Ab initio derivation of Hubbard models for cold atoms in optical lattices

R. Walters, G. Cotugno, T. H. Johnson, S. R. Clark, and D. Jaksch

1Clarendon Laboratory, University of Oxford, Parks Road, Oxford OX1 3PU, United Kingdom
2Max Planck Research Department for Structural Dynamics, University of Hamburg, CFEL, Hamburg, Germany
3Centre for Quantum Technologies, National University of Singapore, 3 Science Drive 2, 117543, Singapore
4Keble College, University of Oxford, Parks Road, Oxford OX1 3PG, United Kingdom

(Dated: March 12, 2013)

We derive ab initio local Hubbard models for several optical lattice potentials of current interest, including the honeycomb and Kagomé lattices, verifying their accuracy on each occasion by comparing the interpolated band structures against the originals. To achieve this, we calculate the maximally-localized generalized Wannier basis by implementing the steepest-descent algorithm of Marzari and Vanderbilt [N. Marzari and D. Vanderbilt, Phys. Rev. B 56, 12847 (1997)] directly in one and two dimensions. To avoid local minima we develop an initialization procedure that is both robust and requires no prior knowledge of the optimal Wannier basis. The MATLAB code that implements our full procedure is freely available online at http://ccpforge.cse.rl.ac.uk/gf/project/mlgws/

PACS numbers: 71.15.Ap, 67.85.-d, 71.10.Fd

Atoms loaded into periodic optical potentials [1,2] can be sufficiently cold to only occupy a small number of lowest energy Bloch bands. The interaction between two atoms occupying the same potential well can be large so that they form a paradigm test-bed for studying the physics of strongly correlated quantum lattice models. To derive accurate microscopic models it is desirable to express the state of the atoms in terms of a basis of highly-localized single-particle states, given by some unitary transformation of the Bloch states forming the lowest energy bands. The reasons for this are two-fold. First, the occupations of localized basis states are measurable through high-resolution imaging [3]. Second, the Hamiltonian rewritten in terms of localized basis states is typically a Hubbard model dominated by a few local terms. Together these two points justify the simulation of local Hubbard models, used to describe many phenomena in condensed matter, using cold atoms in optical lattices [1]. In this article, we develop a procedure to systematically find a set of highly-localized basis states and thereby derive ab initio the parameters of a Hubbard model realized using cold atoms and an optical lattice.

Only in simple cases, e.g., a lattice potential that is orthogonal [4] or leads to an isolated lowest Bloch band [5], have the parameters of Hubbard models realized by cold atoms in optical lattices been derived using a basis of localized single-particle states. The single-particle states used are Fourier transforms of the Bloch states, called Wannier states [6]. For more complicated optical-lattice potentials, Hubbard parameters have been estimated rather than derived from first principles: on-site interaction Hubbard parameters have been estimated by using Gaussians centered at lattice minima as approximations to the single-particle states, and nearest-neighbor hopping parameters found by fitting a tight-binding form to the energy structure of the bands, without a rigorous justification of the tight-binding assumption (see e.g. Refs. [2,7,8]). The approach we take here improves upon such calculations in two ways. We use a class of single-particle states that generalize the Wannier states and can thus be more localized. Also, our procedure calculates Hubbard parameters from first principles, without approximation, and provides a quantitative justification of neglected terms. The necessity of such improvements has recently been noted in the literature [9].

Our procedure is an adaptation of several others already in use in solid-state physics. Specifically, we take as our starting point an algorithm developed by Marzari and Vanderbilt [10]. They consider a basis of generalized Wannier states; Fourier transforms of inter-band mixtures of Bloch states. Choosing some initial basis, a steepest-descent minimization algorithm is used to iteratively generate another set of generalized Wannier states with a smaller spatial spread. The desired end-point of these iterations is the basis corresponding to the global minimum of the spread, the so-called maximally-localized generalized Wannier states (see Ref. [11] and references within for a review on the topic). Once this optimal basis is found, the parameters of the corresponding Hubbard model are easily calculated.

The currently available software packages [12] that implement the steepest-descent minimization algorithm operate in three dimensions. For use with optical lattice potentials, which are often effectively one or two-dimensional, we have implemented the algorithm directly in these lower dimensional spaces, as well as in three dimensions. We find that for the optical-lattice potentials considered here, our implementation in conjunction with commonly used initialization procedures (e.g. that described in [10]) typically fails to converge to the global minimum of the spread and instead becomes trapped in a local minimum; the maximally-localized generalized Wannier states are not obtained.

Therefore, our algorithmic contribution is a new initialization procedure for the Marzari and Vanderbilt steepest-descent algorithm. Our initialization procedure has an additional benefit in that it requires no knowledge of the optimal Wannier states, e.g., their location or approximate form, and therefore

*Electronic address: g.cotugno1@physics.ox.ac.uk
requires no input beyond specifying the lattice potential. The initialization procedure is split into two parts, each minimizing the inter- and intra-band contributions to the spread of the generalized Wannier states, respectively. The former is a method for minimizing the spread in the case of a single band [10]. The latter relates to a procedure devised by Souza, Mazari and Vanderbilt to optimally disentangle a subset of bands from a group of degenerate bands [13]. Our whole procedure, taking the lattice potential as input, and outputting the maximally-localized Wannier states and Hubbard parameters, is combined into a single MATLAB routine. We have made this code freely available online [14].

Note that while in the last stages of preparing this article we became aware of a very recent article [15] in which the authors describe an initialization procedure, taking the lattice potential as input, and outputting the bands from a group of degenerate bands [13]. Our whole procedure devised by Souza, Mazari and Vanderbilt to optimally disentangle a subset of bands. An atom occupying the mode \( w^n_R(r) \) is often said to be in the \( n \)-th excited state or mode of lattice site \( R \).

The expansion thus takes the form

\[
\hat{\Psi}(r) = \sum_R \sum_n w^n_R(r) \hat{b}^n_R,
\]

where \( \hat{b}^n_R \) annihilates a boson in mode \( w^n_R(r) \), such that the Hamiltonian \( \hat{H} \) may be re-expressed as

\[
\hat{H} = -\sum_{mn} i_{RR'}^m \hat{b}^m_{RR'} \hat{b}^n_{RR'},
\]

\[
+ \frac{1}{2} \sum_{mnop} U_{RR'R''}^{mnop} \hat{b}^m_{RR'} \hat{b}^o_{RR''} \hat{b}^n_{RR'} \hat{b}^p_{RR''},
\]

with hopping and interaction parameters

\[
i_{RR'}^{mn} = -\int dr w^m_R(r) \bar{w}^n_R(r),
\]

\[
U_{RR'R''}^{mnop} = g \int dr w^m_R(r) w^o_R(r) \bar{w}^n_R(r) \bar{w}^p_R(r).
\]

Due to Eq. (1), these parameters are invariant under a simultaneous translation in the direct lattice vectors that label them.

The Hamiltonian simplifies in two ways. First, for sufficiently small kinetic \( E_{kin} \) and interaction energies \( E_{int} \approx \rho g \), we can ignore all but some number \( J \) of the bands. Second, \( w^n_R(r) \) are chosen such that they are well localized, meaning that the \( i_{RR'}^{mn} \) and \( U_{RR'R''}^{mnop} \) corresponding to hopping or interaction between distant states are negligible. This leaves the Hubbard model

\[
\hat{H}_{HM} = -\sum_{mn=1} J \sum_{(RR')} i_{RR'}^{mn} \hat{b}^m_{RR'} \hat{b}^n_{RR'},
\]

\[
+ \frac{1}{2} \sum_{mnop=1} J \sum_{(RR'R''R''')} U_{RR'R''}^{mnop} \hat{b}^m_{RR'} \hat{b}^o_{RR''} \hat{b}^n_{RR'} \hat{b}^p_{RR''},
\]

where the angular brackets indicate that the sum is restricted to local terms, e.g., same-site, nearest-neighbor, or next-nearest-neighbor etc. The range of the terms that need to be kept will depend on how local the \( w^n_R(r) \) can be, which in turn is dependent on the form of the potential \( V(r) \).
B. Generalized Wannier states

We now turn our attention to the choice of wavefunctions \( \psi_{\mathbf{R}}^m(\mathbf{r}) \) used in the above procedure. A complete basis of orthonormal functions is provided by the Bloch states \( |\psi_m^k\rangle \), corresponding to eigenfunctions of \( \hat{H} \)

\[
|\psi_m^k(\mathbf{r})\rangle = e^{i\mathbf{k} \cdot \mathbf{r}} u_m^k(\mathbf{r}),
\]

where \( u_m^k(\mathbf{r}) \) are cell-periodic functions \([16]\). The Bloch states of a given band \( m \) are uniquely labeled by a wave-vector \( \mathbf{k} \) that runs over the first Brillouin zone of the reciprocal lattice. Any band \( m \) with energies

\[
E_m^k = \langle \psi_m^k \vert \hat{H} \vert \psi_m^k \rangle,
\]

satisfying \( E_m^k \gg E_{\text{kin}}, E_{\text{int}} \) for all \( \mathbf{k} \) will not contribute to the physics and may be ignored. For all optical lattice potentials we consider here it is possible to focus solely on a small number \( J \) of the lowest-energy bands which may be degenerate amongst themselves but are separated in energy from the others.

To describe local interactions within this \( J \)-band subspace, a good choice of basis are states of the form

\[
|\mathbf{R}n\rangle = \frac{Y}{(2\pi)^D} \int_{BZ} d\mathbf{k} e^{-i\mathbf{k} \cdot \mathbf{R}} \sum_j U_{mn}^j |\psi_m^j\rangle,
\]

where \( Y \) is the volume of the primitive cell of the \( D \)-dimensional direct lattice, and \( U^{(k)} \) is a unitary matrix that mixes the Bloch bands. In the case that \( U^{(k)} \) is diagonal, i.e., there is no band mixing, these states are exactly those first considered by Wannier \([6]\). Thus the states appearing in Eq. \( 2 \) are commonly referred to as generalized Wannier states.

The separation in energy of the \( J \) lowest bands from the others ensures that some states \( |\mathbf{R}n\rangle \) exist with mode functions \( \psi_{\mathbf{R}}^m(\mathbf{r}) \) that are exponentially localized at lattice site \( \mathbf{R} \) in coordinate space \([17][21]\). This exponential localization occurs if and only if the Bloch superpositions

\[
|\tilde{\psi}_n^{(k)}\rangle = \sum_{m=1}^J U_{mn}^{(k)} |\psi_m^{(k)}\rangle,
\]

are analytic (infinitely differentiable) in \( \mathbf{k} \) across the whole Brillouin zone \([22]\). This is a rigorous way of saying that only smoothed-out Bloch superpositions will lead to localization when Fourier transformed. When there are no degeneracies between bands can one simply use the phases of elements of a diagonal \( U^{(k)} \) (representing the freedom in the phase of each \( |\psi_m^{(k)}\rangle \)) to ensure the smoothness of the Bloch states \( |\tilde{\psi}_m^{(k)}\rangle \). Hence simple Wannier states provide an exponentially localized basis in such cases. However, this is no longer the case when degeneracies and crossings in the band structure lead to non-analytic \( |\psi_m^{(k)}\rangle \). In this situation band mixing and therefore a non-diagonal \( U^{(k)} \) are required to obtain smooth Bloch superpositions and an exponentially-localized basis. The ‘only if’ case highlights the importance of the generalization of Wannier states to include non-diagonal \( U^{(k)} \).

Even when exponential localization is possible using simple Wannier states, generalized Wannier states may still significantly improve the localization. We will give examples of this in Sec. \( III \).

C. Maximally-localized generalized Wannier states

Generalized Wannier states therefore have the potential to provide a well-localized basis for the derivation of a Hubbard model. However, generalized Wannier states are highly non-unique and so it remains to find and choose a single exponentially-localized basis.

Several criteria have been proposed as a means of selecting a specific basis of generalized Wannier states \([10][23][24]\). Here, following Ref. \([10]\), we seek the generalized Wannier states with a minimal combined spatial variance, henceforth called spread, defined as

\[
\Omega = \sum_{n=1}^J \left[ \langle 0|n|\mathbf{r}^2|0\rangle - \langle 0|0|\langle 0|\hat{\mathbf{r}}|0\rangle^2 \right] = \sum_{n=1}^J \left[ \langle \mathbf{r}^2 \rangle_n - \mathbf{r}_n^2 \right] = \sum_{n=1}^J \Omega_n. \tag{4}
\]

The minimizing states are called the maximally-localized generalized Wannier states.

It is known that the maximally-localized generalized Wannier states are indeed exponentially localized \([17][19][25][26]\). They do not necessarily provide the optimal approximation to \( \hat{H} \) when restricting the number of terms kept in \( \hat{H}_{\text{int}} \) but can be expected to be close to optimal. Finding a set of generalized Wannier states which makes the Hubbard model approximation optimal is challenging. Hence minimizing the spread of the generalized Wannier functions is both an effective and practical choice with the added benefit of having a straightforward interpretation in terms of a particle occupying a specific lattice site. Note that it has been hypothesized that the maximally-localized generalized Wannier states always correspond to real functions, up to a global phase, when dealing with an isolated group of \( J \) bands \([20][26][27]\). All of our calculations support this hypothesis.

II. METHOD

To obtain maximally localized Wannier states we must first calculate the band structure \( E_m^k \) and Bloch states \( |\psi_m^k\rangle \). Such calculations are well-understood and we include our procedure here only for completeness and to introduce notation required later. Then we must find the gauge \( U^{(k)} \) such that the resulting generalized Wannier states \( |\mathbf{R}n\rangle \), defined by Eq. \( 2 \), minimize the spread, defined by Eq. \( 4 \). This usually consists of several steps: initially, we choose the gauge \( U^{(k)} = \mathbb{I}_J \) to be the \( J \times J \) identity matrix. Then the gauge is transformed iteratively \( U^{(k)} \rightarrow U^{(k)} V^{(k)} \) according to a unitary \( V^{(k)} \). These transformations accumulate until they converge to the desired gauge \( U^{(k)} \), corresponding to the minimum spread. From this, the maximally-localized generalized Wannier states may be
calculated together with the hopping and interaction parameters.

In the strategy devised by Mazari and Vanderbilt \cite{10} an initial unitary $V^{(k)}$ is constructed via projections of $J$ localized trial orbitals onto the Bloch states. This transformation leads to a gauge $U^{(k)} = V^{(k)}$ corresponding to an analytic set of Bloch superpositions $|\tilde{\psi}_n^{(k)}\rangle$ and thus exponentially-localized generalized Wannier states $|\Omega_n\rangle$. Subsequently, other unitary transformations $V^{(k)}$ are iteratively applied as part of a steepest-descent algorithm, in the hope that the cumulative gauge $U^{(k)}$ converges towards the spread-minimizing gauge.

Unfortunately when applying this strategy to investigate common optical-lattice potentials, we found that the gauge corresponding to the maximally-localized generalized Wannier states is rarely obtained. Instead the spread often converges to some non-global minimum. Therefore we adopt a different initialization procedure, to precede the same steepest-descent algorithm. Contrastingly, we find that, for the optical-lattice potentials considered, our strategy converges quite consistently to the global minimum, and thus the maximally-localized generalized Wannier states are reliably obtained. While proving convergence to the global minimum is difficult, we test numerically both that the same solution is found when starting from two Bloch states differing by random permutations of the bands at each $k$, and that the final generalized Wannier states are real.

In this section we begin by outlining the band structure calculation, and the representation of quantities, such as the spread, in reciprocal space. We then briefly describe Mazari and Vanderbilt’s steepest-descent algorithm, before outlining two methods we will use as our initialization procedure. To end the section, we describe our full procedure for calculating the maximally-localized generalized Wannier states.

A. Band structure

We work with the Fourier space representation of the potential and cell-periodic functions

$$v^{(G)}(r) = \frac{1}{\sqrt{V}} \int_{PC} dr \, V(r) e^{-iG \cdot r},$$

$$c_m^{(k,G)} = \frac{1}{\sqrt{V}} \int_{PC} dr \, u_m^{(k)}(r) e^{-iG \cdot r},$$

where the integral is over a primitive cell of the direct lattice. In this representation the single-particle Schrödinger equation $\hat{h} |\psi_m^{(k)}\rangle = E_m^{(k)} |\psi_m^{(k)}\rangle$ may be written as

$$\frac{1}{2M} (G + k)^2 c_m^{(k,G)} + \frac{1}{\sqrt{V}} \sum_{G'} v^{(G-G')} c_m^{(k,G')} c_m^{(k,G')} = E_m^{(k)} c_m^{(k,G)} . \tag{5}$$

The full band structure is obtained by solving this equation for all wave-vectors $k$ in the Brillouin zone and all reciprocal lattice vectors $G$ \cite{16}.

To make this calculation tractable on a computer, we firstly truncate the Fourier expansions to include some finite number $N$ of terms, corresponding to reciprocal lattice vectors $G$ with magnitudes $|G|$ less than $G_{\text{max}}$, as shown in Fig. 1(a). Then we only solve Eq. (5) for a $D$-dimensional uniform discrete mesh of $H^D$ wave-vectors $k = G/H$ contained within some primitive cell of the reciprocal lattice, and interpolate between these wave-vectors. As shown in Fig. 1(b), this primitive cell need not be the first Brillouin zone since the corresponding Bloch states are invariant when translated by a reciprocal lattice vector into the Brillouin zone. For each $k$ in the mesh, solving the set of Eqs. (5) then reduces to an eigenvalue problem. The justification of the truncation and mesh discretization, as well as the values of $G_{\text{max}}$ and $H$ we use are discussed in Appendix A.

Note that various symmetries guarantee certain properties of the coefficients $c_m^{(k,G)}$ \cite{28}. For a given $m$ and $k$, time-reversal symmetry implies that we may choose $c_m^{(k,G)} = c_m^{(k,G)^*}$. The addition of inversion symmetry allows us to set $c_m^{(k,G)}$ and $v^{(G)}$ as real up to a common phase factor and ensures $v^{(G)} = v^{(-G)}$. This implies that, when both symmetries are present, we may both reduce our mesh of reciprocal lattice vectors by nearly half, as the coefficients for $-G$ may be inferred from those for $G$, and restrict all quantities in the eigenvalue equation (5) to be real, thus speeding up computations for each $k$.

B. Contributions to the spread

Following Mazari and Vanderbilt, the spread $\Omega = \Omega_t + \tilde{\Omega}$ can be conveniently decomposed into two positive definite parts, $\Omega_t$ and $\tilde{\Omega}$. The latter depends on the choice of gauge $U^{(k)}$ appearing in Eq. (2), while the former does not. The gauge-independent part $\Omega_t$ depends only on the smoothness in $k$-space of the underlying manifold of Bloch states, while the gauge-dependent part $\tilde{\Omega}$ depends on the additional smoothing achieved by applying phases to and mixing the Bloch states. In preparation for what follows, it is useful to further

![FIG. 1: (Color online) Representing the problem in $k$-space. (a) The reciprocal lattice points for a two-dimensional oblique lattice. The black circle has a radius of $G_{\text{max}}$ and is centered on $\Gamma$. Fourier components corresponding to reciprocal lattice points within the circle (red points) are included in the truncated basis set, while those outside (blue points) are not. (b) A mesh of wave-vectors $k$ for a hexagonal two-dimensional lattice used to interpolate values of functions of $k$ over the Brillouin zone, which is shown by the black hexagon.](image-url)
decompose $\Omega_I = \Omega_{I,D} + \Omega_{I,OD}$ and $\tilde{\Omega} = \Omega_{D} + \Omega_{OD}$ into band-diagonal and band-off-diagonal terms. For the diagonal terms $\Omega_{I,D} = \sum_n \Omega_{I,D}^n$ and $\Omega_D = \sum_n \Omega_D^n$, it makes sense to break them down into positive-definite contributions from each band. The decomposition is expressed neatly as

$$
\Omega = \sum_n \Omega_{I,D}^n \Omega_{I,OD} + \sum_n \Omega_D^n \Omega_{OD}.
$$

Our minimization method will of course leave $\Omega_I$ invariant, while minimizing $\tilde{\Omega}$. The Mazari and Vanderbilt steepest-descent algorithm iteratively minimizes $\tilde{\Omega}$ directly, while our initialization procedure is divided into two stages, one which reduces $\Omega_{OD}$ and another which minimizes $\Omega_{I,D}$.

In terms of generalized Wannier states, the contributions to the spread are written

$$
\begin{align}
\Omega_{I,D}^n &= \langle 0n|\hat{\mathbf{r}}^2|0n \rangle - \sum_{\mathbf{R}} \langle 0n|\hat{\mathbf{r}}|\mathbf{R}n \rangle^2, \\
\Omega_{I,OD} &= -\sum_{n,m \neq n} \sum_{\mathbf{R}} \langle 0m|\hat{\mathbf{r}}|\mathbf{R}n \rangle^2, \\
\Omega_D^n &= \sum_{\mathbf{R} \neq 0} \sum_{n,m \neq n} \langle 0m|\hat{\mathbf{r}}|\mathbf{R}n \rangle^2, \\
\Omega_{OD} &= \sum_{n,m \neq n} \sum_{\mathbf{R}} \langle 0m|\hat{\mathbf{r}}|\mathbf{R}n \rangle^2.
\end{align}
$$

Again, for computational tractability, we move to the truncated Fourier representation with a discretized mesh. In this, all integrals over the Brillouin zone are replaced by summations over the mesh,

$$
\frac{1}{(2\pi)^D} \int_{BZ} \frac{dk}{M^D} = \sum_{\mathbf{k}},
$$

and gradients represented by finite differences (the gradients in reciprocal space arise from moments in position space). We use the finite-difference expressions recommended by Marzari and Vanderbilt [10], which have the property of transforming correctly under translations of the generalized Wannier states by a direct lattice vector. In this way, contributions to the spread are re-expressed as

$$
\begin{align}
\Omega_{I,D}^n &= \frac{1}{M^D} \sum_{\mathbf{k},\mathbf{b}} \omega_b \left(1 - |M_{mn}^{(k,b)}|^2\right), \\
\Omega_{I,OD} &= \frac{1}{M^D} \sum_{\mathbf{k},\mathbf{b}} \omega_b \sum_{n,m \neq n} |M_{mn}^{(k,b)}|^2, \\
\Omega_D^n &= -\frac{1}{M^D} \sum_{\mathbf{k},\mathbf{b}} \omega_b \left[\text{Im}[\ln|M_{mn}^{(k,b)}|] + \mathbf{b} \cdot \mathbf{r}_n\right]^2, \\
\Omega_{OD} &= \frac{1}{M^D} \sum_{\mathbf{k},\mathbf{b}} \omega_b \sum_{n,m \neq n} |M_{mn}^{(k,b)}|^2.
\end{align}
$$

Here the vectors $\mathbf{b}$ connect each wave-vector $\mathbf{k}$ to its nearest-neighbors, $\omega_b$ are factors that depend on the geometry of the mesh [29], and

$$
\mathbf{r}_n = \frac{1}{M^D} \sum_{\mathbf{k},\mathbf{b}} \omega_b \mathbf{b} \text{Im}[\ln|M_{mn}^{(k,b)}|].
$$

It is clear then that all the information about the spread is contained in the matrix elements

$$
M_{mn}^{(k,b)} = \sum_{\mathbf{G}} U_{mn}^{(k,b)} c_p^{(k,G)} c_{m}^{(k+b,G)} U_{in}^{(k)},
$$

which are the truncated Fourier representation of the overlap $\langle \hat{u}_m^{(k)} | \hat{u}_n^{(k+b)} \rangle$, where similarly to Eq. (3),

$$
|\hat{u}_n^{(k)}\rangle = \sum_{m=1}^{J} U_n^{(k)} |u_m^{(k)}\rangle,
$$

with $|u_m^{(k)}\rangle$ the state associated with periodic function $u_m^{(k)}(\mathbf{r})$. These elements are initialized to $M_{mn}^{(k,b)} = \sum_{\mathbf{G}} c_p^{(k,G)} c_{m}^{(k+b,G)}$ when $U(k) = \mathbb{I}_f$. Then under a gauge transformation $U(k) \rightarrow U(k) V(k)$ they undergo the computationally simple transformation $M^{(k,b)} \rightarrow V(k)^T M^{(k,b)} V(k+b)$.

C. Minimizing total spread

The gradient $\Gamma^{(k)} = d\Omega/dW^{(k)}$, embodying the change in spread due to a gauge transformation $V(k) = e^{dW^{(k)}}$, with $dW^{(k)}$ an infinitesimal anti-Hermitian matrix, can be efficiently calculated from the matrices $M^{(k,b)}$ (see Ref. [10] for details). The steepest-descent approach, as used in Refs. [10], [13], then implements the gauge transformation $V^{(k)} = e^{dW^{(k)}}$ with $dW^{(k)} = -\varepsilon \Gamma^{(k)}$ and $\varepsilon$ a small positive number. These steps are repeated until convergence is achieved.

To a large extent, the steepest-descent algorithm is only as good as its initialization procedure, since starting from an arbitrary set of generalized Wannier states, the algorithm is likely to drive the set towards one of the many local minima in the spread, rather than the global minimum. We do not discuss here the commonly used projection-based initialization procedure (see Ref. [11] for information on this) that we found to struggle for optical-lattice potentials. Instead we now discuss two other approaches for reducing the spread, which together will form the initialization procedure we use successfully for optical-lattice potentials.

D. Reducing inter-band spread

We break down the task of finding the maximally-localized Wannier states into two stages. The first stage is to mix the bands to create a new set of pseudo-bands from which a maximally-localized ordinary Wannier states calculation is optimal (i.e. leading to the smallest possible spread). The second stage is to calculate the maximally-localized ordinary Wannier states using these pre-mixed bands as a starting point. The first stage corresponds to minimizing the off-diagonal contribution $\Omega_{OD}$, and the second to minimizing the diagonal term $\Omega_D$. Our initialization procedure is split accordingly: first we reduce (but not necessarily minimize) $\Omega_{OD}$, as described in this subsection; second we minimize $\Omega_D$, as described in the next subsection.
Our first goal is then to reduce the band-off-diagonal term $\Omega_{OD}$, which is equivalent to reducing $\Omega_{ID}$. This equivalence is clear from the interpretation above, that reducing $\Omega_{OD}$ corresponds to optimizing the bands from which to perform a maximally-localized ordinary Wannier states calculation. Mathematically, it follows from observing that the band-off-diagonal parts of the gauge-invariant spread $\Omega$ and the gauge-dependent spread $\Omega^D$ are negative of each other (see Eqs. (20), (21) and (22), (23)): reducing $\Omega_{OD}$ is achieved by increasing the band-off-diagonal part $\Omega_{1,OD}$ of the gauge-invariant spread or, equivalently, reducing its diagonal part $\Omega_{1,D}$.

To reduce $\Omega_{1,D}$, we use a method devised by Souza, Mazari and Vanderbilt 13. For $K$ degenerate bands, their minimizes the contributions to $\Omega$ from a subset $K' < K$ bands obtained through a unitary mixing of these bands. The aim of this approach is then to construct the $K'$ bands with the smoothest $k$-space such that they provide the optimal set of $K'$ bands from which to construct localized generalized Wannier states (optimal in the sense of having the smallest possible gauge-invariant contribution to the spread).

We use their approach to reduce $\Omega_{1,D}$ in the following way: First, we use the Souza et al. method to minimize $\Omega_{1,D}$ and therefore construct, from the $K = J$ bands, a single ($K' = 1$) band whose smoothness in $k$-space is optimum for constructing a localized Wannier state. Then, keeping this band fixed, we use the Souza et al. method again to minimize $\Omega_{2,D}$ and construct from the $K = J - 1$ remaining bands a single ($K' = 1$) band that is optimum for constructing a localized Wannier state. This is repeated in a similar fashion to obtain a third, fourth, etc. band until finally we use the Souza et al. method to minimize $\Omega_{J,D}$ and construct an optimized ($J - 1$)-th band out of the two remaining bands, with all lower bands fixed. Our approach therefore consists of $J - 1$ applications of the Souza et al. method, in each case optimally extracting a single ($K' = 1$) band from $K = J, J - 1, \ldots, 2$ others.

Note that this does not necessarily minimize $\Omega_{ID}$ and therefore $\Omega_{OD}$, but we find that following this procedure $\Omega_{OD}$ is very small. Details of the Souza et al. method and our use of it can be found in Ref. 13 and Appendix B respectively. Here we simply note that the Souza et al. method proceeds via several iterations, each of which applies a transformation $V(k)$ over all $k$-space that would have minimized the contribution to the spread from any given point in $k$-space had there been no transformation applied at the other points in $k$-space. The desired gauge must be left unchanged by such an iteration and thus it is a possible point of convergence. To protect against false convergences, we initialize the whole procedure above by applying a transformation $V(k)$, where at each $k$ we take a $J \times J$ identity matrix and randomly permute its rows. We find that in practice, following this initialization, the desired gauge is nearly always obtained.

E. Reducing intra-band spread

We next present a method that reduces the intra-band contribution $\Omega_D$ to the spread, while leaving the inter-band contribution $\Omega_{OD}$ invariant. To ensure this invariance, in this section we restrict ourselves to gauge transformations $V(k)$ that are diagonal, i.e., while we allow changes to the phases of the Bloch superpositions $|\tilde{\psi}_n(k)\rangle$, we do not allow any transformations that further mix the bands. Hence the task splits into $J$ independent parts, each to reduce $\Omega_D^I$ by applying phases to the Bloch superpositions $|\tilde{\psi}_n(k)\rangle$. One may interpret this as constructing the maximally-localized ordinary Wannier states from a set of bands comprising the mixed Bloch states $|\tilde{\psi}_n(k)\rangle$.

Such single-band tasks are usually described in terms of the Berry connection $A(k) = i\langle \tilde{\psi}_n(k)|\tilde{\psi}_n(k)\rangle C$. Integrals (Berry phases $\varphi_C$) of the Berry connection around closed paths $C$ in the Brillouin zone are invariant under changes to the phases of the Bloch states $|\tilde{\psi}_n(k)\rangle$. This implies that $\tilde{\psi}_n$, equal to the average value of $A(k)$ across the Brillouin zone, is also invariant $\tilde{\psi}_n$. A further invariant quantity is given by $B = \nabla \times A(k)$, called the Berry curvature. Local values of $A(k)$, however, depend on the phases of the Bloch states, which determine the phase-dependent part of the spread $\Omega_D^I$. It is known that this spread is minimized when the divergence of the connection vanishes, $\nabla \cdot A(k) = 0$, and the minimum possible spread depends only on the Berry curvature $B$.

In particular, if $B = 0$, then the minimum possible spread $\Omega_D^I$ is zero. It follows that all Berry phases are zero and it is possible to smooth $A(k)$ such that it is uniform, at which point $\Omega_D^I = 0$. To smooth the connection $A(k)$, we use a progressive phase update method: it consists of taking a succession of closed loops through the Brillouin zone, and, for each, altering the Bloch phases at points along the loop such that the projections of $A(k)$ along it are constant. This constant value is fixed by their integral around the loop, the Berry phase, which is invariant. Adjusting the phases in this way for several loops, given in Appendix C will result in a flattened connection, if possible.

For non-zero $B$, absolute uniformity of the connection $A(k)$ is not possible. However, in an attempt to suppress the divergence of the connection and therefore approach the minimum spread $\Omega_D^I$, we still choose to smooth out $A(k)$ using the progressive phase update method and find this greatly reduces $\Omega_D^I$. To achieve the minimum, we follow the progressive phase updates with the steepest-descent minimization algorithm of Marzari and Vanderbilt (cf. Sec. IIIC), when only terms corresponding to $\Omega_D^I$ contribute to the gradient $\Gamma^I$.

A particular case of interest is a system with inversion symmetry and a current gauge $U(k)$ that is diagonal, i.e., the bands have not been mixed. As a result of the symmetry, the mode functions $\tilde{u}_n(r) \propto \tilde{u}_n(r)$ must be real up to a global phase, at which point the Berry curvature and thus the spread $\Omega_D^I$ vanishes. Note that optical-lattice potentials usually possess inversion symmetry since this is inherited from the lasers that created them; superlattice potentials are required to break this.

We found that even if $U(k)$ is not diagonal, e.g., after the inter-band spread is reduced, the output of the disentangling procedure often still had zero Berry curvature for each band and the progressive phase update method reduced $\Omega_D^I$ to zero.
We hypothesize this is a general feature, also hinted at in the results of Refs. [20, 26, 27]. Specifically, this occurred whenever the degeneracies in our J-band subspace were a result of purely geometric symmetries. We hypothesize this is a general feature, also hinted at in the results of Refs. [20, 26, 27].

**F. Full procedure**

Having described the elements of our computational approach, we now describe how they are pieced together. The full procedure for calculating the maximally-localized generalized Wannier states is shown in Fig. 2.

First we calculate the band structure. Then we minimize the intra-band spread via a progressive phase update, followed by a restricted version of the steepest-descent method (for potentials with inversion symmetry, the steepest-descent part is unnecessary). At this point we are at the gauge corresponding to the interpolated band-structure to the original. Note that this only allows us to determine the accuracy of the hopping parameters, not the interaction terms. Since we are unable to directly verify the accuracy of discarding interaction parameters, we only discard those of magnitude equal to, or less than, that of the discarded hopping parameters for some typical range of interaction strengths $g \lesssim \tilde{g} = E_R \lambda^D$, where

$$E_R = \frac{\hbar}{2 \mu \lambda^2},$$  \hspace{1cm} (9)

is the recoil energy, and $\lambda$ is the ‘averaged’ wavelength of the laser beams creating the optical lattice [40].

We now obtain the maximally-localized Wannier states and related Hubbard parameters for atoms in several one- and two-dimensional optical-lattice potentials. We leave three-dimensional potentials for a future presentation.

**A. One-dimensional systems**

To begin, we find the maximally-localized generalized Wannier states and related Hubbard parameters for bosons in a one-dimensional superlattice potential, given by

$$V(x) = V_0 \left[ (1-s) \sin^2(2\pi x/\lambda) + s \sin^2(4\pi x/\lambda) \right].$$
Such a potential can be produced using two independent pairs of laser beams, each red-detuned from wavelength $\lambda$ and at an angle to each other, as shown in Fig. 3(a). Their total intensities determine the potential depth $V_0$ and their relative intensities determine the superlattice parameter $0 \leq s < 1$. The lattice parameter is $\lambda/2$. This system has been experimentally realized in Refs. [32][33], and was proposed in Ref. [34] as a method for initialising a quantum register on a time-scale that is an order of magnitude smaller than the conventional quantum-freezing of a superfluid to a Mott insulator.

The potentials for three values of $s$ are shown in Fig. 3(b), where $V_0 = 20E_R$. For $s = 0$ the potential is sinusoidal with a minimum at the center of each primitive cell. For $s \neq 0$ there are two minima in each primitive cell, which move either side of the center. As $s \rightarrow 1$ the potential approaches a sinusoid with lattice parameter $\lambda/4$.

The band structures for the same parameters are shown in Fig. 3(c). For all $0 \leq s < 1$ the two lowest lying bands are well separated from the higher bands, and also are not degenerate amongst themselves. Therefore, ordinary Wannier states will provide an exponentially-localized basis. We use this superlattice potential then to demonstrate that using generalized Wannier states can further localize the Wannier states even when there are no inter-band degeneracies.

We expect the benefits of generalized over ordinary Wannier states to be most notable in the $s \rightarrow 1$ limit, where the bands are close together. The maximally-localized generalized Wannier states and ordinary Wannier states for the case $s = 0.999$ and $J = 2$ are presented in Figs. 4(a) and (b), respectively. It is clear from inspection that the generalized Wannier states are more localized. This improved localization occurs for even small $s$ but is very significant for moderate or large $s \gtrsim 0.5$, as is shown in Fig. 4(c), which plots the respective spreads as a function of $s$.

Another way to see the effects of improved localization is to look at the magnitudes of the hopping and interaction parameters. These are shown in Fig. 5(hopping) and Fig. 6(interaction), as a function of $s$, for both the ordinary and generalized Wannier states. For large $s \gtrsim 0.5$, non-local Hubbard parameters are significantly reduced when using generalized Wannier states. This comes at the expense of allowing inter-band hopping.

From these values it is clear that using either ordinary or generalized Wannier states, a tight-binding Hamiltonian

$$\hat{H}_{\text{TB}} = \sum_j \sum_{n=1}^2 \left\{ -t_{nn}^{00} \hat{b}_j^\dagger \hat{b}_{j+1}^{n0} + t_{nn}^{01} \hat{b}_j^\dagger \hat{b}_{j+1}^{n1} + \text{h.c.} \right\}$$
can be derived and justified from first principles. Here, for clarity, we have replaced the label \( R \) by the label \( j = 2|\mathbf{R}|/\lambda \). The model derived using generalized Wannier states is more accurate, as we can demonstrate by comparing interpolated bands to the original. In Fig. 7(a) this is shown for both maximally-localized ordinary and generalized Wannier states, and superlattice parameter \( s = 0.999 \). The generalized Wannier states almost exactly reproduce the band structure, while there are significant deviations for the Hamiltonian derived using ordinary Wannier states. In Fig. 7(b) we show the standard deviation, i.e., the average root mean squared error of the energies averaged over the bands, between the interpolated bands and exact bands as a function of \( s \). This demonstrates that the difference in accuracy between using ordinary and generalized Wannier states is appreciable for \( s \gtrsim 0.5 \). In fact, we should have expected this from the non-sinusoidal nature of the bands for large \( s \). A tight-binding model built from ordinary Wannier states can only ever result in a sinusoidal band structure. Generalized Wannier states and inter-band hopping they describe have no such restriction.

These results confirm that for \( s \gtrsim 0.5 \), the accuracy of the local model found using generalized Wannier states becomes significantly better than that using ordinary Wannier states. The reason for this difference is that the two generalized Wannier states can each break the reflection symmetry in the primitive cell to localize around a different minimum (see Fig. Fig. 8(a)). Meanwhile the maximally-localized ordinary Wannier states cannot break this symmetry and instead are symmetric and antisymmetric combinations of two functions localized at each of the minima (see Fig. Fig. 8(b)). Aside from leading to more accurate local Hubbard models, the use of generalized Wannier states are more relevant for cold atom experiments. In such experiments, it is the presence of a particle at a position in space rather than the symmetry of its wavefunction that is measured through high-resolution imaging [3]. Thus a Hubbard model corresponding to atoms in spatially-separated sites is preferable to atoms in symmetric/antisymmetric superpositions. We similarly expect generalized Wannier states to be important for other lattices that possess more than one potential minimum per primitive cell.

\[ \frac{1}{2} U_0 \hat{b}_j^{\dagger} \hat{b}_j + \sum_{m=1,m \neq n}^{2} \left\{ -t_{mn}^{\mathrm{nn}} \hat{b}_j^{\dagger} \hat{b}_j + t_{mn}^{\mathrm{nn}} \right\} \]

B. Two-dimensional systems

The use of generalized Wannier states is paramount in two dimensions, as degeneracies in the lowest bands are likely to occur as a result of crystallographic point-group symmetries. Hence the maximally-localized ordinary Wannier states could fail to provide an exponentially localized basis due to the resulting non-analyticity of the bands. Further, we will see cases where the maximally-localized generalized Wannier states are not centered around inversion points of the lattice, and do not share the symmetry of the lattice. In these cases, approximating the states using Gaussian functions would lead to a particularly inaccurate estimate of the Hubbard parameters.

To showcase our procedure we now calculate accurately and from first principles the maximally-localized generalized Wannier states and Hubbard parameters for atoms in an optical lattice, with either hexagonal or Kagomé geometries. Both potentials have multiple minima per primitive cell and lead to a degenerate set of lowest bands, thus representing a significant challenge using any other method. Both of these structures also play an important role in condensed-matter physics, see e.g. Refs. [8, 36–39].

1. Hexagonal lattice

Three blue-detuned beams of approximately equal wavelength \( \lambda \), shown in Fig. 8(a), generate a hexagonal optical-
lattice potential, written as
\[ V(x, y) = \frac{V_0}{g} \left[ 3 + 2 \cos \left( \frac{2\sqrt{3}\pi y}{\lambda} \right) \right. \]
\[ \left. + 4 \cos \left( \frac{3\pi x}{\lambda} \right) \cos \left( \frac{\sqrt{3}\pi y}{\lambda} \right) \right], \]
and plotted in Fig. 8(b). The potential exhibits two minima per unit cell, positioned at cell vertices that form a hexagonal (honeycomb) structure. The consequence of there being two potential minima per unit cell is that the two lowest bands are degenerate at the K points of the Brillouin zone, as shown in Figs. 8(c) and (d) for lattice depths \( V_0 = 10E_R \) and \( V_0 = 30E_R \), respectively.

The maximally-localized generalized Wannier states for the two lowest bands, using \( J = 2 \), are shown in Fig. 9 for a lattice depth \( V_0 = 10E_R \). Both states possess three-fold rotational symmetry about their centers and are images of one another through a rotation of \( 60^\circ \) about the center of the Wigner-Seitz unit cell. As for the one-dimensional superlattice, both generalized Wannier states are localized around a potential minimum, rather than at the Wyckoff positions (centers of inversion). As a result \( \Omega_D \neq 0 \) for the pair although the total spread is minimized and the cell-periodic superposed states (cf. Eq. 5) are real. Since inversion symmetry is broken it is clear that a Gaussian function would not adequately describe these Wannier states even in the deep lattice limit.

The magnitudes of the hopping and interaction parameters for the two lowest bands are shown in Fig. 10. Since the maximally-localized generalized Wannier states are related through a symmetry operation the parameters within each band are identical. We therefore label the parameters not by site and band, but by \( j \), the rank of the distance between potential minima, as shown in Fig. 2. Parameters for \( j = 0, 2 \) are intra-band, while those for \( j = 1, 3, 4 \) are inter-band. We calculate these parameters up to a lattice depth of \( V_0 = 200E_R \). We observe that for large \( V_0 \) the significant parameters are the on-site interaction parameter \( U_0 \) and the hopping parameter \( t_1 \). The interaction parameter \( U_1 \) corresponding to the interaction

The Kagomé lattice has received a large degree of interest in recent years because it leads to a highly-frustrated many-body Hamiltonian \[ \{7, 8, 38, 39\} \]. This lattice may be created using
FIG. 13: (Color online) Hopping and density-density interaction parameters for the Kagomé lattice. (a) The magnitudes $|t_j| = |t_{j0}^{mm}|$ of the hopping parameters, as a function of lattice depth $V_0$, where the centers of $[m0]$ and $[00]$ are the $j$-th smallest distance from each other (cf. Fig. 12). The black solid, blue dotted, red dashed, green dot-dashed, and magenta dot-long dashed lines are for $j = 0, 1, 2, 3, 4$ respectively. (b) Similarly for the magnitudes $|U_j| = |U_{j000RR}^{mmnm}|$ of the interaction parameters. (c) The total standard deviation $\sigma$ between the exact lowest bands and the interpolated tight-binding bands as a function of lattice depth.

We have calculated, from first principles, the parameters of nearest-neighbour Hubbard models for several optical lattice potentials, including the honeycomb and Kagomé potentials, demonstrating quantitatively for which lattice depths such models are accurate. Strongly-correlated phenomena probed in optical lattice experiments and quantum simulations depend delicately on the ratios of kinetic and interaction energies. Therefore precisely determining them ab initio, as done here, is essential for diagnosing and interpreting such experimental results and for using optical lattices as quantum simulators.

To perform our calculations we have developed a freely available software package [14] that, given an optical lattice potential, will efficiently calculate the corresponding maximally-localized generalized Wannier states without any prior-knowledge of their form in any spatial dimension. This will allow cold-atom researchers to easily and accurately determine Hubbard models realized by any laser setup. We hope that this tool will be useful for the optical-lattice community.

Acknowledgments

We thank the CCPQ initiative for hosting the code on the CCPForge site and the Oxford Martin School for support through the Programme on Bio-Inspired Quantum Technologies. SRC and DJ thank the National Research Foundation and the Ministry of Education of Singapore for support. This work was completed within the EuroQUAM (EP/E04162/1) project.

Appendix A: Accuracy of the truncated Fourier representation and discretized mesh

The cut-off wave-vector corresponds to a maximum kinetic energy for the plane-wave components given by $E_{\text{cut-off}} = G_{\text{max}}^2/2M$, and introduces a minimum spatial resolution $\lambda_{\text{min}} = 2\pi/G_{\text{max}}$ for describing real space functions in the system, namely, the potential, the Bloch states and the Wannier states. The minimum spatial resolution must be smaller than the spatial variations in these states in order for them to be accurately recreated using the truncated set of coefficients, therefore the cut-off energy must be at least as large as the highest energy Fourier component of the potential. One can then increase the cut-off energy until the energies for each band under consideration have converged, at which point all coefficients $c_{n}^{(k,G)}$
of significant magnitude describing the Bloch periodic functions $u_n^{(k)}(r)$ are included. Typically, this requires the cut-off energy to be an order of magnitude greater than the upper-end of the energy range of interest. In our calculations a cut-off energy of $E_{\text{cut-off}} = 50E_R$ is sufficient for band convergence and suitably limits the total number of coefficients such that even in three dimensions our procedure is not computationally expensive. Here $E_R$ is the recoil energy, defined in Eq. (9).

The discretization of vectors in reciprocal space corresponds to considering a finite real space lattice with periodic boundary conditions and $M$ primitive unit cells in each lattice direction. So we expect it to be valid when $M$ is large and surface effects are negligible.

**Appendix B: Algorithm for reducing inter-band spread**

Our method for reducing the inter-band spread involves taking, for each $n = 1, \ldots, J - 1$ in turn, the $J - n + 1$ bands $n, \ldots, J$ and constructing from them an $n$-th band that is optically smooth in $k$-space, such that the most localized Wannier state possible may be constructed for this band.

For each $n$, the algorithm, based on Ref. [15], proceeds as follows. We calculate the Hermitian matrices

$$Z_{mp}^{(k)} = \sum_b \phi_{b} M_{nn}^{b(k)} M_{pn}^{b(k)*},$$

where $m, p$ run over $n, \ldots, J$. We then apply, for every $k$ in turn, a transformation $V^{(k)} = \mathbb{I} \otimes \chi^{(k)}$, where the unitary $\chi^{(k)}$ diagonalises $Z^{(k)}$, i.e., $Z^{(k)} = \chi^{(k)} \Lambda^{(k)} \chi^{(k)^*}$, with $\Lambda^{(k)}$ diagonal, reducing the spread to

$$\Omega_{L,D}^n = \sum_b \phi_{b} - \frac{1}{M^2} \sum_k \Lambda_{11}^{(k)}.$$  

The $\Lambda^{(k)}$ are always chosen at each $k$ such that $\Lambda_{11}^{(k)}$ is the largest eigenvalue of $Z^{(k)}$, so this spread is as small as possible. The procedure in this paragraph is then applied repeatedly until convergence is achieved. The gauge for which $\Omega_{L,D}^n$ is minimized is a convergence point [13].

On occasion the above procedure can become unstable, and we prevent this by replacing Eq. (B1) by an equal weighting of $Z_{mp}^{(k)}$ calculated during the current and previous iteration. This has no effect on the locations at which the algorithm can converge.

**Appendix C: Algorithm for reducing intra-band spread**

For the progressive phase update method, we smooth the Berry connection over loops consisting of straight lines through the Brillouin zone, in the directions of the reciprocal lattice vectors. In the reciprocal mesh representation, the Berry connection at each $k$ is given by $-\sum_b \phi_{b} \text{Im}[\ln M_{nn}^{b(k)}]$ and so uniformity across a straight loop $C^{(k)}$ going through $k'$ in direction $b$ is achieved by choosing phases such that the projection $\text{Im}[\ln M_{nn}^{b(k)}]$ of the connection onto this line is the same at each point. Specifically, since integrating over the loop must give the Berry phase $\vartheta_{C^{(k)}} = \sum_{k \in C^{(k)}} -\text{Im}[\ln M_{nn}^{b(k)}]$, we set $\text{Im}[\ln M_{nn}^{b(k)}] = \vartheta_{C^{(k)}} / M$ at each point $k$ on the loop. The loops and the order in which we smooth the Berry connection across them is depicted in Fig. [14].

---

[12] See, for example, the Wannier90 algorithm, freely available at http://wannier.org.
[14] The software can be downloaded at http://ccpforge.cse.rl.ac.uk/gf/project/mlgws/.
[40] The ‘averaged’ here acknowledges that the wavelengths of the lasers used to produce the lattice must be slightly detuned from each other to avoid unwanted interference.