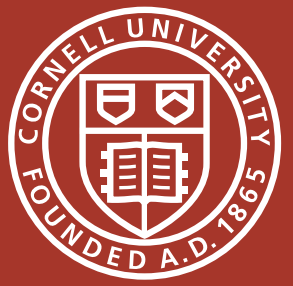
SCDM: A unified approach to Wannier localization

Anil Damle
Computer Science, Cornell University

Wannier90 3.0, March 2020



Acknowledgements

- In collaboration with:
  - Lin Lin — University of California, Berkeley
  - Lexing Ying — Stanford University
  - Antoine Levitt — Inria Paris
- Funding provided by:
  - NSF, Simons Foundation, DOE

Todays talk: SCDM

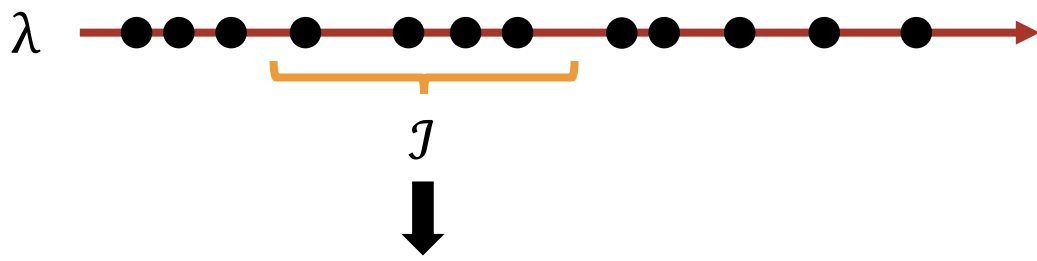
- A “direct” method for computing MLWFs
  - A non-iterative procedure
  - More generally, a robust and automated initialization for an optimization-based approach
  - Not many parameters to choose
- First discuss the isolated case, then the entangled case

Review on MLWFs by Marzari, et. al. [2012]

Our notational perspective for today

- Eigenfunctions $\{\psi_i\}$ of a self-adjoint Hamiltonian $\mathcal{H}$ in interval $\mathcal{I}$

$$\mathcal{H}[\rho]\psi_i(\boldsymbol{r}) = \lambda_i\psi_i(\boldsymbol{r}), \qquad \lambda_i \in \mathcal{I},$$



Associated eigenfunctions: $\{\psi_i(\boldsymbol{r})\}_{i=1}^{N}$

- Localization problem:

minimal set of orthonormal localized $\{\phi_i\}$ (*Wannier functions*) such that

$$\mathrm{span}\{\phi_i\} \subseteq \mathrm{span}\{\psi_i\}_{\lambda_i \in \mathcal{I}}$$

Localized basis set: $\{\phi_i(\boldsymbol{r})\}_{i=1}^{N_w}$

Note: $N_w \leq N$

Notation for this talk

- Consider domain in $\boldsymbol{r}$ to be discretized over N_g uniform gird points
- Naturally pairs with so-called plane wave methods
- Overload ψ_i, ϕ_i, etc. as length N_g vectors, $N_g \times (N \; or \; N_w)$ matrices:

$$\Psi = \begin{bmatrix} | & & | \\ \psi_1 & \cdots & \psi_N \\ | & & | \end{bmatrix} \text{ and } \Phi = \begin{bmatrix} | & & | \\ \phi_1 & \cdots & \phi_{N_w} \\ | & & | \end{bmatrix}$$

Canonical Local

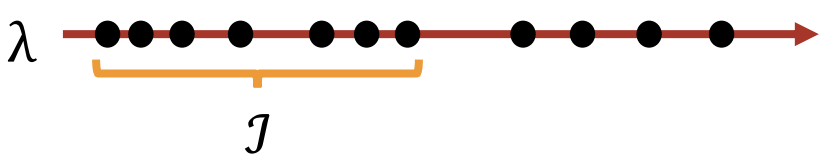
The SCDM methodology

Turns out there is some nice linear algebra under the surface

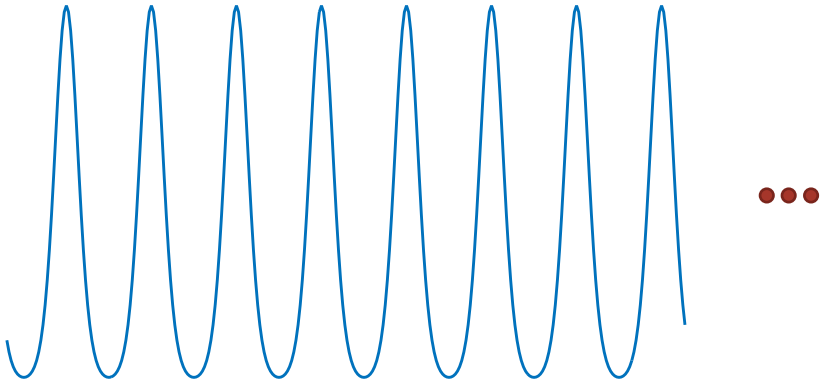
First, simplicity: the isolated case

- Isolated ($N_w = N$):

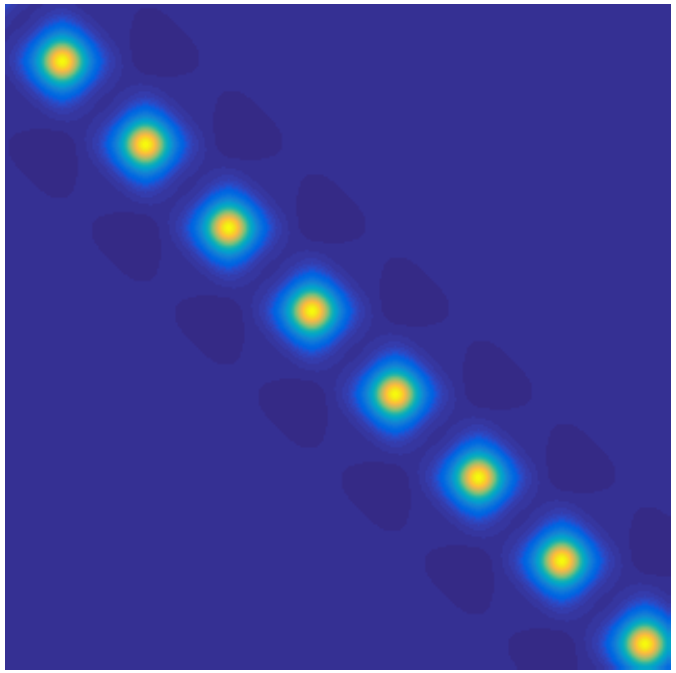
for $\lambda_i \in \mathcal{I}, \lambda_j \notin \mathcal{I}$ $\inf |\lambda_i - \lambda_j| > 0$



$v(\boldsymbol{r})$



$P = \Psi\Psi^* =$



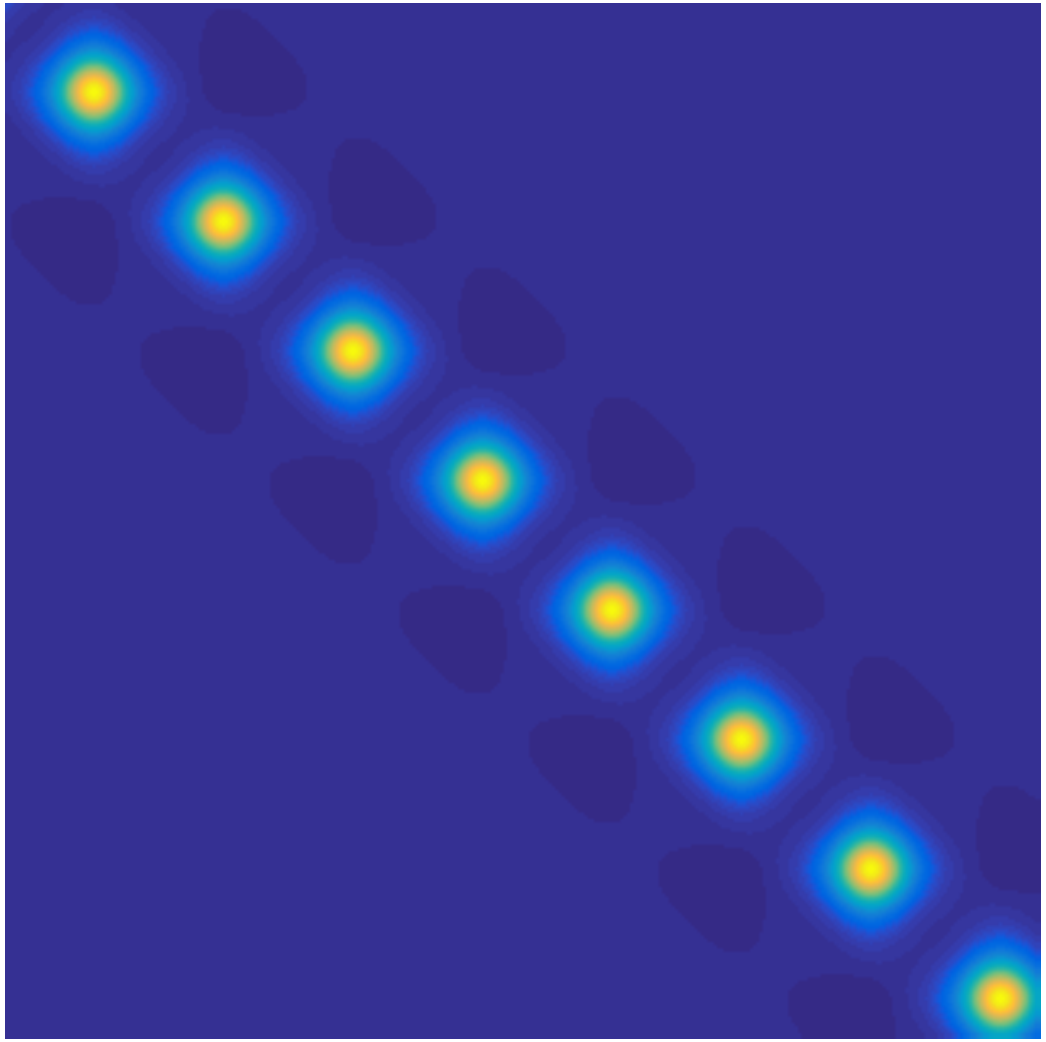
- $P = \Psi\Psi^*$ is the density matrix

SIAM Review reference: Benzi, Boito, Razouk [2013]; So-called Nearsightedness principle in Physics, *e.g.* Kohn [1959]

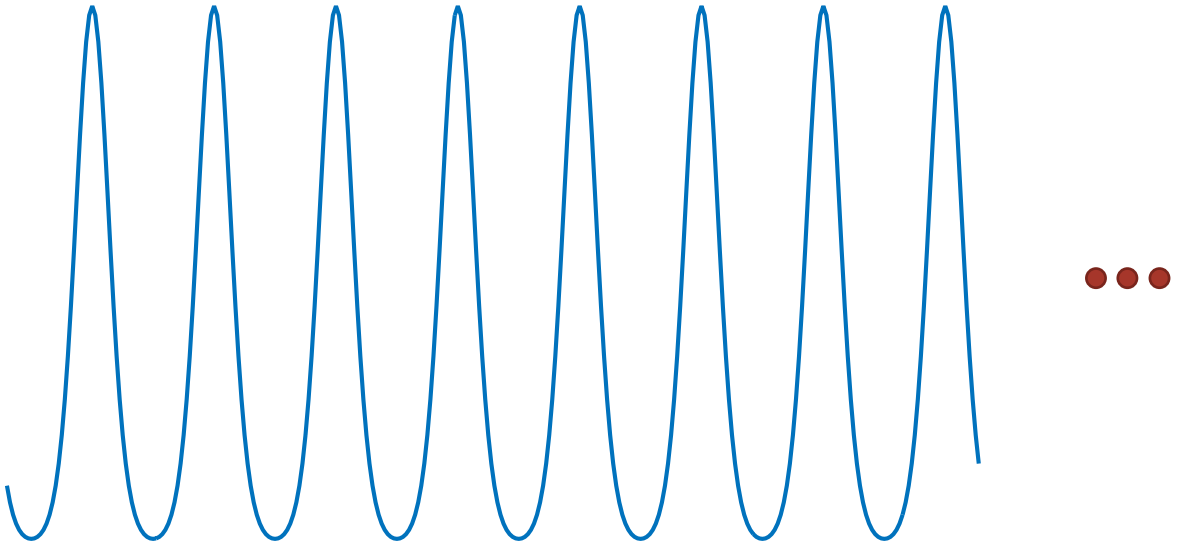
The density matrix

- For insulating systems, $P = \Psi\Psi^*$ has well-localized columns
- In fact, they decay exponentially [Kohn 1959 and 1996]
  - See, *e.g.*, Benzi, Boito, and Razouk [2013] for a nice SIAM Review article
- Columns of P are sparse vectors in the range
  - Use to construct our basis
- Which raises the question: which columns should we select?

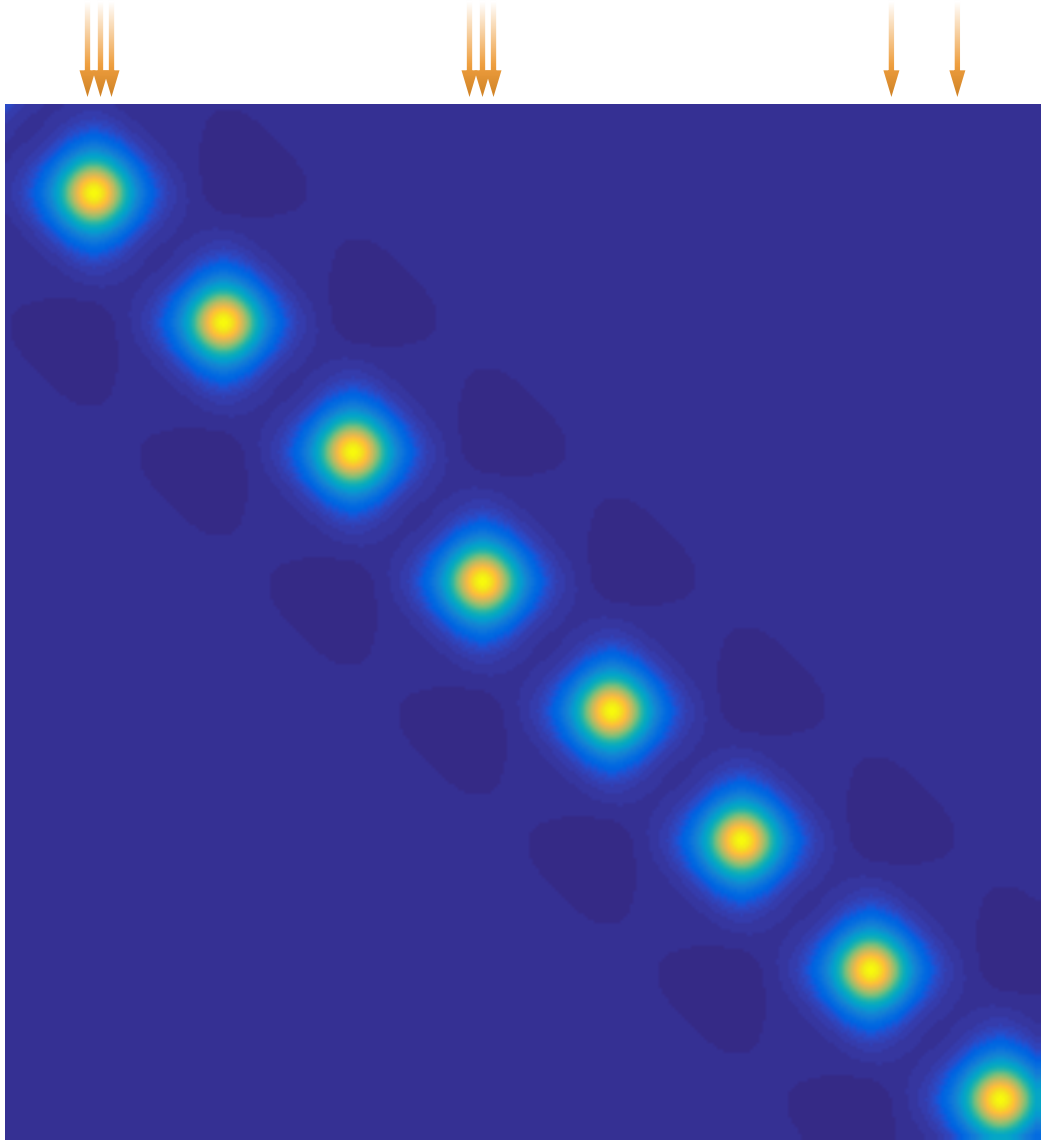
An example density matrix



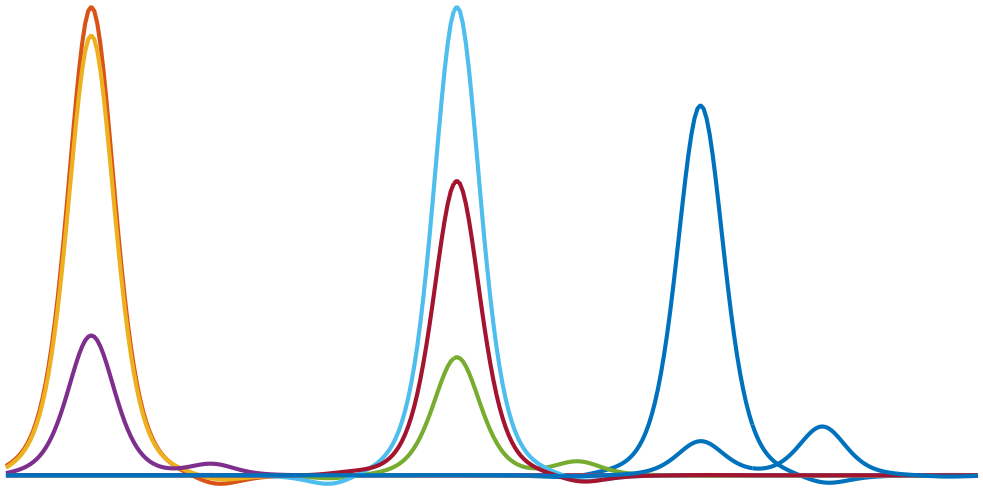
- Diagonal is the electron density, ρ
- Arises from the potential, $V_{ext}(\boldsymbol{r})$



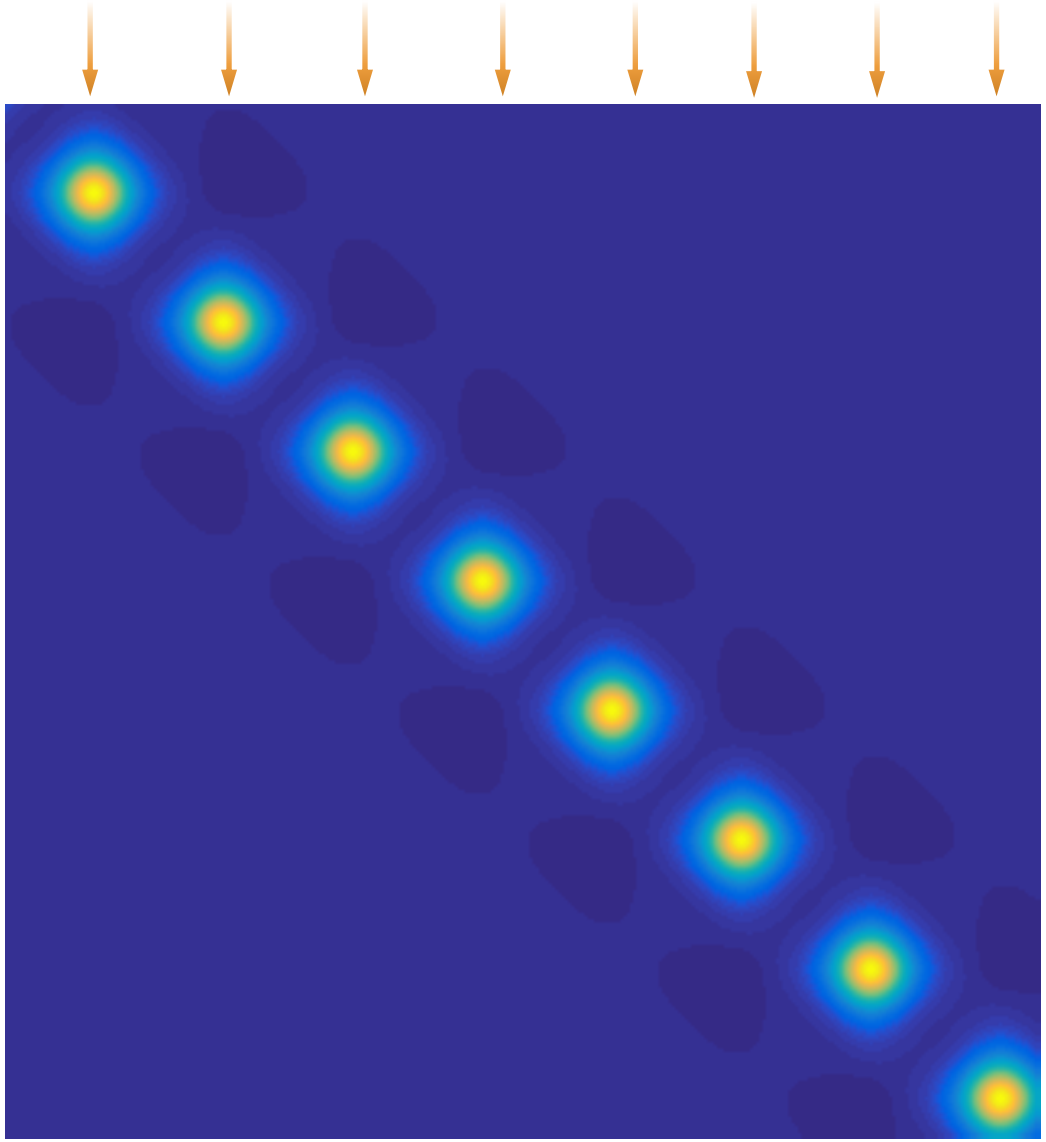
A poor choice of columns



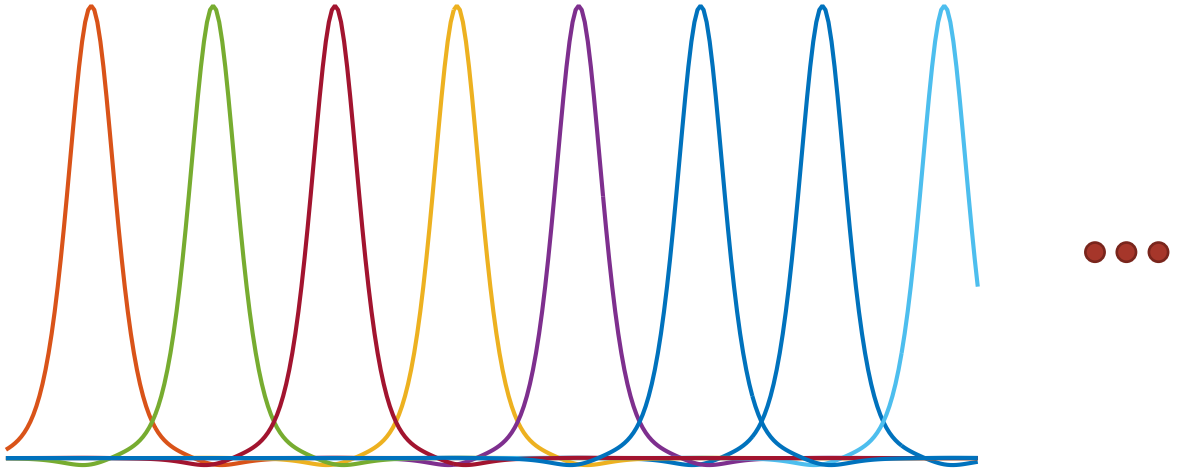
- Could pick any N linearly independent columns
- However, there are potentially poor choices



A good set of columns



- Pick columns that do not overlap much and are well conditioned



Selected columns of the density matrix (SCDM)

- Compute a column pivoted QR factorization:

$$\Psi^*\Pi = Q[R_1 \quad R_2]$$

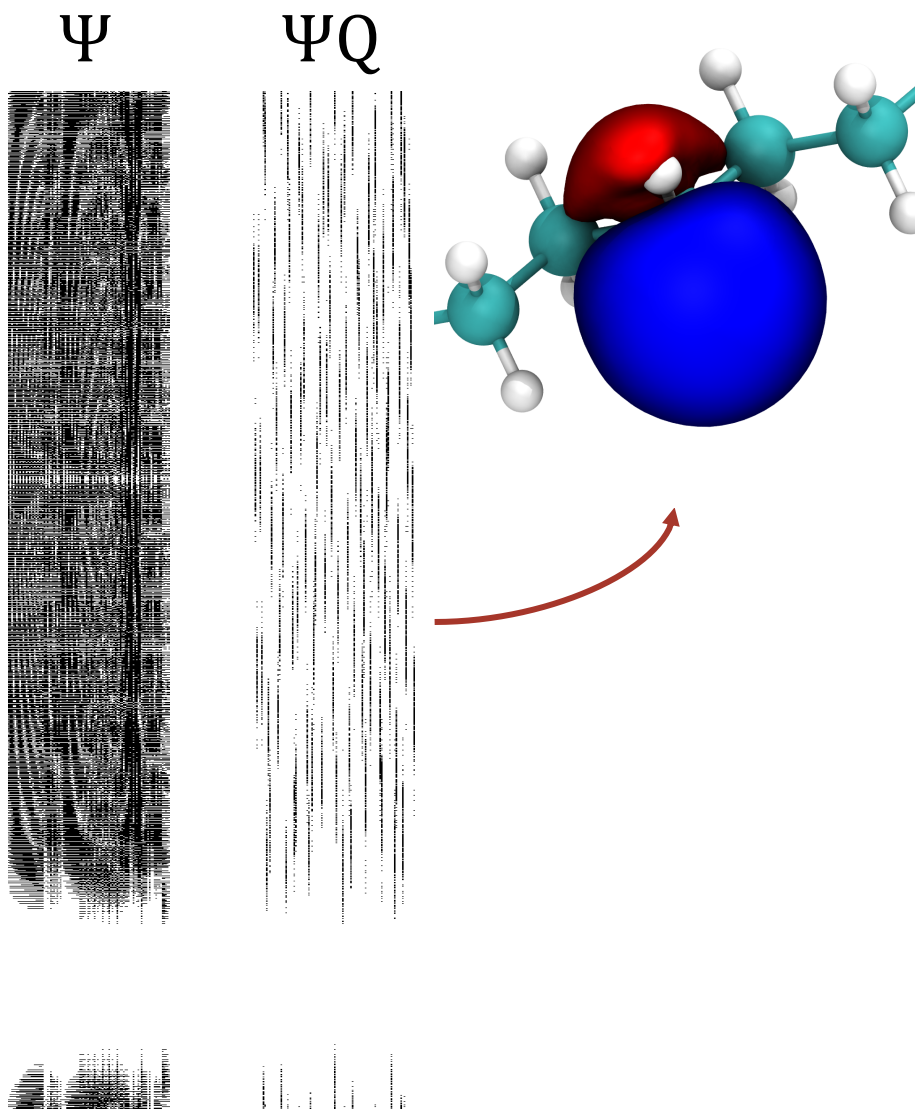
- $\mathcal{C}$ is the first N_w columns selected by Π

Corresponds to N_w columns of P

- Solve (via SVD of $(\Psi^*)_{:,\mathcal{C}}$):

$$\min_{Q^TQ=I} \left\| \Psi Q - (\Psi\Psi^*)_{:,\mathcal{C}} \right\|_F^2$$

- Solution Q yields a well localized basis $\Phi = \Psi Q$



Column pivoted QR

$$P\Pi = Q \begin{pmatrix} R_1 & R_2 \\ 0 & 0 \end{pmatrix}$$

- Π is a permutation, keeps R_1 well conditioned and encodes $\mathcal{C}$
- $\mathrm{T} = R_1^{-1} R_2$ yields $P\Pi = P_{:,\mathcal{C}}(I \quad T)$
- Part of a broader class of so-called rank-revealing QR factorizations
- Apply to Ψ^* instead and get a good set of pivots $\mathcal{C}$
  - Avoid ever having to construct P, identical results when using certain algorithms

QRCP paper [Golub and Businger 1965], textbook [Golub and Van Loan 1996]

An alternative perspective

- Compute a column pivoted QR factorization:

$$\Psi^* \Pi = Q [R_1 \quad R_2]$$

Yields a well-conditioned set of columns of P

- $\mathcal{C}$ is the first N_w columns selected by Π

Corresponds to N_w columns of P

- Solve (via SVD of $(\Psi^*)_{:,\mathcal{C}}$):

$$\min_{Q^T Q = I} \left\| \Psi Q - (\Psi \Psi^*)_{:,\mathcal{C}} \right\|_F^2$$

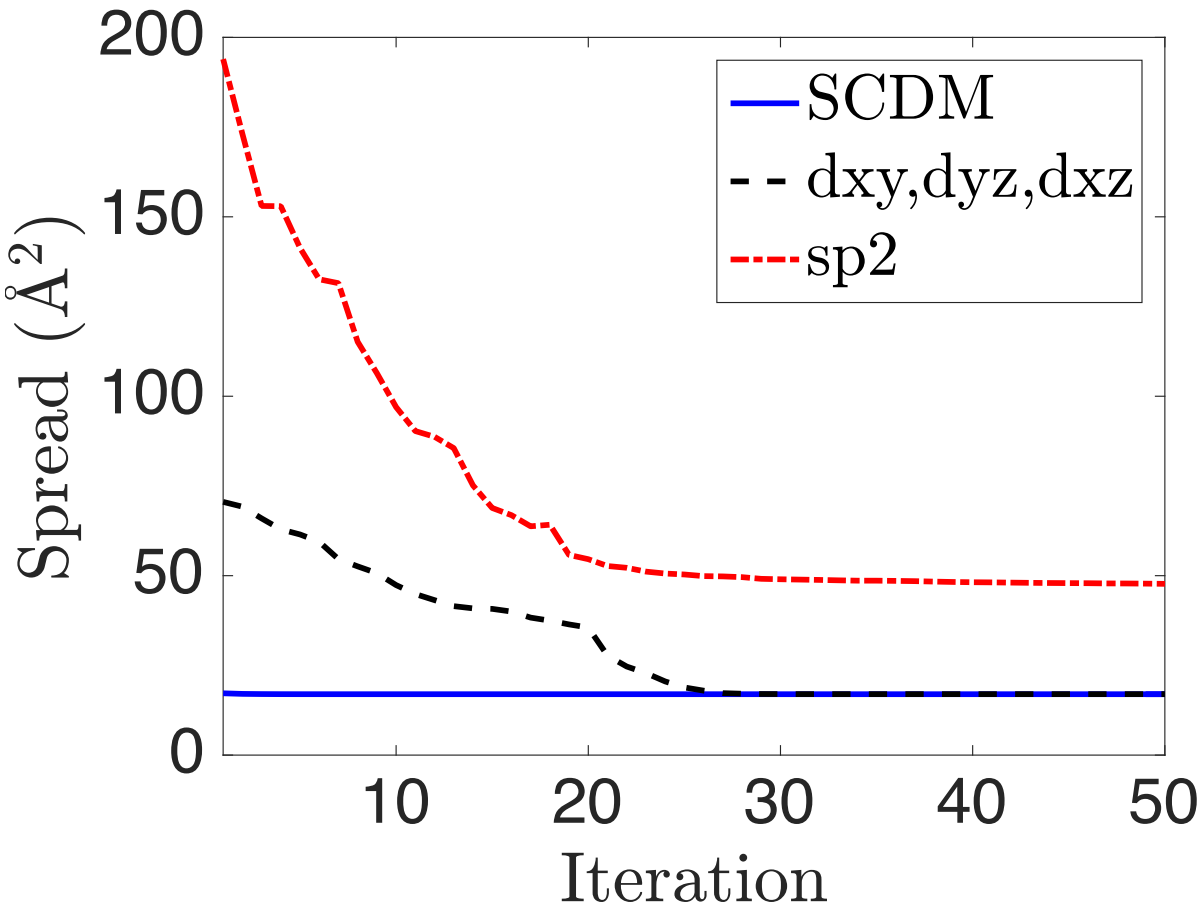
Templates for localized functions

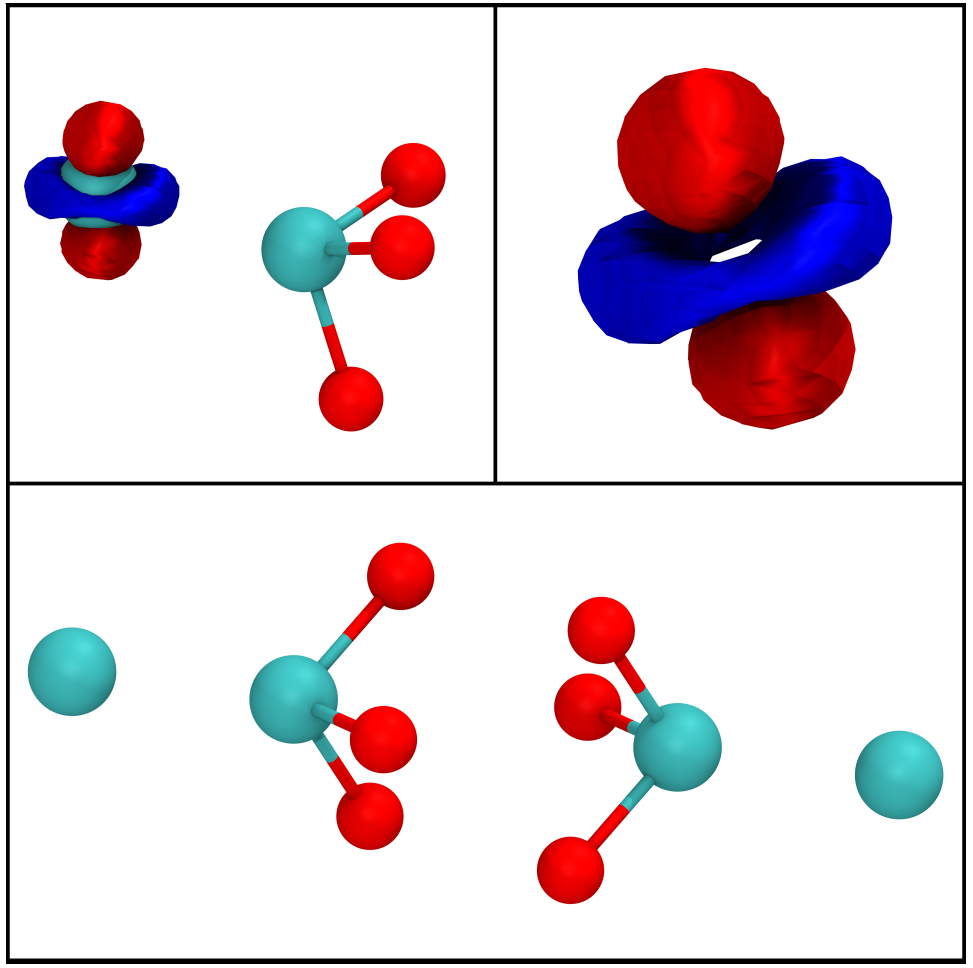
Just my way of writing Löwdin orthogonalization

- Solution Q yields a well localized basis $\Phi = \Psi \mathrm{Q}$

As an initial guess for cr2o3

wannier90 to optimize spread



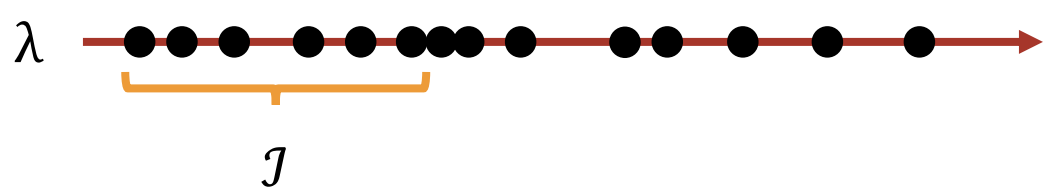


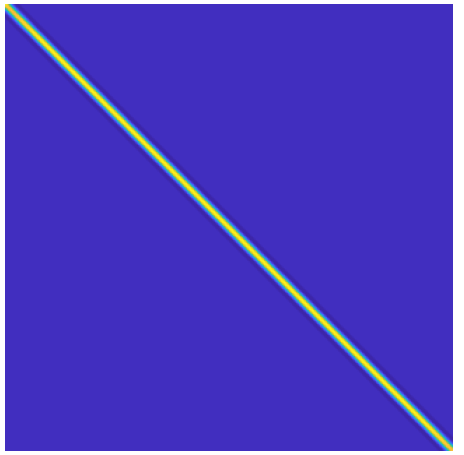
What if $\mathcal{I}$ is not isolated?

- P does not decay nicely
- Use a quasi-density matrix instead

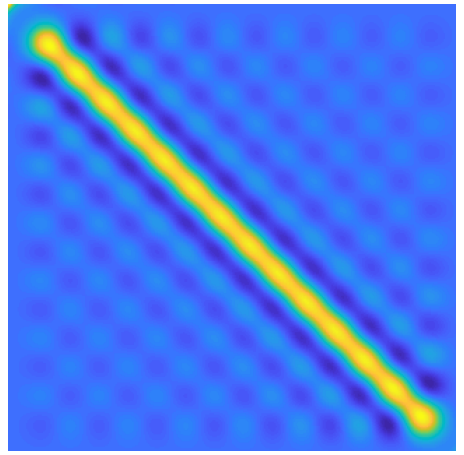
$$\hat{P} = \sum_i \psi_i f(\lambda_i) \psi_i^* = f(H)$$

- Decays rapidly for smooth f
- Want N_w localized functions, start with N eigenfunctions





$\hat{P}$

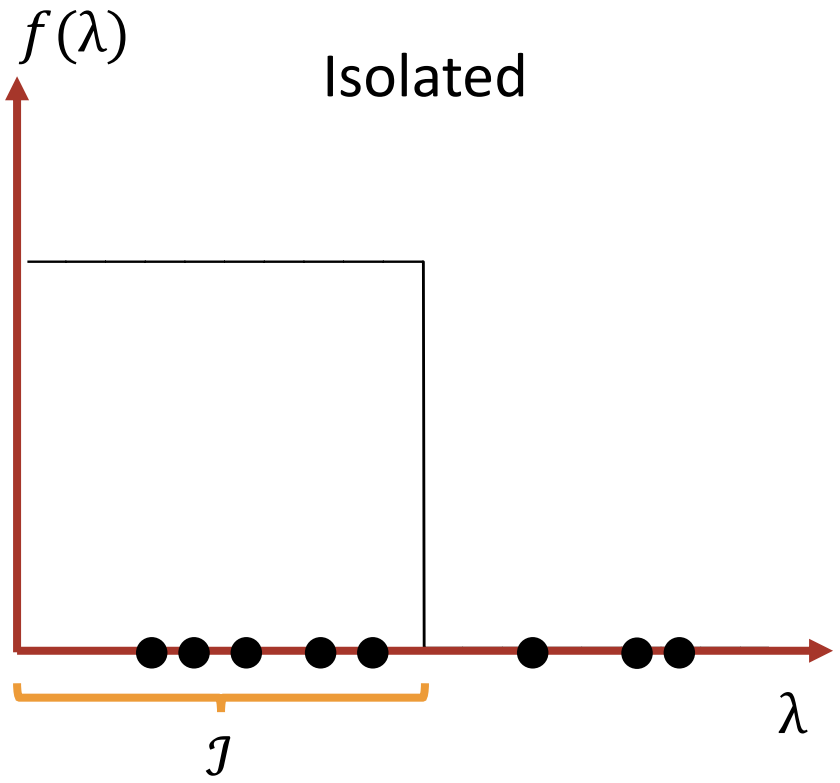


P

To talk about metals and insulators in a unified framework, I have omitted details of the decay rates

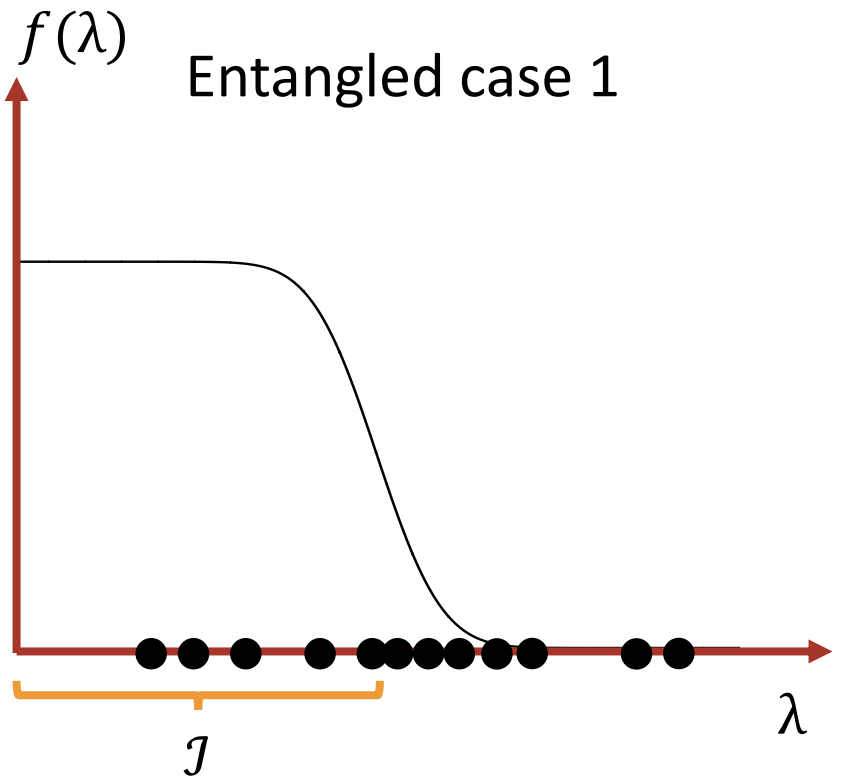
Choices for f, indicates the eigenvalues of interest

$\mathcal{I} = (-\infty, \mu_c)$



Isolated

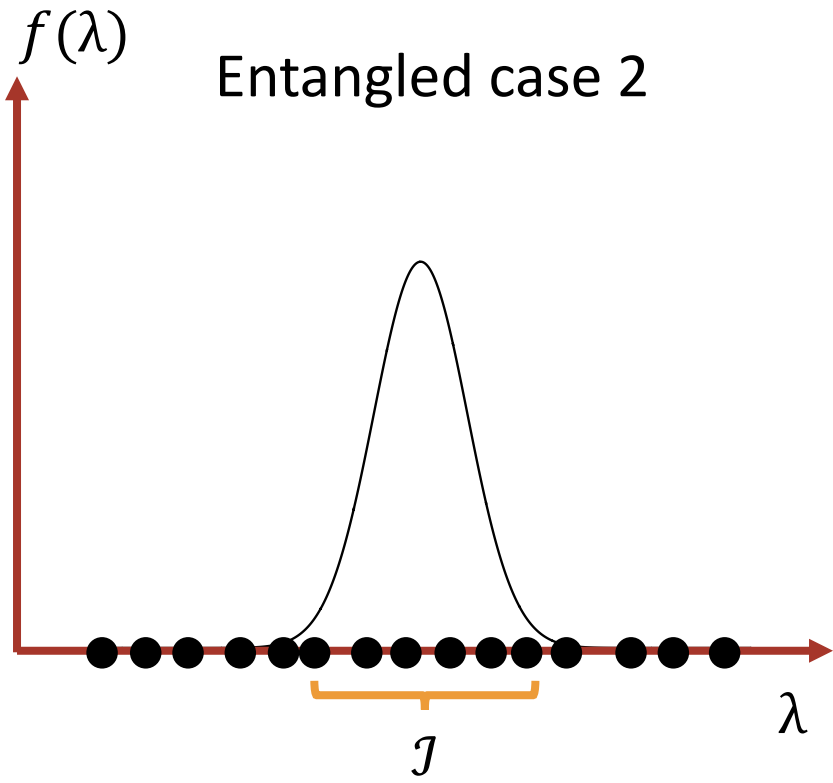
$\mathcal{I} = (-\infty, \mu_c)$



Entangled case 1

$$f(\lambda) = \frac{1}{2}\operatorname{erfc}\left(\frac{\lambda-\mu_c}{\sigma}\right)$$

$\mathcal{I} = (\mu_c - \sigma, \mu_c + \sigma)$



Entangled case 2

$$f(\lambda) = \exp\left(\frac{-(\lambda-\mu_c)^2}{\sigma^2}\right)$$

SCDM in this setting

- Let $\Lambda = \mathrm{diag}(\{\lambda_i\})$ for i such that $|f(\lambda_i)| > \delta$
- Compute a rank-revealing QR factorization:

$$f(\Lambda)\Psi^*\Pi = \mathrm{Q}[R_1 \quad R_2]$$

- $\mathcal{C}$ is the first N_w columns selected by Π
- Solve (via SVD of $(f(\Lambda)\Psi^*)_{:,\mathcal{C}}$)

$$\min_{Q^TQ=I} \left\| \Psi Q - \Psi (f(\Lambda)\Psi^*)_{:,\mathcal{C}} \right\|_F^2$$

- Q yields a reasonably localized basis $\Phi = \Psi Q$

QRCP of $\hat{P} = f(H)$

⬇

A subspace of dimension N_w and N_w localized vectors in it

⬇

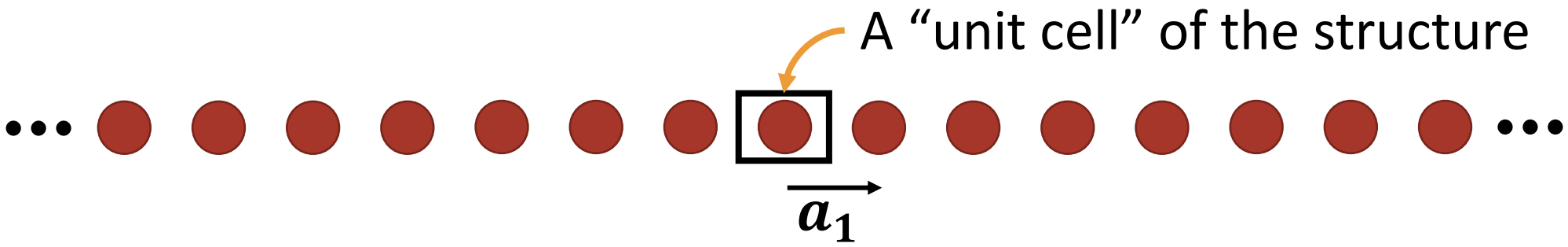
Find Q to align Ψ with these columns (note, now $N \times N_w$)

Simultaneously finds the subspace and localized basis!

Now, crystals

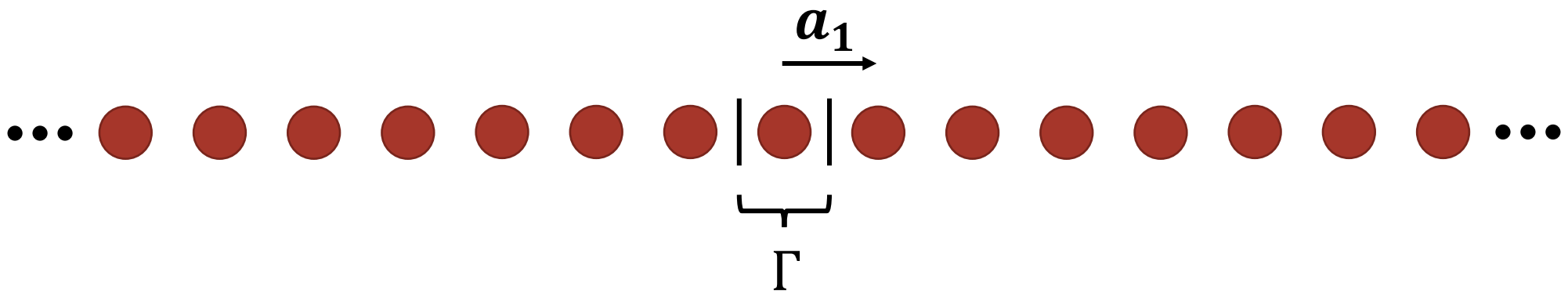
Illustrated in 1D for simplicity

Periodic $v(\boldsymbol{r})$



- Potential $v(\boldsymbol{r})$ satisfies $v(\boldsymbol{r} + \boldsymbol{a}_1 n_1) = v(\boldsymbol{r})\ \forall\ \boldsymbol{r} \in \mathbb{R}, n_1 \in \mathbb{Z}$

Now, necessary notation



- A so-called Bravis lattice with vector $\boldsymbol{a_1}$: $\mathbb{L} = \{\boldsymbol{R} | \boldsymbol{R} = \boldsymbol{a_1} n_1, n_1 \in \mathbb{Z}\}$
- Unit cell: $\Gamma = \{\boldsymbol{r} | \boldsymbol{r} = c_1 \boldsymbol{a_1}, -{}^1/_2 \leq c_1 < {}^1/_2\}$
- Reciprocal lattice $\mathbb{L}^*$, Fourier counterpart of $\mathbb{L}$
  - Its unit cell is the (first) Brillouin zone, denoted Γ^*

Bloch-Floquet notation

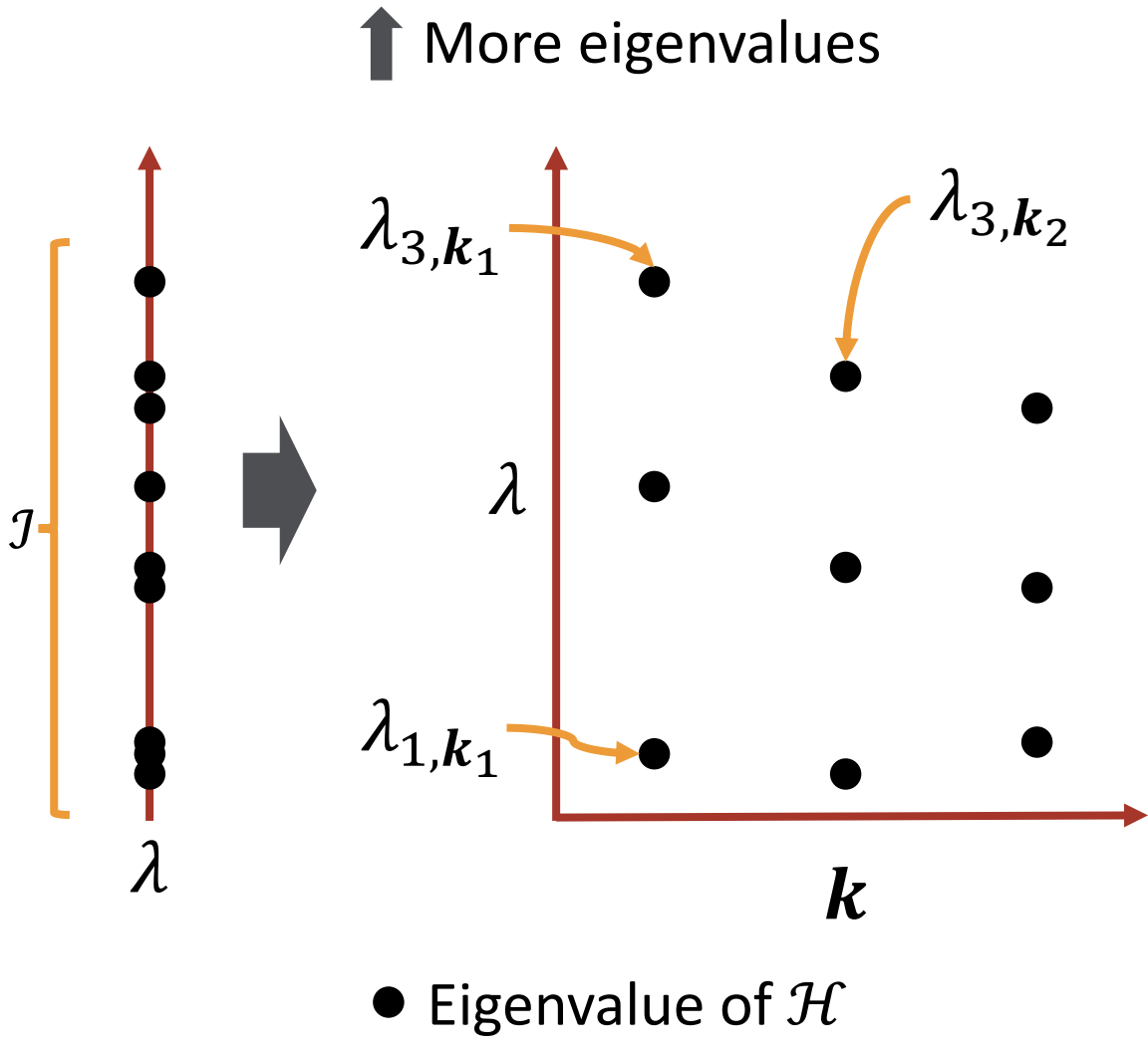
- Solve $\mathcal{H}[\rho]\psi(\boldsymbol{r}) = \lambda\psi(\boldsymbol{r}),\ \lambda \in \mathcal{I}$
- Relabel spectrum of $\mathcal{H}$ via two indices
  - Band $b \in \mathbb{N}$ and Brillouin zone $\boldsymbol{k} \in \Gamma^*$
  - Block orbitals $\psi_{b,\boldsymbol{k}}(\boldsymbol{r})$:

$$\psi_{b,\boldsymbol{k}}(\boldsymbol{r}) = e^{i\boldsymbol{k}\cdot\boldsymbol{r}} u_{b,\boldsymbol{k}}(\boldsymbol{r})$$

  - $u_{b,\boldsymbol{k}}$ is $\mathbb{L}$ periodic and solved for (on $\boldsymbol{r} \in \Gamma$) via:

$$\mathcal{H}(\boldsymbol{k}) u_{b,\boldsymbol{k}}(\boldsymbol{r}) = \lambda_{b,\boldsymbol{k}} u_{b,\boldsymbol{k}}(\boldsymbol{r})$$

$$\mathcal{H}(\boldsymbol{k}) = -\tfrac{1}{2}(\nabla + i\boldsymbol{k})^2 + v(\boldsymbol{r})$$



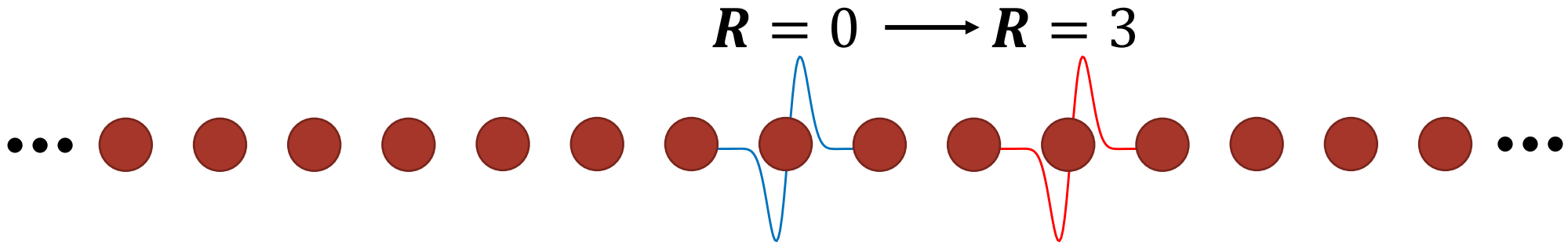
Basis transformations to Wannier functions

- We want a basis $\{w_{b,\boldsymbol{R}}(\boldsymbol{r})\}$ that is spatially localized
- Will construct via an alternative basis for $\{\psi_{b,\boldsymbol{k}}(\boldsymbol{r})\}$ denoted $\{\tilde{\psi}_{b,\boldsymbol{k}}(\boldsymbol{r})\}$
- Express via a set of $N \times N_w$ unitary matrices $Q(\boldsymbol{k})$ (a so-called gauge)

$$\tilde{\Psi}_{\boldsymbol{k}} = \Psi_{\boldsymbol{k}} Q(\boldsymbol{k})$$

Want to identify a subspace and a localized basis

Wannier functions for crystals



- Some $\tilde{\psi}$ yields Wannier functions:

$$w_{b,0}(\boldsymbol{r}) = \frac{1}{|\Gamma^*|}\int_{\boldsymbol{\Gamma}^*} \tilde{\psi}_{b,\boldsymbol{k}}(\boldsymbol{r})d\boldsymbol{k}$$

$$\Updownarrow$$

$$w_{b,0}(\boldsymbol{r}) = \frac{1}{|\Gamma^*|}\int_{\boldsymbol{\Gamma}^*} \tilde{u}_{b,\boldsymbol{k}}(\boldsymbol{r})e^{i\boldsymbol{k}\cdot\boldsymbol{r}}d\boldsymbol{k}$$

- Translates for other $\boldsymbol{R}$:

$$w_{b,\boldsymbol{R}}(\boldsymbol{r}) = \frac{1}{|\Gamma^*|}\int_{\boldsymbol{\Gamma}^*} \tilde{\psi}_{b,\boldsymbol{k}}(\boldsymbol{r})e^{-i\,\boldsymbol{k}\cdot\boldsymbol{R}}d\boldsymbol{k}$$

$\boldsymbol{r} \in \mathbb{R}$ and $\boldsymbol{R} \in \mathbb{L}$

A "good" basis in this sense exists only for some class of $\mathcal{H}$ — we omit, *e.g.*, topological insulators

Which $\tilde{\psi}$, assuming an isolated system ($N_w = N$)

Analytic $\tilde{\psi}_{b,\boldsymbol{k}}$ in $\boldsymbol{k}$ $\Longleftrightarrow$ Localized $w_{b,\boldsymbol{R}}(\boldsymbol{r})$

- Pick a "gauge" so that $\tilde{\psi}_{b,\boldsymbol{k}}$ are smooth with respect to $\boldsymbol{k}$
- $\boldsymbol{k}$ dependent density matrix (mild notational abuse)

$$P(\boldsymbol{k}) = \sum_{\lambda_{b,\boldsymbol{k}} \in \mathcal{I}} \psi_{b,\boldsymbol{k}} \psi^*_{b,\boldsymbol{k}} = \sum_{\lambda_{b,\boldsymbol{k}} \in \mathcal{I}} \tilde{\psi}_{b,\boldsymbol{k}} \tilde{\psi}^*_{b,\boldsymbol{k}}$$

- Gauge invariant and analytic w.r.t $\boldsymbol{k}$
- A fixed (in $\boldsymbol{k}$) set of "columns" of $P(\boldsymbol{k})$ accomplishes this
  - Want them to be well conditioned (singular values uniformly bounded away from zero)

Panati and Pisante [2013], Nenciu [1991], des Cloizeaux [1964], etc.

Now, things become discrete

- Practically, discretize Γ^* (simultaneously truncate $\mathbb{L}$)
  - Via set $\mathcal{K}$ with $N_{\boldsymbol{k}}$ points
- Discretize Γ with N_g uniform points
  - $\Psi_{\boldsymbol{k}} = \begin{bmatrix} | & & | \\ \psi_{1,\boldsymbol{k}} & \cdots & \psi_{N,\boldsymbol{k}} \\ | & & | \end{bmatrix}$ is $N_g \times N$ and $P(\boldsymbol{k})$ is $N \times N$
- $\boldsymbol{k}_0 = 0$ is known as the Γ-point

The isolated case, algorithmically

- Computer the QRCP

$$\Psi^*_{\boldsymbol{k}_0}\Pi = Q[R_1 \quad R_2]$$

- $\mathcal{C}$ is the first N columns selected by Π

  Corresponds to N columns of P

- For each $\boldsymbol{k}$ solve (via SVD of $(\Psi^*_{\mathbf{k}})_{:,\mathcal{C}}$):

$$\min_{Q(\boldsymbol{k})^T Q(\boldsymbol{k})=I} \left\| \Psi_{\boldsymbol{k}} Q(\boldsymbol{k}) - (\Psi_{\boldsymbol{k}}\Psi^*_{\boldsymbol{k}})_{:,\mathcal{C}} \right\|_F^2$$

- Gauge $\{Q(\boldsymbol{k})\}$ yields a smooth $\{\tilde{\psi}_{b,\boldsymbol{k}}\}$

- Need to use a single $\mathcal{C}$ for all $\boldsymbol{k}$
- Prior work shows using the Γ-point suffices

- Independent problem for each $\boldsymbol{k}$

The entangled case

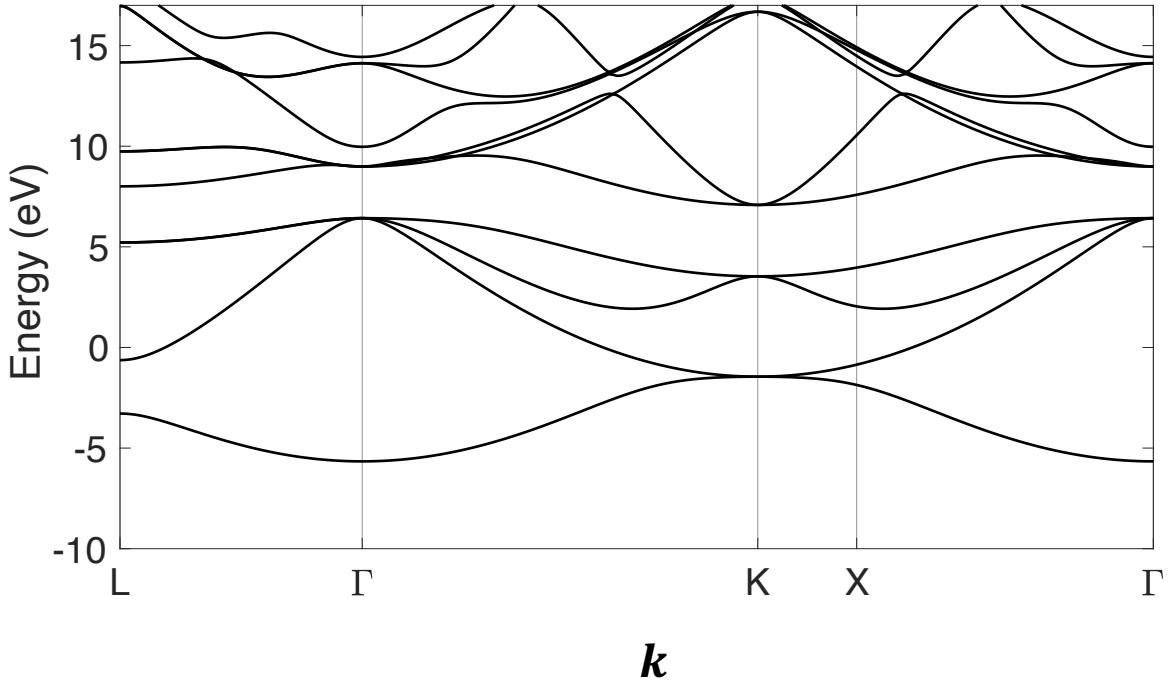
- Isolation condition across $\boldsymbol{k}$, *i.e.*

$$\inf \left| \lambda_{b,\boldsymbol{k}} - \lambda_{b',\boldsymbol{k}'} \right| > 0$$

- Use quasi-density matrices

$$\hat{P}(\boldsymbol{k}) = \sum_{\lambda_{b,\boldsymbol{k}}} \psi_{b,\boldsymbol{k}} f(\lambda_{b,\boldsymbol{k}}) \psi_{b,\boldsymbol{k}}^{*}$$

- Numerically observed smooth in $\boldsymbol{k}$



The entangled case

- Let $\Lambda_{\boldsymbol{k}} = \mathrm{diag}(\{\lambda_{i,\boldsymbol{k}}\})$ for i such that $|f(\lambda_{i,\boldsymbol{k}})| > \delta$
- Compute the QRCP:

$$f(\Lambda_{\boldsymbol{k_0}})\Psi_{\boldsymbol{k_0}}^{*}\Pi = \mathrm{Q}[R_1 \quad R_2]$$

- $\mathcal{C}$ is the first N_w columns selected by Π
- For each $\boldsymbol{k}$ solve (via SVD of $(f(\Lambda_{\boldsymbol{k}})\Psi_{\boldsymbol{k}}^{*})_{:,\mathcal{C}}$)

$$\min_{Q(\boldsymbol{k})^T Q(\boldsymbol{k})=I} \left\| \Psi_{\boldsymbol{k}} Q(\boldsymbol{k}) - \Psi_{\boldsymbol{k}} (f(\Lambda_{\boldsymbol{k}})\Psi_{\boldsymbol{k}}^{*})_{:,\mathcal{C}} \right\|_F^2$$

- Gauge $\{Q(\boldsymbol{k})\}$ yields a smooth $\{\tilde{\psi}_{b,\boldsymbol{k}}\}$

- Observed numerically
- Requires singular values of

$$(f(\Lambda_{\boldsymbol{k}})\Psi_{\boldsymbol{k}}^{*})_{:,\mathcal{C}}$$

uniformly bounded away from zero

SCDM for crystals

- A “direct” method that does not require an initial guess
  - Contrast to existing methods
- Computationally efficient per $\boldsymbol{k}$-point, the QRCP is only done once
- Only 2 parameters, μ_c and σ (to denote the region of interest)
- In some sense, the two “disentanglement” steps [Souza, Marzari and Vanderbilt 2001] are accomplished simultaneously

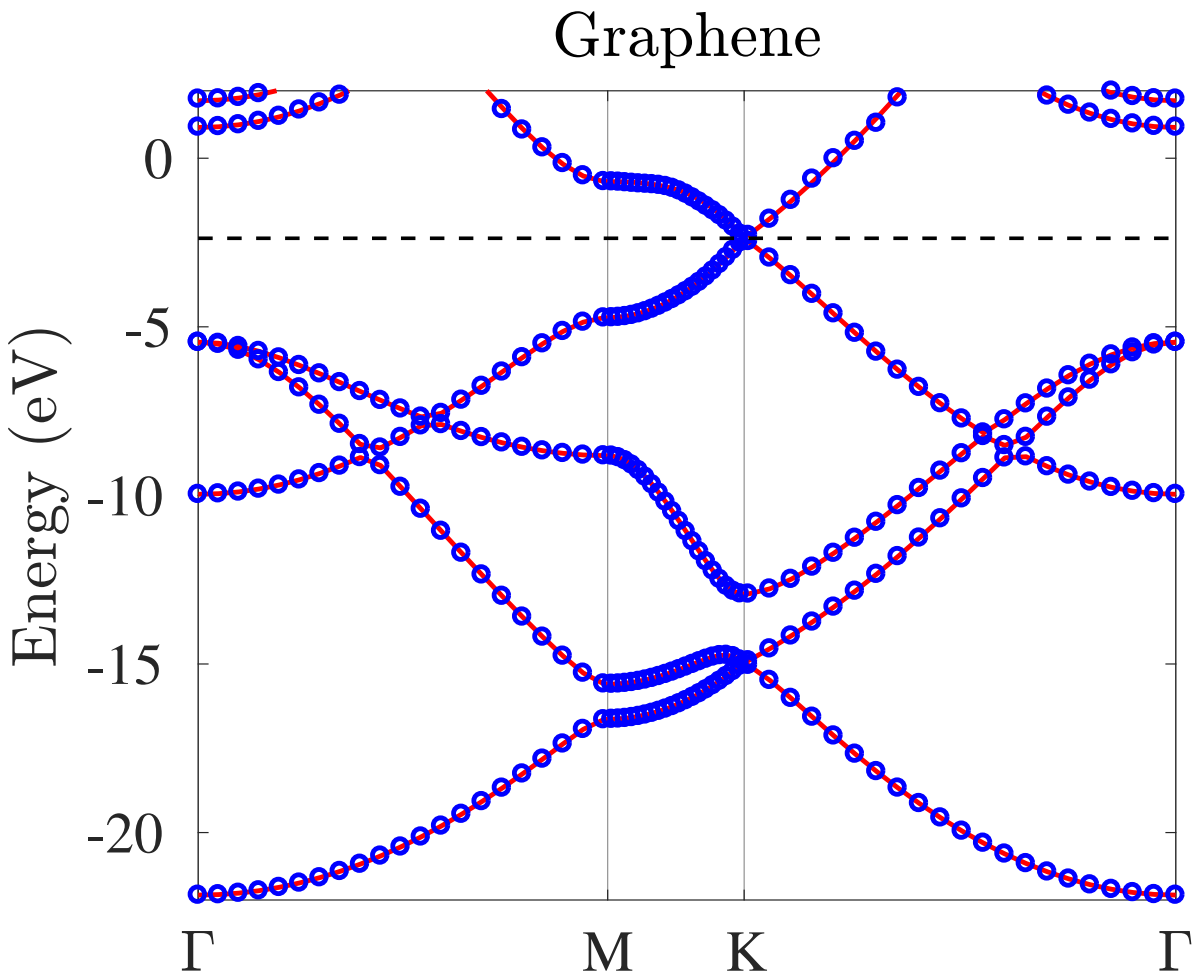
Numerical examples

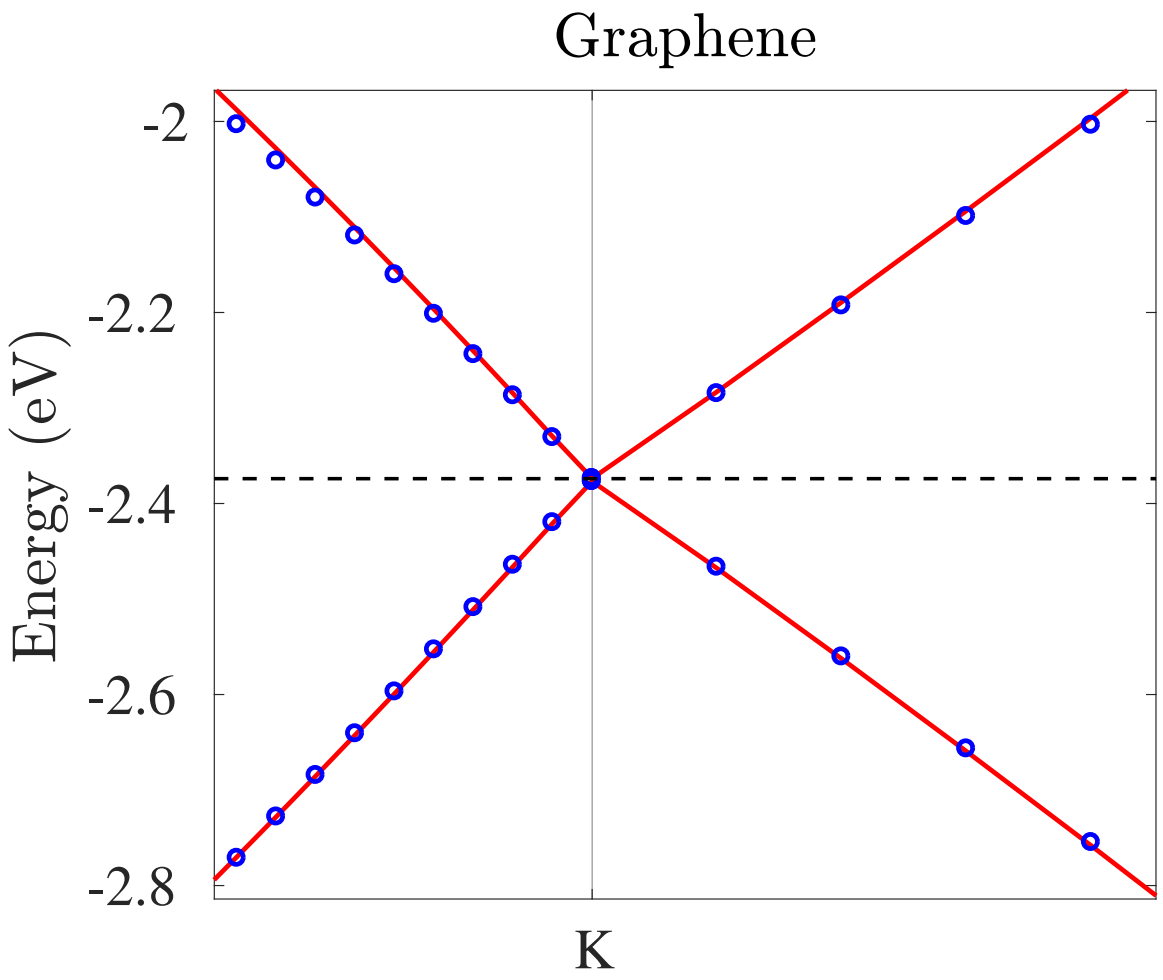
Using Quantum ESPRESSO and wannier90 as needed

SCDM is now in wannier90

https://github.com/asdamle/SCDM

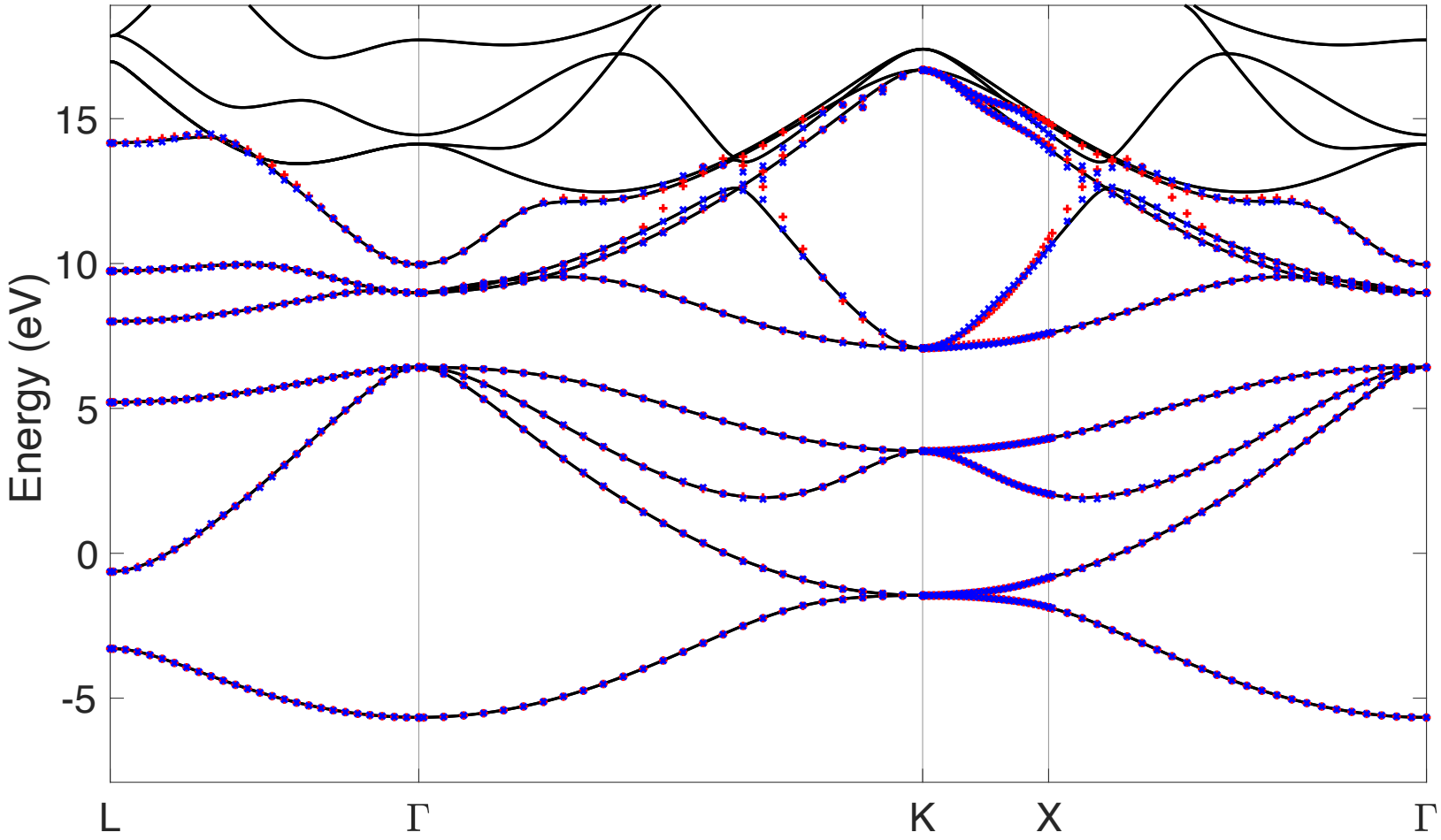
Graphene





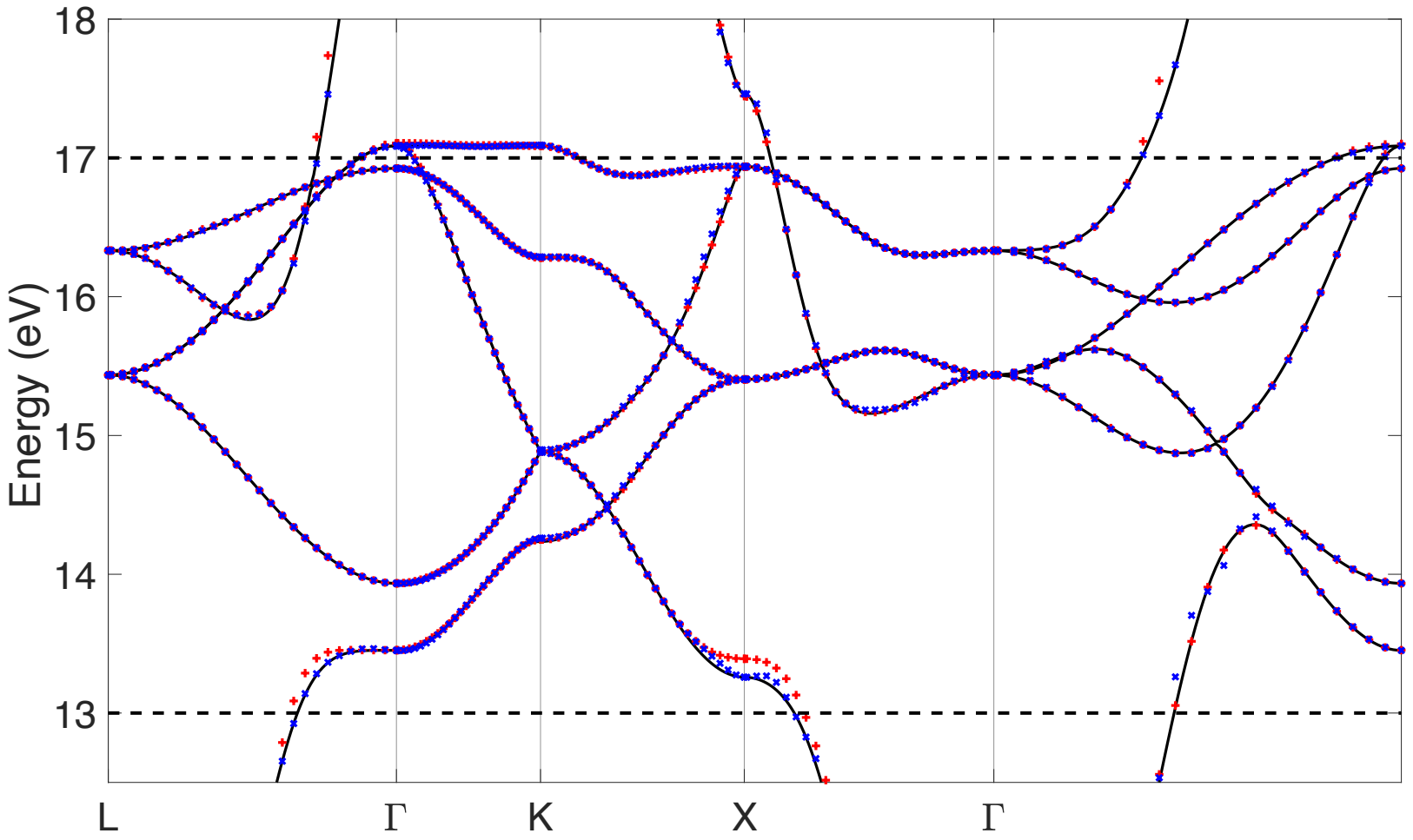
12x12x1 k-point grid, reference calculation in red, and SCDM interpolant in blue

Silicon



10x10x10 k-point grid, reference calculation in black, SCDM in blue, wannier90 in red

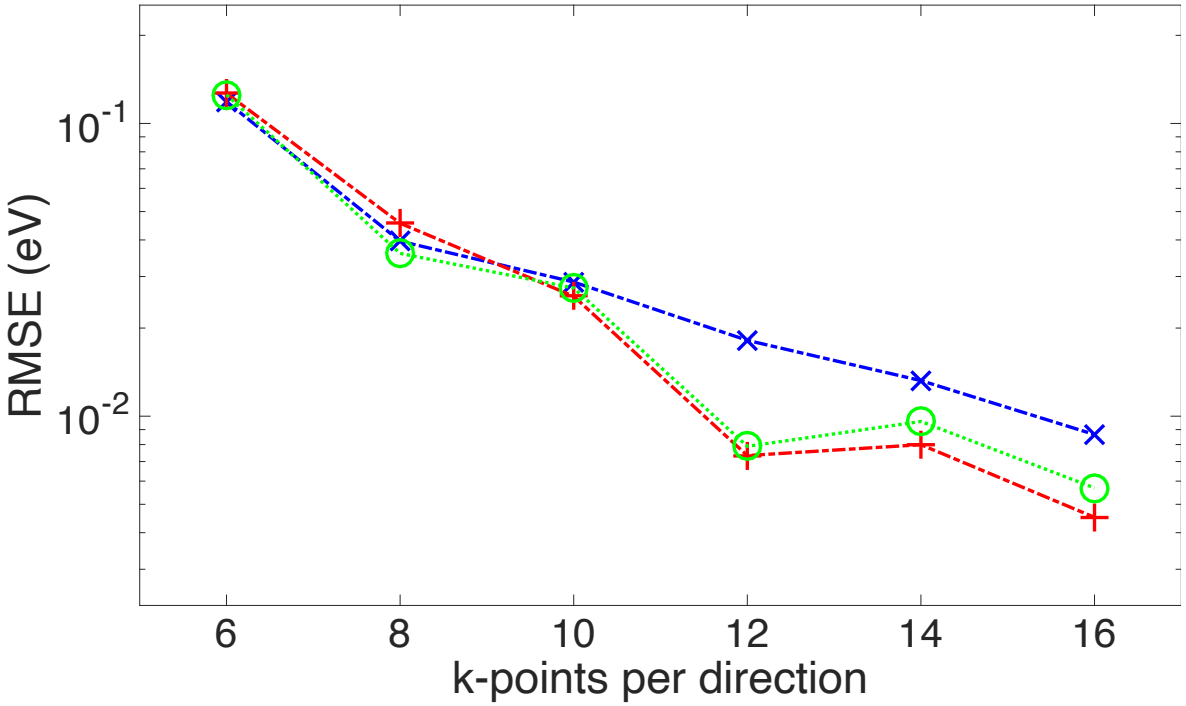
Copper



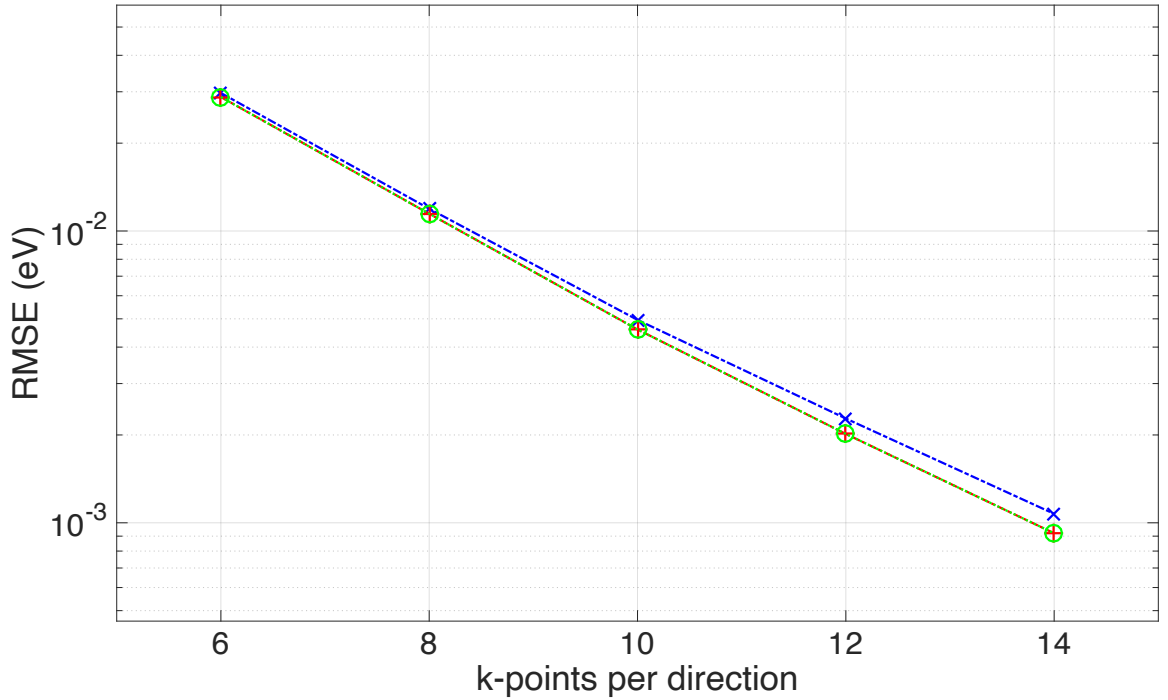
10x10x10 k-point grid, reference calculation in black, SCDM in blue, wannier90 in red

Convergence in interpolation

Al



Si



Interpolation RMSE convergence

SCDM in blue, wannier90 in red, our new optimization method in green

SCDM in Wannier90 3.0

Thank you to Valerio Vitale + all the Wannier90 developers

How to use SCDM in Wannier90 3.0

- In the .win file:
  - Add an "`auto_projections = .true.`" line
- In the .pw2wan file:
  - Add "`scdm_proj = true`"
  - Pick f: "`scdm_entanglement = {'isolated', 'erfc', 'gaussian'}`"
  - If needed set "`scdm_mu =`" and "`scdm_sigma =`"

References

- cs.cornell.edu/~damle

- Anil Damle and Lin Lin "Disentanglement via entanglement: A unified method for Wannier localization," *SIAM Multiscale Modeling and Simulation*, 2018, 16 (3), 1392-1410

- Anil Damle, Antoine Levitt, and Lin Lin "Variational formulation for Wannier functions with entangled band structure," *SIAM Multiscale Modeling and Simulation*, 2019, 17 (1), 167-191

- Anil Damle, Lin Lin, Lexing Ying, "Compressed representation of Kohn–Sham orbitals via selected columns of the density matrix," *J. Chem. Theory Comput.*, 2015, 11 (4), 1463–1469