Details in *GW*-BSE

Jinyuan Wu

October 29, 2023

Table of contents

1. *GW*: a quick course
2. The mystery of `pseudobands`
3. Technical issues in calculating $M_{nn'}(\boldsymbol{k}, \boldsymbol{q}, \boldsymbol{G})$
4. `pseudobands` for χ

To be very concise... $-\mathrm{i}\,\Sigma \approx GW^{\mathrm{RPA}} =$.

Assuming quasiparticles... $v(\boldsymbol{q}+\boldsymbol{G})=\frac{4\pi}{V|\boldsymbol{q}+\boldsymbol{G}|^2}$. $\Sigma \overset{\int_{\mathrm{contour}}}{=} \Sigma^{\mathrm{COH}}+\Sigma^{\mathrm{SEX}}$,
$\epsilon_{\boldsymbol{G}\boldsymbol{G}'}(\boldsymbol{q},\omega)=\delta_{\boldsymbol{G}\boldsymbol{G}'}-v(q+\boldsymbol{G})\chi_{\boldsymbol{G}\boldsymbol{G}'}(\boldsymbol{q},\omega)$, $M_{nn'}(\boldsymbol{k},\boldsymbol{q},\boldsymbol{G})=\langle n\boldsymbol{k}+\boldsymbol{q}|\mathrm{e}^{\mathrm{i}(\boldsymbol{q}+\boldsymbol{G})\cdot\boldsymbol{r}}|n'\boldsymbol{k}\rangle$

$$\begin{aligned}\left\langle n\boldsymbol{k}\middle|\Sigma^{\mathrm{SEX}}(\omega)\middle|n'\boldsymbol{k}\right\rangle = -\sum_{n''}^{\mathrm{occ}}\sum_{\boldsymbol{q}\boldsymbol{G}\boldsymbol{G}'} & M^*_{n''n}(\boldsymbol{k},-\boldsymbol{q},-\boldsymbol{G})M_{n''n'}(\boldsymbol{k},-\boldsymbol{q},-\boldsymbol{G}') \\ & \times \epsilon^{-1}_{\boldsymbol{G}\boldsymbol{G}'}(\boldsymbol{q},\omega-E_{n'',\boldsymbol{k}-\boldsymbol{q}})v(\boldsymbol{q}+\boldsymbol{G}'),\end{aligned} \tag{1}$$

$$\begin{aligned}\left\langle n\boldsymbol{k}\middle|\Sigma^{\mathrm{COH}}(\omega)\middle|n'\boldsymbol{k}\right\rangle = \frac{\mathrm{i}}{2\pi}\sum_{n''}\sum_{\boldsymbol{q},\boldsymbol{G},\boldsymbol{G}'} & M^*_{n''n}(\boldsymbol{k},-\boldsymbol{q},-\boldsymbol{G})M_{n''n'}(\boldsymbol{k},-\boldsymbol{q},-\boldsymbol{G}') \\ & \times\int_0^\infty \mathrm{d}\omega'\,\frac{[\epsilon^{\mathrm{r}}_{\boldsymbol{G}\boldsymbol{G}'}]^{-1}(\boldsymbol{q},\omega')-[\epsilon^{\mathrm{a}}_{\boldsymbol{G}\boldsymbol{G}'}]^{-1}(\boldsymbol{q},\omega')}{\omega-E_{n\boldsymbol{k}}-\omega'+\mathrm{i}\,0^+\,\mathrm{sgn}(E_{n\boldsymbol{k}})}v(\boldsymbol{q}+\boldsymbol{G}').\end{aligned} \tag{2}$$

$$\begin{aligned}\chi^{\mathrm{r/a}}_{\boldsymbol{G}\boldsymbol{G}'}(\boldsymbol{q},\omega) = \sum_{\boldsymbol{k}}\sum_{n}^{\mathrm{occ}}\sum_{n'}^{\mathrm{emp}} & M_{nn'}(\boldsymbol{k},\boldsymbol{q},\boldsymbol{G})M^*_{nn'}(\boldsymbol{k},\boldsymbol{q},\boldsymbol{G}') \\ & \times\left(\frac{1}{\omega+E_{n,\boldsymbol{k}+\boldsymbol{q}}-E_{n'\boldsymbol{k}}\pm\mathrm{i}\,0^+}+\frac{1}{-\omega+E_{n,\boldsymbol{k}+\boldsymbol{q}}-E_{n'\boldsymbol{k}}\mp\mathrm{i}\,0^+}\right).\end{aligned} \tag{3}$$

What's *GW*

What a splendid equation system!
In practice. . . we always use GPP $\epsilon(\omega)$ is assumed to be plasmon model-like, so – we feed

$$\chi_{\boldsymbol{G}\boldsymbol{G}'}(\boldsymbol{q},\omega) = \sum_{\boldsymbol{k}}^{\text{occ}} \sum_{n}^{\text{emp}} \sum_{n'} M_{nn'}(\boldsymbol{k},\boldsymbol{q},\boldsymbol{G}) M^{*}_{nn'}(\boldsymbol{k},\boldsymbol{q},\boldsymbol{G}') \frac{2}{E_{n,\boldsymbol{k}+\boldsymbol{q}} - E_{n'\boldsymbol{k}}} \quad (4)$$

to *analytic* expressions of $\Sigma^{\text{COH, SEX}}$.

Huge simplification. . . Hedin $\overset{\text{assuming } GW}{\rightarrow}$ *GW* $\overset{\text{QP. approx.}}{\rightarrow}$ (1), (2), (3) $\overset{\text{GPP}}{\rightarrow}$ we are here
But still burdensome *Summation over empty bands – 1000-30000 bands!!!*

The problem of summing over many empty bands

Empty states are important – but why?

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

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

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

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

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

Understanding pseudobands: energy averaging

Observation

- $\chi \sim MM^* \times$ some-function$(E_{\text{emp}} - E_{\text{occ}})$;
- $\Sigma^{\text{COH}} \sim MM^* \times$ some-function$(\omega - E)$;
- We are only interested in $\omega \sim E_{\text{F}}$.

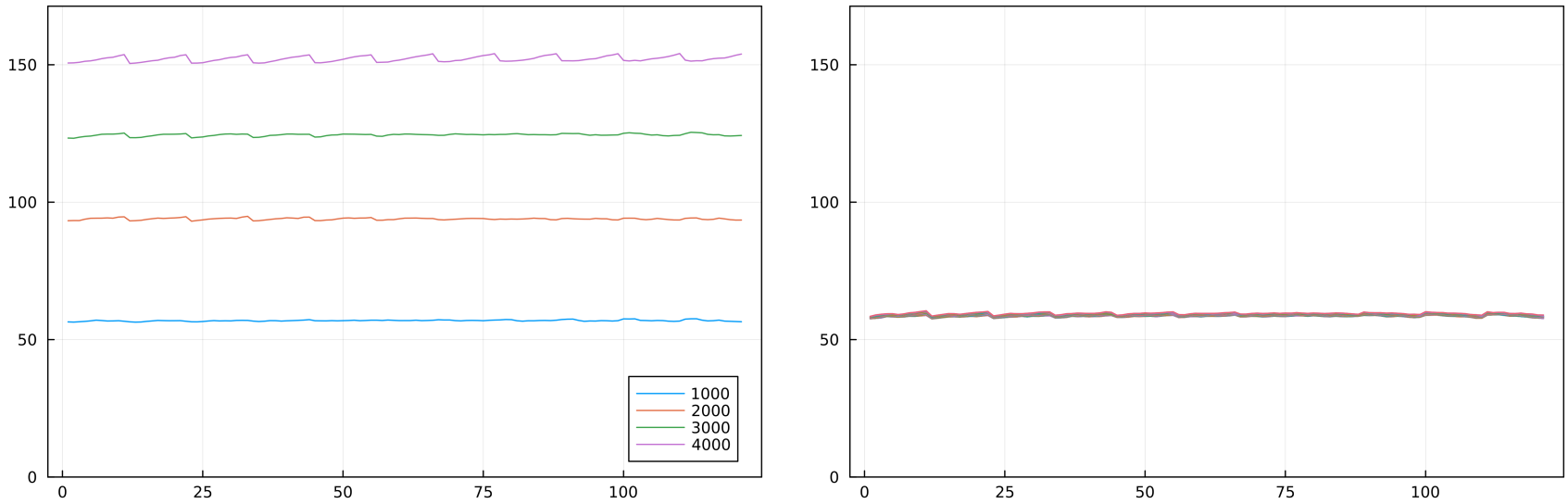
👉For $E_{n\boldsymbol{k}} \gg E_{\text{F}}$ (ω): energy-dependent factors $\sim$ const. for all $\boldsymbol{k}$.

Thus for both χ and Σ^{COH} involving summation over empty bands:

$$\sum_{n''}^{\text{high emp. bands}} M^*_{n''n} M_{n''n'} \times \cdots = \sum_{\text{block } n_{\text{b}}} \cdots \times \sum_{n'' \in \text{block } n_{\text{b}}} M^*_{n''n} M_{n''n'}. \quad (5)$$

In RHS E_{emp} is replaced by the average energy.

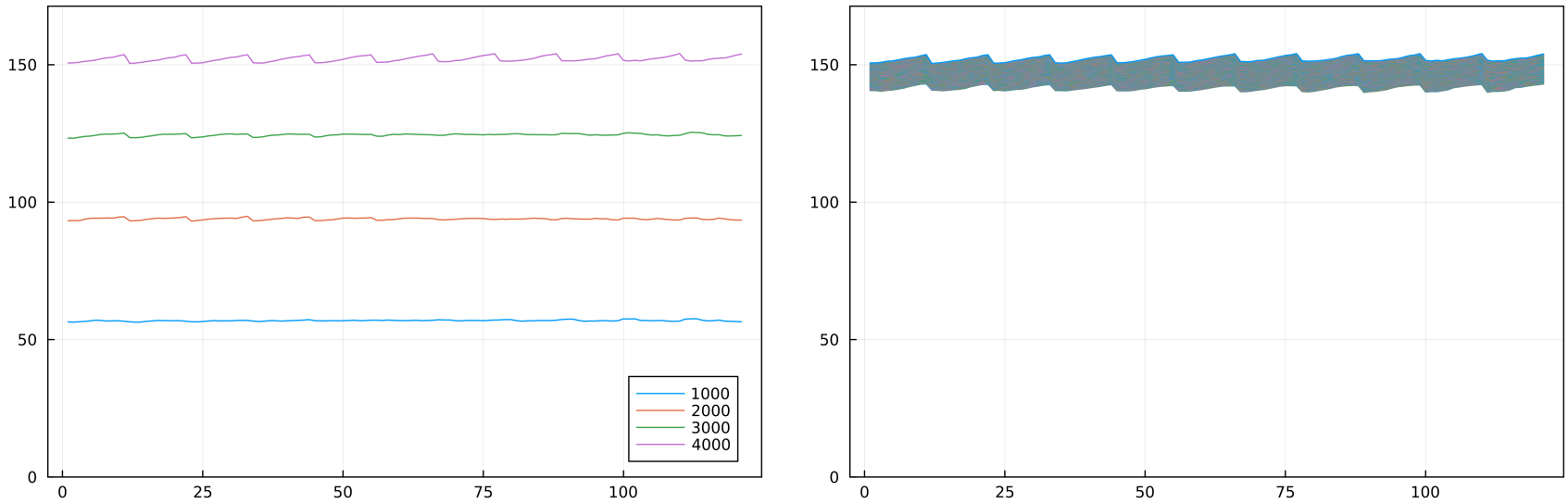
Understanding pseudobands: energy averaging



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

- Bands above 1000 are generally flat (but dispersion $\uparrow$ as $n \uparrow$)
- Low-lying pseudobands blocks are very close in energy (but $1/(E_{\text{emp}} - E_{\text{occ}})$ more sensitive to dispersion)

Understanding pseudobands: energy averaging



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

- Bands above 1000 are generally flat (but dispersion $\uparrow$ as $n \uparrow$)
- High-energy blocks are more dispersive (but $1/(E_{\text{emp}} - E_{\text{occ}})$ is smaller so no worry)

Understanding pseudobands: wave function averaging?

For pseudobands to work, we need

$$\sum_{n'' \in \text{block } n_\text{b}} M^*_{n''n} M_{n''n'} \sim M^*_{\text{averaged band},n} M_{\text{averaged band},n'} = \sum_{n_1'', n_2'' \in \text{block } n_\text{b}} M_{n_1''n'} M^*_{n_2''n}. \tag{6}$$

The problem: it of course isn't the case in general. If $[M_{n_1''n'} M^*_{n_2''n}]_{n_1''n_2''}$ is a random matrix: LHS : RHS ≈ 0.3.

The question: *Then in which case is* (6) *correct in some sense?*

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

Plane wave basis In BerkeleyGW `WFN.h5`,

$$\phi_{n\boldsymbol{k}}(\boldsymbol{r},\sigma) = \frac{1}{\sqrt{V}} \sum_{\boldsymbol{G}} \mathrm{e}^{\mathrm{i}(\boldsymbol{k}+\boldsymbol{G})\cdot\boldsymbol{r}}\, c_{n\boldsymbol{k},\boldsymbol{G}\sigma}. \tag{7}$$

Thus

$$M_{nn'}(\boldsymbol{k},\boldsymbol{q},\boldsymbol{G}) = \langle n\boldsymbol{k}+\boldsymbol{q}|\mathrm{e}^{\mathrm{i}(\boldsymbol{q}+\boldsymbol{G})\cdot\boldsymbol{r}}|n'\boldsymbol{k}\rangle = \sum_{\boldsymbol{G}',\sigma} c^*_{n\boldsymbol{k}+\boldsymbol{q},\boldsymbol{G}+\boldsymbol{G}'\sigma} c_{n\boldsymbol{k},\boldsymbol{G}'\sigma}. \tag{8}$$

Cutoff Each $\boldsymbol{k}$ has its own $\boldsymbol{G}$-grid ($\sim$ 30000 vectors for 80 Ry).

Procedure

Input

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

Procedure

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

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

χ revisited

Under GPP:

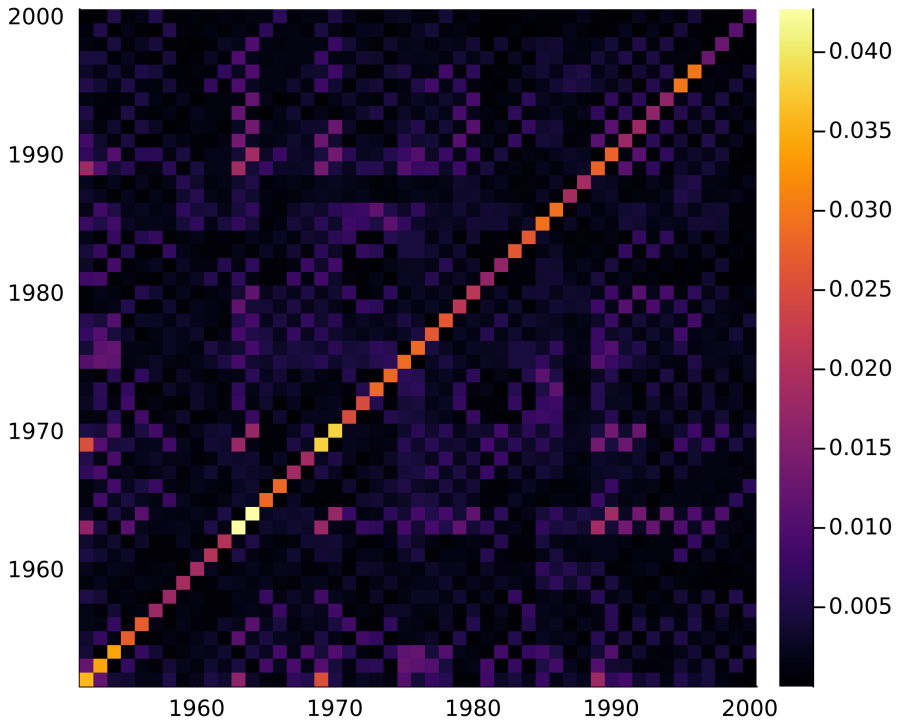
$$\chi_{\boldsymbol{G}\boldsymbol{G}'}(\boldsymbol{q},\omega)^{\text{high band terms}} \approx \sum_{\boldsymbol{k}} \sum_{\text{block } n_\text{b}}^{\text{emp}} \frac{2}{E_{n,\boldsymbol{k}+\boldsymbol{q}} - E_{\text{average in block } n_\text{b}}} \times \sum_{n' \in \text{block } n_\text{b}} \sum_{n}^{\text{occ}} M_{nn'}(\boldsymbol{k},\boldsymbol{q},\boldsymbol{G}) M^*_{nn'}(\boldsymbol{k},\boldsymbol{q},\boldsymbol{G}') \tag{9}$$

Our goal *Finding how diagonal is*

$$\sum_{n}^{\text{occ}} M_{nn'}(\boldsymbol{k},\boldsymbol{q},\boldsymbol{G}) M^*_{nn'}(\boldsymbol{k},\boldsymbol{q},\boldsymbol{G}') \tag{10}$$

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

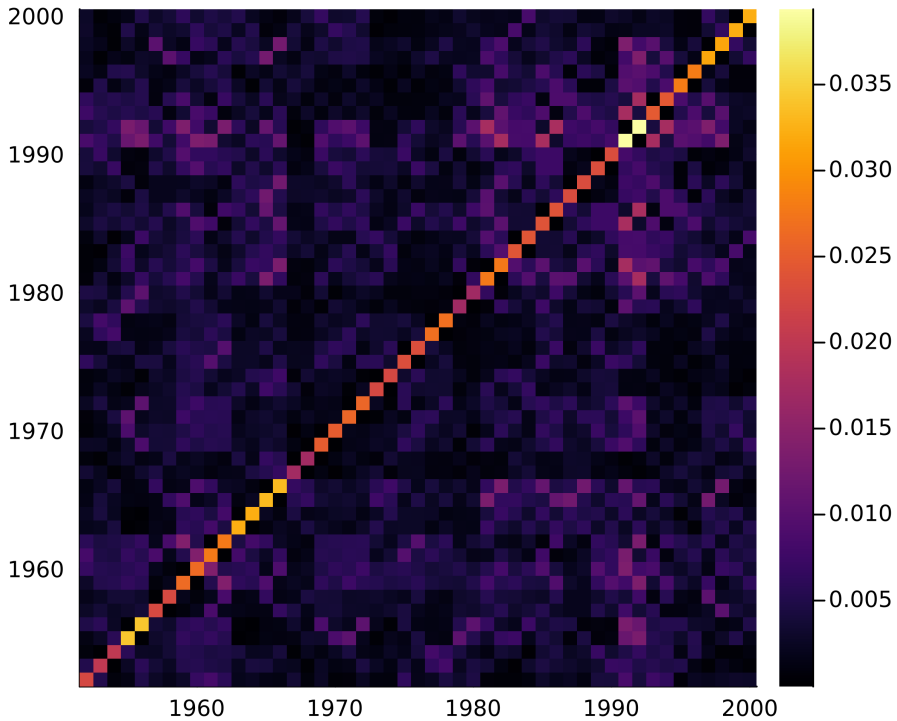
Case 1: $\boldsymbol{G} = \boldsymbol{G}'$ In this case $\sum_n^{\text{occ}} M_{nn'}(\boldsymbol{k},\boldsymbol{q},\boldsymbol{G}) M^*_{nn'}(\boldsymbol{k},\boldsymbol{q},\boldsymbol{G}')$ is large and fairly diagonal



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

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

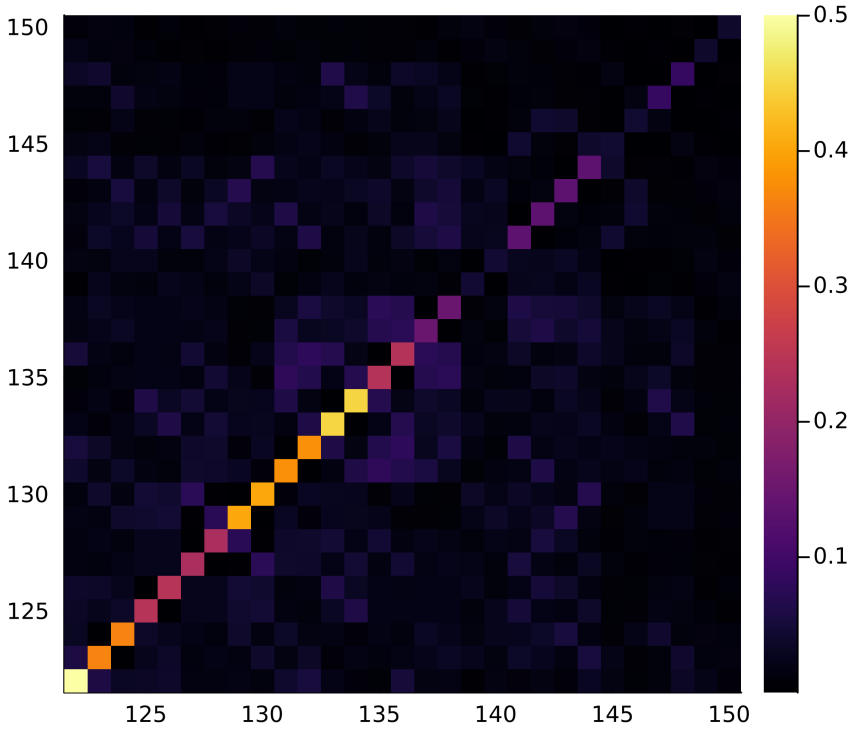
Case 1: $\boldsymbol{G} = \boldsymbol{G}'$ In this case $\sum_n^{\text{occ}} M_{nn'}(\boldsymbol{k},\boldsymbol{q},\boldsymbol{G})M^*_{nn'}(\boldsymbol{k},\boldsymbol{q},\boldsymbol{G}')$ is large and fairly diagonal



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

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

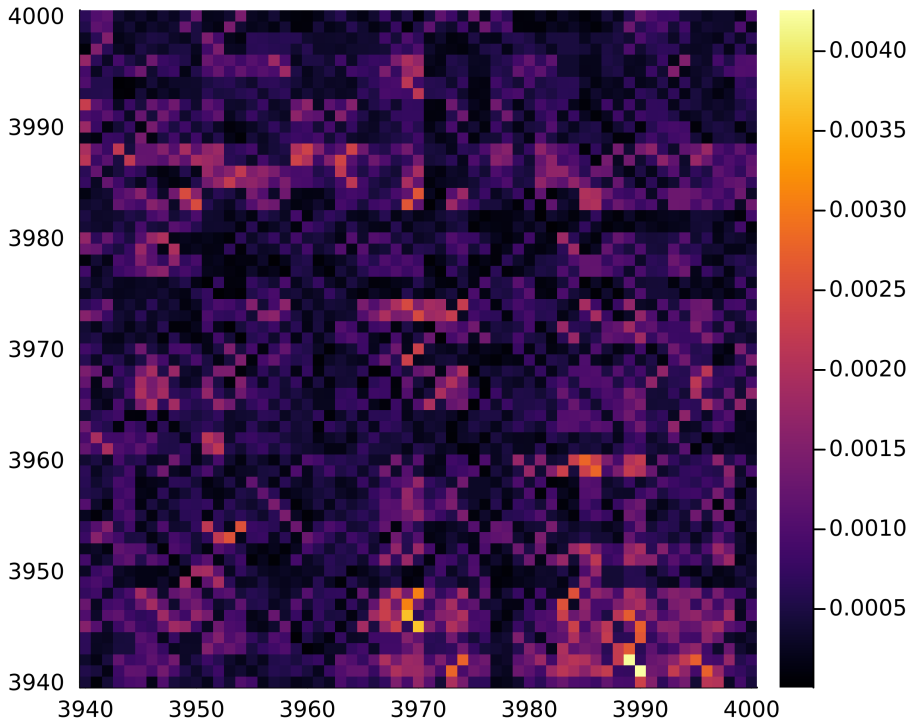
Case 1: $\boldsymbol{G} = \boldsymbol{G}'$ In this case $\sum_n^{\text{occ}} M_{nn'}(\boldsymbol{k},\boldsymbol{q},\boldsymbol{G})M^*_{nn'}(\boldsymbol{k},\boldsymbol{q},\boldsymbol{G}')$ is large and fairly diagonal



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

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

Case 2: $\boldsymbol{G} \neq \boldsymbol{G}'$ $\sum_n^{\text{occ}} M_{nn'}(\boldsymbol{k},\boldsymbol{q},\boldsymbol{G})M^*_{nn'}(\boldsymbol{k},\boldsymbol{q},\boldsymbol{G}')$ is very non-diagonal, but since the terms's random phases cancel each other so it's small



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?