

## APPROXIMATE GAUSSIAN DISTRIBUTION OF THE ATOMS

In order to show how the Gaussian distribution of the atomic positions arises, we consider here an atom of mass  $m$  sitting in a one-dimensional optical trap of frequency  $\omega$ . The results will be straightforwardly generalizable to the three-dimensional case as the three Cartesian coordinates decouple and can be treated independently. We work in a regime of temperatures  $T$  much lower compared to the trap depth, but larger than the trap frequency, i.e.,  $k_B T \gg \hbar\omega$ . The first assumption allows us to treat the trap as an harmonic potential, yielding a Hamiltonian

$$\hat{H}_{\text{trap}} \approx \frac{\hat{p}^2}{2m} + \frac{m}{2}\omega^2 \hat{x}^2, \quad (1)$$

where  $\hat{x}$  and  $\hat{p}$  are the quantum position and momentum, respectively. The thermal state of the system is described by the Gibbs form

$$\rho_{\text{th}} = \frac{1}{Z} e^{-\beta \hat{H}_{\text{trap}}}, \quad (2)$$

where  $\beta = 1/(k_B T)$  and  $Z$  is the partition function

$$Z = \text{tr} \left\{ e^{-\beta \hat{H}_{\text{trap}}} \right\}. \quad (3)$$

Employing the standard mapping

$$\hat{p} = i\sqrt{\frac{\hbar m \omega}{2}}(\hat{b}^\dagger - \hat{b}), \quad \hat{x} = \sqrt{\frac{\hbar}{2m\omega}}(\hat{b}^\dagger + \hat{b}) \quad (4)$$

in terms of bosonic creation ( $\hat{b}^\dagger$ ) and annihilation ( $\hat{b}$ ) operators ( $[\hat{b}, \hat{b}^\dagger] = 1$ ), one readily obtains

$$H = \hbar\omega \left( \hat{b}^\dagger \hat{b} + \frac{1}{2} \right) \quad \text{and} \quad (5a)$$

$$Z = \sum_n e^{-\beta \hbar\omega(n+1/2)} = \frac{1}{2 \sinh\left(\frac{\beta \hbar\omega}{2}\right)}. \quad (5b)$$

Calling  $|x\rangle$  the position eigenvector  $\hat{x}|x\rangle = x|x\rangle$ , the probability density functions of the atomic position is defined as

$$p_{\text{pos}}(x) = \langle x | \rho_{\text{th}} | x \rangle. \quad (6)$$

Its analytical form can be extracted from the Feynman propagator for the harmonic oscillator  $K(x, y, t) = \langle x | e^{-it\hat{H}/\hbar} | y \rangle$ , which, in the time interval  $t \in (0, \pi/\omega)$ , reads (see, e.g., [1, 2] for detailed derivations)

$$K(x, y, t) = \sqrt{\frac{m\omega}{2\pi\hbar i \sin(\omega t)}} \times \exp \left\{ i \frac{m\omega}{2\hbar \sin(\omega t)} [(x^2 + y^2) \cos(\omega t) - 2xy] \right\}. \quad (7)$$

Substituting  $t \rightarrow -i\beta\hbar$  and  $y \rightarrow x$  one finds

$$K(x, x, -i\beta) = \langle x | e^{-\beta H} | x \rangle = \sqrt{\frac{m\omega}{2\pi\hbar \sinh(\omega\beta\hbar)}} \times \exp \left\{ -\frac{m\omega}{\hbar \sinh(\omega\beta\hbar)} (\cosh(\omega\beta\hbar) - 1) x^2 \right\}. \quad (8)$$

Dividing by the partition function (3) one finally finds the Gaussian

$$p_{\text{pos}}(x) = \sqrt{\frac{m\omega(\cosh(\omega\beta\hbar) - 1)}{\pi\hbar \sinh(\omega\beta\hbar)}} \times \exp \left\{ -\frac{m\omega}{\hbar \sinh(\omega\beta\hbar)} (\cosh(\omega\beta\hbar) - 1) x^2 \right\}. \quad (9)$$

The variance  $\sigma$  can be read off directly and amounts to

$$\sigma^2 = \frac{\hbar \sinh(\omega\beta\hbar)}{2m\omega(\cosh(\omega\beta\hbar) - 1)}. \quad (10)$$

Since we assumed  $\beta \ll \hbar\omega$ , i.e.,  $\omega\beta\hbar \ll 1$ , we can expand this expression to lowest order, which yields

$$\sigma^2 = \frac{1}{m\omega^2\beta} = \frac{k_B T}{m\omega^2}, \quad (11)$$

as reported in the main text.

The distribution (9) is straightforwardly generalized to three-dimensions and traps centered along a chain at positions  $k\mathbf{R}_0 = (0, 0, kR_0)$  with  $k$  an integer:

$$p_{\text{pos}}^{(k)}(\mathbf{r}) = \frac{1}{(2\pi)^{3/2} \sigma_1 \sigma_2 \sigma_3} e^{-\frac{r_1^2}{2\sigma_1^2} - \frac{r_2^2}{2\sigma_2^2} - \frac{(r_3 - (k-1) \cdot R_0)^2}{2\sigma_3^2}}. \quad (12)$$

For clarity, we remark here that the indices in the expression above distinguish between Cartesian components only, e.g.,  $r_1$  and  $r_2$  are the components of the same atom along the  $x$  and  $y$  directions. In the following, when necessary the trap index will always appear before the component one, e.g.,  $r_{k,i}$  is the  $i$ -th component of the  $k$ -th atom's position. For a ladder, a second set of position distributions  $p_{\text{pos}}^{(k),2}(\mathbf{r})$  would be added with the same Gaussian form up to  $r_2 \rightarrow r_2 - R_0$ .

## DISTRIBUTION OF THE DISTANCES AND INTERACTIONS FOR A SINGLE CHAIN.

Here we focus on a single one-dimensional chain as most of the properties which affect the results in the main text are due to the presence of an extended longitudinal direction. Still, the considerations made for the marginal distributions for pairs of atoms directly apply to any regular lattice configuration as well. The distribution of differences  $\mathbf{d}_k = \mathbf{r}_{k+1} - \mathbf{r}_k = (d_{k,1}, d_{k,2}, d_{k,3})$  can be found in the Supplemental Material of Ref. [3] and, for isotropic traps, reads

$$\begin{aligned}
p_{\text{diff}}(\mathbf{d}_1, \dots, \mathbf{d}_{L-1}) &= \int \left[ \prod_{k=1}^L d^3 r_k p_{\text{pos}}^{(k)}(\mathbf{r}_k) \right] \left[ \prod_{k'=1}^{L-1} \delta^{(3)}(\mathbf{d}_{k'} - (\mathbf{r}_{k'+1} - \mathbf{r}_{k'})) \right] = \\
&= \left[ \frac{\sigma^{1-L}}{\sqrt{L}(\sqrt{2\pi})^{L-1}} \right]^3 e^{-\frac{1}{2\sigma^2} \sum_{k,q} [d_{k,1} A_{kq} d_{q,1} + d_{k,2} A_{kq} d_{q,2} + (d_{k,3} - R_0) A_{kq} (d_{q,3} - R_0)]},
\end{aligned} \tag{13}$$

where  $A_{kq} = L - \max(k, q) - (L - k)(L - q)/L = (L - \max(k, q)) \min(k, q)/L$ . The correlations between different components  $d_{k,i}$  can be worked out via the inverse [4]

$$C = A^{-1} = \begin{pmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 & \cdots \\ 0 & 0 & -1 & 2 \\ & & \vdots & \ddots \end{pmatrix}, \tag{14}$$

implying,

$$\langle d_{k,i} d_{q,j} \rangle - \langle d_{k,i} \rangle \langle d_{q,j} \rangle = \sigma^2 \delta_{ij} (2\delta_{k,q} - \delta_{k,q+1} - \delta_{k,q-1}). \tag{15}$$

Subsequent distances are therefore (anti-)correlated, and these correlations pass onto any (non-trivial) function of the distances, and in particular the energy displacements  $\delta V_k = V(d_k) - V(R_0)$ .

As a consistency check, we remark that  $C(L)$  is a  $(L-1) \times (L-1)$  matrix, whose determinant satisfies the recursion relation

$$\det C(L) = 2 \det C(L-1) - \det C(L-2) \tag{16}$$

with “seed” (or initial conditions)  $\det C(2) = 2$  and  $\det C(3) = 3$ , which is solved by  $\det C(L) = L$ . Consequently, the factor  $(\sqrt{\det A})^3$  produced by the Gaussian integration over all variables exactly cancels the  $L^{-3/2}$  appearing in the normalization factor, as expected.

### Marginal distribution for a single pair of atoms

The  $\mathbf{d}_k$ s are identically distributed, so we can select any given one (and drop its index for brevity) and integrate over the other  $L-2$  variables from equation (13). This yields

$$\begin{aligned}
p_{\text{diff}}(\mathbf{d}) &= \frac{1}{(4\pi)^{3/2} \sigma^3} e^{-\frac{1}{4\sigma^2} [d_1^2 + d_2^2 + (d_3 - R_0)^2]} = \\
&= \frac{1}{(4\pi)^{3/2} \sigma^3} e^{-\frac{1}{4\sigma^2} [d^2 + R_0^2 - 2d_3 R_0]},
\end{aligned} \tag{17}$$

where  $d = |\mathbf{d}|$  denotes the distance between a pair of neighboring atoms. The distribution for this new variable can be then obtained via a solid angle integration and

reads

$$\begin{aligned}
p_{\text{dist}}(d) &= \frac{d^2}{4(\pi)^{1/2} \sigma^3} e^{-\frac{1}{4\sigma^2} (d^2 + R_0^2)} \int_0^\pi d\theta \sin \theta e^{-\frac{d R_0 \cos \theta}{2\sigma^2}} \\
&= \frac{d}{\sqrt{\pi} \sigma R_0} e^{-\frac{1}{4\sigma^2} (d^2 + R_0^2)} \sinh \left( \frac{d R_0}{2\sigma^2} \right).
\end{aligned} \tag{18}$$

The distribution of an energy shift  $\delta V$  is now just a change of variables ( $d \rightarrow d(\delta V)$ ) away, according to  $P(\delta V) = |d'(\delta V)| p_{\text{dist}}(d(\delta V))$ . For the sake of generality, we keep  $\alpha$  generic in

$$d(\delta V) = \left( \frac{C_\alpha}{V_0 + \delta V} \right)^{\frac{1}{\alpha}}, \tag{19}$$

where  $V_0 = C_\alpha/R_0^\alpha$ , which implies

$$d'(\delta V) = -\frac{1}{\alpha} \frac{C_\alpha^{1/\alpha}}{(V_0 + \delta V)^{1+1/\alpha}}. \tag{20}$$

Hence, the distribution of energy shifts for a pair is

$$\begin{aligned}
P(\delta V|V_0, R_0, \sigma) &= \frac{\frac{R_0}{\sigma}}{\alpha \sqrt{\pi} V_0 \left( 1 + \frac{\delta V}{V_0} \right)^{1+\frac{2}{\alpha}}} \times \\
&\times e^{-\frac{R_0^2}{4\sigma^2} \left[ 1 + \left( 1 + \frac{\delta V}{V_0} \right)^{-\frac{2}{\alpha}} \right]} \times \\
&\times \sinh \left[ \frac{R_0^2}{2\sigma^2} \left( 1 + \frac{\delta V}{V_0} \right)^{-\frac{1}{\alpha}} \right].
\end{aligned} \tag{21}$$

It is relatively simple to see that, if we define the dimensionless quantities  $\delta v = \delta V/V_0$  and  $s = \sigma/R_0$ , we can simplify this expression further:

$$\begin{aligned}
P(\delta v|s) &= \frac{1}{\alpha \sqrt{\pi} s \left( 1 + \delta v \right)^{1+\frac{2}{\alpha}}} \times \\
&\times e^{-\frac{1}{4s^2} \left[ 1 + (1 + \delta v)^{-\frac{2}{\alpha}} \right]} \times \\
&\times \sinh \left[ \frac{1}{2s^2} (1 + \delta v)^{-\frac{1}{\alpha}} \right].
\end{aligned} \tag{22}$$

The probability distribution function in Eq. (22) is defined in the domain  $[-1, +\infty)$ ; for  $\delta v = 1 + \varepsilon$ , in the limit  $\varepsilon \rightarrow 0^+$  it behaves as

$$P(\delta v|s) \propto \varepsilon^{-1-\frac{2}{\alpha}} e^{-\frac{1}{4s^2} \varepsilon^{-\frac{2}{\alpha}}} \sinh \left[ \frac{\varepsilon^{-\frac{1}{\alpha}}}{2s^2} \right] \rightarrow 0, \tag{23}$$

as the (vanishing) exponential factor dominates. In the opposite limit  $\delta v \rightarrow \infty$ , instead, the distribution behaves asymptotically as

$$P(\delta v|s) \approx \frac{1}{2\alpha\sqrt{\pi}s^3} e^{-\frac{1}{4s^2}} \delta v^{-1-3/\alpha}. \quad (24)$$

This shows that this distribution is fat-tailed. In particular, all the distribution moments  $\langle \delta v^\beta \rangle$  with  $\beta \geq 3/\alpha$  are not defined and, for both  $\alpha = 3$  (dipole-dipole interactions) and  $\alpha = 6$  (van der Waals), this includes all integer moments (e.g., the mean and variance). These fat tails are the consequence of the approximation of an atom's position distribution as a Gaussian everywhere in space, i.e., including points much further away from the center of a trap than a few  $\sigma$ s. In other words, it appears to be an artefact of the description, rather than something really physical. The result of this approximation is to allow for an extremely small (but not vanishing) probability that two atoms can be arbitrarily close, which, due to the algebraic scaling of the interactions, produces considerable energy shifts. Moments like the mean and variance are therefore dominated by the very rare events in which two atoms lie very close to each other. The rarity of such events is encoded in the exponential suppression  $e^{-(1/4s^2)}$  in Eq. (24). In principle, these unphysical fat tails could affect our results, as it is known that, in the Anderson problem, the scaling of the localization length is modified when Cauchy-like distributions are chosen instead of more regular ones. However, as mentioned above the fat tails in our case are strongly suppressed and one needs to assess how likely it is to actually probe them in a simulation or an experiment. For that purpose, let us first notice that the asymptotic behavior reported in (24) emerges when the argument of the sinh function in Eq. (22) is small, i.e., still assuming  $\delta v \gg 1$ , for

$$\delta v \gg (2s^2)^{-\alpha}. \quad (25)$$

Let us calculate now the probability of generating an energy shift within the tails, i.e.,

$$\mathbb{P}_s \equiv \mathbb{P}(\delta v > (2s^2)^{-\alpha}) = \int_{(2s^2)^{-\alpha}}^{\infty} d\delta v P(\delta v|s). \quad (26)$$

Employing now the asymptotic expression (24) we obtain

$$\mathbb{P}_s = \mathbb{P}(\delta v > (2s^2)^{-\alpha}) \approx \frac{4s^3}{3\sqrt{\pi}} e^{-\frac{1}{4s^2}}. \quad (27)$$

This result apparently does not depend on  $\alpha$ , but we need to remember that the derivation assumes  $\delta v \gg 1$ , and is therefore only consistent if  $(2s^2)^{-\alpha} \gg 1$ . Considering  $\alpha = 3$  or 6, though, this is satisfied already for rather large disorder amplitudes, e.g.,  $s = 0.3$ , which then yields  $\mathbb{P}_{0.3} \approx 0.0013$ . Due to the exponential factor, these probabilities decrease very fast with  $s$ . For  $s = 0.1$ , for instance, we get  $\mathbb{P}_{0.1} \approx 10^{-14}$  and in the range spanned

in the plots reported in the main text  $s \leq 5 \times 10^{-4}$  this becomes  $\mathbb{P}_s \ll 10^{-400000}$ , which is clearly impossible to observe in any reasonable experiment or numerical procedure. Hence, we can safely assume the unphysical fat tails to be completely irrelevant in the determination of our numerical results in the regime considered.

### Distribution bias towards positive energy shifts

As observed in the main text, there appears to be a bias of the distribution towards positive energy shifts, which makes the features present in the localization length plots bend towards higher energies. Considering the marginal discussed in the previous section, this seems counter-intuitive, since the distribution (22) seems to shift, for increasing  $s$ , towards negative values instead, eventually becoming peaked very close to  $\delta v \approx -1$ . It is also possible to provide an intuitive explanation for this, since, as two neighboring traps becomes wider and wider, it becomes much more likely for two atoms to lie at a larger distance than the one separating the two centers, than it is for them to lie closer than that. The limiting case  $s \gg R_0$  is indicative, as one can imagine that the atoms' positions can be picked uniformly in space on length scales  $\gg R_0$ .

Although we do not hold at the moment a convincing explanation of why the shift seems to point in the opposite direction, we believe it is due to the correlated nature of the full distribution (which indeed would be consistent with not seeing the correct behavior in a marginal) and we provide a physical argument which partially supports this conjecture. For simplicity, we shall work here with a chain, rather than a ladder. We hypothesize that (A) the number  $L+1$  of atoms is large, i.e.,  $L+1 \gg 1$  (this choice is such that the number of distances is  $L$ ); (B) the disorder is extremely small  $s \lll 1$ . A useful simplification from (B) is that we can effectively reduce the dimensionality of the problem and only consider the position displacement along the  $z$  direction: in fact, if we write  $\mathbf{d} = (\delta x, \delta y, R_0 + \delta z)$ , then

$$d = |\mathbf{d}| = R_0 + \delta z + O(\delta x^2, \delta y^2, \delta z^2) \approx R_0 + \delta z. \quad (28)$$

Hence, we can extract the probability of the distances  $d_k$  directly from Eq. (13):

$$p_{\text{dist}}(d_1, \dots, d_L) = \frac{\sigma^{-L}}{\sqrt{L+1} (\sqrt{2\pi})^L} \times \times e^{-\frac{1}{2\sigma^2} \sum_{k,q} (d_k - R_0) A_{kq} (d_q - R_0)}, \quad (29)$$

with the same  $A_{k,q}$  (up to increasing  $L$  by 1). Clearly, the marginal for any given variable  $d_k$  is also Gaussian and its mean and variance can be straightforwardly extracted from the expression above:

$$\langle d_k \rangle = R_0 \quad \text{and} \quad \sqrt{\langle d_k^2 \rangle - R_0^2} = \sqrt{\sigma^2 (A^{-1})_{kk}} = \sqrt{2}\sigma. \quad (30)$$

These variables are therefore identically distributed and, due to the boundedness of the covariance (see matrix  $C$  in Eq. (14)), satisfy a generalized weak law of large numbers, as we demonstrate below for this very special case: let us define  $D = (\sum_k d_k)/L$  and consider the probability  $\mathbb{P}(|D - R_0| > \varepsilon)$  of a fluctuation  $\varepsilon$  around the mean value. By Chebyshev's inequality,

$$\mathbb{P}(|D - R_0| > \varepsilon) \leq \frac{\text{Var}D}{\varepsilon^2}, \quad (31)$$

where

$$\begin{aligned} \text{Var}D &= \frac{\langle (\sum_k (d_k - R_0))^2 \rangle}{L^2} = \\ &= \frac{1}{L^2} \sum_{k,q} \langle (d_k - R_0)(d_q - R_0) \rangle = \frac{\sigma^2}{L^2} \sum_{k,q} C_{k,q}, \end{aligned} \quad (32)$$

where  $C_{k,q}$  are the elements of the matrix  $C = A^{-1}$  in Eq. (14). This sum is not difficult to calculate, since each row but the first and the last totals 0, whereas the first and last contribute 1 each, implying  $\text{Var}D = 2\sigma^2/L^2$ . As a consequence, by choosing a sufficiently large  $L$  the r.h.s. in Eq. (31) can be made arbitrarily small or, more precisely,  $\forall \delta > 0 \exists \bar{L} : \forall L > \bar{L}$

$$\mathbb{P}(|D - R_0| > \varepsilon) \leq \delta, \quad (33)$$

and therefore  $D \rightarrow R_0$  in probability. Note that, had we been dealing with independent variables, we would have retrieved a result where  $\text{Var}D \sim 1/L$  instead of  $1/L^2$ . On a less formal level, this can be understood as follows: consider that, in this effective one-dimensional picture, the sum of all distances corresponds to the distance between the first and last atoms. When generating the positions independently, this variable will not be affected by the random nature of the positions of all the intermediate ones; instead, if one were to generate the distances as independent variables, these would effectuate a random walk (with drift  $R_0$ ) and thus the effective uncertainty in the position of the last atom, assuming knowledge of the first one, would be of order  $O(\sqrt{L})$  and increase with the length of the chain.

Now, consider that, having chosen repulsive interactions ( $V(R) > 0$ ), the interaction potential is a convex function. Hence, we can write down Jensen's inequality as

$$\frac{\sum_k V(d_k)}{L} \geq V\left(\frac{\sum_k d_k}{L}\right). \quad (34)$$

By the weak law of large numbers, for large  $L \gg 1$  we can effectively replace the r.h.s. of the inequality above with  $V(R_0)$ , which leaves us with the approximate statement

$$\sum_k V(d_k) \gtrsim LV(R_0), \quad (35)$$

i.e.,

$$\sum_k (V(d_k) - V(R_0)) \gtrsim 0, \quad (36)$$

i.e.,

$$\sum_k \delta V_k \gtrsim 0. \quad (37)$$

Albeit not a rigorous proof, this argument provides a clear indication that, for sufficiently large system sizes, the correlations among the variables will make positive biases in the energy shifts preferable to negative ones, in agreement with the qualitative features observed in the main text. Accepting this claim, there must be at least a point  $s > 0$  where this bias is strictly positive. It then follows that there exists a right neighborhood of  $s = 0$  in which the bias increases with  $s$ .

## HILBERT SPACE REDUCTIONS AND RESTRICTED HAMILTONIANS

The Hamiltonian introduced in the main text reads

$$\hat{H} = \underbrace{\Omega \sum_k \hat{\sigma}_x^{(k)}}_{\hat{H}_1} + \underbrace{\Delta \sum_k \hat{n}_k + \sum_{\substack{k=1 \\ m \neq k}}^N \frac{1}{2} V(d_{km}) \hat{n}_m \hat{n}_k}_{\hat{H}_0}, \quad (38)$$

where  $d_{km}$  denotes the distance between the  $k$ -th and  $m$ -th atoms. In order to exploit the large energy separations present in the system, we switch to the interaction picture

$$\hat{H}_I(t) = e^{i\hat{H}_0 t} \hat{H}_1 e^{-i\hat{H}_0 t} = \Omega \sum_k e^{i\hat{H}_0 t} \hat{\sigma}_x^{(k)} e^{-i\hat{H}_0 t}. \quad (39)$$

Recalling that  $[\hat{\sigma}_x^{(k)}, \hat{n}_m] = 0$  for every  $k \neq m$  and that  $\hat{\sigma}_x^{(k)} \hat{n}_k = (1 - \hat{n}_k) \hat{\sigma}_x^{(k)}$  we can simplify the  $k$ -th addend in Eq. (39)

$$\begin{aligned} e^{i\hat{H}_0 t} \hat{\sigma}_x^{(k)} e^{-i\hat{H}_0 t} &= e^{it\hat{n}_k(\Delta + \sum_{m \neq k} V(d_{km}) \hat{n}_m)} \hat{\sigma}_x^{(k)} \times \\ &\times e^{-it\hat{n}_k(\Delta + \sum_{m \neq k} V(d_{km}) \hat{n}_m)} = \\ &= e^{it(2\hat{n}_k - 1)(\Delta + \sum_{m \neq k} V(d_{km}) \hat{n}_m)} \hat{\sigma}_x^{(k)}, \end{aligned} \quad (40)$$

where in the first equality we singled out in the exponentials all the terms which depend upon  $\hat{n}_k$ ; all the remaining ones cancel out. The Hamiltonian  $\hat{H}_I$  can then be written as

$$\hat{H}_I(t) = \Omega \sum_k e^{it(2\hat{n}_k - 1)(\Delta + \sum_{m \neq k} V(d_{km}) \hat{n}_m)} \hat{\sigma}_x^{(k)}. \quad (41)$$

We wish now to employ a rotating-wave approximation scheme to discard all terms which oscillate fast in time.

This implies that the oscillation frequency  $\omega$  should be  $\gg \Omega$  for a term to be neglected. Note that the frequency  $\omega$  is however operator-valued:

$$\omega = (2\hat{n}_k - 1)(\Delta + \sum_{m \neq k} V(d_{km})\hat{n}_m). \quad (42)$$

Since the prefactor  $-1 \leq 2\hat{n}_k - 1 \leq 1$  is of order  $O(1)$ , it is the second factor which is decisive for the selection. We introduce now for every site  $k$  a projector  $\hat{P}_k$  over all states where there is a single excitation among the neighbors of  $k$  and no additional one within a radius  $2R_0$ . Its specific structure depends clearly on the structure of the lattice, but if we define by  $\mathcal{F}_k$  the set of nearest-neighboring sites of  $k$  and by  $\mathcal{S}_k$  the set of sites within a distance  $2R_0$  from  $k$  which are neither site  $k$  itself nor one of the sites in  $\mathcal{F}_k$ , then we can give an implicit definition according to

$$\hat{P}_k = \sum_{q \in \mathcal{F}_k} \hat{n}_q \prod_{q' \in \mathcal{F}_k, q' \neq q} (1 - \hat{n}_{q'}) \prod_{q'' \in \mathcal{S}_k} (1 - \hat{n}_{q''}). \quad (43)$$

Checking that the expression above satisfies  $(\hat{P}_k)^2 = \hat{P}_k$  is straightforward if one recalls that  $\hat{n}_q^2 = \hat{n}_q$  and  $(1 - \hat{n}_q)^2 = 1 - \hat{n}_q \forall q$ . The relevance of the projector  $\hat{P}_k$  is that it precisely identifies the constraints – identified in the main text – under which a spin (or atom) is able to flip (or being excited/de-excited). Slightly more formally,

$$(\Delta + \sum_{m \neq k} V(d_{km})\hat{n}_m)\hat{P}_k = (\Delta - V(R_0))\hat{P}_k = 0. \quad (44)$$

Furthermore, note that according to definition (43)  $\hat{P}_k$  acts trivially on site  $k$  and thus commutes with all local operators which instead exclusively act on that site; in particular,  $[\hat{\sigma}_x^{(k)}, \hat{P}_k] = 0$ . Defining for brevity  $\hat{Q}_k = \mathbb{1} - \hat{P}_k$  the projector onto the orthogonal subspace ( $\hat{Q}_k^2 = \hat{Q}_k$ ,  $\hat{Q}_k\hat{P}_k = 0$ ) we thus have

$$\begin{aligned} \hat{\sigma}_x^{(k)} &= (\hat{P}_k + \hat{Q}_k) \hat{\sigma}_x^{(k)} (\hat{P}_k + \hat{Q}_k) = \\ &= \hat{P}_k \hat{\sigma}_x^{(k)} \hat{P}_k + \underbrace{\hat{Q}_k \hat{\sigma}_x^{(k)} \hat{P}_k}_{=0} + \underbrace{\hat{Q}_k \hat{\sigma}_x^{(k)} \hat{P}_k}_{=0} + \hat{Q}_k \hat{\sigma}_x^{(k)} \hat{Q}_k = \\ &= \hat{P}_k \hat{\sigma}_x^{(k)} + \hat{Q}_k \hat{\sigma}_x^{(k)}. \end{aligned} \quad (45)$$

Hence, we can separate the interaction Hamiltonian  $\hat{H}_I$  into two contributions:

$$\begin{aligned} \hat{H}_I(t) &= \Omega \sum_k \hat{P}_k \hat{\sigma}_x^{(k)} + \\ &+ e^{it(2\hat{n}_k - 1)(\Delta + \sum_{m \neq k} V(d_{km})\hat{n}_m)} \hat{Q}_k \hat{\sigma}_x^{(k)}. \end{aligned} \quad (46)$$

The space of configurations onto which  $\hat{Q}_k$  has support can be further split into three classes:

- (A) States where site  $k$  has two or more excited nearest neighbors;
- (B) States where site  $k$  has only one excited neighbor, but there is at least another excitation within a radius  $2R_0$ ;
- (C) States where no neighbors of  $k$  are excited.

In case (A) the interaction potential on site  $k$  is  $\geq 2V(R_0)$ ; accounting for the detuning  $\Delta = -V(R_0)$  we find  $\omega \gtrsim V(R_0) \gg \Omega$ ; these terms are thereby oscillating very fast and can be discarded. Terms of type (B) are facilitated by the single neighboring excitation, but the presence of an additional one within a distance  $2R_0$  implies that

$$\Delta + \sum_{m \neq k} V(d_{km})\hat{n}_m \geq V(2R_0) \quad (47)$$

and therefore  $\omega \gtrsim V(2R_0) \gg \Omega$ , which allows us to neglect all type-(B) contributions as well. Terms belonging to class (C) are instead more delicate, since an appropriate combination of the interactions with many excitations at different distances could approximately cancel out the detuning  $\Delta$ . For instance, for dipole-dipole interactions ( $\alpha = 3$ ) the potential obeys  $V(\gamma R_0) = V(R_0)\gamma^{-3}$ ; considering a honeycomb lattice with 5 excited next-nearest neighbors at distance  $R_1 = \sqrt{3}R_0$  and a single excited next-next-next-next-nearest (or fourth-nearest for brevity) neighbor at distance  $R_4 = 3R_0$  one finds

$$\begin{aligned} \Delta + \sum_{m \notin \mathcal{F}_k} V(d_{km}) &\rightarrow -V(R_0) + 5V(R_1) + V(R_4) = \\ &= V(R_0) \left( -1 + \frac{5}{3\sqrt{3}} + \frac{1}{3^3} \right) \approx -0.00071 V(R_0). \end{aligned} \quad (48)$$

However, configurations such as the one described above always require a large local density of excitations, and hence can only affect Hilbert subspaces at higher energies than the ones considered in the main text, separated at least by some factors of  $V(R_1) \gg \Omega$ . As long as we consider the low-energy Hilbert subspaces, it is thus fine to neglect terms of type (C) as well. Overall, in the subspaces we are interested in we can approximate

$$\hat{H}_I(t) \approx \Omega \sum_k \hat{P}_k \hat{\sigma}_x^{(k)}. \quad (49)$$

Going back to the original Schrödinger representation is now straightforward and yields

$$\hat{H} \approx \Omega \sum_k \hat{P}_k \hat{\sigma}_x^{(k)} + \Delta \sum_k \hat{n}_k + \sum_{\substack{k=1 \\ k \neq k}}^N \frac{1}{2} V(d_{km}) \hat{n}_m \hat{n}_k. \quad (50)$$

Note that in the specific subspace (let us call it  $\mathcal{H}_1$ ) considered in the main text, the one including all possible

one-excitation states plus all possible pairs of neighboring ones, the “diagonal” part  $\hat{H}_0$  acts trivially as the null operator and can thus be discarded, implying

$$\hat{H}_{\mathcal{H}_1} = \Omega \sum_k \hat{P}_k \hat{\sigma}_x^{(k)}. \quad (51)$$

We remark that the same derivation can be followed in the presence of weak disorder by changing the definition of  $\hat{H}_1$  in Eq. (38) to

$$\hat{H}_1 = \Omega \sum_k \hat{\sigma}_x^{(k)} + \frac{1}{2} \sum_{k \neq q} \delta V(d_{kq}) \hat{n}_k \hat{n}_q. \quad (52)$$

Since the second term is diagonal and commutes with  $\hat{H}_0$ , the calculation of the interaction picture is straightforward:

$$\begin{aligned} \hat{H}_I(t) = & \Omega \sum_k e^{it(2\hat{n}_k-1)(\Delta + \sum_{m \neq k} V(d_{km})\hat{n}_m)} \hat{\sigma}_x^{(k)} + \\ & + \frac{1}{2} \sum_{k \neq q} \delta V(d_{kq}) \hat{n}_k \hat{n}_q \end{aligned} \quad (53)$$

and one can follow the same steps outlined above.

### Hilbert space lattice structure

Having derived the restricted Hamiltonian (51) we can now identify the geometric structure of the Hilbert space in the  $\hat{\sigma}_z^{(k)}$  basis. To start with, we introduce the following definitions for the basis itself: we call  $|M_k\rangle$  states with a single excitation present on site  $k$ , whereas we denote by  $|N_{kq}\rangle$  states with a pair of excitations on sites  $k$  and  $q$ . Fixing the number  $N$  of tweezers, the Hilbert subspace we work in is therefore defined as

$$\mathcal{H}_1 = \text{Span} \{ |M_k\rangle, |N_{kq}\rangle \mid k = 1, \dots, N; q \in \mathcal{F}_k \}, \quad (54)$$

where we recall that  $\mathcal{F}_k$  is the set of nearest neighbors of site  $k$ . Note that, since  $|N_{kq}\rangle = |N_{qk}\rangle$  the pair states are doubly counted; however, this clearly still leads to the generation of the same vector space. Alternatively, one can also define an equivalence relation  $|N_{kq}\rangle \sim |N_{ml}\rangle \Leftrightarrow (k = m \wedge q = l) \vee (k = l \wedge q = m)$  and take the quotient of the r.h.s. above. In the following, it is understood that the states  $|N_{kq}\rangle$  are always taken from this space, i.e., we shall never consider states with two isolated excitations at distance  $d > R_0$ .

By construction,  $\hat{H}_{\mathcal{H}_1} \mathcal{H}_1 \subseteq \mathcal{H}_1$ . Furthermore, we know that the action of  $\hat{P}_k \hat{\sigma}_x^{(k)}$  is to flip the spin in site  $k$  conditioned on the presence of a single excitation in  $\mathcal{F}_k$  and no additional one in  $\mathcal{S}_k$ . This implies

$$\hat{P}_k \hat{\sigma}_x^{(k)} |M_l\rangle = \begin{cases} 0 & \text{if } l = k, \\ |N_{kq}\rangle & \text{if } l \in \mathcal{F}_k, \\ 0 & \text{otherwise.} \end{cases} \quad (55)$$

Considering that  $l \in \mathcal{F}_k \Leftrightarrow k \in \mathcal{F}_l$ , one can see that

$$\hat{H}_{\mathcal{H}_1} |M_l\rangle = \Omega \sum_{k \in \mathcal{F}_l} |N_{kl}\rangle. \quad (56)$$

Similarly,

$$\hat{P}_k \hat{\sigma}_x^{(k)} |N_{ql}\rangle = \begin{cases} |M_l\rangle & \text{if } q = k, \\ |M_q\rangle & \text{if } l = k, \\ 0 & \text{otherwise,} \end{cases} \quad (57)$$

since by construction the only facilitated spins are in sites  $q$  and  $l$ . Hence,

$$\hat{H}_{\mathcal{H}_1} |N_{ql}\rangle = \Omega (|M_q\rangle + |M_l\rangle). \quad (58)$$

Collecting these considerations, we can find the Hamiltonian matrix elements:

$$\langle M_q | \hat{H}_{\mathcal{H}_1} | M_k \rangle = 0 \quad (59a)$$

$$\langle N_{ml} | \hat{H}_{\mathcal{H}_1} | N_{kq} \rangle = 0 \quad (59b)$$

$$\langle N_{ml} | \hat{H}_{\mathcal{H}_1} | M_k \rangle = \begin{cases} \Omega & \text{if } l = k, \\ \Omega & \text{if } m = k, \\ 0 & \text{otherwise.} \end{cases} \quad (59c)$$

Now, there are as many states  $|M_k\rangle$  as there are sites, so it is natural to connect the two: starting from the geometry of the tweezer array, which defines the original lattice structure, we place for visual aid each state  $|M_k\rangle$  on the corresponding site  $k$ . Crucially, each pair state  $|N_{kq}\rangle$  is exclusively connected (via the Hamiltonian) to the two one-excitation states  $|M_k\rangle$  and  $|M_q\rangle$ , so it is quite natural to place it in the middle point between sites  $k$  and  $q$ , changing the structure to a generalized Lieb lattice. Now, by drawing a link between any pair of sites every time the corresponding states yield a non-zero Hamiltonian matrix element one precisely reconstructs the kind of lattices we displayed in Fig. 1 in the main text. Note that this construction is general and it is much more easily implemented pictorially than it is by following the steps described in this section.

### BOUND ON THE NUMBER OF FLAT BANDS

We provide here an account of the bound  $n_{\text{flat}} \geq |n_1 - n_2|$  mentioned in the main text, where we recall that  $n_{\text{flat}}$  denotes the number of flat bands in the model,  $n_1$  the number of one-particle states per unit cell and  $n_2$  the corresponding number of pair states per unit cell. Before doing that, however, we briefly comment on the fact that the spectrum of the hopping Hamiltonians (51) is always symmetric with respect to  $\epsilon = 0$ . In fact, one can define the parity transformation

$$\hat{U} = \hat{U}^\dagger = (-1)^{\sum_k \hat{n}_k} \quad (60)$$

which, in the subspace  $\mathcal{H}_1$  act according to  $\hat{U} |M_k\rangle = -|M_k\rangle$  on all one-excitation states and  $\hat{U} |N_{kq}\rangle = |N_{kq}\rangle$

on all pair states. Combined with Eqs. (59a)-(59c), this implies  $\hat{U}^\dagger \hat{H}_{\mathcal{H}_1} \hat{U} = -\hat{H}_{\mathcal{H}_1}$ . Hence, if  $|\epsilon\rangle$  is an eigenvector of the Hamiltonian at energy  $\epsilon$ , then  $\hat{U}|\epsilon\rangle$  is also an eigenvector, but at eigenvalue  $-\epsilon$ , proving the symmetry of the spectrum under reflection  $\epsilon \rightarrow -\epsilon$ .

We start directly from the synthetic lattice reconstructed in the Hilbert space according to the procedure described in the previous section. This structure is not in general a Bravais lattice and needs, as a first step, to be reduced to one by identifying an appropriate “basis”. This is a standard procedure in crystallography and solid state physics and we refer the reader to any good introductory textbook (see e.g., [5]). For the reader’s convenience, we however recall here just a few of the most basic concepts: a Bravais lattice is a lattice structure where the positions  $\vec{l}$  of the lattice sites correspond to discrete translations

$$\vec{l} = \sum_{i=1}^d z_i \vec{a}_i \quad \text{with } z_i \in \mathbb{Z}. \quad (61)$$

generated by a set of  $d$  linearly-independent *primitive lattice vectors*  $\vec{a}_i$  ( $i = 1 \dots d$ ), where  $d$  is the dimensionality of the system. If a site is located at the origin, all sites can be found this way and all points at positions  $\vec{l}$  are lattice sites. Any lattice is, by definition, a periodically repeating pattern, and is therefore invariant under a certain set of translations by  $\vec{l}$  for some specific choice of the primitive lattice vectors. However, in many cases an additional set of  $B$  vectors  $\{\vec{b}_1, \dots, \vec{b}_B\}$ , called “basis”, is required. In such cases, and fixing conventionally  $\vec{b}_1 = 0$  which can be done without loss of generality, if one lattice point is located at the origin, every point at a position  $\vec{l}$  is also a lattice site, but not all lattice sites are at positions  $\vec{l}$ . All of them are instead found at positions  $\vec{l} + \vec{b}_j$  with  $j = 1, \dots, B$ . We also remark that distances between sites in the synthetic lattice are not meaningful, being just a convenient way to visualize the structure of the Hilbert space. Hence, we are free to rescale the length of all (dimensionless) vectors  $\vec{a}_i$ ,  $\vec{b}_i$  by a common factor. In all the examples discussed below the primitive lattice vectors have the same length and we shall chose to normalize them to unit length ( $|\vec{a}_i| = 1$ ). Also, for brevity in the following we refer to the  $\mathbb{R}^d$  space where these vectors live as the *direct space*.

We also introduce the reciprocal lattice vectors  $\vec{a}_i^*$ ,  $i = 1 \dots d$  which satisfy the defining relations

$$\vec{a}_i^* \cdot \vec{a}_j = 2\pi \delta_{ij}. \quad (62)$$

The reciprocal Bravais lattice is then reconstructed by taking integer combinations of these vectors, i.e.,

$$\vec{G} = \sum_{i=1}^d z_i^* \vec{a}_i^* \quad \text{with } z_i^* \in \mathbb{Z}. \quad (63)$$

Exactly like in direct space, one can define a unit cell  $\mathcal{U}^*$  like a portion of  $\mathbb{R}^d$  which contains only one reciprocal lattice point and, repeated periodically in space under all possible translations  $\vec{G}$  tiles  $\mathbb{R}^d$  entirely and without overlaps. From a slightly different (but equivalent) perspective, one can define the equivalence relation between vectors  $\vec{k}, \vec{q} \in \mathbb{R}^d$  living in reciprocal space

$$\vec{k} \sim \vec{q} \Leftrightarrow \exists \vec{G} \mid \vec{k} = \vec{q} + \vec{G} \quad (64)$$

with  $\vec{G}$  a reciprocal *lattice* vector. Hence, the unit cell may be defined as a representative from the quotient  $\mathbb{R}^d / \sim$ . By defining quasi-momenta  $\vec{k}$  as reciprocal space vectors belonging to a unit cell  $\mathcal{U}^*$ , one can define a Fourier series in the usual way for any generic quantity  $A_{\vec{l}}$  living on the direct-space Bravais lattice

$$\tilde{A}_{\vec{k}} = \sum_{\vec{l}} e^{-i\vec{k} \cdot \vec{l}} A_{\vec{l}}. \quad (65)$$

The corresponding inverse transform is also standard:

$$A_{\vec{l}} = \int_{\mathcal{U}^*} \frac{d^d k}{(2\pi)^d} e^{i\vec{k} \cdot \vec{l}} \tilde{A}_{\vec{k}}, \quad (66)$$

as can be shown remembering that

$$\frac{\vec{G} \cdot \vec{l}}{2\pi} \in \mathbb{Z} \quad (67)$$

and using the Poisson-summation-derived distributional identity

$$\sum_{z \in \mathbb{Z}} e^{-i\alpha z} = \sum_{m \in \mathbb{Z}} 2\pi \delta(\alpha + 2\pi m), \quad (68)$$

with  $\alpha \in \mathbb{R}$  and  $\delta$  the Dirac delta. The choice of the unit cell is not unique; in the following we assume to be working in the *first Brillouin zone*  $\mathcal{B}$  [5].

Clearly, the definitions above do not hinge upon working in a specific space and, indeed, one can analogously define a unit cell in direct space which contains a single Bravais lattice point. Hence, such a unit cell includes  $B$  synthetic lattice points. It is quite natural to subdivide them according to whether they are of the “one-excitation” or “pair” kind. As done in the main text, we define  $n_1$  the number of one-excitation states in a unit cell and  $n_2 = B - n_1$  the number of pair ones. For example,

- Synthetic square lattice (Lieb lattice):  $n_1 = 1$ ,  $n_2 = 2$ ,  $B = 3$ .
- Synthetic triangular lattice:  $n_1 = 1$ ,  $n_2 = 3$ ,  $B = 4$ .
- Synthetic honeycomb lattice:  $n_1 = 2$ ,  $n_2 = 3$ ,  $B = 5$ .

Since each synthetic lattice point can be uniquely associated to a given primitive lattice vector  $\vec{l}$  and basis vector  $\vec{b}_i$ , we can unambiguously denote each state in the Hilbert subspace  $\mathcal{H}_1$  as a tensor product  $|\vec{l}\rangle \otimes |\vec{b}_i\rangle$ . For later convenience, we now wish to distinguish between the basis vectors identifying one-excitation states  $(|\vec{b}_i\rangle \rightarrow |\mu_j\rangle, j = 1, \dots, n_1)$  and pair states  $(|\vec{b}_i\rangle \rightarrow |\nu_j\rangle, j = 1, \dots, n_2)$ , so that the space of basis states is equivalently generated as

$$\text{Span}\{|\mu_1\rangle, \dots, |\mu_{n_1}\rangle, |\nu_1\rangle, \dots, |\nu_{n_2}\rangle\}. \quad (69)$$

Consequently, there is a bijective correspondence between states  $|M_k\rangle$  and states  $|\vec{l}\rangle \otimes |\mu_i\rangle$  and between states  $|N_{kq}\rangle$  and states  $|\vec{l}\rangle \otimes |\nu_i\rangle$ .

For later convenience, we define the lattice translation operator  $T_{\vec{j}}$ , where  $\vec{j}$  is a Bravais lattice vector, which acts on the positional degrees of freedom according to

$$T_{\vec{j}}|\vec{l}\rangle = |\vec{l} + \vec{j}\rangle. \quad (70)$$

By the straightforward quasi-momentum states definition

$$|\vec{k}\rangle = \sum_{\vec{l}} e^{-i\vec{k}\cdot\vec{l}} |\vec{l}\rangle \quad (71)$$

one also gets

$$T_{\vec{j}}|\vec{k}\rangle = e^{i\vec{k}\cdot\vec{j}} |\vec{k}\rangle. \quad (72)$$

The Hamiltonian can now be generically characterized as

a sum of terms

$$\hat{H}_{\mathcal{H}_1} = \Omega \sum_{\vec{l}} \sum_{\vec{j}} \sum_{m=1}^{n_1} \sum_{n=1}^{n_2} \left( C_{\vec{j},m,n} |\mu_m\rangle \langle \nu_n| + D_{\vec{j},m,n} |\nu_n\rangle \langle \mu_m| \right) |\vec{l} + \vec{j}\rangle \langle \vec{l}|, \quad (73)$$

where  $C_{\vec{j}}$  and  $D_{\vec{j}}$  are collections of connectivity matrices with elements 1 (if two states are linked) and 0 (if the two states are not). For instance, if the Hamiltonian can cause a hop from  $\vec{l}$  to  $\vec{l} + \vec{a}_1$  accompanied by a change  $|\mu_1\rangle \rightarrow |\nu_1\rangle$ , then  $D_{\vec{a}_1,1,1} = 1$ . Note that these are, in general, rectangular matrices of size  $n_1 \times n_2$ . Furthermore, to ensure that  $H$  is hermitian they must satisfy

$$C_{-\vec{j},m,n} = D_{\vec{j},m,n}^* = D_{\vec{j},m,n}, \quad (74)$$

where the last equality comes from the fact that they are defined to be real (their elements being either 0 or 1). Note that no terms  $\propto |\mu_m\rangle \langle \mu_n|$  or  $|\nu_m\rangle \langle \nu_n|$  appear, as one-excitation states are exclusively connected to pair ones and vice versa (see Eqs. (59a)-(59c)). Neither  $C$  nor  $D$  depends explicitly on  $\vec{l}$ , as otherwise the form of the Hamiltonian would be dependent upon the specific choice of which Bravais lattice point should act as the origin of coordinates, whereas by construction all points separated by a Bravais lattice vector should be equivalent. In this form, it is not difficult to exploit this symmetry of the Hamiltonian under discrete lattice translations to partially diagonalize it in terms of Fourier modes:

$$\begin{aligned} \hat{H}_{\mathcal{H}_1} &= \Omega \sum_{\vec{l}} \sum_{\vec{j}} \sum_{m=1}^{n_1} \sum_{n=1}^{n_2} \left( C_{\vec{j},m,n} |\mu_m\rangle \langle \nu_n| + C_{-\vec{j},m,n} |\nu_n\rangle \langle \mu_m| \right) T_{\vec{j}} |\vec{l}\rangle \langle \vec{l}| = \\ &= \Omega \sum_{\vec{j}} \sum_{m=1}^{n_1} \sum_{n=1}^{n_2} \left( C_{\vec{j},m,n} |\mu_m\rangle \langle \nu_n| + C_{-\vec{j},m,n} |\nu_n\rangle \langle \mu_m| \right) T_{\vec{j}} \sum_{\vec{l}} |\vec{l}\rangle \langle \vec{l}| = \\ &= \Omega \sum_{\vec{j}} \sum_{m=1}^{n_1} \sum_{n=1}^{n_2} \left( C_{\vec{j},m,n} |\mu_m\rangle \langle \nu_n| + C_{-\vec{j},m,n} |\nu_n\rangle \langle \mu_m| \right) T_{\vec{j}} \int_{\mathcal{B}} \frac{d^d k}{(2\pi)^d} |\vec{k}\rangle \langle \vec{k}| = \\ &= \Omega \int_{\mathcal{B}} \frac{d^d k}{(2\pi)^d} \sum_{\vec{j}} \sum_{m=1}^{n_1} \sum_{n=1}^{n_2} \left( C_{\vec{j},m,n} |\mu_m\rangle \langle \nu_n| + C_{-\vec{j},m,n} |\nu_n\rangle \langle \mu_m| \right) e^{i\vec{k}\cdot\vec{j}} |\vec{k}\rangle \langle \vec{k}| = \\ &= \Omega \int_{\mathcal{B}} \frac{d^d k}{(2\pi)^d} \sum_{m=1}^{n_1} \sum_{n=1}^{n_2} \left[ \left( \sum_{\vec{j}} C_{\vec{j},m,n} e^{i\vec{k}\cdot\vec{j}} \right) |\mu_m\rangle \langle \nu_n| + \left( \sum_{\vec{j}} C_{\vec{j},m,n} e^{i\vec{k}\cdot\vec{j}} \right)^* |\nu_n\rangle \langle \mu_m| \right] |\vec{k}\rangle \langle \vec{k}| = \\ &= \Omega \int_{\mathcal{B}} \frac{d^d k}{(2\pi)^d} \sum_{m=1}^{n_1} \sum_{n=1}^{n_2} \left[ \tilde{C}_{-\vec{k},m,n} |\mu_m\rangle \langle \nu_n| + \left( \tilde{C}_{-\vec{k},m,n} \right)^* |\nu_n\rangle \langle \mu_m| \right] |\vec{k}\rangle \langle \vec{k}|, \end{aligned} \quad (75)$$

where again

$$\tilde{C}_{-\vec{k},m,n} = \left( \sum_{\vec{j}} C_{\vec{j},m,n} e^{i\vec{k}\cdot\vec{j}} \right) \quad (76)$$

is, for every  $\vec{k} \in \mathcal{B}$ , a rectangular  $n_1 \times n_2$  matrix. Calling



now

$$\hat{M}_{\vec{k}} = \sum_{m=1}^{n_1} \sum_{n=1}^{n_2} \left[ \tilde{C}_{-\vec{k},m,n} |\mu_m\rangle \langle \nu_n| + h.c. \right], \quad (77)$$

we can represent it as a matrix in the basis  $\{|\mu_1\rangle, \dots, |\mu_{n_1}\rangle, |\nu_1\rangle, \dots, |\nu_{n_2}\rangle\}$ , which yields

$$M_{\vec{k}} = \left( \begin{array}{c|c} 0 & \tilde{C}_{-\vec{k}} \\ \hline \tilde{C}_{-\vec{k}}^\dagger & 0 \end{array} \right). \quad (78)$$

Due to this particular block structure,

$$\text{Rank} \{M_{\vec{k}}\} = \text{Rank} \{\tilde{C}_{-\vec{k}}\} + \text{Rank} \{\tilde{C}_{-\vec{k}}^\dagger\}. \quad (79)$$

Furthermore, the rank of a rectangular matrix is never greater than its shortest side. In this case,

$$\text{Rank} \{\tilde{C}_{-\vec{k}}\} \leq \min \{n_1, n_2\}, \quad (80)$$

which in turn implies that the rank of the square matrix  $M_{\vec{k}}$  is  $\leq 2 \min \{n_1, n_2\}$ . This means that the size of the kernel of  $M_{\vec{k}}$  has a lower bound

$$\begin{aligned} \dim(\text{Ker } M_{\vec{k}}) &= B - \text{Rank} \{M_{\vec{k}}\} \geq \\ &\geq (n_1 + n_2) - 2 \min \{n_1, n_2\} = \\ &= \max \{n_1, n_2\} - \min \{n_1, n_2\} = \\ &= |n_1 - n_2|. \end{aligned} \quad (81)$$

Hence, if  $|n_1 - n_2| \geq 1$  then for every  $\vec{k}$  one can find a kernel vector  $|v_{\vec{k}}\rangle$  in the basis such that  $\hat{M}_{\vec{k}} |v_{\vec{k}}\rangle = 0$ . Correspondingly,  $\hat{H}_{\mathcal{H}_1} |\vec{k}\rangle \otimes |v_{\vec{k}}\rangle = 0 \ \forall \vec{k}$  and the set of all these states forms a zero-energy flat band. Clearly, if  $|n_1 - n_2| > 1$  then more than one choice of  $|v\rangle_{|\vec{k}\rangle}$  can be made per each quasi-momentum  $\vec{k}$ , each identifying an independent flat band. Hence, calling the number of flat bands in the model  $n_{\text{flat}}$ , consistently with the main text notation,

$$n_{\text{flat}} = \dim(\text{Ker } M_{\vec{k}}) \geq |n_1 - n_2|, \quad (82)$$

which proves the bound.

The general rules for filling the matrix elements of  $\tilde{C}_{\vec{k}}$  are the following:

- Choose a lattice basis. It is typically convenient to do so in a relatively compact fashion, i.e., keeping the basis vectors “short”.
- Fix a column  $1 \leq n \leq n_2$ .
- Consider the *two* possible ways in which a particle can hop from the intermediate state  $|\nu_n\rangle$  within the basis to its neighbors  $|\mu_m\rangle$  and  $|\mu_p\rangle$ .

- Add  $e^{i\vec{k} \cdot \vec{j}_n}$  to  $C_{-\vec{k},m,n}$  and  $e^{i\vec{k} \cdot \vec{j}_p}$  to  $C_{-\vec{k},p,n}$ , where  $\vec{j}_{m/p}$  are the lattice vectors pointing to the *arrival* lattice sites.

In the next sections we work out some examples among the ones displayed in the main text. For simplicity, we set  $\Omega = 1$ .

### Example: the triangular lattice

The triangular lattice (also called “hexagonal”, although here we shall not adopt this convention to avoid confusion with the honeycomb lattice) is a two-dimensional Bravais lattice with primitive lattice vectors

$$\vec{a}_1 = a(1, 0)^\top \quad \text{and} \quad \vec{a}_2 = a \left( \cos \frac{\pi}{3}, \sin \frac{\pi}{3} \right)^\top, \quad (83)$$

with  $a$  the real-space lattice spacing. In the Hilbert space, we have again a triangular structure where a new site is added on each link.

It is not difficult to see that this reduces to a pure triangular lattice by choosing a basis of 4 sites, a single one-excitation one ( $n_1 = 1$ ) and 3 pair ones ( $n_2 = 3$ ). The primitive lattice vectors will be the same as above, where we fix for simplicity  $a = 1$ . The basis states can be chosen according to:

$|\mu_1\rangle$  : a one-excitation site at  $\vec{b} = 0$ .

$|\nu_1\rangle$  : a pair site at  $\vec{b} = \vec{a}_1/2$ .

$|\nu_2\rangle$  : a pair site at  $\vec{b} = \vec{a}_2/2$ .

$|\nu_3\rangle$  : a pair site at  $\vec{b} = (\vec{a}_1 - \vec{a}_2)/2$ .

The matrix  $\tilde{C}_{-\vec{k}}$  is now a  $1 \times 3$  matrix whose elements can be computed via the procedure outlined above:

$\tilde{C}_{-\vec{k},1,1}$  : from basis state  $|\nu_1\rangle$  one can reach state  $|\mu_1\rangle$  within the same Bravais lattice site ( $\Rightarrow +1$ ) or state  $|\mu_1\rangle$  at the neighboring site  $\vec{j} = \vec{a}_1$  ( $\Rightarrow +e^{i\vec{k} \cdot \vec{a}_1}$ ).

$\tilde{C}_{-\vec{k},1,2}$  : from basis state  $|\nu_2\rangle$  one can reach state  $|\mu_1\rangle$  within the same site ( $\Rightarrow +1$ ) or state  $|\mu_1\rangle$  at the neighboring site  $\vec{j} = \vec{a}_2$  ( $\Rightarrow +e^{i\vec{k} \cdot \vec{a}_2}$ ).

$\tilde{C}_{-\vec{k},1,3}$  : from state  $|\nu_3\rangle$  one can reach state  $|\mu_1\rangle$  within the same site ( $\Rightarrow +1$ ) or state  $|\mu_1\rangle$  at the neighboring site  $\vec{j} = \vec{a}_1 - \vec{a}_2$  ( $\Rightarrow +e^{i\vec{k} \cdot (\vec{a}_1 - \vec{a}_2)}$ ).

Collecting all terms, the matrix  $\tilde{C}_{-\vec{k}}$  reads

$$\tilde{C}_{-\vec{k}} = \left( 1 + e^{i\vec{k} \cdot \vec{a}_1}, 1 + e^{i\vec{k} \cdot \vec{a}_2}, 1 + e^{i\vec{k} \cdot (\vec{a}_1 - \vec{a}_2)} \right) \equiv \vec{w}_{\vec{k}}^\dagger \quad (84)$$

and is equivalent to a three-dimensional vector  $\vec{w}_{\vec{k}}$ . Thus, the total matrix  $M_{\vec{k}}$  can be expressed as

$$M_{\vec{k}} = \left( \begin{array}{c|c} 0 & \vec{w}_{\vec{k}}^\dagger \\ \hline \vec{w}_{\vec{k}} & 0 \end{array} \right) \quad (85)$$

There are two kernel states corresponding to four-dimensional vectors  $(0, v_{\vec{k},1})$  and  $(0, v_{\vec{k},2})$  with  $\vec{w}_{\vec{k}}^\dagger \cdot \vec{v}_{\vec{k},1/2} = 0$ . These states thus reconstruct two flat bands,

in line with the bound  $n_{\text{flat}} \geq 2$  of this case.

The remaining two bands can be calculated instead by squaring  $M_{\vec{k}}$ :

$$M_{\vec{k}}^2 = \left( \begin{array}{c|c} \vec{w}_{\vec{k}}^\dagger \cdot \vec{w}_{\vec{k}} & 0 \\ \hline 0 & \vec{w}_{\vec{k}} \otimes \vec{w}_{\vec{k}}^\dagger \end{array} \right). \quad (86)$$

From the symmetric structure of the spectrum and the presence of two flat bands, we can simply infer the non-zero ones as

$$\begin{aligned} \pm \sqrt{\vec{w}_{\vec{k}}^\dagger \cdot \vec{w}_{\vec{k}}} &= \pm \sqrt{\left|1 + e^{i\vec{k} \cdot \vec{a}_1}\right|^2 + \left|1 + e^{i\vec{k} \cdot \vec{a}_2}\right|^2 + \left|1 + e^{i\vec{k} \cdot (\vec{a}_1 - \vec{a}_2)}\right|^2} \\ &= \pm \sqrt{2} \sqrt{3 + \cos(\vec{k} \cdot \vec{a}_1) + \cos(\vec{k} \cdot \vec{a}_2) + \cos(\vec{k} \cdot (\vec{a}_1 - \vec{a}_2))} \end{aligned} \quad (87)$$

Choosing the reciprocal lattice vectors as

$$\vec{a}_1^* = \frac{4\pi}{\sqrt{3}} \left( \cos \frac{\pi}{6}, -\sin \frac{\pi}{6} \right)^\top \quad \text{and} \quad \vec{a}_2^* = \frac{4\pi}{\sqrt{3}} (0, 1)^\top \quad (88)$$

the first Brillouin zone  $\mathcal{B}$  is an hexagon in  $\vec{k}$  space identified by the conditions

$$\begin{aligned} &\left( \left| \vec{k} \cdot \vec{a}_1^* \right| \leq \frac{1}{2} |\vec{a}_1^*|^2 \right) \cap \left( \left| \vec{k} \cdot \vec{a}_2^* \right| \leq \frac{1}{2} |\vec{a}_2^*|^2 \right) \cap \\ &\cap \left( \left| \vec{k} \cdot (\vec{a}_1^* - \vec{a}_2^*) \right| \leq \frac{1}{2} |\vec{a}_1^* - \vec{a}_2^*|^2 \right). \end{aligned} \quad (89)$$

#### Example: the honeycomb lattice

The honeycomb lattice is a triangular Bravais lattice with primitive lattice vectors

$$\vec{a}_1 = a (1, 0)^\top \quad \text{and} \quad \vec{a}_2 = a \left( \cos \frac{\pi}{3}, \sin \frac{\pi}{3} \right)^\top, \quad (90)$$

where the lattice spacing  $a$  is  $\sqrt{3}$  times the edge of the hexagons, plus a basis of two vectors

$$\vec{b}_1 = 0 \quad \text{and} \quad \vec{b}_2 = \frac{2\vec{a}_2 - \vec{a}_1}{3}. \quad (91)$$

In the synthetic lattice, this gives rise to a structure with a basis of 5 elements: 2 one-excitation sites ( $n_1 = 2$ ) and 3 pair ones ( $n_2 = 3$ ), which we choose as follows:

$|\mu_1\rangle$  : a one-excitation site at  $\vec{b} = 0$ .

$|\mu_2\rangle$  : a one-excitation site at  $\vec{b} = \frac{2\vec{a}_2 - \vec{a}_1}{3}$ .

$|\nu_1\rangle$  : a pair site at  $\vec{b} = \frac{2\vec{a}_2 - \vec{a}_1}{6}$ .

$|\nu_2\rangle$  : a pair site at  $\vec{b} = \frac{2\vec{a}_1 - \vec{a}_2}{6}$ .

$|\nu_3\rangle$  : a pair site at  $\vec{b} = -\frac{\vec{a}_1 + \vec{a}_2}{6}$ .

We thus see that the  $\tilde{C}_{-\vec{k}}$  are  $2 \times 3$  matrices and that there is at least a flat zero-energy band. The matrix elements can be identified column by column as follows:

$|\nu_1\rangle$  : From  $|\nu_1\rangle$  one can jump to  $|\mu_1\rangle$  or to  $|\mu_2\rangle$  remaining in the same Bravais lattice site.

$|\nu_2\rangle$  : From  $|\nu_2\rangle$  one can jump to  $|\mu_1\rangle$  in the same site or to  $|\mu_2\rangle$  changing site by  $\vec{j} = \vec{a}_1 - \vec{a}_2$  ( $\Rightarrow e^{i\vec{k} \cdot (\vec{a}_1 - \vec{a}_2)}$ ).

$|\nu_3\rangle$  : From  $|\nu_3\rangle$  one can jump to  $|\mu_1\rangle$  in the same site or to  $|\mu_2\rangle$  changing site by  $\vec{j} = -\vec{a}_2$  ( $\Rightarrow e^{-i\vec{k} \cdot \vec{a}_2}$ ).

Hence,

$$\tilde{C}_{-\vec{k}} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & e^{i\vec{k} \cdot (\vec{a}_1 - \vec{a}_2)} & e^{-i\vec{k} \cdot \vec{a}_2} \end{pmatrix} \equiv \begin{pmatrix} \vec{w}_{\vec{k},1}^\dagger \\ \vec{w}_{\vec{k},2}^\dagger \end{pmatrix}, \quad (92)$$

with  $\vec{w}_{\vec{k},1/2}$  three-dimensional vectors. The matrix  $M_{\vec{k}}$  is thus

$$M_{\vec{k}} = \left( \begin{array}{cc|c} 0 & 0 & \vec{w}_{\vec{k},1}^\dagger \\ 0 & 0 & \vec{w}_{\vec{k},2}^\dagger \\ \hline \vec{w}_{\vec{k},1} & \vec{w}_{\vec{k},2} & 0 \end{array} \right). \quad (93)$$

The kernel state is a five-dimensional vector  $(0, 0, \vec{v}_{\vec{k}})$  which satisfies  $\vec{w}_{\vec{k},1/2}^\dagger \cdot \vec{v}_{\vec{k}} = 0$ .

To identify the remaining non-zero bands, we again take the square of the total matrix  $M_{\vec{k}}$ :

$$M_{\vec{k}}^2 = \left( \begin{array}{cc|cc} \vec{w}_{\vec{k},1}^\dagger \cdot \vec{w}_{\vec{k},1} & \vec{w}_{\vec{k},1}^\dagger \cdot \vec{w}_{\vec{k},2} & & 0 \\ \vec{w}_{\vec{k},2}^\dagger \cdot \vec{w}_{\vec{k},1} & \vec{w}_{\vec{k},2}^\dagger \cdot \vec{w}_{\vec{k},2} & & 0 \\ \hline 0 & 0 & \vec{w}_{\vec{k},1} \otimes \vec{w}_{\vec{k},1}^\dagger + \vec{w}_{\vec{k},2} \otimes \vec{w}_{\vec{k},2}^\dagger & \end{array} \right), \quad (94)$$

where the first block is  $2 \times 2$  and the second one  $3 \times 3$ .

We can now diagonalize the first block to find

$$\lambda_{\vec{k},\pm} = \frac{1}{2} \left[ \left( |\vec{w}_{\vec{k},1}|^2 + |\vec{w}_{\vec{k},2}|^2 \right) \pm \sqrt{\left( |\vec{w}_{\vec{k},1}|^2 - |\vec{w}_{\vec{k},2}|^2 \right)^2 + 4 |\vec{w}_{\vec{k},2}^\dagger \cdot \vec{w}_{\vec{k},1}|^2} \right], \quad (95)$$

with  $\lambda_{\vec{k},\pm} \geq 0$ . The four non-trivial bands will thus correspond to  $\pm\sqrt{\lambda_{\vec{k},+}}$  and  $\pm\sqrt{\lambda_{\vec{k},-}}$ . Working out the scalar products

$$|\vec{w}_{\vec{k},1}|^2 = |\vec{w}_{\vec{k},2}|^2 = 3 \quad (96)$$

and

$$|\vec{w}_{\vec{k},2}^\dagger \cdot \vec{w}_{\vec{k},1}| = \left| 1 + e^{i\vec{k} \cdot (\vec{a}_1 - \vec{a}_2)} + e^{-i\vec{k} \cdot \vec{a}_2} \right| \quad (97)$$

we obtain by substitution

$$\lambda_{\vec{k},\pm} = 3 \pm \left| 1 + e^{i\vec{k} \cdot (\vec{a}_1 - \vec{a}_2)} + e^{-i\vec{k} \cdot \vec{a}_2} \right| = 3 \pm \sqrt{3 + 2 \cos(\vec{k} \cdot (\vec{a}_1 - \vec{a}_2)) + 2 \cos(\vec{k} \cdot \vec{a}_2) + 2 \cos(\vec{k} \cdot \vec{a}_1)}. \quad (98)$$

From the first equality we see that the second addend is always  $\leq 3$ . It is 3 only when  $\vec{k} = 0$  (up to reciprocal lattice translations  $\vec{G}$ , see (63)). Hence,  $\lambda_-(\vec{k} = 0) = 0$  is a minimum and  $\lambda_+(\vec{k} = 0) = 6$  is a maximum. The bands  $\pm\sqrt{\lambda_{\vec{k},-}}$  touch at  $\vec{k} = 0$  with linear dispersion. Second, the argument of the absolute value will vanish when

$$\vec{k} \cdot (\vec{a}_1 - \vec{a}_2) = \pm \frac{2\pi}{3} + 2\pi n, \quad -\vec{k} \cdot \vec{a}_2 = \pm \frac{4\pi}{3} + 2\pi m \quad (99)$$

where the signs must be chosen consistently. Up to reciprocal lattice translations, one can choose

$$\vec{k} = \pm \frac{1}{3} (\vec{a}_2^* - \vec{a}_1^*), \quad (100)$$

identifying the points at the vertices of the hexagonal first Brillouin zone (one can verify this point lies at the

boundary of two of the conditions in (89)). Therefore, the two upper bands  $\sqrt{\lambda_{\vec{k},+}}$  and  $\sqrt{\lambda_{\vec{k},-}}$  touch at the vertices of the first Brillouin zone with linear dispersion and similarly do the lower bands  $-\sqrt{\lambda_{\vec{k},+}}$  and  $-\sqrt{\lambda_{\vec{k},-}}$ .

- 
- [1] B. R. Holstein, American Journal of Physics **66**, 583 (1998).
  - [2] F. A. Barone, H. Boschi-Filho, and C. Farina, American Journal of Physics **71**, 483 (2003).
  - [3] M. Marcuzzi, J. Minář, D. Barredo, S. de Léséleuc, H. Labuhn, T. Lahaye, A. Browaeys, E. Levi, and I. Lesanovsky, Phys. Rev. Lett. **118**, 063606 (2017).
  - [4] C. da Fonseca and J. Petronilho, Linear Algebra and its Applications **325**, 7 (2001).
  - [5] G. Grosso and G. Pastori Parravicini, *Solid State Physics* (Elsevier Science, 2000).