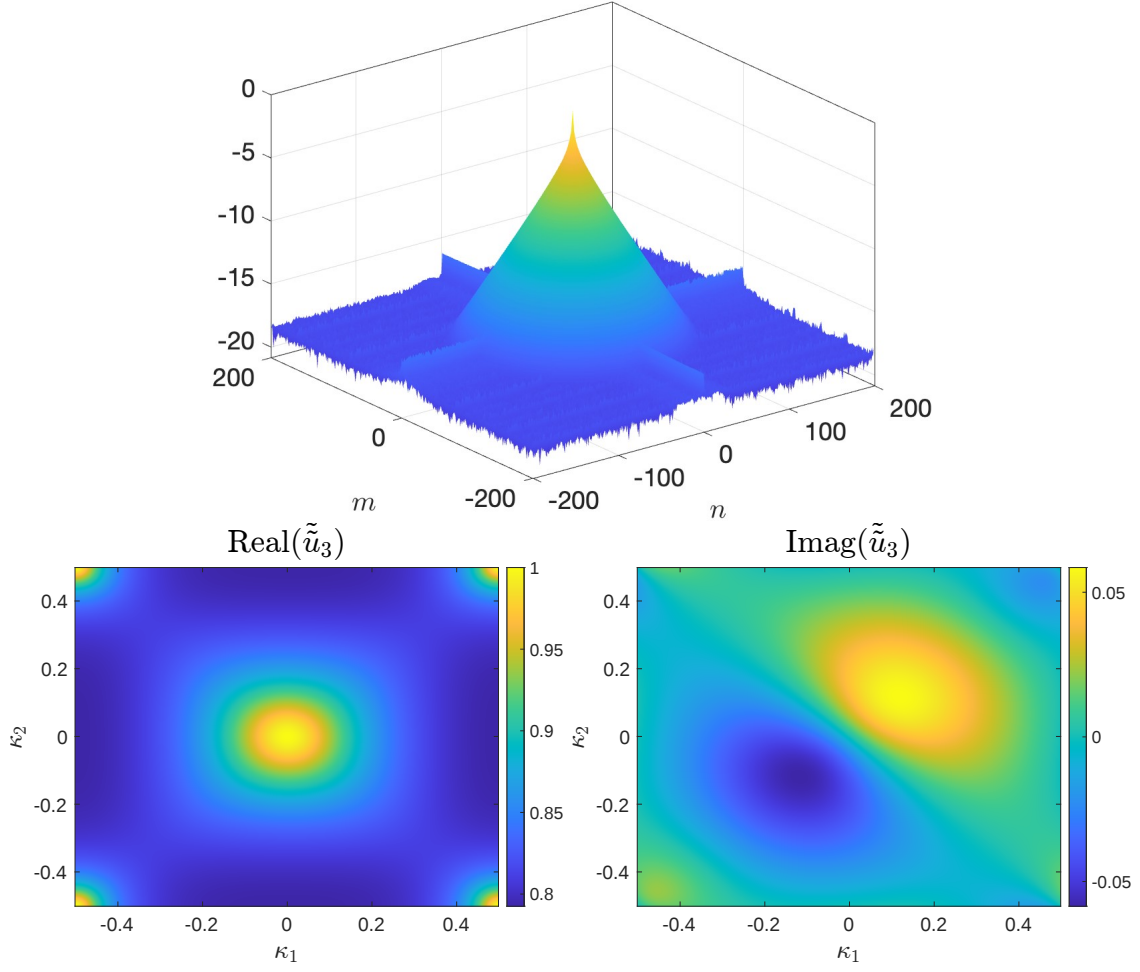


Figure 7: Plot of the same quantities as in Figure 6 after eliminating the divergence of the Berry connection.

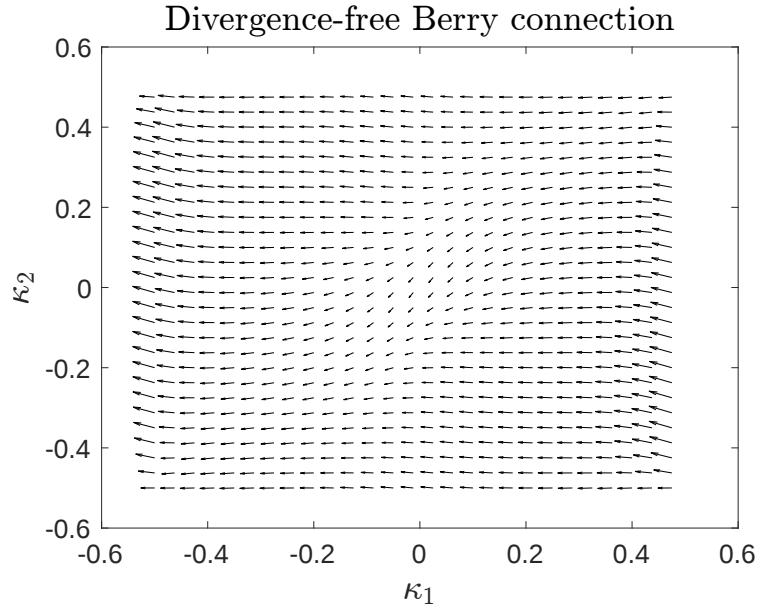
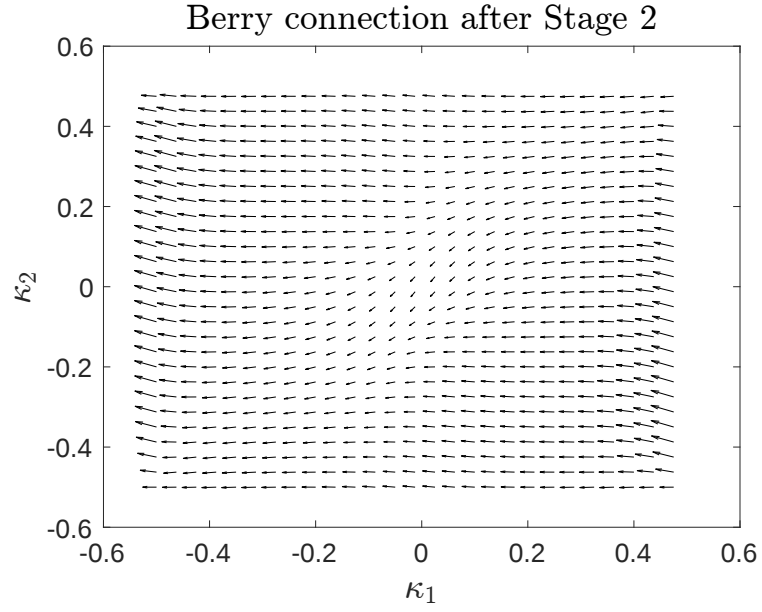


Figure 8: (a) Plot of the Berry connection after Stage 2 for Example 1. (b) Plot of the Berry connection in (a) after eliminating its curl-free component in Stage 3. All vectors shown are in the $\mathbf{e}_x, \mathbf{e}_y$ basis.

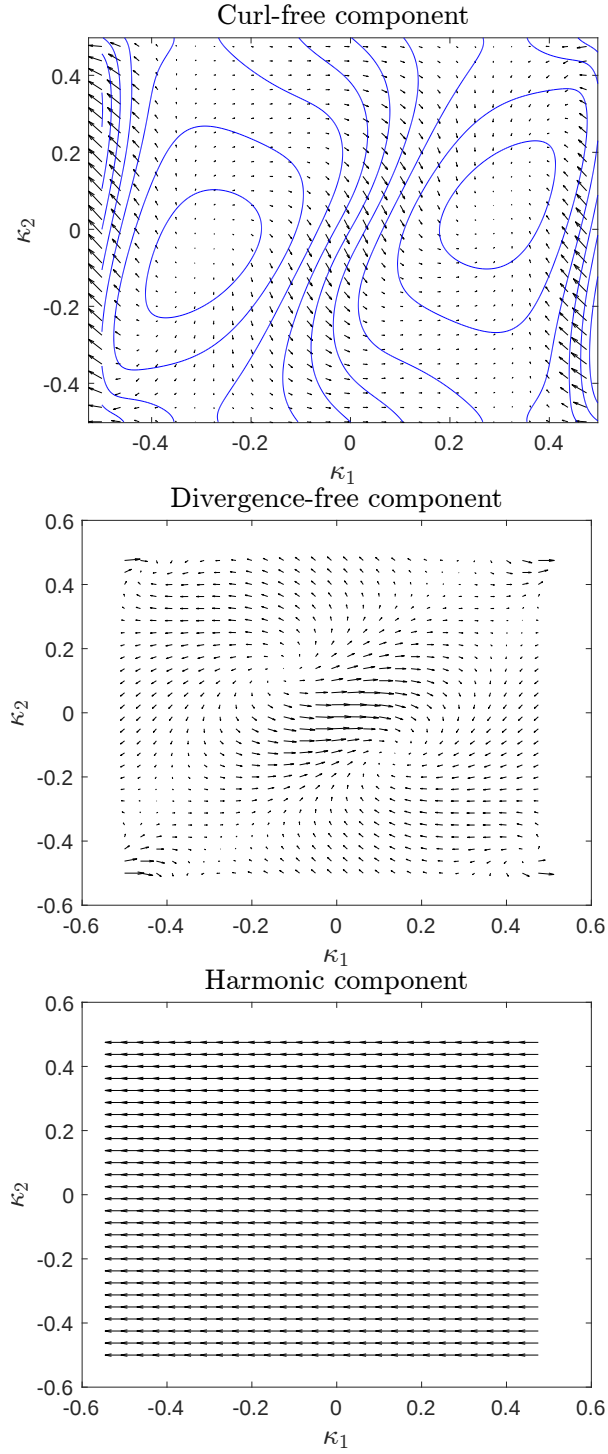


Figure 9: The Helmholtz-Hodge decomposition of the Berry connection in Figure 8(a). All vectors shown are in the $\mathbf{e}_x, \mathbf{e}_y$ basis. The equipotential lines are shown for the curl-free vector field, which is eliminated to obtain the optimal solution in Figure 8(b).

8.2 Example 2: Haldane model (topologically trivial case)

In this example, we consider a topologically trivial version of the Haldane model [14] on a Honeycomb lattice. It is 2×2 matrix model with time-reversal symmetry, thus topologically trivial (i.e. $C_1 = 0$). The real space lattice vectors are given by $\mathbf{a}_1 = \frac{a}{2}(\sqrt{3}, 1)$ and $\mathbf{a}_2 = \frac{a}{2}(\sqrt{3}, -1)$ with the reciprocal lattice vectors $\mathbf{b}_1 = \frac{2\pi}{\sqrt{3}a}(1, \sqrt{3})$ and $\mathbf{b}_2 = \frac{2\pi}{\sqrt{3}a}(1, -\sqrt{3})$. We choose $a = 1$ here. The matrix H is given by

$$H(\mathbf{k}) = \begin{bmatrix} V_0 & t_1(1 + e^{-i\mathbf{k} \cdot \mathbf{a}_1} + e^{-i\mathbf{k} \cdot \mathbf{a}_2}) \\ t_1(1 + e^{i\mathbf{k} \cdot \mathbf{a}_1} + e^{i\mathbf{k} \cdot \mathbf{a}_2}) & -V_0 \end{bmatrix}, \quad (254)$$

where the constants are chosen to be $t_1 = 1$ and $V_0 = 0.5$. The eigenvalues as a functions of the parameterization κ_1 and κ_2 are shown Figure 10(a) and the top band is chosen for the construction. This model has time-reversal symmetry and its constructed φ_2 (see (181)) is shown in Figure 10(b), which is an even function as proved in Lemma 5.7.

Table 5 shows the timings and errors listed at the beginning of Section 8. The overall trends is very similar to those in Example 1. The error in the divergence E_{div} shows that, for $N \leq 200$, the sampling in D^* is not sufficient to resolve frequency content of the Berry connection. As a result, for achieving 10-digit accuracy, the optimal choice is roughly $N = 200$.

Similar to Example 1, Table 6 shows that the Wannier centers and variance of the solution obtained by parallel transport and the optimal one after the divergence of the Berry connection is eliminated. The Wannier center is not changed and the variance is reduced. The solution from parallel transport, although not optimal, is also close to the optimal one in terms of the variance.

The two components of the assignment $\tilde{\mathbf{u}}$ after Stage 2 in Section 6.2 is shown in Figure 11 – 12, together with the absolute value of their Fourier coefficients in the log scale with base 10. The Fourier coefficients decay exponentially asymptotically. Their real and imaginary part are even and odd, respectively, under the transform $\kappa_1 \rightarrow -\kappa_1$ and $\kappa_2 \rightarrow -\kappa_2$, as proved in Theorem 6.11. After eliminating the divergence of the Berry connection of $\tilde{\mathbf{u}}$ in Stage 3 in Section 6.3, the first component of $\tilde{\mathbf{u}}$ is shown in Figure 13. The other components are not shown as they are visually very similar to those of $\tilde{\mathbf{u}}$ in Figure 11. The Berry connection of $\tilde{\mathbf{u}}$ is shown in Figure 14(a), whose Helmholtz-Hodge decomposition is shown in Figure 15. After eliminating its divergence, the Berry connection of $\tilde{\mathbf{u}}$ is shown in Figure 14(b).

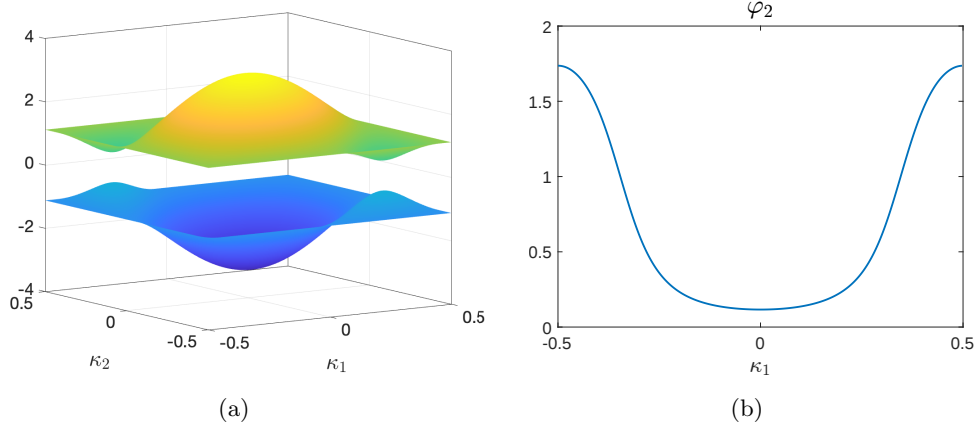


Figure 10: (a) Plot of eigenvalues of H in Example 2. The band on top is picked. (b) The phase φ_2 in (181) for Example 2. It is an even function as proved in Lemma 5.7.

N	t_{para} (s)	t_{div} (s)	E_{evc}	E_{Ch}	E_{div}
50	1.30	0.040	5.66e-10	4.56e-13	1.14e-4
100	4.51	0.091	7.09e-12	1.50e-17	1.68e-9
200	16.6	0.289	1.07e-13	2.93e-17	2.53e-12
400	66.3	1.07	2.10e-14	5.00e-18	1.14e-11

Table 5: Timings and errors for Example 2. The error E_{evc} shows six-order convergence and the error E_{div} indicates if the sampling in D^* is sufficient.

After Stage 2			After Stage 3 (Optimal solution)		
$\langle \mathbf{R}_x \rangle$	$\langle \mathbf{R}_y \rangle$	$\langle \ \mathbf{R}\ ^2 \rangle - \ \langle \mathbf{R} \rangle\ ^2$	$\langle \mathbf{R}_x \rangle$	$\langle \mathbf{R}_y \rangle$	$\langle \ \mathbf{R}\ ^2 \rangle - \ \langle \mathbf{R} \rangle\ ^2$
-0.184913	4.29e-16	0.270171	-0.184913	-2.23e-16	0.233954

Table 6: The Wannier center and variance for Example 2 before and after eliminating the divergence of the Berry connection computed for $N = 400$. The solution after Stage 2 is already relatively close to the optimal one.

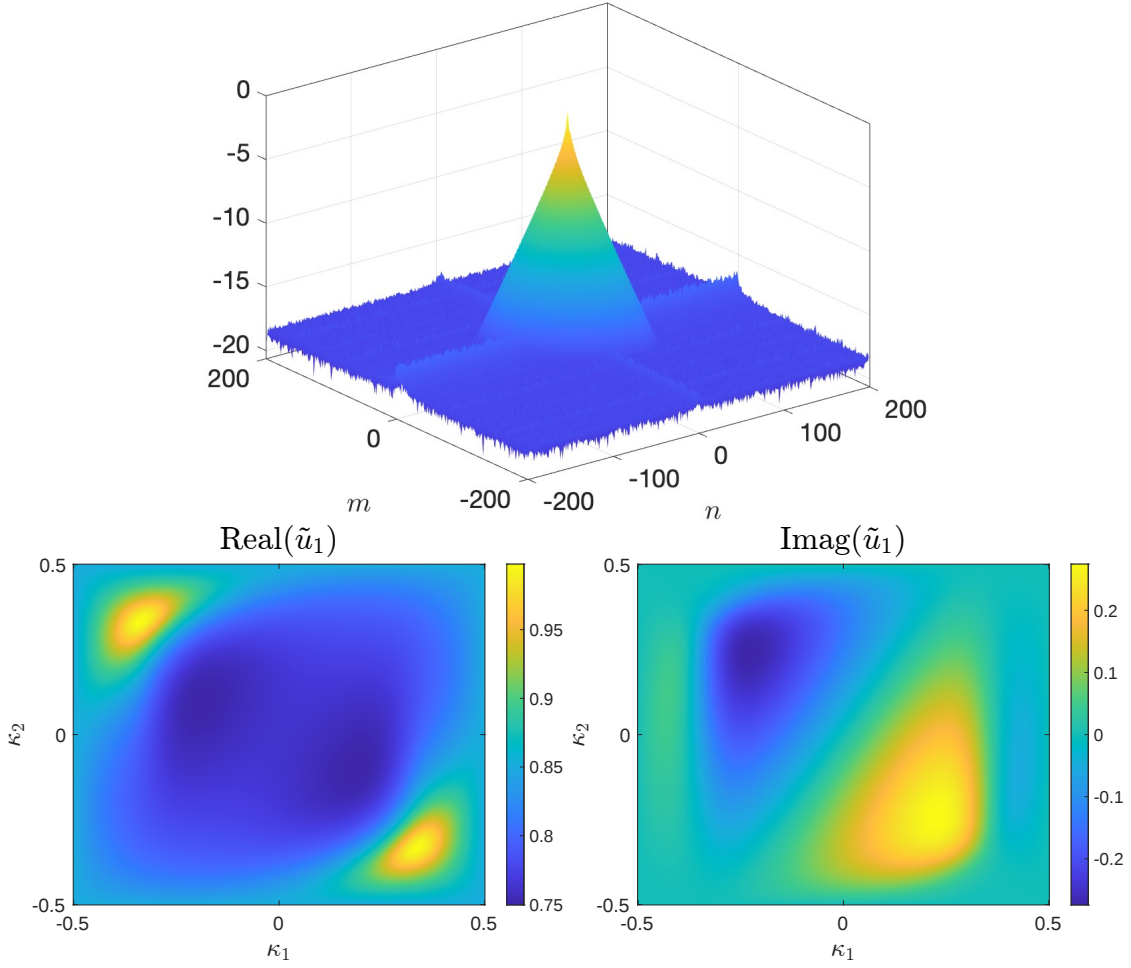


Figure 11: Plot of the real and imaginary part of the component $\tilde{u}_1$ with the absolute value of the Fourier coefficients of $\tilde{u}_1$ in the log scale with base 10 for in Example 2.

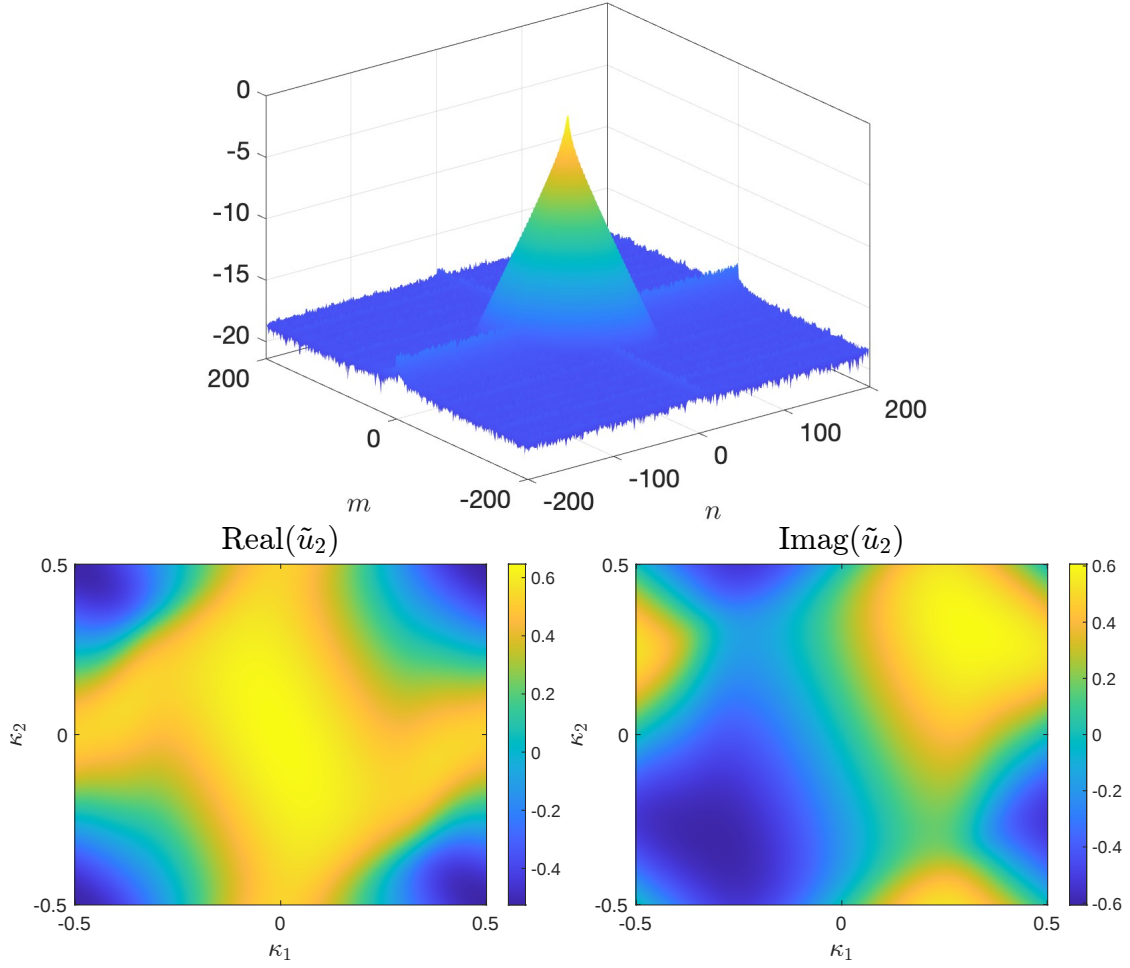


Figure 12: Plot of the real and imaginary part of the component $\tilde{u}_2$ with the absolute value of the Fourier coefficients of $\tilde{u}_1$ in the log scale with base 10 for in Example 2.

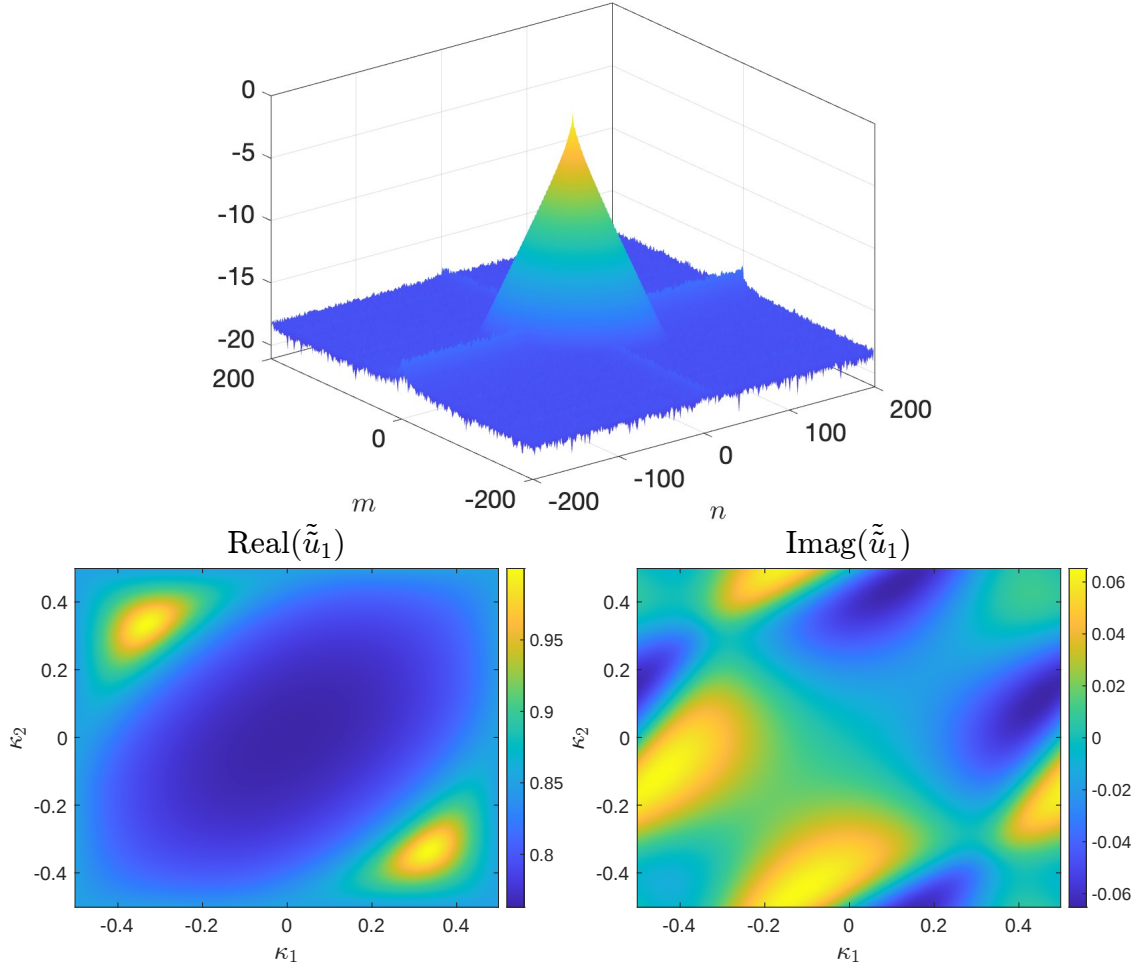


Figure 13: Plot of the quantities in Figure 11 after eliminating the divergence of the Berry connection. Other components is not shown as they are visually very similar to that in Figure 11

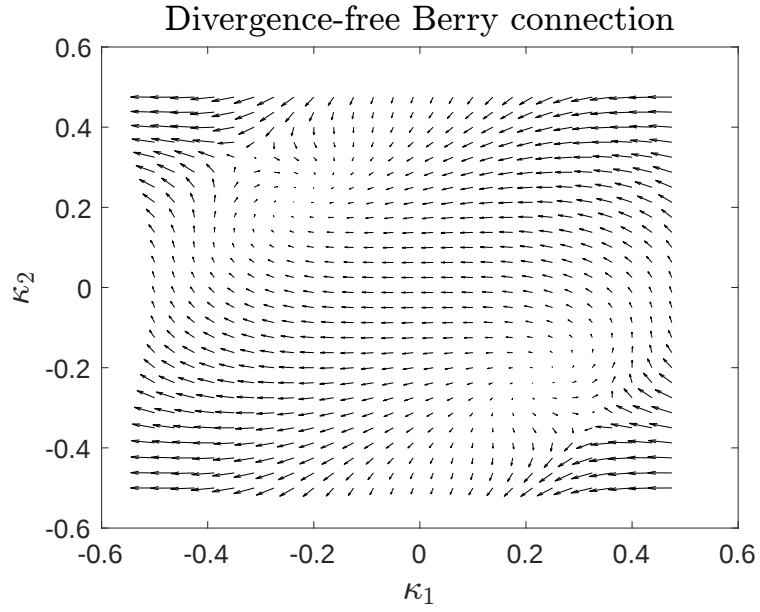
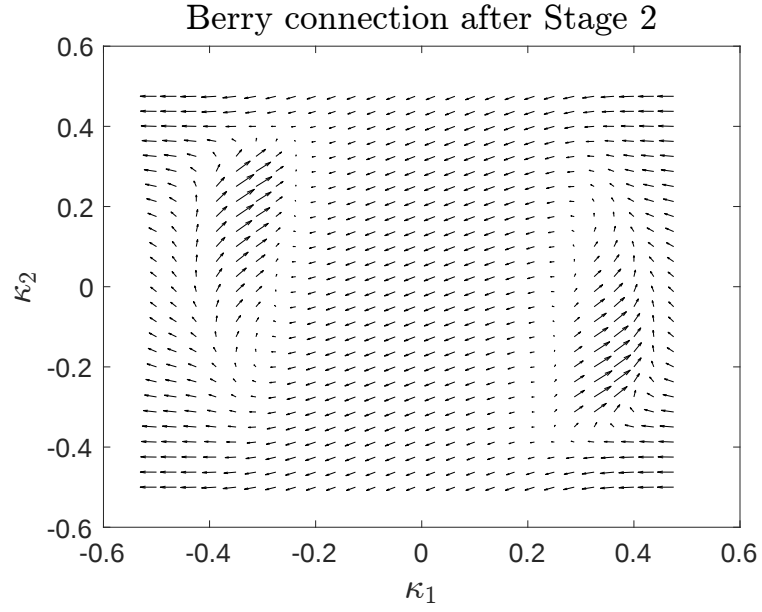


Figure 14: (a) Plot of the Berry connection after Stage 2 for Example 2. (b) Plot of the Berry connection in after eliminating its curl-free component in Stage 3. All vectors shown are in the $\mathbf{e}_x, \mathbf{e}_y$ basis.

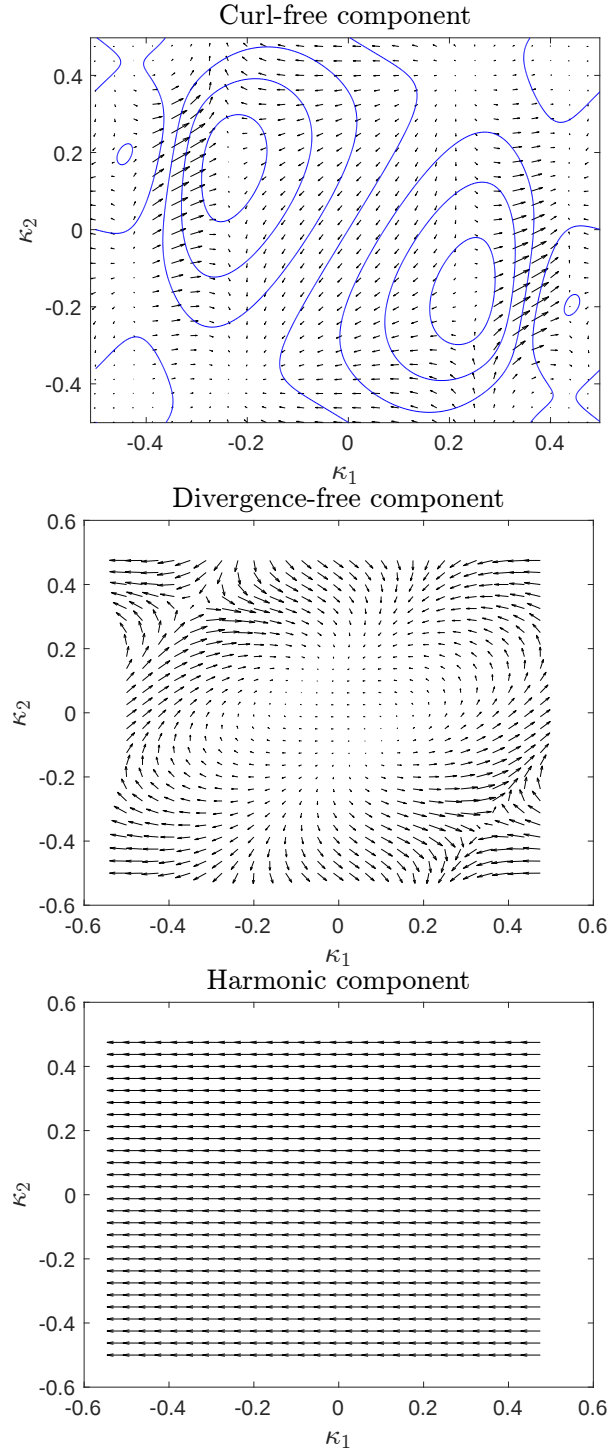


Figure 15: The Helmholtz-Hodge decomposition of the Berry connection in Figure 14(a). All vectors shown are in the $\mathbf{e}_x, \mathbf{e}_y$ basis. The equipotential lines are shown for the curl-free vector field, which is eliminated to obtain the optimal solution in Figure 14(b).

8.3 Example 3: Haldane model (topologically non-trivial case)

In this example, we consider a topologically non-trivial version of Haldane model in Section 8.2. All the lattice vectors are identical to those in Section 8.2 with a modified matrix H given by the formula

$$H(\mathbf{k}) = \begin{bmatrix} V_0 & t_1(1 + e^{-i\mathbf{k}\cdot\mathbf{a}_1} + e^{-i\mathbf{k}\cdot\mathbf{a}_2}) \\ t_1(1 + e^{i\mathbf{k}\cdot\mathbf{a}_1} + e^{i\mathbf{k}\cdot\mathbf{a}_2}) & -V_0 \end{bmatrix} + t_2(\sin(\mathbf{k}\cdot\mathbf{a}_1) - \sin(\mathbf{k}\cdot\mathbf{a}_2) - \sin(\mathbf{k}\cdot(\mathbf{a}_1 - \mathbf{a}_2))) \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \quad (255)$$

where the constants are chosen to be $t_1 = 1$, $t_2 = -0.45$ and $V_0 = 0.5$. The eigenvalues as a functions of the parameterization κ_1 and κ_2 are shown Figure 16(a) and the top band is chosen for the construction. The first Chern number $C_1 = 1$, so the topological obstruction is present (see the discussion below Theorem 6.6). As a result, in Figure 16(b), its constructed φ_2 (see (181)) is not periodic, as opposed to that in Figure 2(b) and 10(b) for Example 1 and 2.

The two components of the assignment $\tilde{\mathbf{u}}$ after Stage 2 in Section 6.2 is shown in Figure 17 – 18, together with the absolute value of their Fourier coefficients in the log scale with base 10. The Fourier coefficients decay exponentially asymptotically in one direction but very slowly in the other one since by construction $\tilde{\mathbf{u}}(\mathbf{k}(\kappa_1, \kappa_2))$ is analytic and periodic in κ_1 , but only analytic in κ_2 , which can be seen from Figure 17 – 18.

Since this example computationally similar to the trivial case in Section 8.2, we only show the convergence of the computed Chern number in Table 7.

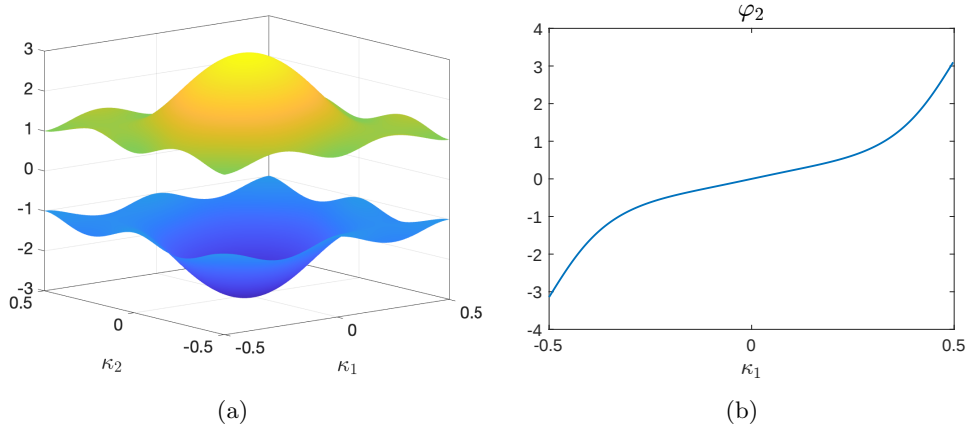


Figure 16: (a) Plot of eigenvalues of H in Example 3. The band on top is picked. (b) The phase φ_2 in (181) for Example 3. It is discontinuous due to the non-zero Chern number.

N	E_{Ch}
50	2.69e-14
100	1.11e-16

Table 7: Errors in the computed first Chern number C_1 for Example 3 ($C_1 = 1$).

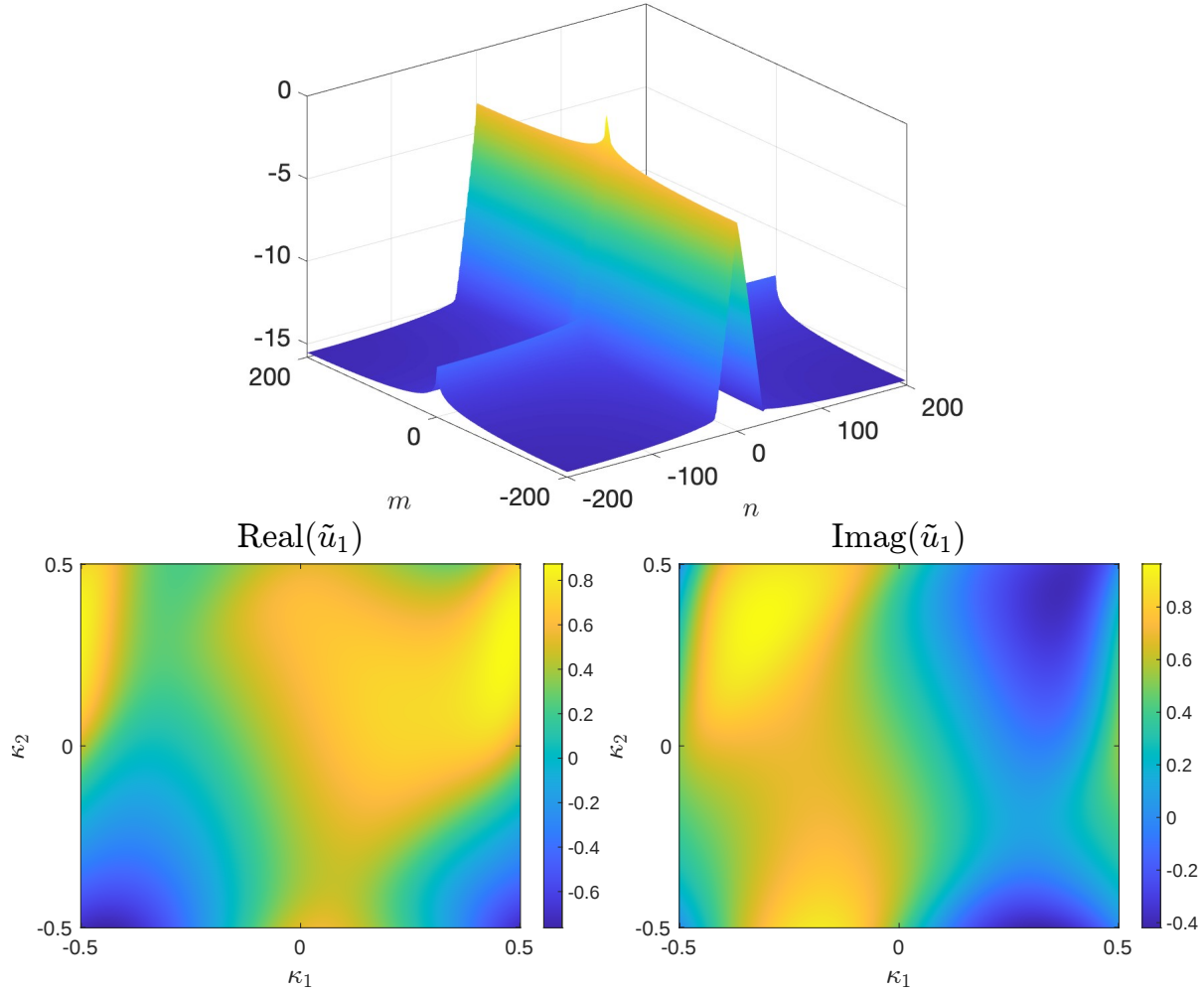


Figure 17: Plot of the real and imaginary part of the component $\tilde{u}_1$ with the absolute value of the Fourier coefficients of $\tilde{u}_1$ in the log scale with base 10 for in Example 3. The Fourier coefficients decays exponentially only in one direction since $\tilde{u}_1$ is analytic and periodic in κ_1 but only analytic in κ_2 .

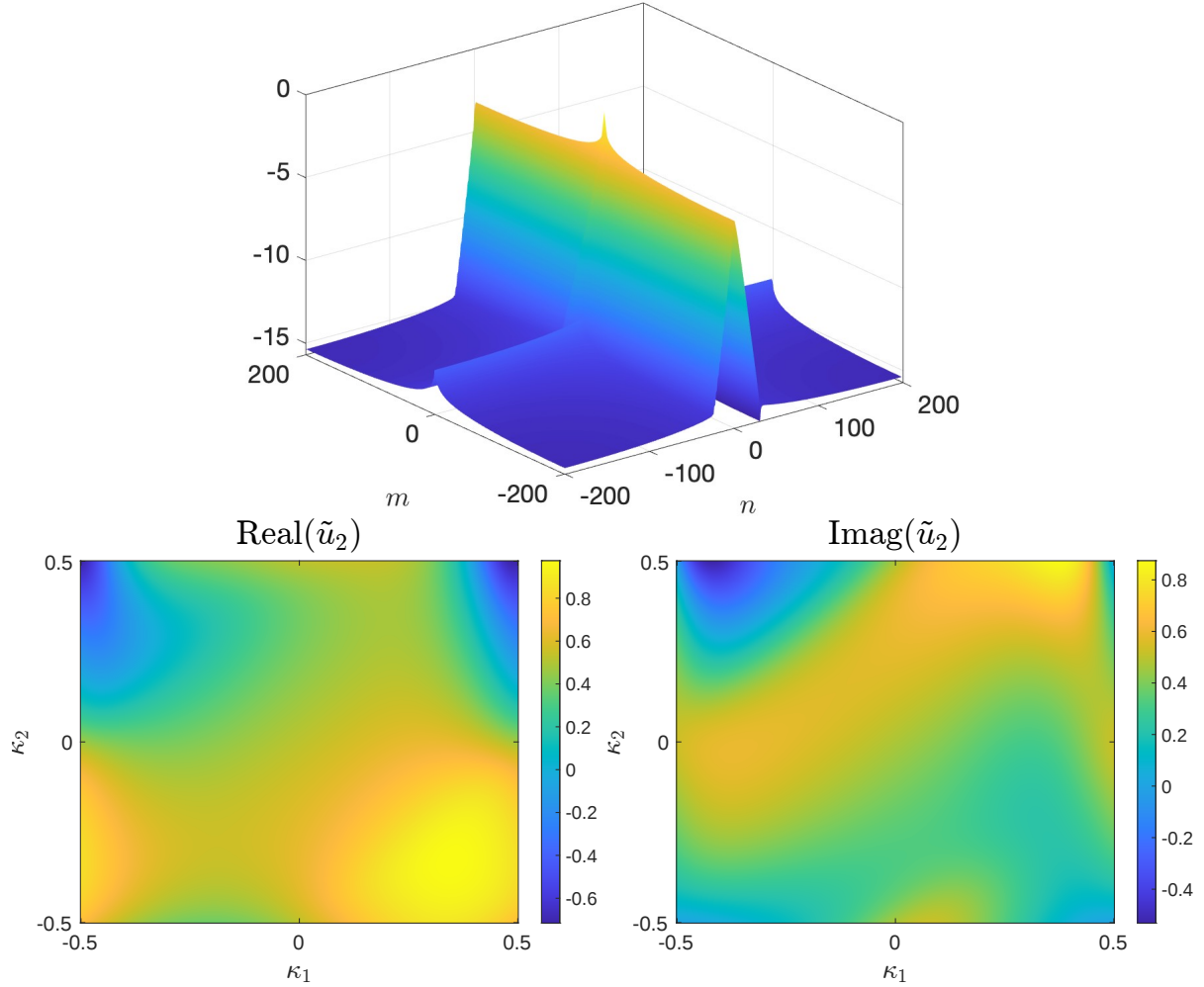


Figure 18: Plot of the real and imaginary part of the component $\tilde{u}_2$ with the absolute value of the Fourier coefficients of $\tilde{u}_2$ in the log scale with base 10 for in Example 3. The Fourier coefficients decays exponentially only in one direction since $\tilde{u}_2$ is analytic and periodic in κ_1 but only analytic in κ_2 .

9 Conclusion and generalization

In this paper, we presented rapidly convergent schemes for computing globally optimal Wannier functions for the isolated single band case in two dimensions. In the absence of topological obstruction, we proved that parallel transport (with simple corrections) alone leads to assignments of eigenvectors corresponding to exponentially localized Wannier functions. Furthermore, when the model possesses time reversal symmetry, we proved that the resulting Wannier functions are automatically real. Then a single gauge transform can be applied to eliminate the divergence of the Berry connections of the corresponding assignments, obtaining the globally optimal assignments for Wannier functions with minimum spread. We illustrated the efficiency and accuracy of the schemes with several examples. We compared the results above to those obtained via the parallel transport approach in Remark 7.2, which decouples the computation of band structure information and Wannier functions. This alternative approach produces the same Wannier functions of lower accuracy, but it has the advantage of being simple and computationally cheap once the band structure information is known.

The schemes in this paper can be further accelerated by exploiting symmetries of the lattice (and the time-reversal symmetry if available); at least partial symmetry information can be easily utilized to avoid unnecessary computation for parallel transport and band structure information without affecting the equispaced grids scheme in this paper. Moreover, better time-stepping schemes for parallel transport can be used to further improve convergence or to reduce the number of matrix inversions.

Since the schemes in this paper build two-dimensional assignments of eigenvectors based on one-dimensional ones, the analysis and algorithms can be easily generalized to three-dimensional cases. The three-dimensional construction will be made of two-dimensional slices obtained by the schemes in this paper, and all Chern-like numbers will appear in the construction.

If we replace the construction of one-dimensional Wannier functions in this paper with the Fourier approach in [12] for Schrödinger operators, the analysis in this paper carries over to the operator case almost verbatim. For the algorithms, when only low accuracy is required, a similar parallel transport scheme to that in Remark 7.2 can be applied, followed by eliminating the divergence of Berry connections as in this paper. When higher accuracy is needed, better parallel transport schemes are needed. The scheme in this paper and [12] can be applied, but matrices to be inverted in the operator case during the parallel transport computation can be large. As a result, better numerical schemes for the inversion are needed. These questions are under vigorous investigation.

For extensions to the isolated multi-band case, although the overall idea remains the same, there are notable differences. For example, the Berry connections in the multi-band case become Lie-algebra valued; in two dimension, the elimination of divergence cannot be done in a single step, but the potential theory approach will significantly reduce the number of optimization

steps. Extensions of [12] and this paper to the isolated multi-band case have been worked out and are now in preparation for publication.

Acknowledgement

The author thanks John Schotland for drawing their attention to this subject. The author also thanks Vladimir Rokhlin for his support and helpful discussions pertaining to this work.

References

- [1] Michael V Berry. The quantum phase, five years after, 1989.
- [2] Michael Victor Berry. Quantal phase factors accompanying adiabatic changes. *Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences*, 392(1802):45–57, 1984.
- [3] EI Blount. Formalisms of band theory. In *Solid state physics*, volume 13, pages 305–373. Elsevier, 1962.
- [4] William L Briggs and Van Emden Henson. *The DFT: an owner’s manual for the discrete Fourier transform*. SIAM, 1995.
- [5] Christian Brouder, Gianluca Panati, Matteo Calandra, Christophe Mourougane, and Nicola Marzari. Exponential localization of wannier functions in insulators. *Physical review letters*, 98(4):046402, 2007.
- [6] Éric Cancès, Antoine Levitt, Gianluca Panati, and Gabriel Stoltz. Robust determination of maximally localized wannier functions. *Physical Review B*, 95(7):075114, 2017.
- [7] Philippe G Ciarlet. *Linear and nonlinear functional analysis with applications*. SIAM, 2013.
- [8] Germund Dahlquist and Åke Björck. *Numerical methods*. Courier Corporation, 2003.
- [9] Anil Damle, Antoine Levitt, and Lin Lin. Variational formulation for Wannier functions with entangled band structure. *Multiscale Modeling & Simulation*, 17(1):167–191, 2019.
- [10] Anil Damle, Lin Lin, and Lexing Ying. Scdm-k: Localized orbitals for solids via selected columns of the density matrix. *Journal of Computational Physics*, 334:1–15, 2017.
- [11] Alok Dutt, Leslie Greengard, and Vladimir Rokhlin. Spectral deferred correction methods for ordinary differential equations. *BIT Numerical Mathematics*, 40:241–266, 2000.
- [12] Abinand Gopal and Hanwen Zhang. A highly accurate procedure for computing globally optimal wannier functions in one-dimensional crystalline insulators. *arXiv preprint arXiv:2409.04369*, 2024.
- [13] Dominik Gresch, Gabriel Autes, Oleg V Yazyev, Matthias Troyer, David Vanderbilt, B Andrei Bernevig, and Alexey A Soluyanov. Z2pack: Numerical implementation of hybrid wannier centers for identifying topological materials. *Physical Review B*, 95(7):075146, 2017.

- [14] F Duncan M Haldane. Model for a quantum hall effect without landau levels: Condensed-matter realization of the” parity anomaly”. *Physical review letters*, 61(18):2015, 1988.
- [15] Tosio Kato. *Perturbation theory for linear operators*, volume 132. Springer Science & Business Media, 2013.
- [16] Efthimios Kaxiras and John D Joannopoulos. *Quantum theory of materials*. Cambridge university press, 2019.
- [17] Nicola Marzari, Arash A Mostofi, Jonathan R Yates, Ivo Souza, and David Vanderbilt. Maximally localized wannier functions: Theory and applications. *Reviews of Modern Physics*, 84(4):1419–1475, 2012.
- [18] Nicola Marzari and David Vanderbilt. Maximally localized generalized Wannier functions for composite energy bands. *Physical review B*, 56(20):12847, 1997.
- [19] Mikio Nakahara. *Geometry, topology and physics*. CRC press, 2018.
- [20] Gianluca Panati. Triviality of bloch and bloch–dirac bundles. In *Annales Henri Poincaré*, volume 8, pages 995–1011. Springer, 2007.
- [21] Dimitris A Papaconstantopoulos et al. *Handbook of the Band Structure of Elemental Solids: From Z*. Springer, 2015.
- [22] Lev Semenovich Pontryagin. *Ordinary Differential Equations: Adiwes International Series in Mathematics*. Elsevier, 2014.
- [23] Barry Simon. Holonomy, the quantum adiabatic theorem, and berry’s phase. *Physical Review Letters*, 51(24):2167, 1983.
- [24] DJ Thouless. Wannier functions for magnetic sub-bands. *Journal of Physics C: Solid State Physics*, 17(12):L325, 1984.
- [25] Lloyd N Trefethen and JAC Weideman. The exponentially convergent trapezoidal rule. *SIAM review*, 56(3):385–458, 2014.
- [26] David Vanderbilt. *Berry phases in electronic structure theory: electric polarization, orbital magnetization and topological insulators*. Cambridge University Press, 2018.
- [27] Joshua Zak. Berry’s phase for energy bands in solids. *Physical review letters*, 62(23):2747, 1989.

10 Appendix

10.1 Alternative approach: direct computation of optimal gauge

In this section, we describe an alternative approach to the method in section 6. In this approach, we directly compute the gauge-invariant parts of the Berry connection in (156), which uniquely determines the globally optimal assignment of eigenvectors (and the Wannier function). It proceeds by first computing the Berry curvature, from which a Poisson's equation solve is done to obtain the potential F in (156). The potential F gives the Berry connection satisfying the integrability condition in Theorem 5.1, so that a Pfaffian system of the form in (114) can be solved to define an analytic (not necessarily Λ^* -periodic) assignment of eigenvectors in D^* . In order to make the assignment Λ^* -periodic, the constant harmonic component (h_x, h_y) is obtained, thus completing the construction. We do not provide any numerical procedure for this approach since the key ingredients required are essentially identical to those in Section 7.

It should be observed that, although the method in this section is conceptually simpler, it is more expensive than the first method in Section 6 due to the computation of the Berry curvature without doing the assignment. Furthermore, since it uses the gauge-invariant property of the Berry curvature for a single band, it cannot be generalized to the multi-band case, where the Berry curvature tensor is only gauge-covariant.

10.1.1 Stage 1: computing the Berry curvature

In order to compute the Berry curvature Ω_{xy} , we compute the family of eigenvalues E and eigenvectors $\mathbf{u}$:

$$H(\mathbf{k})\mathbf{u}(\mathbf{k}) = E(\mathbf{k})\mathbf{u}(\mathbf{k}), \quad \mathbf{k} \in D^*. \quad (256)$$

It should be observed that the phase choice of $\mathbf{u}$ can be arbitrary since the Berry curvature Ω_{xy} is gauge-invariant (see Remark 5.2). Next, we compute Ω_{xy} by the formula

$$\Omega_{xy}(\mathbf{k}) = i \frac{\partial}{\partial k_x} \mathbf{u}^*(\mathbf{k}) \frac{\partial}{\partial k_y} \mathbf{u}(\mathbf{k}) - i \frac{\partial}{\partial k_y} \mathbf{u}^*(\mathbf{k}) \frac{\partial}{\partial k_x} \mathbf{u}(\mathbf{k}), \quad \mathbf{k} \in D^*, \quad (257)$$

where the derivatives are given by

$$\frac{\partial}{\partial k_x} \mathbf{u} = -(H - E)^\dagger \frac{\partial}{\partial k_x} H \mathbf{u}, \quad \frac{\partial}{\partial k_y} \mathbf{u} = -(H - E)^\dagger \frac{\partial}{\partial k_y} H \mathbf{u}. \quad (258)$$

10.1.2 Stage 2: computing the divergence-free component

Next, based on the decomposition in (140), we compute the gauge-invariant vector fields. More explicitly, this involves finding an analytic and Λ^* -periodic F that generates the divergence-free component and the harmonic component (h_x, h_y) . By Theorem 5.8, they uniquely (up to a lattice constant) determine the assignment of eigenvectors.

In order to obtain F , by (259), we solve the following Poisson's equation

$$\frac{\partial^2 F}{\partial k_x^2} + \frac{\partial^2 F}{\partial k_y^2} = -g, \quad F \text{ is } \Lambda^* \text{-periodic in } D^*, \quad (259)$$

where g is given by the Berry curvature given in (257):

$$g = \Omega_{xy}. \quad (260)$$

However, by Theorem 3.3, the equation is solvable if and only if the integral of Ω_{xy} over D^* is zero. By the definition of the first Chern number C_1 in (132), this corresponds to the condition

$$\int_{D^*} d\mathbf{k} \Omega_{xy}(\mathbf{k}) = 2\pi C_1 = 0. \quad (261)$$

Hence, we conclude that (259) is solvable if and only if $C_1 = 0$. In other words, when $C_1 \neq 0$, we encounter the topological obstruction in Theorem 5.3 and no such F can be found. If we assume that $C_1 = 0$, we observe that (259) is easily solvable for the following reason. Since the projector P is analytic and periodic in D^* , so is Ω_{xy} by its definition in (125). Hence, the Fourier coefficients of Ω_{xy} decays exponentially. By (54), the solution F is given by its Fourier series as

$$F(\mathbf{k}) = \sum_{\substack{\mathbf{R} \in \Lambda \\ \mathbf{R} \neq \mathbf{0}}} \frac{g_{\mathbf{R}}}{\|\mathbf{R}\|^2} \cdot e^{i\mathbf{R} \cdot \mathbf{k}}, \quad \mathbf{k} \in D^*, \quad (262)$$

where $g_{\mathbf{R}}$ is the Fourier coefficient of g at $\mathbf{R} \in \Lambda$:

$$g_{\mathbf{R}} = \frac{V_{\text{puc}}}{(2\pi)^2} \int_{D^*} d\mathbf{k} e^{-i\mathbf{R} \cdot \mathbf{k}} g(\mathbf{k}) = \frac{V_{\text{puc}}}{(2\pi)^2} \int_{D^*} d\mathbf{k} e^{-i\mathbf{R} \cdot \mathbf{k}} \Omega_{xy}(\mathbf{k}). \quad (263)$$

After obtaining F , we compute the divergence-free Berry connection by the formula

$$A_x(\mathbf{k}) = \frac{\partial}{\partial k_y} F(\mathbf{k}) = \sum_{\substack{\mathbf{R} \in \Lambda \\ \mathbf{R} \neq \mathbf{0}}} iR_y \frac{g_{\mathbf{R}}}{\|\mathbf{R}\|^2} \cdot e^{i\mathbf{R} \cdot \mathbf{k}}, \quad (264)$$

$$A_y(\mathbf{k}) = -\frac{\partial}{\partial k_x} F(\mathbf{k}) = -\sum_{\substack{\mathbf{R} \in \Lambda \\ \mathbf{R} \neq \mathbf{0}}} iR_x \frac{g_{\mathbf{R}}}{\|\mathbf{R}\|^2} \cdot e^{i\mathbf{R} \cdot \mathbf{k}}, \quad (265)$$

where $\mathbf{R} = (R_x, R_y)$.

By Theorem 5.1, with the Berry connection (A_x, A_y) in (265), we have the following integrable Pfaffian system in D^* :

$$\frac{\partial}{\partial k_x} \tilde{\mathbf{u}} = \frac{\partial P}{\partial k_x} \tilde{\mathbf{u}} - iA_x \tilde{\mathbf{u}}, \quad \frac{\partial}{\partial k_y} \tilde{\mathbf{u}} = \frac{\partial P}{\partial k_y} \tilde{\mathbf{u}} - iA_y \tilde{\mathbf{u}}, \quad (266)$$

subject to the initial condition

$$\tilde{\mathbf{u}}(\mathbf{k}^0) = \mathbf{u}(\mathbf{k}^0), \quad (267)$$

where $\mathbf{k}^0$ can be any point in D^* and $\mathbf{u}(\mathbf{k}^0)$ is computed in (256). However, we observe that the solution $\tilde{\mathbf{u}}$ to (266) only defines an analytic assignment in D^* viewed as a subset of $\mathbb{R}^2$ since D^* is not simply connected when viewed as a torus. Hence, the solution $\tilde{\mathbf{u}}$ is not necessarily Λ^* -periodic in D^* .

10.1.3 Stage 3: compute the harmonic component

Next, we compute the harmonic component (h_x, h_y) in (140) in order to turn $\tilde{\mathbf{u}}$ into a Λ^* -periodic assignment. To do so, we parameterize D^* by $(\kappa_1, \kappa_2) \in T$ so that we compute (h_x, h_y) by its components h_1 and h_2 in the $\mathbf{b}_1$ and $\mathbf{b}_2$ direction via the formula

$$(h_x, h_y) = h_1 \frac{\mathbf{b}_1}{\|\mathbf{b}_1\|} + h_2 \frac{\mathbf{b}_2}{\|\mathbf{b}_2\|}. \quad (268)$$

By (46) and the definition of the Berry connection, the Berry connection A_1, A_2 in the $\mathbf{b}_1$ and $\mathbf{b}_2$ directions are given by the formulas

$$A_1 = \mathbf{b}_1 \cdot \mathbf{e}_x A_x + \mathbf{b}_1 \cdot \mathbf{e}_y A_y, \quad A_2 = \mathbf{b}_2 \cdot \mathbf{e}_x A_x + \mathbf{b}_2 \cdot \mathbf{e}_y A_y, \quad (269)$$

where A_x, A_y are given in (265). Then the system in (266) in the $\mathbf{b}_1$ and $\mathbf{b}_2$ directions are given by

$$\frac{\partial}{\partial \kappa_1} \tilde{\mathbf{u}}(\mathbf{k}(\kappa_1, \kappa_2)) = \frac{\partial P(\mathbf{k}(\kappa_1, \kappa_2))}{\partial \kappa_1} \tilde{\mathbf{u}}(\mathbf{k}(\kappa_1, \kappa_2)) - i A_1(\mathbf{k}(\kappa_1, \kappa_2)) \tilde{\mathbf{u}}(\mathbf{k}(\kappa_1, \kappa_2)), \quad (270)$$

$$\frac{\partial}{\partial \kappa_2} \tilde{\mathbf{u}}(\mathbf{k}(\kappa_1, \kappa_2)) = \frac{\partial P(\mathbf{k}(\kappa_1, \kappa_2))}{\partial \kappa_2} \tilde{\mathbf{u}}(\mathbf{k}(\kappa_1, \kappa_2)) - i A_2(\mathbf{k}(\kappa_1, \kappa_2)) \tilde{\mathbf{u}}(\mathbf{k}(\kappa_1, \kappa_2)). \quad (271)$$

In order to compute h_1 , we solve (270) along the line γ_1 across T in Figure 19, where κ_2 can be any value in $[-\frac{1}{2}, \frac{1}{2}]$ and κ_1 runs from $-\frac{1}{2}$ to $\frac{1}{2}$. We specify the initial condition as

$$\tilde{\mathbf{u}}(\mathbf{k}(-1/2, \kappa_2)) = \mathbf{u}(\mathbf{k}(-1/2, \kappa_2)), \quad (272)$$

where $\mathbf{u}(\mathbf{k}(-1/2, \kappa_2))$ is computed in (256). Since $\tilde{\mathbf{u}}$ is not guaranteed to be Λ^* -periodic in D^* , the solution $\tilde{\mathbf{u}}(\mathbf{k}(\kappa_1, \kappa_2))$ to (270) may not be the same at $\kappa_1 = -\frac{1}{2}$ and $\kappa_1 = \frac{1}{2}$. However,

$\tilde{\mathbf{u}}(\mathbf{k}(-1/2, \kappa_2))$ and $\tilde{\mathbf{u}}(\mathbf{k}(1/2, \kappa_2))$ are eigenvectors to the same eigenvalue by periodicity, so they can only differ by a phase factor e^{ih_1} :

$$\tilde{\mathbf{u}}(\mathbf{k}(-1/2, \kappa_2)) = e^{ih_1} \tilde{\mathbf{u}}(\mathbf{k}(1/2, \kappa_2)), \quad (273)$$

from which we compute h_1 by the formula

$$h_1 = -i \log(\tilde{\mathbf{u}}^*(\mathbf{k}(-1/2, \kappa_2)) \tilde{\mathbf{u}}(\mathbf{k}(1/2, \kappa_2))). \quad (274)$$

Before we show that adding h_1 to A_1 will turn $\tilde{\mathbf{u}}(\mathbf{k}(\kappa_1, \kappa_2))$ into a periodic function in $\kappa_1 \in [-\frac{1}{2}, \frac{1}{2}]$ for any κ_2 , we prove that h_1 in (274) is independent of κ_1 . Consider solving the system defined by (270) and (271), subject to the initial condition in (275), along the rectangle in Figure 19, where γ_1 is the path from which we obtain (274) and γ_3 is parallel to γ_1 at κ_1' . Since system is integrable by construction, the solution after going around the square is the same as the initial condition in (274). Furthermore, (271) on γ_2 and γ_4 are identical due to the periodicity of $\frac{\partial P}{\partial \kappa_2}$ and A_2 in (271). Since γ_2 and γ_4 are in the opposite direction, their contribution cancels. Consequently, the contribution from solving (270) on γ_1 and γ_3 also cancels. Thus, we conclude that h_1 in (274) is independent κ_1 ; the same h_1 would be obtained by solving (270) on any line parallel to γ_1 .

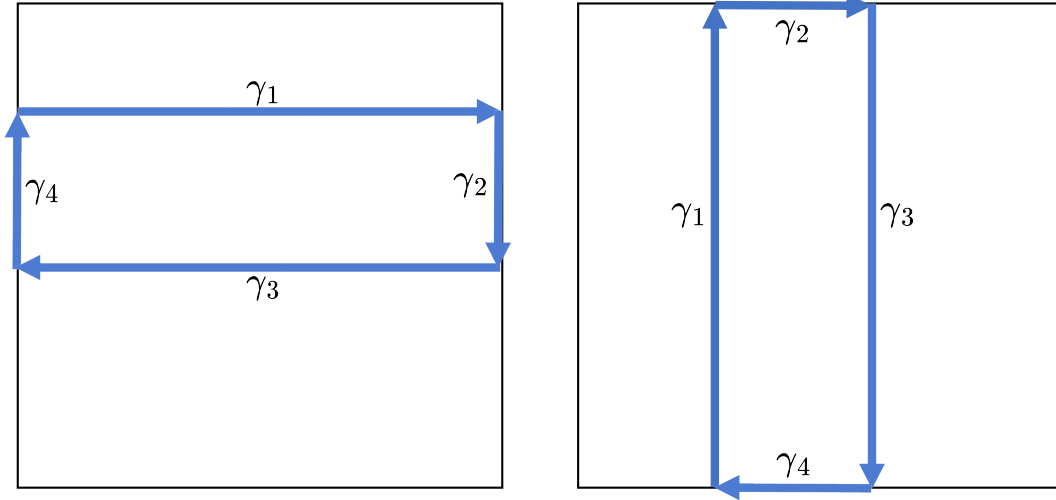


Figure 19: The path for computing the harmonic component h_1 (left) and h_2 (right).

The procedures for computing h_2 is similar to those above for h_1 by switching the role of κ_1 and κ_2 . We solve (271) along the line γ_1 , where κ_2 runs from $-\frac{1}{2}$ to $\frac{1}{2}$ for any $\kappa_1 \in [-\frac{1}{2}, \frac{1}{2}]$ with the initial condition

$$\tilde{\mathbf{u}}(\mathbf{k}(\kappa_1, -1/2)) = \mathbf{u}(\mathbf{k}(\kappa_1, -1/2)), \quad (275)$$

where $\mathbf{u}(\mathbf{k}(\kappa_1, -1/2))$ is computed in (256). By periodicity, $\tilde{\mathbf{u}}(\mathbf{k}(\kappa_1, -1/2))$ and $\tilde{\mathbf{u}}(\mathbf{k}(\kappa_1, 1/2))$ can only differ by a phase factor e^{ih_2} :

$$\tilde{\mathbf{u}}(\mathbf{k}(\kappa_1, -1/2)) = e^{ih_2} \tilde{\mathbf{u}}(\mathbf{k}(\kappa_1, 1/2)), \quad (276)$$

from which we compute h_2 by the formula

$$h_2 = -i \log(\tilde{\mathbf{u}}^*(\mathbf{k}(\kappa_1, -1/2)) \tilde{\mathbf{u}}(\mathbf{k}(\kappa_1, 1/2))). \quad (277)$$

Similar to showing that h_1 is independent of κ_2 , we show that h_2 in (277) is independent κ_1 .

Next, we apply the following gauge transformation to the solution $\tilde{\mathbf{u}}(\mathbf{k}(\kappa_1, \kappa_2))$ to (270) subject to (275)

$$\tilde{\tilde{\mathbf{u}}}(\mathbf{k}(\kappa_1, \kappa_2)) = e^{-ih_1\kappa_1} e^{-ih_2\kappa_2} \tilde{\mathbf{u}}(\mathbf{k}(\kappa_1, \kappa_2)) \quad (278)$$

so that A_1 and A_2 in (270) and (271) are modified accordingly as

$$\tilde{A}_1 = A_1 + h_1, \quad \tilde{A}_2 = A_2 + h_2, \quad (279)$$

and the new $\tilde{\tilde{\mathbf{u}}}$ satisfies

$$\frac{\partial}{\partial \kappa_1} \tilde{\tilde{\mathbf{u}}}(\mathbf{k}(\kappa_1, \kappa_2)) = \frac{\partial P(\mathbf{k}(\kappa_1, \kappa_2))}{\partial \kappa_1} \tilde{\tilde{\mathbf{u}}}(\mathbf{k}(\kappa_1, \kappa_2)) - i \tilde{A}_1(\mathbf{k}(\kappa_1, \kappa_2)) \tilde{\tilde{\mathbf{u}}}(\mathbf{k}(\kappa_1, \kappa_2)), \quad (280)$$

$$\frac{\partial}{\partial \kappa_2} \tilde{\tilde{\mathbf{u}}}(\mathbf{k}(\kappa_1, \kappa_2)) = \frac{\partial P(\mathbf{k}(\kappa_1, \kappa_2))}{\partial \kappa_2} \tilde{\tilde{\mathbf{u}}}(\mathbf{k}(\kappa_1, \kappa_2)) - i \tilde{A}_2(\mathbf{k}(\kappa_1, \kappa_2)) \tilde{\tilde{\mathbf{u}}}(\mathbf{k}(\kappa_1, \kappa_2)). \quad (281)$$

By (278) and (273), it is easy to see that $\tilde{\tilde{\mathbf{u}}}$ satisfies

$$\tilde{\tilde{\mathbf{u}}}(\mathbf{k}(-1/2, \kappa_2)) = \tilde{\tilde{\mathbf{u}}}(\mathbf{k}(1/2, \kappa_2)), \quad \kappa_2 \in \left[-\frac{1}{2}, \frac{1}{2}\right]. \quad (282)$$

The periodicity of $\partial_{\kappa_1} P$ and $\tilde{A}_1$ in (280) show that

$$\frac{\partial}{\partial \kappa_1} \tilde{\tilde{\mathbf{u}}}(\mathbf{k}(-1/2, \kappa_2)) = \frac{\partial}{\partial \kappa_1} \tilde{\tilde{\mathbf{u}}}(\mathbf{k}(1/2, \kappa_2)), \quad \kappa_2 \in \left[-\frac{1}{2}, \frac{1}{2}\right]. \quad (283)$$

By repetitively differentiating (280), we conclude that the derivative of order $n = 2, 3, \dots$ satisfies

$$\frac{\partial^n}{\partial \kappa_1^n} \tilde{\tilde{\mathbf{u}}}(\mathbf{k}(-1/2, \kappa_2)) = \frac{\partial^n}{\partial \kappa_1^n} \tilde{\tilde{\mathbf{u}}}(\mathbf{k}(1/2, \kappa_2)), \quad \kappa_2 \in \left[-\frac{1}{2}, \frac{1}{2}\right]. \quad (284)$$

By (278), (276), (281) and the periodicity of $\partial_{\kappa_2} P$ and $\tilde{A}_2$, we conclude similarly that

$$\tilde{\mathbf{u}}(\mathbf{k}(\kappa_1, -1/2)) = \tilde{\mathbf{u}}(\mathbf{k}(\kappa_1, 1/2)), \quad \kappa_1 \in \left[-\frac{1}{2}, \frac{1}{2}\right], \quad (285)$$

and

$$\frac{\partial^n}{\partial \kappa_2^n} \tilde{\mathbf{u}}(\mathbf{k}(\kappa_1, -1/2)) = \frac{\partial^n}{\partial \kappa_2^n} \tilde{\mathbf{u}}(\mathbf{k}(\kappa_1, 1/2)), \quad \kappa_1 \in \left[-\frac{1}{2}, \frac{1}{2}\right], \quad (286)$$

for $n = 1, 2, \dots$. For simplicity, suppose that the initial condition for the system defined by (280) and (281) is specified at $\kappa_1 = \kappa_2 = -\frac{1}{2}$ by the formula

$$\tilde{\mathbf{u}}(\mathbf{k}(-1/2, -1/2)) = \mathbf{u}(\mathbf{k}(-1/2, -1/2)), \quad (287)$$

where $\mathbf{u}(\mathbf{k}(-1/2, -1/2))$ is computed in (256). Then, by (282-286), together with the integrability of the system defined by (280) and (281), we conclude that the solution $\tilde{\mathbf{u}}$ defines a analytic and Λ^* -periodic assignment in D^* .

Remark 10.1. *Since the system defined by (280) and (281) is integrable, the paths for which they are solved for finding $\tilde{\mathbf{u}}$ over D^* is irrelevant. For example, the paths can be taken to be the same as those in Section 6.*

10.1.4 Realty of Wannier functions

In the previous section, we have constructed the optimal assignment $\tilde{\mathbf{u}}$ in D^* . There is one remaining degree of freedom – the phase choice for the initial condition in (287). In this section, we show that, when H has time-reversal symmetry (see (27)), $\tilde{\mathbf{u}}$ will have the symmetry

$$\tilde{\mathbf{u}}(\mathbf{k}) = \tilde{\mathbf{u}}(-\mathbf{k}), \quad \mathbf{k} \in D^* \quad (288)$$

provided the initial condition in (287) is chosen to be real. More explicitly, we have

$$\overline{\mathbf{u}}(\mathbf{k}(-1/2, -1/2)) = \mathbf{u}(\mathbf{k}(-1/2, -1/2)). \quad (289)$$

First, we observe that (289) is always possible by considering the complex conjugate of (256) at $\kappa_1 = \kappa_2 = -\frac{1}{2}$:

$$\begin{aligned} \overline{H}(\mathbf{k}(-1/2, -1/2)) \overline{\mathbf{u}}(\mathbf{k}(-1/2, -1/2)) &= E(\mathbf{k}(-1/2, -1/2)) \overline{\mathbf{u}}(\mathbf{k}(-1/2, -1/2)) \\ &= H(\mathbf{k}(-1/2, -1/2)) \overline{\mathbf{u}}(\mathbf{k}(-1/2, -1/2)), \end{aligned} \quad (290)$$

where we used (27) and the periodicity of H to obtain

$$\overline{H}(\mathbf{k}(-1/2, -1/2)) = H(\mathbf{k}(1/2, 1/2)) = H(\mathbf{k}(-1/2, -1/2)). \quad (291)$$

The relation in (290) shows that both $\mathbf{u}(\mathbf{k}(-1/2, -1/2))$ and $\bar{\mathbf{u}}(\mathbf{k}(-1/2, -1/2))$ are eigenvectors of the same (non-degenerate) eigenvalue, so they can only differ by a phase factor of the form $\mathbf{u}^*(\mathbf{k}(-1/2, -1/2)) = \mathbf{u}(\mathbf{k}(-1/2, -1/2))e^{2i\varphi_0}$ for some real φ_0 . If (289) does not hold, we can pick $\mathbf{u}(\mathbf{k}(-1/2, -1/2))e^{i\varphi_0}$ in stead of $\mathbf{u}(\mathbf{k}(-1/2, -1/2))$ as the vector in (287) so that (289) holds.

We also observe that, since the Berry curvature satisfies $\Omega_{xy}(\mathbf{k}) = -\Omega_{xy}(-\mathbf{k})$ if H has time-reversal symmetry (see Remark 5.4), formulas in (265) show that A_x and A_y stratify

$$A_x(-\mathbf{k}) = A_x(\mathbf{k}), \quad A_y(-\mathbf{k}) = A_y(\mathbf{k}). \quad (292)$$

Combining (292) with (269) and (279) shows that

$$\tilde{A}_1(-\mathbf{k}) = \tilde{A}_1(\mathbf{k}), \quad \tilde{A}_2(-\mathbf{k}) = \tilde{A}_2(\mathbf{k}). \quad (293)$$

Furthermore, since $\bar{P}(\mathbf{k}) = P(-\mathbf{k})$, we have

$$\frac{\partial}{\partial \kappa_1} \bar{P}(\mathbf{k}) = -\frac{\partial}{\partial \kappa_1} P(-\mathbf{k}), \quad \frac{\partial}{\partial \kappa_2} \bar{P}(\mathbf{k}) = -\frac{\partial}{\partial \kappa_2} P(-\mathbf{k}). \quad (294)$$

By (294) and (293), we observe that (280) and (281) are unchanged under complex conjugation followed by replacing $\mathbf{k}$ by $-\mathbf{k}$. This implies that, if we transform by the same operations the solution $\tilde{\mathbf{u}}$ in the following form

$$\tilde{\mathbf{u}}(\mathbf{k}) \rightarrow \tilde{\tilde{\mathbf{u}}}(-\mathbf{k}), \quad (295)$$

the transformed solution will still satisfy (280) and (281). Furthermore, by periodicity, the vector on the right of (287) remains an eigenvector at $\kappa_1 = \kappa_2 = \frac{1}{2}$. By the realty of the initial condition in (289), the transformed solutions has the initial condition as the original one. Namely, we have

$$\tilde{\tilde{\mathbf{u}}}(\mathbf{k}(-1/2, -1/2)) = \mathbf{u}(\mathbf{k}(-1/2, -1/2)) = \tilde{\tilde{\mathbf{u}}}(\mathbf{k}(1/2, 1/2)). \quad (296)$$

As a result, the transformed solution satisfies the same equations as $\tilde{\mathbf{u}}$ with the same initial condition. By the uniqueness theorem of initial value problems, we conclude that $\tilde{\tilde{\mathbf{u}}}(\mathbf{k}) = \tilde{\tilde{\mathbf{u}}}(-\mathbf{k})$. By complex conjugation, we conclude that

$$\tilde{\tilde{\mathbf{u}}}(\mathbf{k}) = \tilde{\mathbf{u}}(-\mathbf{k}), \quad \mathbf{k} \in D^*. \quad (297)$$

Hence the Fourier coefficients of $\tilde{\tilde{\mathbf{u}}}$ are real, so is its corresponding Wannier function defined by (33).

10.2 Derivation of formulas in Lemma 5.5

The following contains the derivation for the moment functions in Lemma 5.5. By (114), we have

$$\begin{aligned}
\langle \mathbf{R} \rangle &= i \frac{V_{\text{puc}}}{(2\pi)^2} \int_{D^*} d\mathbf{k} \tilde{\mathbf{u}}^*(\mathbf{k}) \nabla_{\mathbf{k}} \tilde{\mathbf{u}}(\mathbf{k}) \\
&= \frac{V_{\text{puc}}}{(2\pi)^2} \int_{D^*} d\mathbf{k} \mathbf{A}(\mathbf{k}) \\
&= (h_x, h_y),
\end{aligned} \tag{298}$$

and

$$\begin{aligned}
\langle \|\mathbf{R}\|^2 \rangle &= \frac{V_{\text{puc}}}{(2\pi)^2} \int_{D^*} d\mathbf{k} \|\nabla_{\mathbf{k}} \tilde{\mathbf{u}}(\mathbf{k})\|^2 \\
&= \frac{V_{\text{puc}}}{(2\pi)^2} \int_{D^*} d\mathbf{k} \left(\left\| \frac{\partial P(\mathbf{k})}{\partial k_x} \tilde{\mathbf{u}}(\mathbf{k}) \right\|^2 + \left\| \frac{\partial P(\mathbf{k})}{\partial k_y} \tilde{\mathbf{u}}(\mathbf{k}) \right\|^2 + \|\mathbf{A}(\mathbf{k})\|^2 \right).
\end{aligned} \tag{299}$$

We observe that we have used (142) so that there is no cross terms that contain both the derivatives of the projector and $\mathbf{A}$. Moreover, we have used (50) to obtain (299) so that integrals of the derivatives of ψ, F vanishes. Applying (50) to the last term in (299) gives

$$\begin{aligned}
&\frac{V_{\text{puc}}}{(2\pi)^2} \int_{D^*} d\mathbf{k} \|\mathbf{A}(\mathbf{k})\|^2 \\
&= \frac{V_{\text{puc}}}{(2\pi)^2} \int_{D^*} d\mathbf{k} \left[\left(\frac{\partial \psi}{\partial k_x} \right)^2 + \left(\frac{\partial \psi}{\partial k_y} \right)^2 + \left(\frac{\partial F}{\partial k_x} \right)^2 + \left(\frac{\partial F}{\partial k_y} \right)^2 \right. \\
&\quad \left. + 2 \left(\frac{\partial \psi}{\partial k_x} \frac{\partial F}{\partial k_y} - \frac{\partial \psi}{\partial k_y} \frac{\partial F}{\partial k_x} \right) \right] + h_x^2 + h_y^2 \\
&= \frac{V_{\text{puc}}}{(2\pi)^2} \int_{D^*} d\mathbf{k} \left[\left(\frac{\partial \psi}{\partial k_x} \right)^2 + \left(\frac{\partial \psi}{\partial k_y} \right)^2 + \left(\frac{\partial F}{\partial k_x} \right)^2 + \left(\frac{\partial F}{\partial k_y} \right)^2 \right] + h_x^2 + h_y^2,
\end{aligned} \tag{300}$$

(301)

where we applied integration by parts to $\frac{\partial \psi}{\partial k_x}$ and $\frac{\partial \psi}{\partial k_y}$ in the second line of (300) and the fact that both ψ and F are Λ^* -periodic to make the cross term between derivatives of ψ and F vanish.