Supporting Information for Fast Generation of Pipek–Mezey Wannier Functions via the Co-Iterative Augmented Hessian Method

Gengzhi Yang[1,2] and Hong-Zhou Ye[3,4,a)]

[1)] *Joint Center for Quantum Information and Computer Science, University of Maryland, College Park, Maryland, 20742*

[2)] *Department of Mathematics, University of Maryland, College Park, Maryland, 20742*

[3)] *Department of Chemistry and Biochemistry, University of Maryland, College Park, Maryland, 20742*

[4)] *Institute for Physical Science and Technology, University of Maryland, College Park, Maryland, 20742*

(Dated: 12 February 2026)

[a)] Electronic mail: hzye@umd.edu

CONTENTS

I. STRUCTURE FILES

The structure files for all solid-state systems studied in this work are available in the following GitHub repository:

https://github.com/hongzhouye/supporting_data/tree/main/2026/PMWF

where

- the structure of h-BN, diamond, MgO, silicon, SiO_2, *trans*-$(C_2H_2)_\infty$, C-nanotube, and graphene is taken from J. Phys. Chem. A **128**, 8570 (2024),
- the structure of CO/MgO(001) is taken from Faraday Discuss. **254**, 628 (2024),
- and the structure of aluminum (*fcc*) is generated with a lattice constant of 4.05 Å.

II. DERIVATION OF ANALYTICAL GRADIENTS, HESSIAN–VECTOR PRODUCT, AND HESSIAN DIAGONALS

A. Parameterization

As discussed in the main text, the k-space unitaries can be parameterized using k-dependent generators, each of which is a complex anti-Hermitian matrix. Let $\kappa_{\boldsymbol{k}} = X_{\boldsymbol{k}} + \mathrm{i}\, Y_{\boldsymbol{k}}$ where both $X_{\boldsymbol{k}}$ and $Y_{\boldsymbol{k}} \in \mathbb{R}^{n_{\mathrm{orb}} \times n_{\mathrm{orb}}}$. The unique parameters are the lower-triangular part of $X_{\boldsymbol{k}}$, excluding the diagonals, and the the lower-triangular part of $Y_{\boldsymbol{k}}$, including the diagonals.

B. Basic derivatives

Define $(E_{mn})_{ij} = \delta_{mi}\delta_{nj}$. The first- and second-order derivatives of the unitary with respect to the generators evaluated at $\kappa_{\boldsymbol{k}} = 0$ are

$$\begin{aligned} \frac{\partial U_{\boldsymbol{k}'}}{\partial X_{\boldsymbol{k}mn}} &= \delta_{\boldsymbol{k}\boldsymbol{k}'}(E_{mn} - E_{nm}), \qquad (m \neq n) \\ \frac{\partial U_{\boldsymbol{k}'}}{\partial Y_{\boldsymbol{k}mn}} &= \mathrm{i}\delta_{\boldsymbol{k}\boldsymbol{k}'}(E_{mn} + E_{nm}), \qquad (m \neq n) \\ \frac{\partial U_{\boldsymbol{k}'}}{\partial Y_{\boldsymbol{k}mm}} &= \mathrm{i}\delta_{\boldsymbol{k}\boldsymbol{k}'}E_{mm}. \end{aligned} \tag{1}$$

$$\frac{\partial^2 U_{\boldsymbol{k}''}}{\partial A_{\boldsymbol{k}mn} B_{\boldsymbol{k}'pq}} = \frac{1}{2}\left\{ \frac{\partial U_{\boldsymbol{k}''}}{\partial A_{\boldsymbol{k}mn}}, \frac{\partial U_{\boldsymbol{k}''}}{\partial B_{\boldsymbol{k}'mn}} \right\}, \qquad (A, B = X, Y) \tag{2}$$

C. Cost function

The PM objective function defined in the main text here is recapped here. A minus sign is added to the objective function to turn the maximization a minimization problem.

$$L = -\sum_{\boldsymbol{T}A,i} Q^p_{\boldsymbol{T}A,\mathbf{0}i} \tag{3}$$

$$Q_{\boldsymbol{T}A,\mathbf{0}i} = P_{\boldsymbol{T}A,\mathbf{0}i,\mathbf{0}i} = \sum_{\boldsymbol{k}\boldsymbol{k}'} \left(U^\dagger_{\boldsymbol{k}} P_{\boldsymbol{T}A,\boldsymbol{k}\boldsymbol{k}'} U_{\boldsymbol{k}'} \right)_{ii} \tag{4}$$

$$P_{\boldsymbol{T}A,\boldsymbol{k}\boldsymbol{k}'} = \frac{1}{N_k} \sum_{\boldsymbol{R}\boldsymbol{R}'} \theta^*_{\boldsymbol{R}\boldsymbol{k}} P_{\boldsymbol{T}A,\boldsymbol{R}\boldsymbol{R}'} \theta_{\boldsymbol{R}'\boldsymbol{k}'} \tag{5}$$

D. Gradient

Consider the first-order variation of L,

$$\delta L = -p \sum_{\boldsymbol{T}A,i} Q^{p-1}_{\boldsymbol{T}A,\mathbf{0}i} \delta Q_{\boldsymbol{T}A,\mathbf{0}i} \tag{6}$$

which depends on the first-order variation of Q,

$$\delta Q_{\boldsymbol{T}A,\mathbf{0}i} = 2\Re \sum_{\boldsymbol{k}'} \left(P_{\boldsymbol{T}A,\mathbf{0}\boldsymbol{k}'} \delta U_{\boldsymbol{k}'} \right)_{ii}, \qquad \left(P_{\boldsymbol{T}A,\mathbf{0}\boldsymbol{k}'} := \sum_{\boldsymbol{k}} P_{\boldsymbol{T}A,\boldsymbol{k}\boldsymbol{k}'} \right) \tag{7}$$

For the real and imaginary part of the generator, the derivatives are

$$\begin{aligned} \frac{\partial Q_{\boldsymbol{T}A,\mathbf{0}i}}{\partial X_{\boldsymbol{k}mn}} &= -2 \left(P^{\Re}_{\boldsymbol{T}A,\mathbf{0}m,\boldsymbol{k}n} \delta_{mi} - P^{\Re}_{\boldsymbol{T}A,\mathbf{0}n,\boldsymbol{k}m} \delta_{ni} \right) \\ \frac{\partial Q_{\boldsymbol{T}A,\mathbf{0}i}}{\partial Y_{\boldsymbol{k}mn}} &= -2 \left(P^{\Im}_{\boldsymbol{T}A,\mathbf{0}m,\boldsymbol{k}n} \delta_{mi} + P^{\Im}_{\boldsymbol{T}A,\mathbf{0}n,\boldsymbol{k}m} \delta_{ni} \right) \end{aligned} \tag{8}$$

The two can be combined conveniently into a complex-valued expression,

$$\frac{\partial Q_{\boldsymbol{T}A,\mathbf{0}i}}{\partial X_{\boldsymbol{k}mn}} + \mathrm{i} \frac{\partial Q_{\boldsymbol{T}A,\mathbf{0}i}}{\partial Y_{\boldsymbol{k}mn}} = -2 \left(P_{\boldsymbol{T}A,\mathbf{0}m,\boldsymbol{k}n} \delta_{mi} - P^*_{\boldsymbol{T}A,\mathbf{0}n,\boldsymbol{k}m} \delta_{ni} \right) \tag{9}$$

which leads to the following expression for the gradient,

$$\boxed{\begin{aligned} G_{\boldsymbol{k}} &= \tilde{G}_{\boldsymbol{k}} - \tilde{G}^\dagger_{\boldsymbol{k}}, \\ \tilde{G}_{\boldsymbol{k}mn} &= 2p \sum_{\boldsymbol{T}A} Q^{p-1}_{\boldsymbol{T}A,\mathbf{0}m} P_{\boldsymbol{T}A,\mathbf{0}m,\boldsymbol{k}n} \end{aligned}} \tag{10}$$

Using $P_{\boldsymbol{T}A,\boldsymbol{k}\mathbf{0}} = P^\dagger_{\boldsymbol{T}A,\mathbf{0}\boldsymbol{k}}$, one can convert eq. (10) to Eqn M22 in the main text. We prefer to put $\boldsymbol{k}$ index first as in $P_{\boldsymbol{T}A,\boldsymbol{k}\mathbf{0}}$ due to the C-ordering of matrices and tensors in our code.

E. Hessian–vector product

Let $v_{\boldsymbol{k}}$ denote the vector to be applied by the Hessian. Let $v^X_{\boldsymbol{k}}$ and $v^Y_{\boldsymbol{k}}$ denote its real and imaginary parts, respectively. The Hessian–vector product is defined as follows.

$$\begin{aligned}
\Sigma^X_{\boldsymbol{k}mn} &= \frac{1}{2}\sum_{\boldsymbol{k}'pq} \frac{\partial^2 L}{\partial X_{\boldsymbol{k}mn}\partial X_{\boldsymbol{k}'pq}} v^X_{\boldsymbol{k}'pq} + \frac{\partial^2 L}{\partial X_{\boldsymbol{k}mn}\partial Y_{\boldsymbol{k}'pq}} v^Y_{\boldsymbol{k}'pq}, \\
\Sigma^Y_{\boldsymbol{k}mn} &= \frac{1}{2}\sum_{\boldsymbol{k}'pq} \frac{\partial^2 L}{\partial Y_{\boldsymbol{k}mn}\partial X_{\boldsymbol{k}'pq}} v^X_{\boldsymbol{k}'pq} + \frac{\partial^2 L}{\partial Y_{\boldsymbol{k}mn}\partial Y_{\boldsymbol{k}'pq}} v^Y_{\boldsymbol{k}'pq},
\end{aligned} \tag{11}$$

where at a high-level, we need to evaluate the second-order variation of L,

$$\delta^2 L = -p(p-1)\sum_{\boldsymbol{T}Ai} Q^{p-2}_{\boldsymbol{T}A,\boldsymbol{0}i}\left(\delta Q_{\boldsymbol{T}A,\boldsymbol{0}i}\right)^2 - p\sum_{\boldsymbol{T}Ai} Q^{p-1}_{\boldsymbol{T}A,\boldsymbol{0}i}\delta^2 Q_{\boldsymbol{T}A,\boldsymbol{0}i} \tag{12}$$

The first term only depends on the first-order variation of Q, whose expression has already been derived above. We call this term *disconnected*. The second term depends on the second-order variation of Q,

$$\delta^2 Q_{\boldsymbol{T}A,\boldsymbol{0}i} = 2\Re\sum_{\boldsymbol{k}'}\left(P_{\boldsymbol{T}A,\boldsymbol{0}\boldsymbol{k}'}\delta^2 U_{\boldsymbol{k}'}\right)_{ii} + 2\Re\sum_{\boldsymbol{k}\boldsymbol{k}'}\left(\delta U^\dagger_{\boldsymbol{k}} P_{\boldsymbol{T}A,\boldsymbol{k}\boldsymbol{k}'}\delta U_{\boldsymbol{k}'}\right)_{ii} \tag{13}$$

and therefore is *connected*. We further differentiate the first and second term in $\delta^2 Q$ to be *asymmetric* and *symmetric* connected terms.

1. Disconnected term

$$\begin{aligned}
\Sigma^X_{\boldsymbol{k}mn} &= -\frac{1}{2}p(p-1)\sum_{\boldsymbol{T}Ai} Q^{p-2}_{\boldsymbol{T}A,\boldsymbol{0}i}\frac{\partial Q_{\boldsymbol{T}A,\boldsymbol{0}i}}{\partial X_{\boldsymbol{k}mn}}\sum_{\boldsymbol{k}'pq}\left(\frac{\partial Q_{\boldsymbol{T}A,\boldsymbol{0}i}}{\partial X_{\boldsymbol{k}'pq}} v^X_{\boldsymbol{k}'pq} + \frac{\partial Q_{\boldsymbol{T}A,\boldsymbol{0}i}}{\partial Y_{\boldsymbol{k}'pq}} v^Y_{\boldsymbol{k}'pq}\right), \\
\Sigma^Y_{mn} &= -\frac{1}{2}p(p-1)\sum_{\boldsymbol{T}Ai} Q^{p-2}_{\boldsymbol{T}A,\boldsymbol{0}i}\frac{\partial Q_{\boldsymbol{T}A,\boldsymbol{0}i}}{\partial Y_{\boldsymbol{k}mn}}\sum_{\boldsymbol{k}'pq}\left(\frac{\partial Q_{\boldsymbol{T}A,\boldsymbol{0}i}}{\partial X_{\boldsymbol{k}'pq}} v^X_{\boldsymbol{k}'pq} + \frac{\partial Q_{\boldsymbol{T}A,\boldsymbol{0}i}}{\partial Y_{\boldsymbol{k}'pq}} v^Y_{\boldsymbol{k}'pq}\right)
\end{aligned} \tag{14}$$

where the term in the parenthesis can be simplified to

$$\sum_{\boldsymbol{k}'pq}\left(\frac{\partial Q_{\boldsymbol{T}A,\boldsymbol{0}i}}{\partial X_{\boldsymbol{k}'pq}} v^X_{\boldsymbol{k}'pq} + \frac{\partial Q_{\boldsymbol{T}A,\boldsymbol{0}i}}{\partial Y_{\boldsymbol{k}'pq}} v^Y_{\boldsymbol{k}'pq}\right) = 4\Re\sum_{\boldsymbol{k}'}\left(P_{\boldsymbol{T}A,\boldsymbol{0}\boldsymbol{k}'}v_{\boldsymbol{k}'}\right)_{ii} \tag{15}$$

Combining this with the gradient of Q [eq. (8)] gives

$$\boxed{\begin{aligned}
\Sigma_{\boldsymbol{k}} &= \tilde{\Sigma}_{\boldsymbol{k}} - \tilde{\Sigma}_{\boldsymbol{k}}, \\
\tilde{\Sigma}_{\boldsymbol{k}mn} &= 4p(p-1)\sum_{\boldsymbol{T}A} Q^{p-2}_{\boldsymbol{T}A,\boldsymbol{0}m}\sum_{\boldsymbol{k}'}\Re(P_{\boldsymbol{T}A,\boldsymbol{0}\boldsymbol{k}'}v_{\boldsymbol{k}'})_{mm}P_{\boldsymbol{T}A,\boldsymbol{0}m,\boldsymbol{k}n}
\end{aligned}} \tag{16}$$

2. *Connected symmetric term*

$$\begin{aligned}
\Sigma^X_{\boldsymbol{k}mn} &= -p \sum_{\boldsymbol{T}A,i} Q^{p-1}_{\boldsymbol{T}A,\boldsymbol{0}i} \Re \left[\frac{\partial U^\dagger_{\boldsymbol{k}}}{\partial X_{\boldsymbol{k}mn}} P_{\boldsymbol{T}A,\boldsymbol{k}\boldsymbol{k}'} \sum_{\boldsymbol{k}'pq} \left(\frac{\partial U_{\boldsymbol{k}'}}{\partial X_{\boldsymbol{k}'pq}} v^X_{\boldsymbol{k}'pq} + \frac{\partial U_{\boldsymbol{k}'}}{\partial Y_{\boldsymbol{k}'pq}} v^Y_{\boldsymbol{k}'pq} \right) \right]_{ii}, \\
\Sigma^Y_{\boldsymbol{k}mn} &= -p \sum_{\boldsymbol{T}A,i} Q^{p-1}_{\boldsymbol{T}A,\boldsymbol{0}i} \Re \left[\frac{\partial U^\dagger_{\boldsymbol{k}}}{\partial Y_{\boldsymbol{k}mn}} P_{\boldsymbol{T}A,\boldsymbol{k}\boldsymbol{k}'} \sum_{\boldsymbol{k}'pq} \left(\frac{\partial U_{\boldsymbol{k}'}}{\partial X_{\boldsymbol{k}'pq}} v^X_{\boldsymbol{k}'pq} + \frac{\partial U_{\boldsymbol{k}'}}{\partial Y_{\boldsymbol{k}'pq}} v^Y_{\boldsymbol{k}'pq} \right) \right]_{ii},
\end{aligned} \tag{17}$$

where the term in the parenthesis can be simplified to

$$\sum_{\boldsymbol{k}'pq} \frac{\partial U_{\boldsymbol{k}'}}{\partial X_{\boldsymbol{k}'pq}} v^X_{\boldsymbol{k}'pq} + \frac{\partial U_{\boldsymbol{k}'}}{\partial Y_{\boldsymbol{k}'pq}} v^Y_{\boldsymbol{k}'pq} = 2 \sum_{\boldsymbol{k}'pq} E_{pq} v_{\boldsymbol{k}'pq} \tag{18}$$

Combining this with the gradient of U [eq. (1)] gives

$$\boxed{\begin{aligned}
\Sigma_{\boldsymbol{k}} &= \tilde{\Sigma}_{\boldsymbol{k}} - \tilde{\Sigma}^\dagger_{\boldsymbol{k}}, \\
\tilde{\Sigma}_{\boldsymbol{k}mn} &= -2p \sum_{\boldsymbol{T}A} Q^{p-1}_{\boldsymbol{T}A,\boldsymbol{0}n} \sum_{\boldsymbol{k}'} \left(P_{\boldsymbol{T}A,\boldsymbol{k}\boldsymbol{k}'} v_{\boldsymbol{k}'} \right)_{mn}
\end{aligned}} \tag{19}$$

3. *Connected asymmetric term*

$$\begin{aligned}
\Sigma^X_{\boldsymbol{k}mn} &= -\frac{1}{2} p \sum_{\boldsymbol{T}A,i} Q^{p-1}_{\boldsymbol{T}A,\boldsymbol{0}i} \Re \left[P_{\boldsymbol{T}A,\boldsymbol{0}\boldsymbol{k}} \left\{ \frac{\partial U_{\boldsymbol{k}}}{\partial X_{\boldsymbol{k}mn}}, \sum_{pq} \frac{\partial U_{\boldsymbol{k}}}{\partial X_{\boldsymbol{k}pq}} v^X_{\boldsymbol{k}pq} + \frac{\partial U_{\boldsymbol{k}}}{\partial Y_{\boldsymbol{k}pq}} v^Y_{\boldsymbol{k}pq} \right\} \right]_{ii} \\
\Sigma^Y_{\boldsymbol{k}mn} &= -\frac{1}{2} p \sum_{\boldsymbol{T}A,i} Q^{p-1}_{\boldsymbol{T}A,\boldsymbol{0}i} \Re \left[P_{\boldsymbol{T}A,\boldsymbol{0}\boldsymbol{k}} \left\{ \frac{\partial U_{\boldsymbol{k}}}{\partial Y_{\boldsymbol{k}mn}}, \sum_{pq} \frac{\partial U_{\boldsymbol{k}}}{\partial X_{\boldsymbol{k}pq}} v^X_{\boldsymbol{k}pq} + \frac{\partial U_{\boldsymbol{k}}}{\partial Y_{\boldsymbol{k}pq}} v^Y_{\boldsymbol{k}pq} \right\} \right]_{ii}
\end{aligned} \tag{20}$$

Using eq. (18) to simplify terms in the parenthesis leads to

$$\begin{aligned}
\Sigma^X_{\boldsymbol{k}mn} &= -p \sum_{\boldsymbol{T}A,i} Q^{p-1}_{\boldsymbol{T}A,\boldsymbol{0}i} \Re \left[P_{\boldsymbol{T}A,\boldsymbol{0}\boldsymbol{k}} \left\{ \frac{\partial U_{\boldsymbol{k}}}{\partial X_{\boldsymbol{k}mn}}, \sum_{pq} E_{pq} v_{\boldsymbol{k}pq} \right\} \right]_{ii} \\
\Sigma^Y_{\boldsymbol{k}mn} &= -p \sum_{\boldsymbol{T}A,i} Q^{p-1}_{\boldsymbol{T}A,\boldsymbol{0}i} \Re \left[P_{\boldsymbol{T}A,\boldsymbol{0}\boldsymbol{k}} \left\{ \frac{\partial U_{\boldsymbol{k}}}{\partial Y_{\boldsymbol{k}mn}}, \sum_{pq} E_{pq} v_{\boldsymbol{k}pq} \right\} \right]_{ii}
\end{aligned} \tag{21}$$

Using eq. (1) to further simplify the expression leads to

$$\boxed{\begin{aligned}
\Sigma_{\boldsymbol{k}} &= \tilde{\Sigma}_{\boldsymbol{k}} - \tilde{\Sigma}^\dagger_{\boldsymbol{k}}, \\
\tilde{\Sigma}_{\boldsymbol{k}mn} &= p \sum_i v_{\boldsymbol{k}mi} \left[\sum_{\boldsymbol{T}A} Q^{p-1}_{\boldsymbol{T}A,\boldsymbol{0}i} (P_{\boldsymbol{T}A,\boldsymbol{0}\boldsymbol{k}})_{in} \right] + p \sum_{\boldsymbol{T}A} Q^{p-1}_{\boldsymbol{T}A,\boldsymbol{0}m} (P_{\boldsymbol{T}A,\boldsymbol{0}\boldsymbol{k}} v_{\boldsymbol{k}})_{mn}
\end{aligned}} \tag{22}$$

Finally, using $P_{\boldsymbol{T}A,\boldsymbol{k}\boldsymbol{0}} = P^\dagger_{\boldsymbol{T}A,\boldsymbol{0}\boldsymbol{k}}$, one can convert eqs. (16), (19) and (22) to Eqn M24–26 in the main text.

F. Hessian diagonal

The Hessian diagonal is defined as:

$$\begin{aligned}
D^X_{\boldsymbol{k}mn} &= \frac{\partial^2 L}{\partial X^2_{\boldsymbol{k}mn}} = \tilde{D}^X_{\boldsymbol{k}} + \tilde{D}^{X\top}_{\boldsymbol{k}} \\
D^Y_{\boldsymbol{k}mn} &= \frac{\partial^2 L}{\partial Y^2_{\boldsymbol{k}mn}} = \tilde{D}^Y_{\boldsymbol{k}} + \tilde{D}^{Y\top}_{\boldsymbol{k}}
\end{aligned} \tag{23}$$

where the second equality explicitly shows its symmetric nature.

1. Disconnected term

$$\begin{aligned}
D^X_{\boldsymbol{k}mn} &= -p(p-1)\sum_{\boldsymbol{T}A,i} Q^{p-2}_{\boldsymbol{T}A,\boldsymbol{0}i}\left(\frac{\partial Q_{\boldsymbol{T}A,\boldsymbol{0}i}}{\partial X_{\boldsymbol{k}mn}}\right)^2 \\
&= -4p(p-1)\sum_{\boldsymbol{T}A,i} Q^{p-2}_{\boldsymbol{T}A,\boldsymbol{0}i}\left[\Re(P_{\boldsymbol{T}A,\boldsymbol{0k}})_{mn}\right]^2 \delta_{mi} + (m \leftrightarrow n) \\
&= -4p(p-1)\sum_{\boldsymbol{T}A} Q^{p-2}_{\boldsymbol{T}A,\boldsymbol{0}m}\left[\Re(P_{\boldsymbol{T}A,\boldsymbol{0k}})_{mn}\right]^2 + (m \leftrightarrow n)
\end{aligned} \tag{24}$$

$$\begin{aligned}
D^Y_{\boldsymbol{k}mn} &= -p(p-1)\sum_{\boldsymbol{T}A,i} Q^{p-2}_{\boldsymbol{T}A,\boldsymbol{0}i}\left(\frac{\partial Q_{\boldsymbol{T}A,\boldsymbol{0}i}}{\partial Y_{\boldsymbol{k}mn}}\right)^2 \\
&= -4p(p-1)\sum_{\boldsymbol{T}A,i} Q^{p-2}_{\boldsymbol{T}A,\boldsymbol{0}i}\left[\Im(P_{\boldsymbol{T}A,\boldsymbol{0k}})_{mn}\right]^2 \delta_{mi} + (m \leftrightarrow n) \\
&= -4p(p-1)\sum_{\boldsymbol{T}A} Q^{p-2}_{\boldsymbol{T}A,\boldsymbol{0}m}\left[\Im(P_{\boldsymbol{T}A,\boldsymbol{0k}})_{mn}\right]^2 + (m \leftrightarrow n)
\end{aligned} \tag{25}$$

$$\boxed{\begin{aligned}
\tilde{D}^X_{mn} &= -4p(p-1)\sum_{\boldsymbol{T}A} Q^{p-2}_{\boldsymbol{T}A,\boldsymbol{0}m}\left[\left(\Re P_{\boldsymbol{T}A,\boldsymbol{0k}}\right)_{mn}\right]^2 \\
\tilde{D}^Y_{mn} &= -4p(p-1)\sum_{\boldsymbol{T}A} Q^{p-2}_{\boldsymbol{T}A,\boldsymbol{0}m}\left[\left(\Im P_{\boldsymbol{T}A,\boldsymbol{0k}}\right)_{mn}\right]^2
\end{aligned}} \tag{26}$$

2. Connected asymmetric term

$$
\begin{aligned}
D^{X}_{\boldsymbol{k}mn} &= -2p\sum_{\boldsymbol{T}A,i} Q^{p-1}_{\boldsymbol{T}A,\boldsymbol{0}i}\Re\left(P_{\boldsymbol{T}A,\boldsymbol{0k}}\frac{\partial^2 U_{\boldsymbol{k}}}{\partial X^2_{\boldsymbol{k}mn}}\right)_{ii} \\
&= -2p\sum_{\boldsymbol{T}A,i} Q^{p-1}_{\boldsymbol{T}A,\boldsymbol{0}i}\Re\left[P_{\boldsymbol{T}A,\boldsymbol{0k}}(E_{mn}-E_{nm})^2\right]_{ii} \\
&= 2p\sum_{\boldsymbol{T}A,i} Q^{p-1}_{\boldsymbol{T}A,\boldsymbol{0}i}\Re\left(P_{\boldsymbol{T}A,\boldsymbol{0k}}E_{mm}\right)_{ii} + (m\leftrightarrow n)
\end{aligned}
\tag{27}
$$

$$
\begin{aligned}
D^{Y}_{\boldsymbol{k}mn} &= -2p\sum_{\boldsymbol{T}A,i} Q^{p-1}_{\boldsymbol{T}A,\boldsymbol{0}i}\Re\left(P_{\boldsymbol{T}A,\boldsymbol{0k}}\frac{\partial^2 U_{\boldsymbol{k}}}{\partial Y^2_{\boldsymbol{k}mn}}\right)_{ii} \\
&= 2p\sum_{\boldsymbol{T}A,i} Q^{p-1}_{\boldsymbol{T}A,\boldsymbol{0}i}\Re\left[P_{\boldsymbol{T}A,\boldsymbol{0k}}(E_{mn}+E_{nm})^2\right]_{ii} \\
&= 2p\sum_{\boldsymbol{T}A,i} Q^{p-1}_{\boldsymbol{T}A,\boldsymbol{0}i}\Re\left(P_{\boldsymbol{T}A,\boldsymbol{0k}}E_{mm}\right)_{ii} + (m\leftrightarrow n)
\end{aligned}
\tag{28}
$$

Therefore

$$
\boxed{\tilde{D}^{X}_{\boldsymbol{k}mn} = \tilde{D}^{Y}_{\boldsymbol{k}mn} = 2p\sum_{\boldsymbol{T}A} Q^{p-1}_{\boldsymbol{T}A,\boldsymbol{0}m}\left(\Re P_{\boldsymbol{T}A,\boldsymbol{0k}}\right)_{mm}}
\tag{29}
$$

3. Connected symmetric term

$$
\begin{aligned}
D^{X}_{\boldsymbol{k}mn} &= -2p\sum_{\boldsymbol{T}A,i} Q^{p-1}_{\boldsymbol{T}A,\boldsymbol{0}i}\Re\left(\frac{\partial U^{\dagger}_{\boldsymbol{k}}}{\partial X_{\boldsymbol{k}mn}}P_{\boldsymbol{T}A,\boldsymbol{kk}}\frac{\partial U_{\boldsymbol{k}}}{\partial X_{\boldsymbol{k}mn}}\right)_{ii} \\
&= 2p\sum_{\boldsymbol{T}A,i} Q^{p-1}_{\boldsymbol{T}A,\boldsymbol{0}i}\Re\left[(E_{mn}-E_{nm})P_{\boldsymbol{T}A,\boldsymbol{kk}}(E_{mn}-E_{nm})\right]_{ii} \\
&= -2p\sum_{\boldsymbol{T}A,i} Q^{p-1}_{\boldsymbol{T}A,\boldsymbol{0}i}\Re\left(E_{mn}P_{\boldsymbol{T}A,\boldsymbol{kk}}E_{nm}\right)_{ii} + (m\leftrightarrow n)
\end{aligned}
\tag{30}
$$

$$
\begin{aligned}
D^{Y}_{\boldsymbol{k}mn} &= -2p\sum_{\boldsymbol{T}A,i} Q^{p-1}_{\boldsymbol{T}A,\boldsymbol{0}i}\Re\left(\frac{\partial U^{\dagger}_{\boldsymbol{k}}}{\partial Y_{\boldsymbol{k}mn}}P_{\boldsymbol{T}A,\boldsymbol{kk}}\frac{\partial U_{\boldsymbol{k}}}{\partial Y_{\boldsymbol{k}mn}}\right)_{ii} \\
&= -2p\sum_{\boldsymbol{T}A,i} Q^{p-1}_{\boldsymbol{T}A,\boldsymbol{0}i}\Re\left[(E_{mn}+E_{nm})P_{\boldsymbol{T}A,\boldsymbol{kk}}(E_{mn}+E_{nm})\right]_{ii} \\
&= -2p\sum_{\boldsymbol{T}A,i} Q^{p-1}_{\boldsymbol{T}A,\boldsymbol{0}i}\Re\left(E_{mn}P_{\boldsymbol{T}A,\boldsymbol{kk}}E_{nm}\right)_{ii} + (m\leftrightarrow n)
\end{aligned}
\tag{31}
$$

Therefore

$$
\boxed{\tilde{D}^{X}_{\boldsymbol{k}mn} = \tilde{D}^{Y}_{\boldsymbol{k}mn} = -2p\sum_{\boldsymbol{T}A} Q^{p-1}_{\boldsymbol{T}A,\boldsymbol{0}m}\left(\Re P_{\boldsymbol{T}A,\boldsymbol{kk}}\right)_{nn}}
\tag{32}
$$

III. JACOBI SWEEP

A. General consideration

A general real-valued, translationally symmetric Jacobi rotation can be written as

$$\begin{bmatrix} \tilde{w}_{\boldsymbol{R}_0,i} & \tilde{w}_{\boldsymbol{R}_0+\boldsymbol{R},j} \end{bmatrix} = \begin{bmatrix} w_{\boldsymbol{R}_0,i} & w_{\boldsymbol{R}_0+\boldsymbol{R},j} \end{bmatrix} \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}, \tag{33}$$

which mixes $w_{\mathbf{0},i}$ with $w_{\boldsymbol{R},j}$ and, by translational covariance, applies the same rotation to all their lattice translates. Fourier transforming eq. (33) yields the corresponding k-space formulation,

$$\begin{bmatrix} \tilde{\phi}_{\boldsymbol{k}i} & \tilde{\phi}_{\boldsymbol{k}j} \end{bmatrix} = \begin{bmatrix} \phi_{\boldsymbol{k}i} & \phi_{\boldsymbol{k}j} \end{bmatrix} \begin{bmatrix} \cos\theta & \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{R}}\sin\theta \\ -\mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{R}}\sin\theta & \cos\theta \end{bmatrix}, \tag{34}$$

which corresponds to a *k-dependent* 2×2 rotation between Bloch orbitals i and j.

B. Working equations for $p = 2$

To determine the optimal rotation angle θ, consider the PM objective for a pair of WFs $(\mathbf{0}i, \boldsymbol{R}j)$ after applying the Jacobi rotation in eq. (33):

$$\tilde{L}_{ij} = \sum_{\boldsymbol{T}A} \left[(\tilde{P}_{\boldsymbol{T}A,\mathbf{00}})_{ii}^2 + (\tilde{P}_{\boldsymbol{T}A,\mathbf{00}})_{jj}^2 \right]. \tag{35}$$

Using

$$\begin{aligned} \tilde{w}_{\mathbf{0}i} &= w_{\mathbf{0}i}\cos\theta - w_{\boldsymbol{R}j}\sin\theta, \\ \tilde{w}_{\mathbf{0}j} &= w_{-\boldsymbol{R}i}\sin\theta + w_{\mathbf{0}j}\cos\theta, \end{aligned} \tag{36}$$

the diagonal projection matrix elements after rotation become

$$\begin{aligned} (\tilde{P}_{\boldsymbol{T}A,\mathbf{00}})_{ii} &= (P_{\boldsymbol{T}A,\mathbf{00}})_{ii}\cos^2\theta + (P_{\boldsymbol{T}A,\boldsymbol{RR}})_{jj}\sin^2\theta - 2\Re(P_{\boldsymbol{T}A,\mathbf{0}\boldsymbol{R}})_{ij}\cos\theta\sin\theta, \\ (\tilde{P}_{\boldsymbol{T}A,\mathbf{00}})_{jj} &= (P_{\boldsymbol{T}A,(-\boldsymbol{R})(-\boldsymbol{R})})_{ii}\sin^2\theta + (P_{\boldsymbol{T}A,\mathbf{00}})_{jj}\cos^2\theta + 2\Re(P_{\boldsymbol{T}A,(-\boldsymbol{R})\mathbf{0}})_{ij}\cos\theta\sin\theta \\ &= (P_{\boldsymbol{T}A,\mathbf{00}})_{ii}\sin^2\theta + (P_{\boldsymbol{T}A,\boldsymbol{RR}})_{jj}\cos^2\theta + 2\Re(P_{\boldsymbol{T}A,\mathbf{0}\boldsymbol{R}})_{ij}\cos\theta\sin\theta. \end{aligned} \tag{37}$$

In the second line for $(\tilde{P}_{\boldsymbol{T}A,\mathbf{00}})_{jj}$, we have used translational symmetry together with the fact that $\boldsymbol{T}$ is summed over in the final objective. The change in the PM objective due to the $(\mathbf{0}i, \boldsymbol{R}j)$ rotation is therefore

$$\Delta L_{ij}(\theta) = \tilde{L}_{ij}(\theta) - L_{ij} = -\frac{1}{2}\sin 2\theta \left[(A_{\mathbf{0}\boldsymbol{R}})_{ij}\cos 2\theta + (B_{\mathbf{0}\boldsymbol{R}})_{ij}\sin 2\theta \right], \tag{38}$$

where we have introduced the intermediates

$$\begin{aligned}
(A_{\mathbf{0R}})_{ij} &= 4\sum_{\boldsymbol{T}A}\left[(P_{\boldsymbol{T}A,\mathbf{00}})_{ii} - (P_{\boldsymbol{T}A,\boldsymbol{RR}})_{jj}\right]\Re(P_{\boldsymbol{T}A,\mathbf{0R}})_{ij}, \\
(B_{\mathbf{0R}})_{ij} &= \sum_{\boldsymbol{T}A}\left[(P_{\boldsymbol{T}A,\mathbf{00}})_{ii} - (P_{\boldsymbol{T}A,\boldsymbol{RR}})_{jj}\right]^2 - [2\,\Re(P_{\boldsymbol{T}A,\mathbf{0R}})_{ij}]^2.
\end{aligned} \tag{39}$$

Setting $\Delta L'_{ij}(\theta) = 0$ yields

$$\tan 4\theta = -\frac{(A_{\mathbf{0R}})_{ij}}{(B_{\mathbf{0R}})_{ij}}. \tag{40}$$

Equation (40) can be interpreted as follows:

1. Since the input orbitals are already at a stationary point of the PM functional, $\theta = 0$ is a trivial solution of eq. (40), implying that the right-hand side vanishes. The non-trivial solutions therefore satisfy $\tan 4\theta = 0$, i.e., $4\theta = n\pi$ with $n \neq 0$.

2. Furthermore, it is sufficient to restrict θ to the interval $[0, \pi)$, because a rotation by $\theta = \pi$ corresponds to a simultaneous sign change of the two orbitals, which leaves the PM objective invariant, and any $\theta > \pi$ can be mapped back into $[0, \pi)$.

We therefore conclude that the only distinct non-trivial solutions are

$$\theta = \frac{\pi}{4}, \frac{\pi}{2}, \frac{3\pi}{4}. \tag{41}$$

In practice, we can calculate ΔL_{ij} for all possible θ values in eq. (41) to find one that maximizes the PM objective.

C. Extension for $p > 2$

The derivation above strictly assumes $p = 2$. A similar derivation for $p = 3$ leads to the same condition eq. (40) for determining θ except that the two intermediates are defined as

$$\begin{aligned}
(A_{\mathbf{0R}})_{ij} &= 4\sum_{\boldsymbol{T}A}\left[(P_{\boldsymbol{T}A,\mathbf{0R}})^2_{ii} - (P_{\boldsymbol{T}A,\mathbf{0R}})^2_{jj}\right]\Re(P_{\boldsymbol{T}A,\mathbf{0R}})_{ij}, \\
(B_{\mathbf{0R}})_{ij} &= \sum_{\boldsymbol{T}A}\left[(P_{\boldsymbol{T}A,\mathbf{0R}})_{ii} + (P_{\boldsymbol{T}A,\mathbf{0R}})_{jj}\right]\left\{\left[(P_{\boldsymbol{T}A,\mathbf{0R}})_{ii} - (P_{\boldsymbol{T}A,\mathbf{0R}})_{jj}\right]^2 - \left[2\Re(P_{\boldsymbol{T}A,\mathbf{0R}})_{ij}\right]^2\right\},
\end{aligned} \tag{42}$$

For general $p > 3$, an analytic derivation becomes cumbersome. However, the second argument above (periodicity in θ) always holds, and we have numerically validated that the first argument holds for $p \leq 6$ (which arguably already covers most use cases of PM localization). We therefore *postulate* that the first argument remains valid for all $p \geq 2$.

D. Efficient implementation and computational cost

In practice, we evaluate $\Delta L_{ij}(\theta)$ numerically for the three candidate angles in eq. (41) and check whether any of them increases the PM objective. The computational cost of this Jacobi stability check scales as

$$O(N_k^2\, n_{\text{proj}}\, n_{\text{orb}}^2) \tag{43}$$

for building the $(P_{\boldsymbol{T}A,\boldsymbol{k}\boldsymbol{R}})_{ij}$ intermediate. This cost is modest and comparable to that of a single gradient evaluation.

However, one can reasonably expect that pairwise instabilities primarily involve orbital pairs that are not too far apart in real space. We therefore can optionally restrict the real-space shifts in eq. (34) to satisfy

$$|\boldsymbol{R}| < R_{\max}, \tag{44}$$

for a chosen cutoff $R_{\max}$, which provides a controllable reduction in computational cost. Let N_R denote the number of lattice vectors $\boldsymbol{R}$ satisfying eq. (44). Under this restriction, the cost of the Jacobi stability check becomes

$$O(N_k N_R\, n_{\text{proj}}\, n_{\text{orb}}^2), \tag{45}$$

which is linear in N_k when N_R is independent of N_k. We choose $R_{\max} = 10$ Bohr in this work, which covers WF pairs that are separated by roughly 5 chemical bonds.