Why `pseudobands` works?

Jinyuan Wu

January 8, 2024

A review of GW

GW in momentum space

$-\mathrm{i}\,\Sigma \approx GW^{\text{RPA}} =$  . In $\mathbf{k}$ space, $M_{nn'}(\mathbf{k},\mathbf{q},\mathbf{G}) = \underbrace{\langle n\mathbf{k}+\mathbf{q}|\mathrm{e}^{\mathrm{i}(\mathbf{q}+\mathbf{G})\cdot\mathbf{r}}|n'\mathbf{k}\rangle}_{\text{quasiparticle approx.}}$; $v(\mathbf{q}+\mathbf{G}) = \frac{4\pi}{V|\mathbf{q}+\mathbf{G}|^2}$.

In `epsilon` $\epsilon_{\mathbf{G}\mathbf{G}'}(\mathbf{q},\omega) = \delta_{\mathbf{G}\mathbf{G}'} - v(q+\mathbf{G})\chi_{\mathbf{G}\mathbf{G}'}(\mathbf{q},\omega)$,

$$\chi^{\text{r/a}}_{\mathbf{G}\mathbf{G}'}(\mathbf{q},\omega) = \sum_{\mathbf{k}} \sum_{n}^{\text{occ}} \sum_{n'}^{\text{emp}} M_{nn'}(\mathbf{k}\mathbf{q}\mathbf{G}) M^*_{nn'}(\mathbf{k},\mathbf{q},\mathbf{G}') \left(\frac{1}{\omega + E_{n,\mathbf{k}+\mathbf{q}} - E_{n'\mathbf{k}} \pm \mathrm{i}\,0^+} + \frac{1}{-\omega + E_{n,\mathbf{k}+\mathbf{q}} - E_{n'\mathbf{k}} \mp \mathrm{i}\,0^+} \right).$$

In `sigma` $\Sigma \overset{\int\text{contour}}{=} \Sigma^{\text{CH}} + \Sigma^{\text{SX}}$,

$$\langle n\mathbf{k}|\Sigma^{\text{SX}}(\omega)|n'\mathbf{k}\rangle = -\sum_{n''}^{\text{occ}} \sum_{\mathbf{q}\mathbf{G}\mathbf{G}'} M^*_{n''n}(\mathbf{k},-\mathbf{q},-\mathbf{G}) M_{n''n'}(\mathbf{k},-\mathbf{q},-\mathbf{G}') \epsilon^{-1}_{\mathbf{G}\mathbf{G}'}(\mathbf{q},\omega - E_{n'',\mathbf{k}-\mathbf{q}}) v(\mathbf{q}+\mathbf{G}')$$

$$\langle n\mathbf{k}|\Sigma^{\text{CH}}(\omega)|n'\mathbf{k}\rangle = \frac{\mathrm{i}}{2\pi} \sum_{n''} \sum_{\mathbf{q}\mathbf{G}\mathbf{G}'} M^*_{n''n}(\mathbf{k},-\mathbf{q},-\mathbf{G}) M_{n''n'}(\mathbf{k},-\mathbf{q},-\mathbf{G}') \int_0^\infty \frac{[\epsilon^{\text{r}}(\mathbf{q},\omega') - \epsilon^{\text{a}}(\mathbf{q},\omega')]^{-1}_{\mathbf{G}\mathbf{G}'}\,\mathrm{d}\omega'}{\omega - E_{n\mathbf{k}} - \omega' + \mathrm{i}\,0^+\,\mathrm{sgn}(E_{n\mathbf{k}})} v(\mathbf{q}+\mathbf{G}').$$

Infinite $\sum^{\text{emp.}}$ in χ and $\Sigma^{\text{CH}}_{nn'}$: **A major bottleneck of GW in momentum space**

The problem of summing over many empty bands

One further simplification trick: `pseudobands`. For $E_{n\mathbf{k}} \gg E_\mathrm{F}$:

$$\{\phi_{n\mathbf{k}}\}_{\text{adjacent energy block } n_\mathrm{b}} \to \sum_{n \in \text{block } n_\mathrm{b}} \phi_{n\mathbf{k}},$$

$$\{E_{n\mathbf{k}}\}_{\text{adjacent energy block } n_\mathrm{b}} \to \frac{1}{|n_\mathrm{b}|} \sum_{n \in \text{block } n_\mathrm{b}} E_{n\mathbf{k}}.$$

4000 bands $\to \sim 400$ to ~ 1000 bands.

...but is this snake oil? It makes no sense!!!

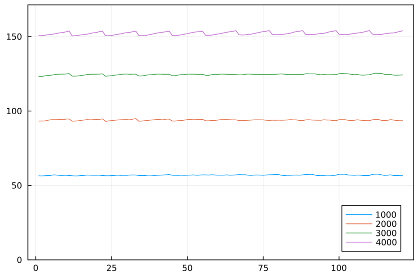
`pseudobands`, and when it works

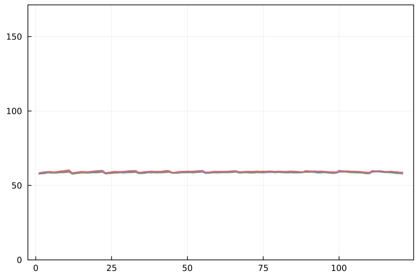
Understanding `pseudobands`: energy averaging

Flatness of high energy bands

High-energy DFT bands of WTe_2 monolayer (ONCV-SG15, 120 electrons; 80 Ry cutoff, $20 \times 20 \times 1$ grid). x axis $=$ $\mathbf{k}$ index in irreducible 1BZ; fastest varying coordinate $= k_y$

☛ $n \sim 1000$: bands are generally flat

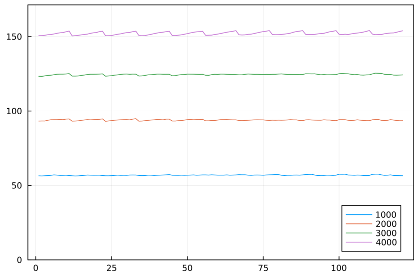


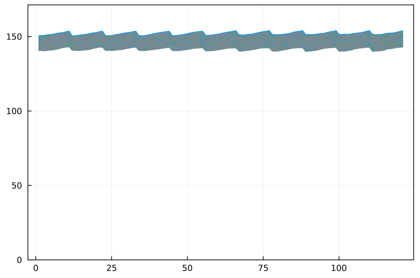


Flatness of high energy bands

High-energy DFT bands of WTe_2 monolayer (ONCV-SG15, 120 electrons; 80 Ry cutoff, $20 \times 20 \times 1$ grid). x axis $=$ $\mathbf{k}$ index in irreducible 1BZ; fastest varying coordinate $= k_y$

👉High-lying bands are more dispersive but since they are high, $1/(E_{\mathrm{emp}} - E_{\mathrm{occ}})$ is still flat



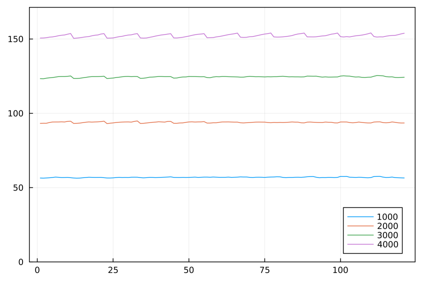


Understanding `pseudobands`: energy averaging

Flatness of high energy bands

- $\chi \sim MM^* \times f_1(E_{\mathrm{emp}} - E_{\mathrm{occ}})$;
- $\Sigma^{\mathrm{CH}} \sim MM^* \times f_2(\omega - E), \quad \omega \sim E_{\mathrm{F}}$.

For $E_{n\mathbf{k}} \gg E_{\mathrm{F}}$ (ω): $f_{1,2} \sim$ const. for all $\mathbf{k}$.



Thus in both χ and Σ^{CH}:

$$
\begin{aligned}
&\sum_{n''}^{\text{high emp.}} M^*_{n''n} M_{n''n'} \times f(E_{n''\mathbf{k}}) \\
= &\sum_{\text{block } n_{\mathrm{b}}}^{\text{high emp.}} f(\bar{E}_{n_{\mathrm{b}}}) \sum_{n''}^{\text{block } n_{\mathrm{b}}} M^*_{n''n} M_{n''n'}.
\end{aligned}
$$

Understanding `pseudobands`: wave function averaging?

For `pseudobands` to work, we need

$$
\begin{aligned}
\underbrace{\sum_{n'' \in \text{block } n_\text{b}} M^*_{n'' n} M_{n'' n'}}_{\text{normal}} &\sim M^*_{\text{averaged band},n} M_{\text{averaged band},n'} \\
&= \underbrace{\sum_{n''_1, n''_2 \in \text{block } n_\text{b}} M_{n''_1 n'} M^*_{n''_2 n}}_{\text{pseudobands}} .
\end{aligned} \tag{1}
$$

The question: is $[M_{n''_1 n'} M^*_{n''_2 n}]_{n''_1 n''_2}$ **diagonal?** Easy to verify: it's not.

Then in which case is (1) **correct in some sense?**

Technical issues in calculating $M_{nn'}(\mathbf{k}, \mathbf{q}, \mathbf{G})$

The structure of $\phi_{n\mathbf{k}}$

Plane wave basis In BerkeleyGW `WFN.h5`,

$$\phi_{n\mathbf{k}}(\mathbf{r}, \sigma) = \frac{1}{\sqrt{V}} \sum_{\mathbf{G}} \mathrm{e}^{\mathrm{i}(\mathbf{k}+\mathbf{G})\cdot\mathbf{r}} \, c_{n\mathbf{k},\mathbf{G}\sigma}.$$

Thus

$$M_{nn'}(\mathbf{k}, \mathbf{q}, \mathbf{G}) = \langle n\mathbf{k} + \mathbf{q} | \mathrm{e}^{\mathrm{i}(\mathbf{q}+\mathbf{G})\cdot\mathbf{r}} | n'\mathbf{k} \rangle = \sum_{\mathbf{G}', \sigma} c^*_{n\mathbf{k}+\mathbf{q},\mathbf{G}+\mathbf{G}'\sigma} c_{n\mathbf{k},\mathbf{G}'\sigma}.$$

Cutoff Each $\mathbf{k}$ has its own $\mathbf{G}$-grid (~ 30000 vectors for $80\,\mathrm{Ry}$).

Procedure

Input

- indices of $\mathbf{k}, \mathbf{q}$ in $\mathbf{k}$-grid;
- index of $\mathbf{G}$ in $\mathbf{G}$-grid of $\mathbf{k}$ (expect a $\mathbf{G}$ in GW $\mathbf{G}$-grid, cutoff $=$ say 30 Ry, not 80 Ry);
- n, n'.

Procedure

1. find index of $\mathbf{k}$
2. find index of $\mathbf{G} + \mathbf{G}'$ in $\mathbf{G}$-grid of $\mathbf{k} + \mathbf{q}$, for each $\mathbf{G}'$ in $\mathbf{G}$-grid of $\mathbf{k}$
3. do summation $\sum_{\mathbf{G}',\sigma} c^*_{n\mathbf{k}+\mathbf{q},\mathbf{G}+\mathbf{G}'\sigma} c_{n\mathbf{k},\mathbf{G}'\sigma}$.

Performance Main bottleneck: finding $\mathbf{G} + \mathbf{G}'$. Using `StaticArrays.jl` helps a lot!

`pseudobands` for χ

χ revisited

Under GPP:

$$\chi_{\mathbf{G}\mathbf{G}'}(\mathbf{q}, \omega = 0)^{\text{high band terms}} \approx \sum_{\mathbf{k}} \sum_{\text{block } n_\mathrm{b}}^{\text{emp}} \frac{2}{E_{n,\mathbf{k}+\mathbf{q}} - E_{\text{average in block } n_\mathrm{b}}} \times \sum_{n' \in \text{block } n_\mathrm{b}} \sum_{n}^{\text{occ}} M_{nn'}(\mathbf{k}, \mathbf{q}, \mathbf{G}) M^*_{nn'}(\mathbf{k}, \mathbf{q}, \mathbf{G}')$$

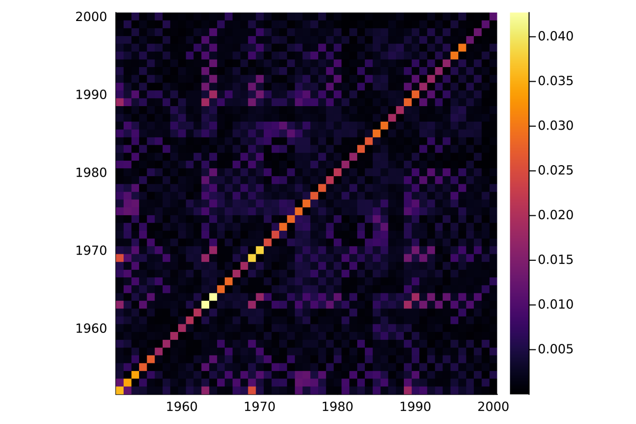
Our goal *Finding how diagonal is*

$$\sum_{n}^{\text{occ}} M_{nn'_1}(\mathbf{k}, \mathbf{q}, \mathbf{G}) M^*_{nn'_2}(\mathbf{k}, \mathbf{q}, \mathbf{G}')$$

It should be diagonal for `pseudobands` to work.

Numerical charaterization of $\sum_n^{\text{occ}} M_{nn'_1}(\mathbf{k},\mathbf{q},\mathbf{G}) M^*_{nn'_2}(\mathbf{k},\mathbf{q},\mathbf{G}')$

Case 1: $\mathbf{G} = \mathbf{G}'$ In this case $\sum_n^{\text{occ}} M_{nn'_1}(\mathbf{k},\mathbf{q},\mathbf{G}) M^*_{nn'_2}(\mathbf{k},\mathbf{q},\mathbf{G}')$ is large and fairly diagonal

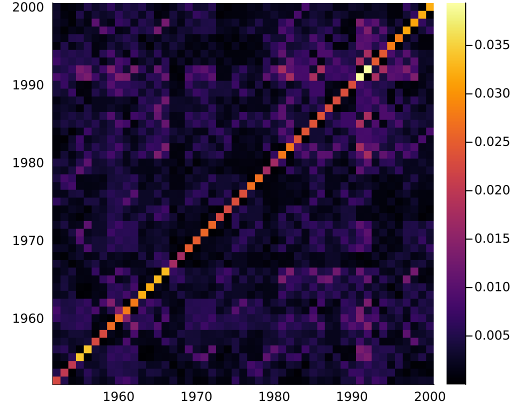


Numerical charaterization of $\sum_n^{\mathrm{occ}} M_{nn'_1}(\mathbf{k},\mathbf{q},\mathbf{G}) M^*_{nn'_2}(\mathbf{k},\mathbf{q},\mathbf{G}')$

Parameters $\mathbf{G} = (0, 1, -14), \mathbf{k} = \mathbf{k}_2 = (0, 0.00, 0), \mathbf{q} = (0, 0.00, 0)$

Numerical charaterization of $\sum_n^{\text{occ}} M_{nn_1'}(\mathbf{k},\mathbf{q},\mathbf{G}) M^*_{nn_2'}(\mathbf{k},\mathbf{q},\mathbf{G}')$

Case 1: $\mathbf{G} = \mathbf{G}'$ In this case $\sum_n^{\text{occ}} M_{nn_1'}(\mathbf{k},\mathbf{q},\mathbf{G}) M^*_{nn_2'}(\mathbf{k},\mathbf{q},\mathbf{G}')$ is large and fairly diagonal

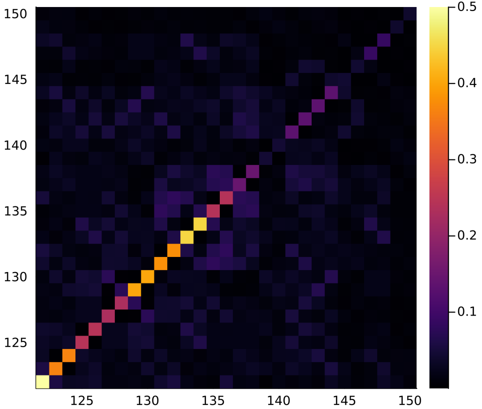


Numerical charaterization of $\sum_n^{\mathbf{occ}} M_{nn'_1}(\mathbf{k},\mathbf{q},\mathbf{G}) M^*_{nn'_2}(\mathbf{k},\mathbf{q},\mathbf{G}')$

Parameters $\mathbf{G} = (0, 1, -14), \mathbf{k} = \mathbf{k}_2 = (0, 0.05, 0), \mathbf{q} = (0, 0.10, 0)$

Numerical charaterization of $\sum_n^{\text{occ}} M_{nn_1'}(\mathbf{k},\mathbf{q},\mathbf{G}) M^*_{nn_2'}(\mathbf{k},\mathbf{q},\mathbf{G}')$

Case 1: $\mathbf{G} = \mathbf{G}'$ In this case $\sum_n^{\text{occ}} M_{nn_1'}(\mathbf{k},\mathbf{q},\mathbf{G}) M^*_{nn_2'}(\mathbf{k},\mathbf{q},\mathbf{G}')$ is large and fairly diagonal



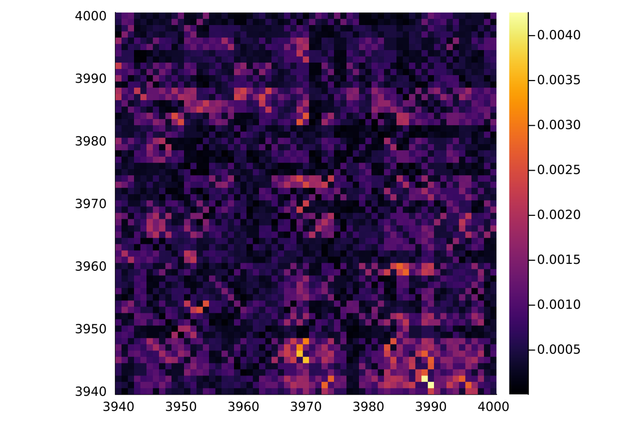
Numerical charaterization of $\sum_n^{\mathbf{occ}} M_{nn'_1}(\mathbf{k},\mathbf{q},\mathbf{G}) M^*_{nn'_2}(\mathbf{k},\mathbf{q},\mathbf{G}')$

Strikingly, $\sum_n^{\mathrm{occ}} M_{nn'_1}(\mathbf{k},\mathbf{q},\mathbf{G}) M^*_{nn'_2}(\mathbf{k},\mathbf{q},\mathbf{G}')$ is still very diagonal for bands near Fermi surface!

Numerical charaterization of $\sum_n^{\mathrm{occ}} M_{nn'_1}(\mathbf{k},\mathbf{q},\mathbf{G}) M^*_{nn'_2}(\mathbf{k},\mathbf{q},\mathbf{G}')$

Case 2: $\mathbf{G} \neq \mathbf{G}'$ $\sum_n^{\mathrm{occ}} M_{nn'_1}(\mathbf{k},\mathbf{q},\mathbf{G}) M^*_{nn'_2}(\mathbf{k},\mathbf{q},\mathbf{G}')$ is very non-diagonal, but since the terms's random phases cancel each other so the overall sum after $\sum_{n'}^{\mathrm{emp}}$ is small

Numerical charaterization of $\sum_n^{\mathrm{occ}} M_{nn_1'}(\mathbf{k},\mathbf{q},\mathbf{G}) M^*_{nn_2'}(\mathbf{k},\mathbf{q},\mathbf{G}')$



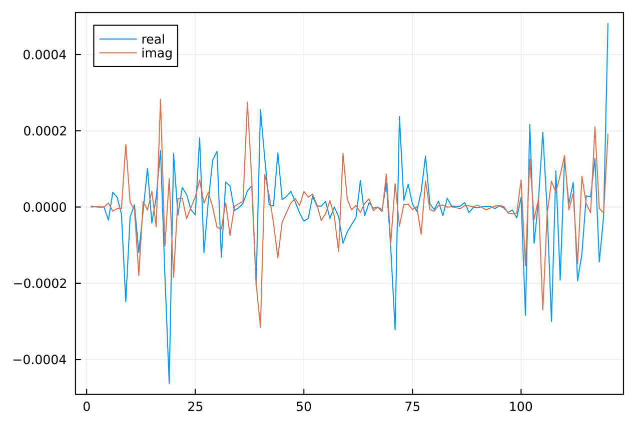
Half-way generalization about `pseudobands` in χ

- `pseudobands` works when it's necessary to do so
- What prevents `pseudobands` from working around Fermi surface is the energy dispersion

. . . but do we have any theoretical explanation for this?

Cancellation of cross-terms in χ

$M_{nn'_1}(\mathbf{k},\mathbf{q},\mathbf{G}) M^*_{nn'_2}(\mathbf{k},\mathbf{q},\mathbf{G}')$ and n (occupied band index), when $n'_1 \neq n'_2$:



Cancellation of cross-terms in χ

☛Although for a single n, the value can be large, as we sum over n the terms cancel each other $\Rightarrow$ diagonal $\sum_n^{\text{occ}} M_{nn'_1}(\mathbf{k},\mathbf{q},\mathbf{G}) M^*_{nn'_2}(\mathbf{k},\mathbf{q},\mathbf{G}')$

Why is this the case?

Case 1: $|n\mathbf{k}\rangle \propto \mathrm{e}^{\mathrm{i}(\mathbf{G}_n+\mathbf{k})\cdot\mathbf{r}}$ For high-energy bands primarily containing only one **G** components, unwanted terms are:

$$\sum_n^{\text{occ}} M_{nn_1}(\mathbf{k},\mathbf{q},\mathbf{G}) M^*_{nn_2}(\mathbf{k},\mathbf{q},\mathbf{G}) = \sum_n^{\text{occ}} c_{n\mathbf{k}+\mathbf{q}}(\mathbf{G}+\mathbf{G}_1)^* c_{n\mathbf{k}+\mathbf{q}}(\mathbf{G}+\mathbf{G}_2)$$

The terms rapid fast when $\mathbf{G}_1 \neq \mathbf{G}_2$:

Cancellation of cross-terms in χ

Case 2: $|n\mathbf{k}\rangle$ contains several $\mathbf{G}$ components suppose high-energy bands $n_{1,2}$ primarily contain two $\mathbf{G}$ components:

$$|n_i\mathbf{k}\rangle = (c_{i1}\,|\mathbf{G}_1\rangle + c_{i2}\,|\mathbf{G}_2\rangle)\,\mathrm{e}^{\mathrm{i}\,\mathbf{k}\cdot\mathbf{r}}, \quad i = 1, 2. \tag{2}$$

In this way

$$\begin{aligned}
&\sum_n^{\text{occ}} M_{nn_1}(\mathbf{k},\mathbf{q},\mathbf{G}) M^*_{nn_2}(\mathbf{k},\mathbf{q},\mathbf{G}) \\
={}& \sum_n^{\text{occ}} (c_{11}\, c^*_{n\mathbf{k}+\mathbf{q}}(\mathbf{G}+\mathbf{G}_1) + c_{12}\, c^*_{n\mathbf{k}+\mathbf{q}}(\mathbf{G}+\mathbf{G}_2)) \times \\
& (c^*_{21}\, c_{n\mathbf{k}+\mathbf{q}}(\mathbf{G}+\mathbf{G}_1) + c^*_{22}\, c_{n\mathbf{k}+\mathbf{q}}(\mathbf{G}+\mathbf{G}_2)).
\end{aligned}$$

Cancellation of cross-terms in χ

The following terms vanish:

$$\sum_{n}^{\text{occ}} (c_{11} c_{21}^* |c_{n\mathbf{k}+\mathbf{q}}(\mathbf{G} + \mathbf{G}_1)|^2 + c_{12} c_{22}^* |c_{n\mathbf{k}+\mathbf{q}}(\mathbf{G} + \mathbf{G}_2)|^2)$$

1. After summation over n, $\sum_n |c_{n\mathbf{k}+\mathbf{q}}(\mathbf{G} + \mathbf{G}_1)|^2 = \sum_n |c_{n\mathbf{k}+\mathbf{q}}(\mathbf{G} + \mathbf{G}_2)|^2$
2. Orthogonal relation: $c_{11} c_{21}^* + c_{12} c_{22}^* = 0$.

The other terms

$$\sum_{n}^{\text{occ}} (c_{11} c_{22}^* c_{n\mathbf{k}+\mathbf{q}}^*(\mathbf{G} + \mathbf{G}_1) c_{n\mathbf{k}+\mathbf{q}}(\mathbf{G} + \mathbf{G}_2) + c_{12} c_{21}^* c_{n\mathbf{k}+\mathbf{q}}^*(\mathbf{G} + \mathbf{G}_2) c_{n\mathbf{k}+\mathbf{q}}(\mathbf{G} + \mathbf{G}_1))$$

oscillate in the same way in the first case.

`pseudobands` for Σ

Only Σ^{CH} involves $\sum^{\text{emp}}$:

$$\langle n\mathbf{k}|\Sigma^{\text{CH, GPP}}|n'\mathbf{k}\rangle = \frac{1}{2}\sum_{n''}\sum_{\mathbf{q}\mathbf{G}\mathbf{G}'} M^*_{n''n}(\mathbf{k}, -\mathbf{q}, -\mathbf{G}) M_{n''n'}(\mathbf{k}, -\mathbf{q}, -\mathbf{G}') \times$$
$$\text{something about } \mathbf{q}, \mathbf{k}, \mathbf{G}, \mathbf{G}'$$

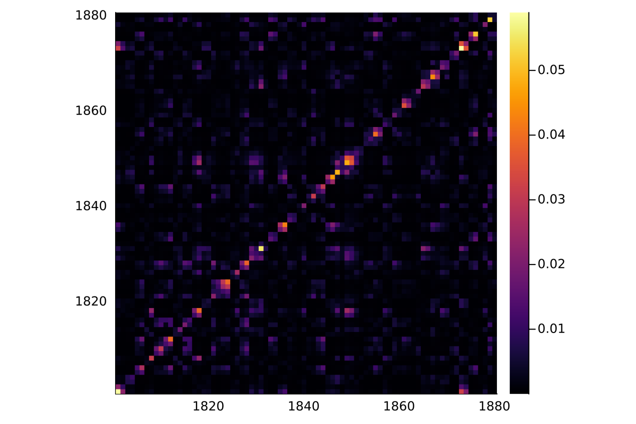
Problems

- No summation over occupied states $\Rightarrow$ ✗the aforementioned cancellation mechanism
- Summation over $\mathbf{G}$ very complicated $\Rightarrow$ analysis based on $\mathbf{G}$ very hard.

Some tentative ideas

- It has been verified that a single $M^*_{n_1'' n} M_{n_2'' n}$ term is never diagonal
- But it seems naively summing over $\mathbf{G} = \mathbf{G}'$ (without considering the weight factor) makes it diagonal...

Some tentative ideas



- $M^*_{n_1'' n}(\mathbf{G}_1) M_{n_2'' n}(\mathbf{G}_2)$ is not diagonal at all when $\mathbf{G}_1 \neq \mathbf{G}_2$;

Some tentative ideas

- but probably only the $\mathbf{G}_1 = \mathbf{G}_2$ terms matter?

Comments from Felipe H. Jornada's recent paper

“Stochastic” pseudobands

Energy averaging is also done in their paper:

$$G = \sum_{n \text{ near F.S.}} \frac{|\phi_{n\mathbf{k}}\rangle\langle\phi_{n\mathbf{k}}|}{\omega - E_{n\mathbf{k}}} + \sum_{\text{pseudobands block}} \frac{1}{\omega - \bar{E}} \sum_{n \in \text{block}} |\phi_{n\mathbf{k}}\rangle\langle\phi_{n\mathbf{k}}| \,.$$

Their procedure Replace each pseudobands block by *several bands*:

$$\{\phi_{n\mathbf{k}}\}_{n \text{ in block}} \longrightarrow \left\{ |\phi_{\xi\mathbf{k}}\rangle = \sum_{n \text{ in block}} \mathrm{e}^{\mathrm{i}\,\theta_{n\xi}} \, |\phi_{n\mathbf{k}}\rangle \right\}_{\xi}$$

Justification When $N_\xi \to \infty$, $\left\langle \mathrm{e}^{-\mathrm{i}\,\theta_{n'\xi}} \, \mathrm{e}^{\mathrm{i}\,\theta_{n\xi}} \right\rangle_{\xi} = \delta_{nn'}$

Comparison with the current script When $N_\xi = 1$, $\mathrm{e}^{\mathrm{i}\,\theta_{n\xi}}$ comes from DFT diagonalization

FHJ's justification of pseudobands

Convergence conditions When $N_{\text{pseudobands blocks}} \to \infty$ and in each block $N_\xi \to \infty$,

- the expectations of χ and Σ goes to the true value;
- the standard error goes to zero.

Is this result enough as a justification? My answer: no; because the limits = not doing pseudobands at all. (And their recommended values are not large enough anyway.)

FHJ's justification of pseudobands

Convergence conditions When $N_{\text{pseudobands blocks}} \to \infty$ and in each block $N_\xi \to \infty$,

- the expectations of χ and Σ goes to the true value;
- the standard error goes to zero.

Is this result enough as a justification? My answer: no; because the limits = not doing pseudobands at all. (And their recommended values are not large enough anyway.)

Additional justification "the polarizability tends to converge much faster than G, partially due to the *rapidly oscillating nature* of the matrix elements involving Kohn-Sham states used in the evaluation of the polarizability"

Compressing the bands even harder: size of blocks

Recall that band dispersion is the main problem...

$$\Delta\chi \sim \frac{\Delta E}{E^2}, \quad \chi \sim \frac{1}{E} \Rightarrow \frac{\Delta\chi}{\chi} \simeq \frac{\Delta E}{E}$$

$$\Rightarrow \frac{\Delta\chi}{\chi} \lesssim \text{const} \Leftrightarrow \boxed{\frac{\Delta E}{E} \lesssim \text{const}}$$

Exponential growth of the energy spread of blocks!

Compressing the bands even harder: pseudobands near E_F?

They claim that for `epsilon`, bands near Fermi surface can also be pseudo-ized

It somehow goes against my own numerical experiments; but they are quite sure about that hmm...

They also claim that for `sigma`, only the bands near Fermi surface shouldn't be pseudo-ized.

Important claim: protection window is not a convergence parameter.

Discussion

Discussion

Implications to GW acceleration There are lots of garbage in the giant input files we feed to GW

Implications to machine learning There exists a *analytic* relation between Σ^{GW} and $\{\bar{E}, |\phi_{\mathbf{k}}\rangle^{\text{pseudo}}\}_{\text{blocks}}$

- Starting from pseudobands is *not* feature engineering
- Pseudobands (with further compression of autoencoders) should be the starting point of all ML tasks