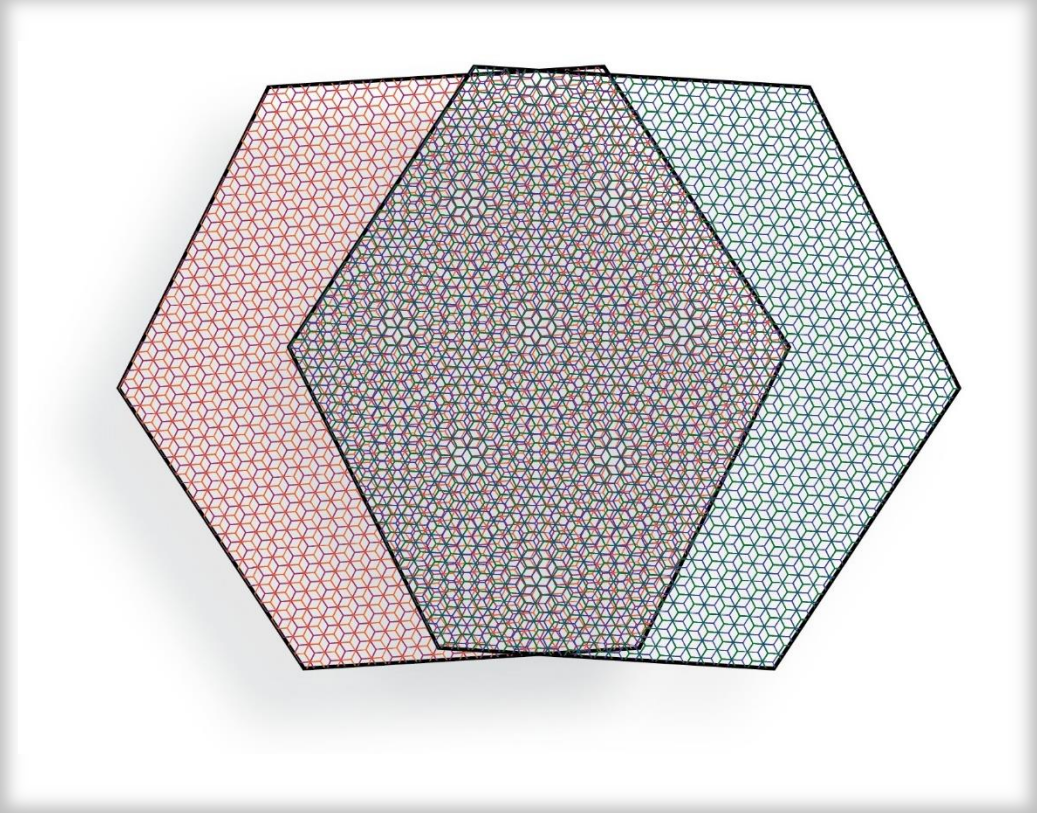
Low-scaling *GW* algorithms: overview & applications using CP2K



University of Regensburg

Jan Wilhelm

Institute of Theoretical Physics
Regensburg Center for Ultrafast Nanoscopy (RUN)

1. **Standard *GW* workflow [$O(N^4)$]**

2. ***GW* space-time method [$O(N^3)$]**
   Rojas, Godby, Needs, Phys. Rev. Lett. **74**, 1827 (1995)

3. **Why do we need large-scale *GW* calculations?**

4. **Low-scaling *GW* algorithm for 2D materials [$O(N^3)$]**
   Graml, Zollner, Hernangomez, Faria, Wilhelm, arXiv:2306.16066 (2023)

1. **Standard *GW* workflow [$O(N^4)$]**

2. ***GW* space-time method [$O(N^3)$]**
   Rojas, Godby, Needs, Phys. Rev. Lett. **74**, 1827 (1995)

3. **Why do we need large-scale *GW* calculations?**

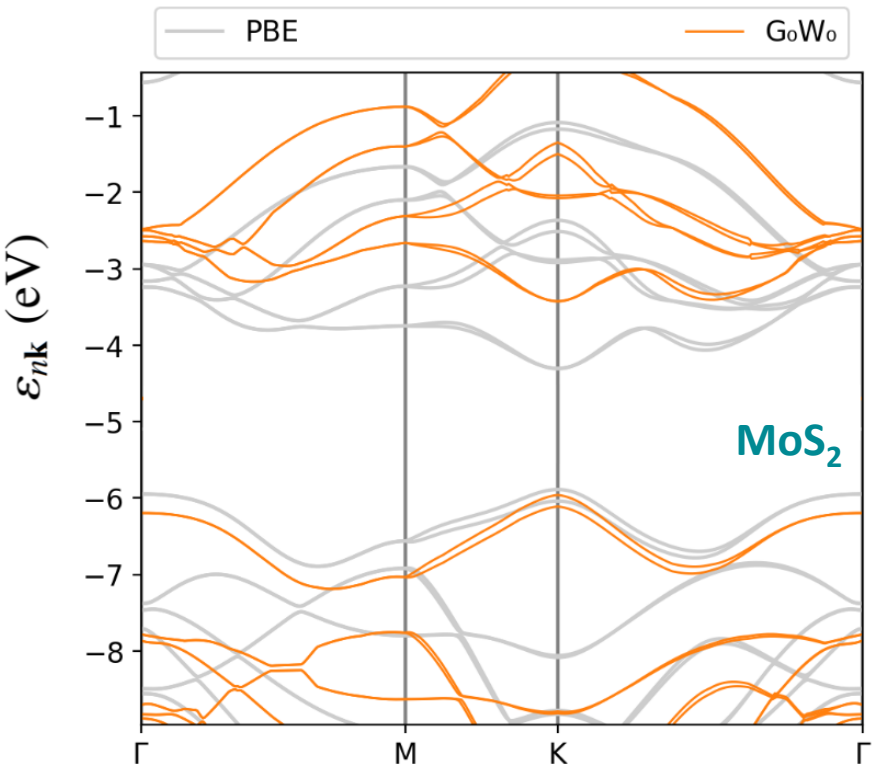
4. **Low-scaling *GW* algorithm for 2D materials**
   Graml, Zollner, Hernangomez, Faria, Wilhelm, arXiv:2306.16066 (2023)

GW band structure

$$\left[-\frac{\hbar^2}{2m}\nabla^2 + v_{\text{ext}}(\mathbf{r}) + v_{\text{H}}[\rho](\mathbf{r}) + v_{\text{xc}}[\rho](\mathbf{r})\right]\psi_{n\mathbf{k}}(\mathbf{r}) = \varepsilon_{n\mathbf{k}}^{\text{DFT}}\psi_{n\mathbf{k}}(\mathbf{r})$$

G_0W_0@DFT band structure:

$$\varepsilon_{n\mathbf{k}}^{G_0W_0} = \varepsilon_{n\mathbf{k}}^{\text{DFT}} + \Sigma_{n\mathbf{k}}^{G_0W_0}(\varepsilon_{n\mathbf{k}}^{G_0W_0}) - v_{n\mathbf{k}}^{\text{xc}}$$

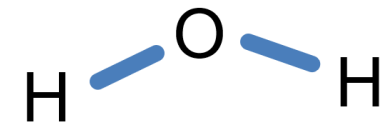


Standard *GW* workflow

$$\chi(\mathbf{r},\mathbf{r}',i\omega) = \sum_{i}^{\text{occ}} \sum_{a}^{\text{unocc}} \psi_a(\mathbf{r}')\psi_i(\mathbf{r}')\psi_i(\mathbf{r})\psi_a(\mathbf{r}) \frac{2(\varepsilon_i - \varepsilon_a)}{\omega^2 + (\varepsilon_i - \varepsilon_a)^2}$$

$$N^{\chi}_{\text{FLOP}} = N_\omega N_{\text{occ}} N_{\text{unocc}} N^2_{\mathbf{r}}$$

$O(N^4)$



$$\epsilon(\mathbf{r},\mathbf{r}',i\omega) = \delta(\mathbf{r},\mathbf{r}') - \int d\mathbf{r}'' v(\mathbf{r},\mathbf{r}'')\chi(\mathbf{r}'',\mathbf{r}',i\omega)$$

$$N^{\epsilon}_{\text{FLOP}} = N_\omega N^3_{\mathbf{r}}$$

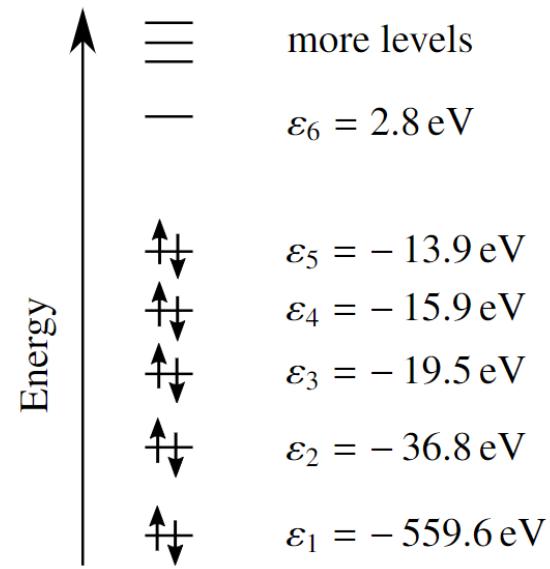
$O(N^3)$

$$W(\mathbf{r},\mathbf{r}',i\omega) = \int d\mathbf{r}'' \epsilon^{-1}(\mathbf{r},\mathbf{r}'',i\omega) v(\mathbf{r}'',\mathbf{r}')$$

$$G(\mathbf{r},\mathbf{r}',i\omega) = \sum_{m}^{\text{all}} \frac{\psi_m(\mathbf{r}')\psi_m(\mathbf{r})}{i\omega + \varepsilon_{\text{F}} - \varepsilon_m}$$

$$\Sigma(\mathbf{r},\mathbf{r}',i\omega) = -\int_{-\infty}^{\infty} \frac{d\omega'}{2\pi} G(\mathbf{r},\mathbf{r}',i\omega - i\omega') W(\mathbf{r},\mathbf{r}',i\omega')$$

$$\varepsilon^{G_0W_0}_{n\mathbf{k}} = \varepsilon^{\text{DFT}}_{n\mathbf{k}} + \Sigma^{G_0W_0}_{n\mathbf{k}}(\varepsilon^{G_0W_0}_{n\mathbf{k}}) - v^{\text{xc}}_{n\mathbf{k}}$$



$N_{\text{occ}} = 5$ $\quad N_{\text{unocc}} = 30$

$N_{\mathbf{r}} = 1000$

Standard *GW* workflow

$$\chi(\mathbf{r},\mathbf{r}',i\omega) = \sum_i^{\text{occ}} \sum_a^{\text{unocc}} \psi_a(\mathbf{r}')\psi_i(\mathbf{r}')\psi_i(\mathbf{r})\psi_a(\mathbf{r}) \frac{2(\varepsilon_i - \varepsilon_a)}{\omega^2 + (\varepsilon_i - \varepsilon_a)^2}$$

$$O(N^4)$$

$$N^{\chi}_{\text{FLOP}} = N_\omega N_{\text{occ}} N_{\text{unocc}} N_{\mathbf{r}}^2 = 40 \cdot 5 \cdot 30 \cdot 1000^2 = 6 \cdot 10^9$$

exascale: FLOP/s = 10^{18}/s

$$N^{\chi}_{\text{FLOP}} = N_\omega N_{\text{occ}} N_{\text{unocc}} N_{\mathbf{r}}^2 = 40 \cdot 5 \cdot 10^3 \cdot 3 \cdot 10^4 \cdot (10^6)^2 = 6 \cdot 10^{21}$$

$$\epsilon(\mathbf{r},\mathbf{r}',i\omega) = \delta(\mathbf{r},\mathbf{r}') - \int d\mathbf{r}'' v(\mathbf{r},\mathbf{r}'')\chi(\mathbf{r}'',\mathbf{r}',i\omega)$$

$$O(N^3)$$

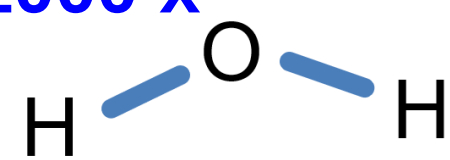
$$N^{\epsilon}_{\text{FLOP}} = N_\omega N_{\mathbf{r}}^3 = 40 \cdot 1000^3 = 4 \cdot 10^{10}$$

$$N^{\epsilon}_{\text{FLOP}} = N_\omega N_{\mathbf{r}}^3 = 40 \cdot (10^6)^3 = 4 \cdot 10^{19}$$

$$W(\mathbf{r},\mathbf{r}',i\omega) = \int d\mathbf{r}'' \epsilon^{-1}(\mathbf{r},\mathbf{r}'',i\omega) v(\mathbf{r}'',\mathbf{r}')$$

$$G(\mathbf{r},\mathbf{r}',i\omega) = \sum_m^{\text{all}} \frac{\psi_m(\mathbf{r}')\psi_m(\mathbf{r})}{i\omega + \varepsilon_{\text{F}} - \varepsilon_m}$$

1000 x



$$\Sigma(\mathbf{r},\mathbf{r}',i\omega) = -\int_{-\infty}^{\infty} \frac{d\omega'}{2\pi} G(\mathbf{r},\mathbf{r}',i\omega - i\omega') W(\mathbf{r},\mathbf{r}',i\omega')$$

$$N_{\text{occ}} = 5000 \qquad N_{\text{unocc}} = 30000$$

$$N_{\mathbf{r}} = 10^6$$

$$\varepsilon^{G_0W_0}_{n\mathbf{k}} = \varepsilon^{\text{DFT}}_{n\mathbf{k}} + \Sigma^{G_0W_0}_{n\mathbf{k}}(\varepsilon^{G_0W_0}_{n\mathbf{k}}) - v^{\text{xc}}_{n\mathbf{k}}$$

1. Standard *GW* workflow [$O(N^4)$]

2. ***GW* space-time method [$O(N^3)$]**
   Rojas, Godby, Needs, Phys. Rev. Lett. **74**, 1827 (1995)

3. Why do we need large-scale *GW* calculations?

4. Low-scaling *GW* algorithm for 2D materials
   Graml, Zollner, Hernangomez, Faria, Wilhelm, arXiv:2306.16066 (2023)

GW space-time method: Algorithm

$O(N^4)$

$$\chi(\mathbf{r},\mathbf{r}',i\omega) = \sum_i^{\text{occ}} \sum_a^{\text{unocc}} \psi_a(\mathbf{r}')\psi_i(\mathbf{r}')\psi_i(\mathbf{r})\psi_a(\mathbf{r}) \frac{2(\varepsilon_i - \varepsilon_a)}{\omega^2 + (\varepsilon_i - \varepsilon_a)^2}$$

$$\epsilon(\mathbf{r},\mathbf{r}',i\omega) = \delta(\mathbf{r},\mathbf{r}') - \int d\mathbf{r}'' v(\mathbf{r},\mathbf{r}'')\chi(\mathbf{r}'',\mathbf{r}',i\omega)$$

$$W(\mathbf{r},\mathbf{r}',i\omega) = \int d\mathbf{r}'' \epsilon^{-1}(\mathbf{r},\mathbf{r}'',i\omega) v(\mathbf{r}'',\mathbf{r}')$$

$$G(\mathbf{r},\mathbf{r}',i\omega) = \sum_m^{\text{all}} \frac{\psi_m(\mathbf{r}')\psi_m(\mathbf{r})}{i\omega + \varepsilon_{\text{F}} - \varepsilon_m}$$

$$\Sigma(\mathbf{r},\mathbf{r}',i\omega) = -\int_{-\infty}^{\infty} \frac{d\omega'}{2\pi} G(\mathbf{r},\mathbf{r}',i\omega - i\omega') W(\mathbf{r},\mathbf{r}',i\omega')$$

$$\varepsilon_{n\mathbf{k}}^{G_0W_0} = \varepsilon_{n\mathbf{k}}^{\text{DFT}} + \Sigma_{n\mathbf{k}}^{G_0W_0}(\varepsilon_{n\mathbf{k}}^{G_0W_0}) - v_{n\mathbf{k}}^{\text{xc}}$$

$O(N^3)$

$$\chi(\mathbf{r},\mathbf{r}',i\tau) = \sum_{ia} \psi_a(\mathbf{r}')\psi_i(\mathbf{r}')\psi_i(\mathbf{r})\psi_a(\mathbf{r}) e^{-|\varepsilon_i - \varepsilon_a|\tau}$$

$$\overset{\tau>0}{=} \sum_i^{\text{occ}} \psi_i(\mathbf{r}')\psi_i(\mathbf{r}) e^{\varepsilon_i \tau} \sum_a^{\text{unocc}} \psi_a(\mathbf{r}')\psi_a(\mathbf{r}) e^{-\varepsilon_a \tau}$$

$$\chi(\mathbf{r},\mathbf{r}',i\omega) = \int d\tau\, e^{i\omega\tau} \chi(\mathbf{r},\mathbf{r}',i\tau)$$

$$W(\mathbf{r},\mathbf{r}',i\tau) = \int d\omega\, e^{-i\omega\tau} W(\mathbf{r},\mathbf{r}',i\omega)$$

$$\Sigma(\mathbf{r},\mathbf{r}',i\tau) = G(\mathbf{r},\mathbf{r}',i\tau)\, W(\mathbf{r},\mathbf{r}',i\tau)$$

$$\Sigma(\mathbf{r},\mathbf{r}',i\omega) = \int d\tau\, e^{i\omega\tau} \Sigma(\mathbf{r},\mathbf{r}',i\tau)$$

Rojas, Godby, Needs, Phys. Rev. Lett. **74**, 1827 (1995)

GW space-time method: Algorithm

$O(N^4)$

$$\chi(\mathbf{r},\mathbf{r}',i\omega) = \sum_i^{\text{occ}} \sum_a^{\text{unocc}} \psi_a(\mathbf{r}')\psi_i(\mathbf{r}')\psi_i(\mathbf{r})\psi_a(\mathbf{r}) \frac{2(\varepsilon_i - \varepsilon_a)}{\omega^2 + (\varepsilon_i - \varepsilon_a)^2}$$

$$N^{\chi}_{\text{FLOP}} = N_\omega N_{\text{occ}} N_{\text{unocc}} N_{\mathbf{r}}^2$$
$$= 40 \cdot 5 \cdot 10^3 \cdot 3 \cdot 10^4 \cdot (10^6)^2$$
$$\boxed{= 6 \cdot 10^{21}}$$

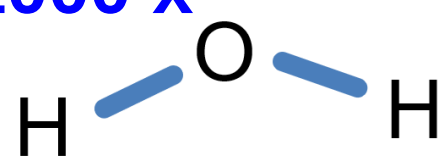
$O(N^3)$

$$\chi(\mathbf{r},\mathbf{r}',i\tau) = \sum_{ia} \psi_a(\mathbf{r}')\psi_i(\mathbf{r}')\psi_i(\mathbf{r})\psi_a(\mathbf{r}) e^{-|\varepsilon_i - \varepsilon_a|\tau}$$
$$\overset{\tau>0}{=} \sum_i^{\text{occ}} \psi_i(\mathbf{r}')\psi_i(\mathbf{r}) e^{\varepsilon_i \tau} \sum_a^{\text{unocc}} \psi_a(\mathbf{r}')\psi_a(\mathbf{r}) e^{-\varepsilon_a \tau}$$

Algorithm:

1. $D_1(\mathbf{r},\mathbf{r}',i\tau) = \sum_i^{\text{occ}} \psi_i(\mathbf{r}')\psi_i(\mathbf{r}) e^{\varepsilon_i \tau}$ $\quad N^{D_1}_{\text{FLOP}} = N_\tau N_{\text{occ}} N_{\mathbf{r}}^2 = 40 \cdot 5 \cdot 10^3 \cdot (10^6)^2 = 2 \cdot 10^{16}$
2. $D_2(\mathbf{r},\mathbf{r}',i\tau) = \sum_a^{\text{unocc}} \psi_a(\mathbf{r}')\psi_a(\mathbf{r}) e^{-\varepsilon_a \tau}$ $\quad N^{D_2}_{\text{FLOP}} = N_\tau N_{\text{unocc}} N_{\mathbf{r}}^2 = 40 \cdot 3 \cdot 10^4 \cdot (10^6)^2 = 1.2 \cdot 10^{17}$
3. $\chi(\mathbf{r},\mathbf{r}',i\tau) = D_1(\mathbf{r},\mathbf{r}',i\tau)\, D_2(\mathbf{r},\mathbf{r}',i\tau)$ $\quad N^{\chi}_{\text{FLOP}} = N_\tau N_{\mathbf{r}}^2 = 40 \cdot (10^6)^2 = 4 \cdot 10^{13}$

1000 x



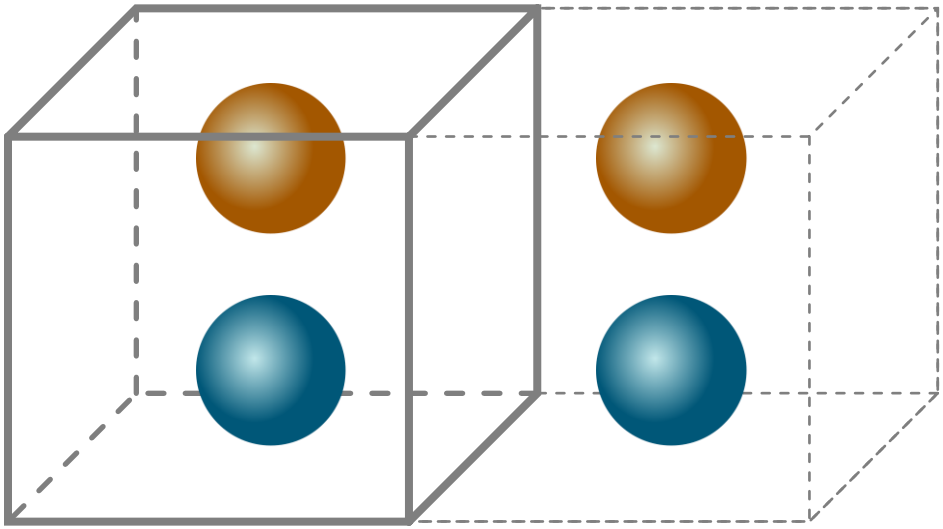
$N_{\text{occ}} = 5000 \qquad N_{\text{unocc}} = 30000 \qquad N_{\mathbf{r}} = 10^6$

GW space-time method: long-range Coulomb interaction

$$\epsilon(\mathbf{r}, \mathbf{r}', i\omega) = \delta(\mathbf{r}, \mathbf{r}') - \int d\mathbf{r}'' \; \underbrace{v(\mathbf{r}, \mathbf{r}'')}_{= \frac{1}{|\mathbf{r} - \mathbf{r}''|}} \; \chi(\mathbf{r}'', \mathbf{r}', i\omega)$$

$\in$ cell $\quad \in \mathbb{R}^3 \quad \in \mathbb{R}^3$

$$N^{\epsilon}_{\text{FLOP}} = N_{\mathbf{r}} \cdot (N_{\mathbf{r}} \cdot N_{\text{cells}})^2 = N^3_{\mathbf{r}} \cdot N^2_{\text{cells}}$$



GW space-time method: long-range Coulomb interaction

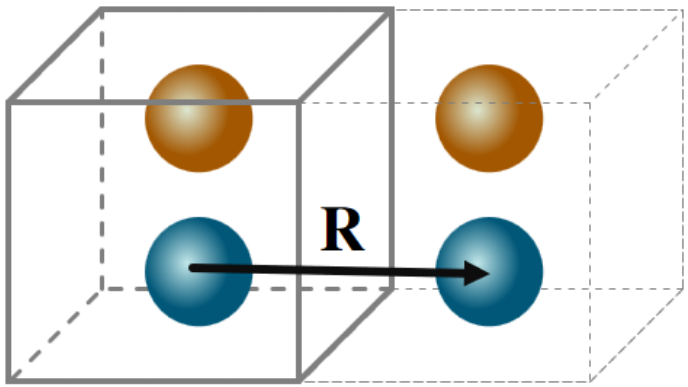
$$\epsilon(\mathbf{r}, \mathbf{r}', i\omega) = \delta(\mathbf{r}, \mathbf{r}') - \int_{\mathbb{R}^3} d\mathbf{r}'' v(\mathbf{r}, \mathbf{r}'') \chi(\mathbf{r}'', \mathbf{r}', i\omega)$$

$\in$ cell $\quad \in \mathbb{R}^3$

$$N^{\epsilon}_{\text{FLOP}} = N_{\mathbf{r}} \cdot (N_{\mathbf{r}} \cdot N_{\text{cells}})^2 = N^3_{\mathbf{r}} \cdot N^2_{\text{cells}}$$

$$\epsilon(\mathbf{r}, \mathbf{r}' + \mathbf{R}', i\omega) = \delta(\mathbf{r}, \mathbf{r}')\delta_{\mathbf{R}',\mathbf{0}} - \int d\mathbf{r}'' v(\mathbf{r}, \mathbf{r}'') \chi(\mathbf{r}'', \mathbf{r}' + \mathbf{R}', i\omega)$$

$\in$ cell



$$\epsilon(\mathbf{r}, \mathbf{r}', \mathbf{q}, i\omega) = \sum_{\mathbf{R}'} e^{i\mathbf{q}\mathbf{R}'} \epsilon(\mathbf{r}, \mathbf{r}' + \mathbf{R}', i\omega)$$

$$\epsilon_{\mathbf{G}\mathbf{G}'}(\mathbf{q}, i\omega) \overset{\text{FFT}}{=} \int_{\text{cell}} e^{i\mathbf{G}\mathbf{r}} \, \epsilon(\mathbf{r}, \mathbf{r}', \mathbf{q}, i\omega) \, e^{i\mathbf{G}'\mathbf{r}'} \, d\mathbf{r} \, d\mathbf{r}'$$

$$\epsilon_{\mathbf{G}\mathbf{G}'}(\mathbf{q}, i\omega) = \delta_{\mathbf{G},\mathbf{G}'} - V_{\mathbf{G}\mathbf{G}'}(\mathbf{q}) \chi_{\mathbf{G}\mathbf{G}'}(\mathbf{q}, i\omega), \qquad V_{\mathbf{G}\mathbf{G}'}(\mathbf{q}) = \frac{4\pi}{|\mathbf{G} + \mathbf{q}| \, |\mathbf{G}' + \mathbf{q}|}$$

$$N^{\epsilon}_{\text{FLOP}} = N_{\mathbf{q}} \, N^2_{\mathbf{G}}$$

GW space-time method: whole algorithm

$$D^1_{\mathbf{GG'}}(\mathbf{k}, i\tau) = \sum_i^{\text{occ}} C^*_{i\mathbf{G}}(\mathbf{k})\, C_{i\mathbf{G'}}(\mathbf{k})\, e^{\varepsilon_i(\mathbf{k})\tau} \qquad D^2_{\mathbf{GG'}}(\mathbf{k}, i\tau) = \sum_a^{\text{unocc}} C^*_{a\mathbf{G}}(\mathbf{k})\, C_{a\mathbf{G'}}(\mathbf{k})\, e^{-\varepsilon_a(\mathbf{k})\tau}$$

FFT

$$D_1(\mathbf{r}, \mathbf{r}', i\tau) \quad D_2(\mathbf{r}, \mathbf{r}', i\tau)$$

$$\chi(\mathbf{r}, \mathbf{r}', i\tau) = D_1(\mathbf{r}, \mathbf{r}', i\tau)\, D_2(\mathbf{r}, \mathbf{r}', i\tau)$$

FFT

$$\chi_{\mathbf{GG'}}(\mathbf{k}, i\tau)$$

$$\chi_{\mathbf{GG'}}(\mathbf{k}, i\omega) = \int d\tau\, e^{i\omega\tau}\, \chi_{\mathbf{GG'}}(\mathbf{k}, i\tau)$$

$$\epsilon_{\mathbf{GG'}}(\mathbf{k}, i\omega) = \delta_{\mathbf{G},\mathbf{G'}} - V_{\mathbf{GG'}}(\mathbf{k})\, \chi_{\mathbf{GG'}}(\mathbf{k}, i\omega)$$

$$W_{\mathbf{GG'}}(\mathbf{k}, i\omega) = V_{\mathbf{GG'}}(\mathbf{k})\, \epsilon^{-1}_{\mathbf{GG'}}(\mathbf{k}, i\omega)$$

$$W_{\mathbf{GG'}}(\mathbf{k}, i\tau) = \int d\omega\, e^{-i\omega\tau}\, W_{\mathbf{GG'}}(\mathbf{k}, i\omega)$$

FFT

$$W(\mathbf{r}, \mathbf{r}', i\tau)$$

$$\Sigma(\mathbf{r}, \mathbf{r}', i\tau) = G(\mathbf{r}, \mathbf{r}', i\tau)\, W(\mathbf{r}, \mathbf{r}', i\tau)$$

$$\Sigma(\mathbf{r}, \mathbf{r}', i\omega) = \int d\tau\, e^{i\omega\tau}\, \Sigma(\mathbf{r}, \mathbf{r}', i\tau)$$

$$\varepsilon^{G_0W_0}_{n\mathbf{k}} = \varepsilon^{\text{DFT}}_{n\mathbf{k}} + \Sigma_{n\mathbf{k}}(\varepsilon^{G_0W_0}_{n\mathbf{k}}) - v^{\text{xc}}_{n\mathbf{k}}$$

Rojas, Godby, Needs, Phys. Rev. Lett. **74**, 1827 (1995)
Liu, Kaltak, Klimes, Kresse, Phys. Rev. B **94**, 165109 (2016)

Challenge 1: 3x τ-ω Fourier transform

$$D^1_{\mathbf{GG'}}(\mathbf{k}, i\tau) = \sum_i^{\text{occ}} C^*_{i\mathbf{G}}(\mathbf{k})\, C_{i\mathbf{G'}}(\mathbf{k})\, e^{\varepsilon_i(\mathbf{k})\tau} \qquad D^2_{\mathbf{GG'}}(\mathbf{k}, i\tau) = \sum_a^{\text{unocc}} C^*_{a\mathbf{G}}(\mathbf{k})\, C_{a\mathbf{G'}}(\mathbf{k})\, e^{-\varepsilon_a(\mathbf{k})\tau}$$

FFT

$$D_1(\mathbf{r}, \mathbf{r}', i\tau)$$
$$D_2(\mathbf{r}, \mathbf{r}', i\tau)$$

$$\chi(\mathbf{r}, \mathbf{r}', i\tau) = D_1(\mathbf{r}, \mathbf{r}', i\tau)\, D_2(\mathbf{r}, \mathbf{r}', i\tau)$$

FFT

$$\chi_{\mathbf{GG'}}(\mathbf{k}, i\tau)$$

$$\chi_{\mathbf{GG'}}(\mathbf{k}, i\omega) = \int d\tau\, e^{i\omega\tau}\, \chi_{\mathbf{GG'}}(\mathbf{k}, i\tau)$$

$$\epsilon_{\mathbf{GG'}}(\mathbf{k}, i\omega) = \delta_{\mathbf{G},\mathbf{G'}} - V_{\mathbf{GG'}}(\mathbf{k})\, \chi_{\mathbf{GG'}}(\mathbf{k}, i\omega)$$

$$W_{\mathbf{GG'}}(\mathbf{k}, i\omega) = V_{\mathbf{GG'}}(\mathbf{k})\, \epsilon^{-1}_{\mathbf{GG'}}(\mathbf{k}, i\omega)$$

$$W_{\mathbf{GG'}}(\mathbf{k}, i\tau) = \int d\omega\, e^{-i\omega\tau}\, W_{\mathbf{GG'}}(\mathbf{k}, i\omega)$$

FFT

$$W(\mathbf{r}, \mathbf{r}', i\tau)$$

$$\Sigma(\mathbf{r}, \mathbf{r}', i\tau) = G(\mathbf{r}, \mathbf{r}', i\tau)\, W(\mathbf{r}, \mathbf{r}', i\tau)$$

$$\Sigma(\mathbf{r}, \mathbf{r}', i\omega) = \int d\tau\, e^{i\omega\tau}\, \Sigma(\mathbf{r}, \mathbf{r}', i\tau)$$

$$\varepsilon^{G_0W_0}_{n\mathbf{k}} = \varepsilon^{\text{DFT}}_{n\mathbf{k}} + \Sigma_{n\mathbf{k}}(\varepsilon^{G_0W_0}_{n\mathbf{k}}) - v^{\text{xc}}_{n\mathbf{k}}$$

Rojas, Godby, Needs, Phys. Rev. Lett. **74**, 1827 (1995)
Liu, Kaltak, Klimes, Kresse, Phys. Rev. B **94**, 165109 (2016)

Challenge 2: 3x G-r Fast Fourier Transform (FFT)

$$D^1_{\mathbf{GG'}}(\mathbf{k}, i\tau) = \sum_i^{\text{occ}} C^*_{i\mathbf{G}}(\mathbf{k})\, C_{i\mathbf{G'}}(\mathbf{k})\, e^{\varepsilon_i(\mathbf{k})\tau} \qquad D^2_{\mathbf{GG'}}(\mathbf{k}, i\tau) = \sum_a^{\text{unocc}} C^*_{a\mathbf{G}}(\mathbf{k})\, C_{a\mathbf{G'}}(\mathbf{k})\, e^{-\varepsilon_a(\mathbf{k})\tau}$$

FFT

$$D_1(\mathbf{r}, \mathbf{r}', i\tau) \quad D_2(\mathbf{r}, \mathbf{r}', i\tau)$$

$$\chi(\mathbf{r}, \mathbf{r}', i\tau) = D_1(\mathbf{r}, \mathbf{r}', i\tau)\, D_2(\mathbf{r}, \mathbf{r}', i\tau)$$

FFT

$$\chi_{\mathbf{GG'}}(\mathbf{k}, i\tau)$$

$$\chi_{\mathbf{GG'}}(\mathbf{k}, i\omega) = \int d\tau\, e^{i\omega\tau}\, \chi_{\mathbf{GG'}}(\mathbf{k}, i\tau)$$

$$\epsilon_{\mathbf{GG'}}(\mathbf{k}, i\omega) = \delta_{\mathbf{G},\mathbf{G'}} - V_{\mathbf{GG'}}(\mathbf{k})\, \chi_{\mathbf{GG'}}(\mathbf{k}, i\omega)$$

$$W_{\mathbf{GG'}}(\mathbf{k}, i\omega) = V_{\mathbf{GG'}}(\mathbf{k})\, \epsilon^{-1}_{\mathbf{GG'}}(\mathbf{k}, i\omega)$$

$$W_{\mathbf{GG'}}(\mathbf{k}, i\tau) = \int d\omega\, e^{-i\omega\tau}\, W_{\mathbf{GG'}}(\mathbf{k}, i\omega)$$

FFT

$$W(\mathbf{r}, \mathbf{r}', i\tau)$$

$$\Sigma(\mathbf{r}, \mathbf{r}', i\tau) = G(\mathbf{r}, \mathbf{r}', i\tau)\, W(\mathbf{r}, \mathbf{r}', i\tau)$$

$$\Sigma(\mathbf{r}, \mathbf{r}', i\omega) = \int d\tau\, e^{i\omega\tau}\, \Sigma(\mathbf{r}, \mathbf{r}', i\tau)$$

$$\varepsilon^{G_0W_0}_{n\mathbf{k}} = \varepsilon^{\text{DFT}}_{n\mathbf{k}} + \Sigma_{n\mathbf{k}}(\varepsilon^{G_0W_0}_{n\mathbf{k}}) - v^{\text{xc}}_{n\mathbf{k}}$$

Rojas, Godby, Needs, Phys. Rev. Lett. **74**, 1827 (1995)

Liu, Kaltak, Klimes, Kresse, Phys. Rev. B **94**, 165109 (2016)

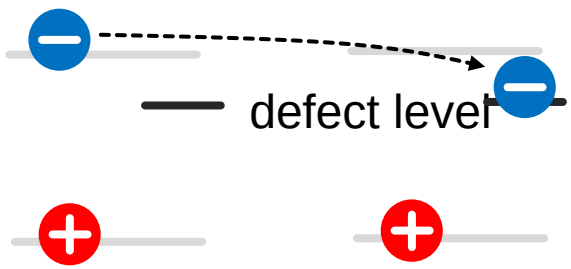
1. Standard *GW* workflow [$O(N^4)$]

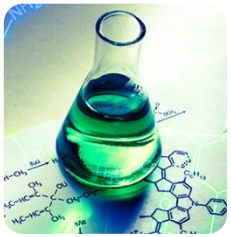
2. *GW* space-time method [$O(N^3)$]
   Rojas, Godby, Needs, Phys. Rev. Lett. **74**, 1827 (1995)

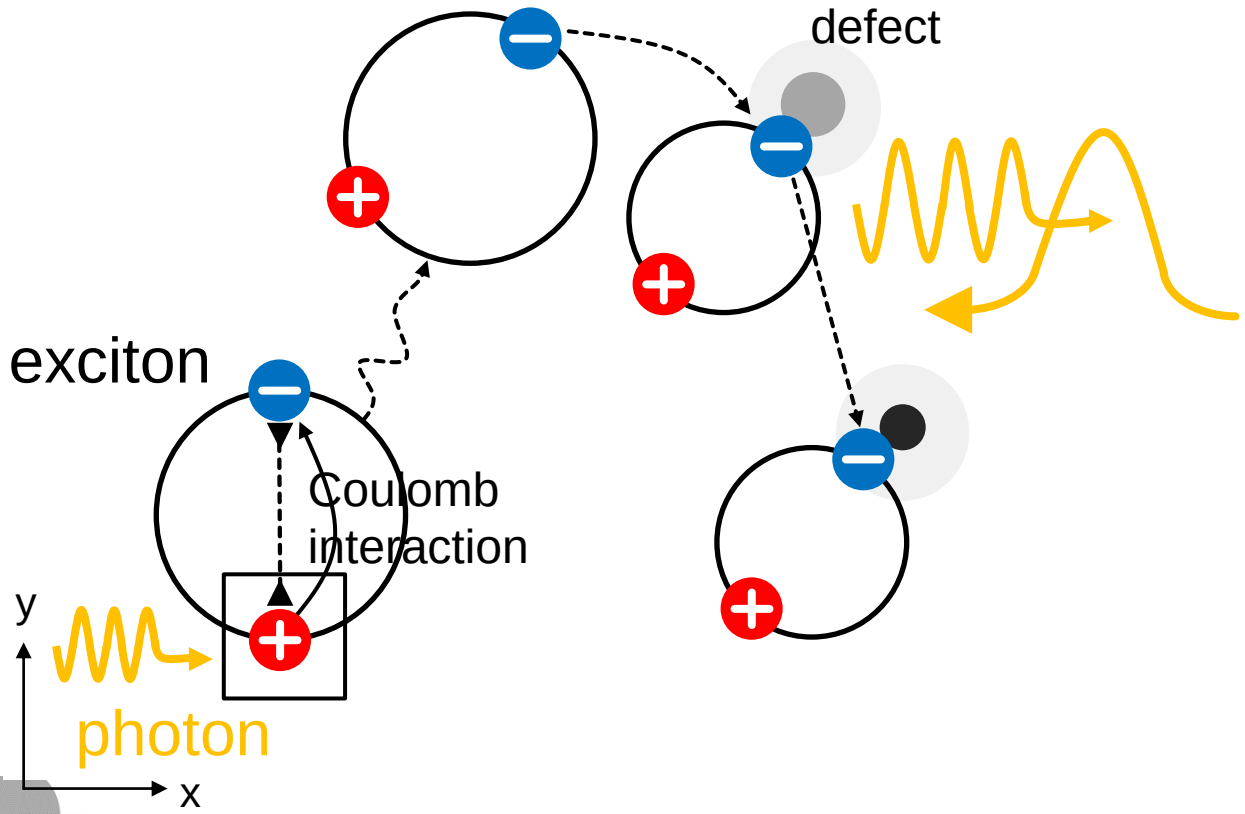
3. **Why do we need large-scale *GW* calculations?**

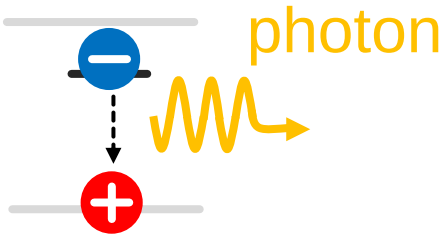
4. Low-scaling *GW* algorithm for 2D materials
   Graml, Zollner, Hernangomez, Faria, Wilhelm, arXiv:2306.16066 (2023)

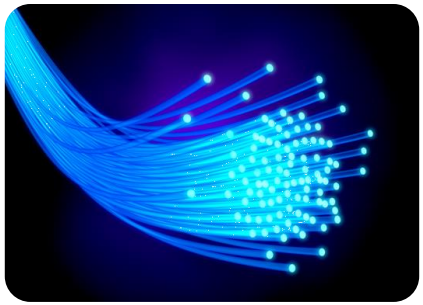
Dynamics of an electronic excitation



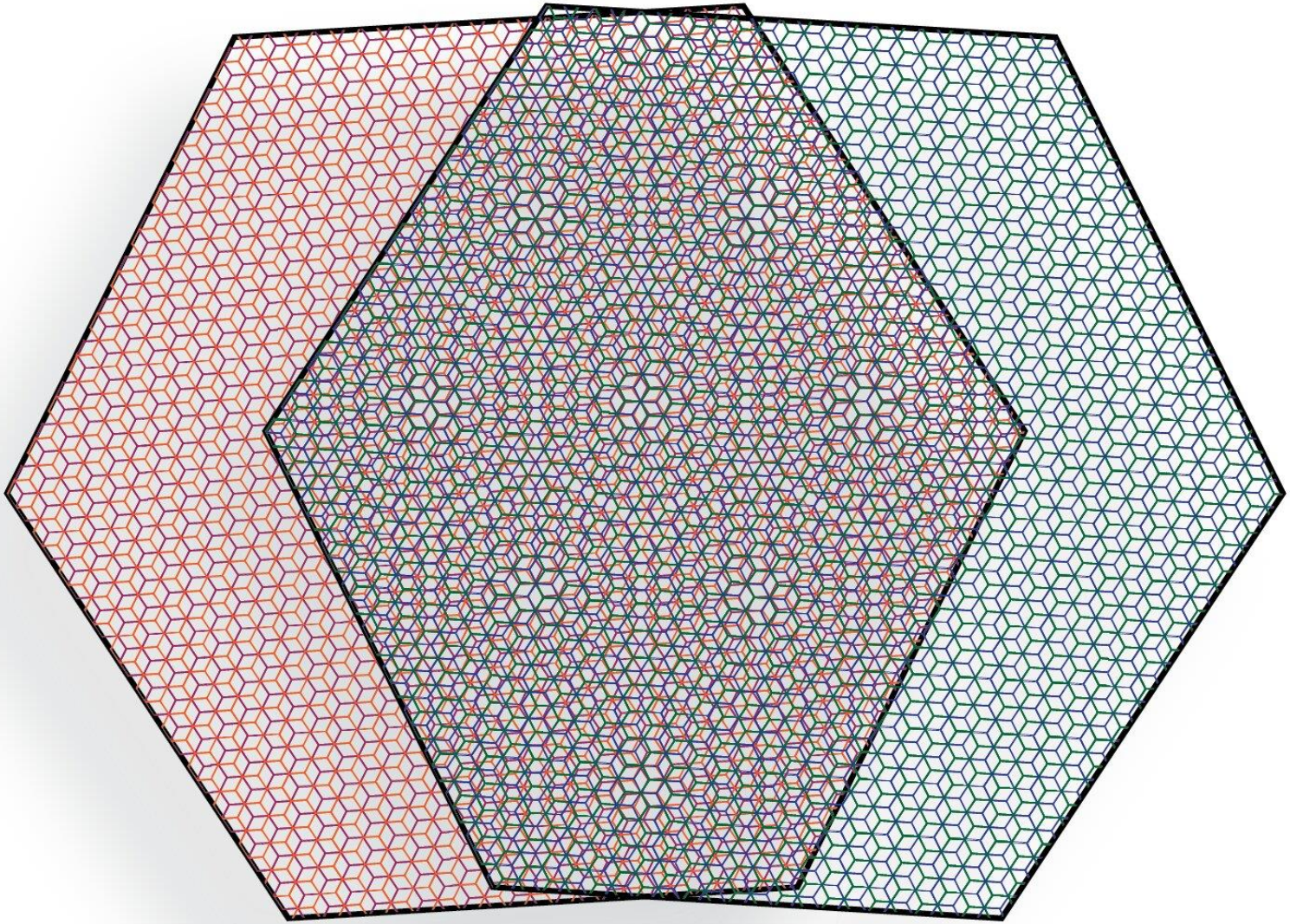




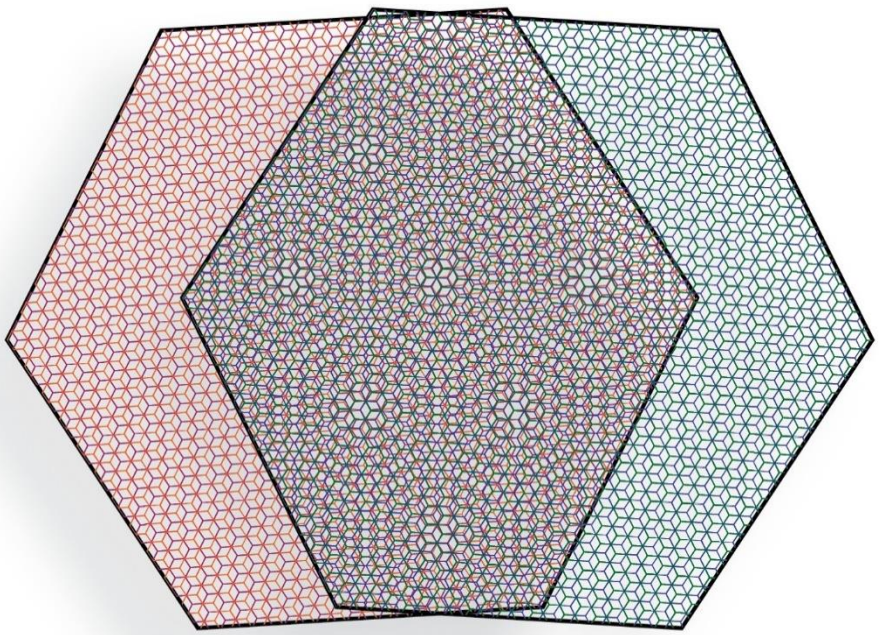




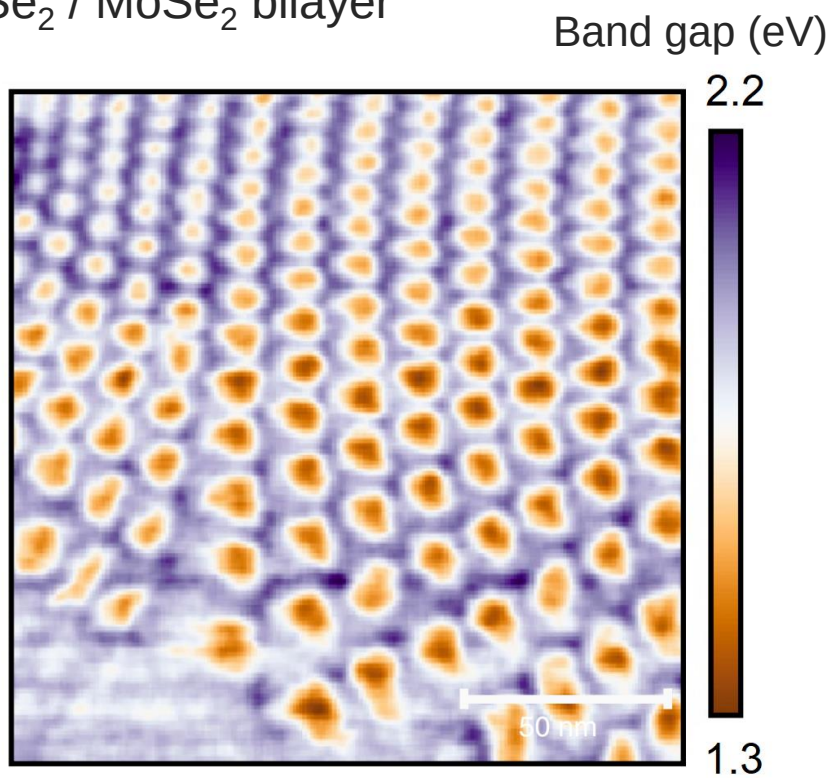
Moiré structures



Band structure of moiré structures

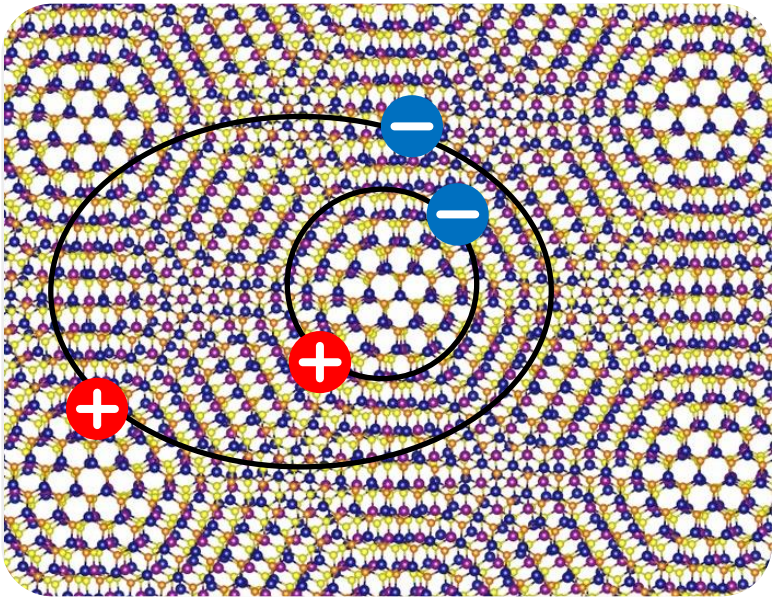


WSe_2 / $MoSe_2$ bilayer

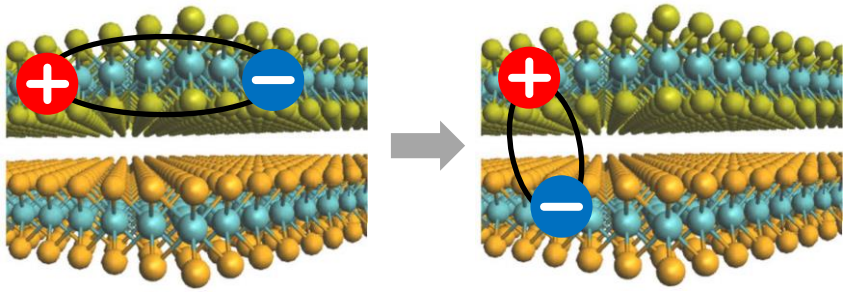


Shabani et al., Nat. Phys. ***17****, 720 (2021)*

Dynamics of electronic excitations in moiré structures



photon



*D. Schmitt, S. Mathias et al., Nature **608**, 499 (2022)*
"Formation of moiré interlayer excitons in space and time"

*O. Karnis, T. F. Heinz, et al., Nature **603**, 247 (2022), "Structure of the moiré exciton captured by imaging its electron and hole"*

RUN. Regensburg Center for Ultrafast Nanoscopy

1. Standard *GW* workflow [$O(N^4)$]

2. *GW* space-time method [$O(N^3)$]
   Rojas, Godby, Needs, Phys. Rev. Lett. **74**, 1827 (1995)

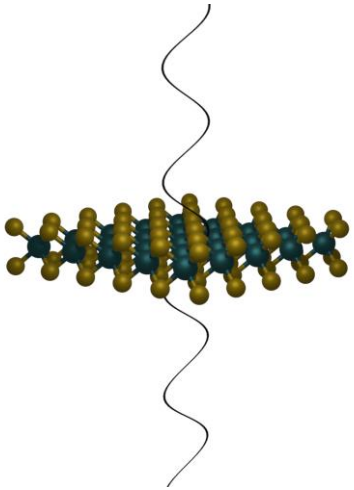
3. Why do we need large-scale *GW* calculations?

4. **Low-scaling *GW* algorithm for 2D materials [$O(N^3)$]**
   Graml, Zollner, Hernangomez, Faria, Wilhelm, arXiv:2306.16066 (2023)

Large-scale *GW* calculations for 2D materials – a challenge

Large prefactor of space-time method for 2D materials:

1. **large plane-wave basis, large real-space grid**
2. **dense k-point mesh due to** $W_0(\mathbf{q}) \sim \frac{1}{|\mathbf{q}|}$



Desired: small prefactor in the space-time method for 2D materials:

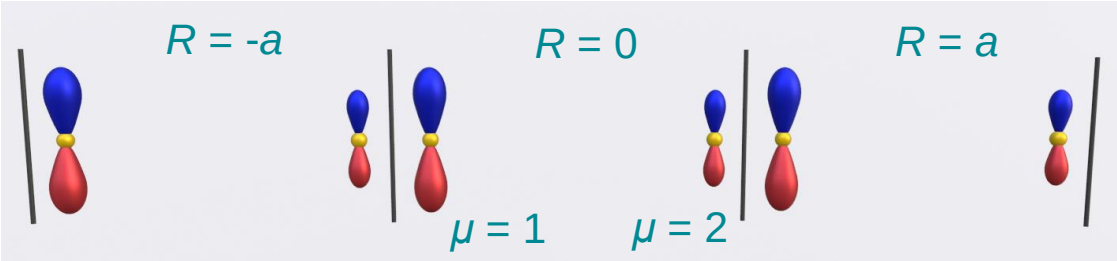
1. **local, atomic-orbital like basis (e.g. Gaussians)**
2. **Γ-point-only whenever possible**

Gaussian basis

Expand Bloch states in Gaussian basis:

Blum *et al.*, Comput. Phys. Commun. **180**, 2175 (2009)

$$\psi_{n\mathbf{k}}(\mathbf{r}) = \sum_{\mathbf{R}} \sum_{\mu} C_{\mu n}(\mathbf{k})\, e^{i\mathbf{k}\mathbf{R}}\, \phi_\mu(\mathbf{r}-\mathbf{R})$$



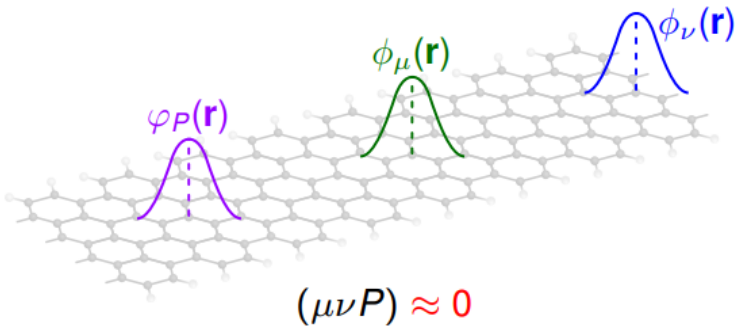
Four-center Coulomb integrals:

$$(\mu\nu|\lambda\sigma) = \int d^3\mathbf{r}\, d^3\mathbf{r}'\, \frac{\phi_\nu(\mathbf{r})\phi_\mu(\mathbf{r})\phi_\lambda(\mathbf{r}')\phi_\sigma(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}$$

Resolution of the identity: auxiliary Gaussian basis functions $\varphi_P(\mathbf{r})$: $\mathrm{Id} \approx \sum_{PQ} |P\rangle\, S^{-1}_{PQ}\, \langle Q|$

$$(\mu\nu|\lambda\sigma)_{\mathrm{RI}} = \sum_{PQRT} (\mu\nu P) S^{-1}_{PQ} V_{QR} S^{-1}_{RT} (T\lambda\sigma)$$

$$(\mu\nu P) = \int d^3\mathbf{r}\, \phi_\mu(\mathbf{r})\phi_\nu(\mathbf{r})\varphi_P(\mathbf{r})$$



GW space-time method for molecules in Gaussian basis

$$G_{\mu\nu}(i\tau) = \begin{cases} \sum_i C_{\mu i} C_{\nu i} e^{\varepsilon_i \tau}, & \text{if } \tau < 0 \\ -\sum_a C_{\mu a} C_{\nu a} e^{-\varepsilon_a \tau}, & \text{if } \tau > 0 \end{cases}$$

$$\tilde{\chi}_{PQ}(i\tau) = \sum_{\mu\sigma} \sum_{\lambda} (\lambda\sigma P) G_{\mu\lambda}(i\tau) \times \sum_{\nu} (\mu\nu Q) G_{\nu\sigma}(-i\tau)$$

$\mathcal{O}(N^2)$

$$\chi(i\tau) = \mathbf{S}^{-1} \tilde{\chi}(i\tau) \mathbf{S}^{-1}$$

$$\chi(i\omega)$$

$$\varepsilon(i\omega) = \mathbf{1} - \mathbf{V}\chi(i\omega)$$

$$\mathbf{W}(i\omega) = \varepsilon^{-1}(i\omega)\mathbf{V}$$

$$\tilde{\mathbf{W}}(i\omega) = \mathbf{S}^{-1}\mathbf{W}(i\omega)\mathbf{S}^{-1}$$

$$\tilde{\mathbf{W}}(i\tau)$$

$\mathcal{O}(N^2)$

$$\Sigma_n(i\tau) = -\sum_{\mu\nu PQ} G_{\mu\nu}(i\tau)(n\mu P) \times \tilde{W}_{PQ}(i\tau)(Q\nu n)$$

$$\Sigma_n(i\omega)$$

$$\Sigma_n(\varepsilon)$$

$$\varepsilon_n^{GW} = \varepsilon_n + \mathrm{Re}\,\Sigma_n(\varepsilon_n^{GW}) - v_n^{xc}$$

RI with overlap metric: $(\mu\nu|\lambda\sigma) = \sum_{PQRT} (\mu\nu P) S^{-1}_{PQ} V_{QR} S^{-1}_{RT} (T\lambda\sigma)$

GW space-time method

$$D^1_{\mathbf{GG'}}(\mathbf{k}, i\tau) = \sum_i^{\text{occ}} C^*_{i\mathbf{G}}(\mathbf{k})\, C_{i\mathbf{G'}}(\mathbf{k})\, e^{\varepsilon_i(\mathbf{k})\tau} \qquad D^2_{\mathbf{GG'}}(\mathbf{k}, i\tau) = \sum_a^{\text{unocc}} C^*_{a\mathbf{G}}(\mathbf{k})\, C_{a\mathbf{G'}}(\mathbf{k})\, e^{-\varepsilon_a(\mathbf{k})\tau}$$

FFT

$$D_1(\mathbf{r}, \mathbf{r}', i\tau) \quad D_2(\mathbf{r}, \mathbf{r}', i\tau)$$

$$\chi(\mathbf{r}, \mathbf{r}', i\tau) = D_1(\mathbf{r}, \mathbf{r}', i\tau)\, D_2(\mathbf{r}, \mathbf{r}', i\tau)$$

FFT

$$\chi_{\mathbf{GG'}}(\mathbf{k}, i\tau)$$

$$\chi_{\mathbf{GG'}}(\mathbf{k}, i\omega) = \int d\tau\, e^{i\omega\tau}\, \chi_{\mathbf{GG'}}(\mathbf{k}, i\tau)$$

$$\epsilon_{\mathbf{GG'}}(\mathbf{k}, i\omega) = \delta_{\mathbf{G},\mathbf{G'}} - V_{\mathbf{GG'}}(\mathbf{k})\, \chi_{\mathbf{GG'}}(\mathbf{k}, i\omega)$$

$$W_{\mathbf{GG'}}(\mathbf{k}, i\omega) = V_{\mathbf{GG'}}(\mathbf{k})\, \epsilon^{-1}_{\mathbf{GG'}}(\mathbf{k}, i\omega)$$

$$W_{\mathbf{GG'}}(\mathbf{k}, i\tau) = \int d\omega\, e^{-i\omega\tau}\, W_{\mathbf{GG'}}(\mathbf{k}, i\omega)$$

FFT

$$W(\mathbf{r}, \mathbf{r}', i\tau)$$

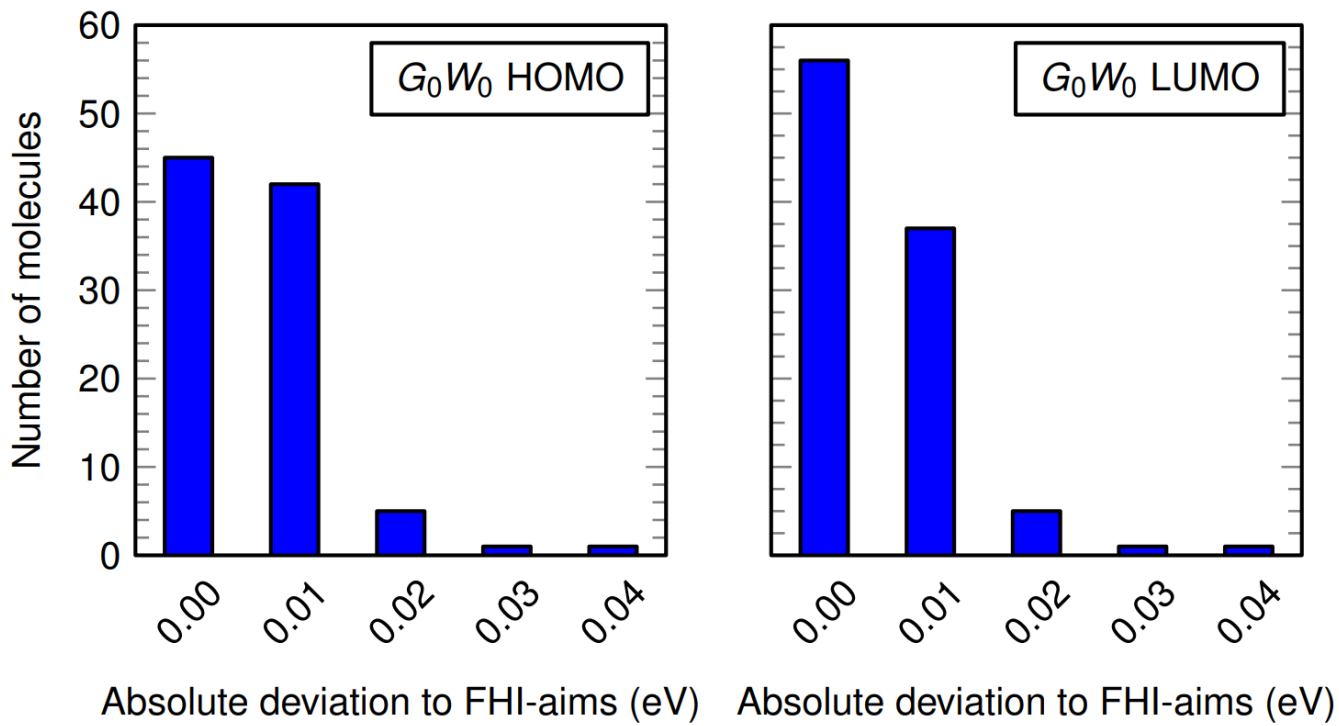
$$\Sigma(\mathbf{r}, \mathbf{r}', i\tau) = G(\mathbf{r}, \mathbf{r}', i\tau)\, W(\mathbf{r}, \mathbf{r}', i\tau)$$

$$\Sigma(\mathbf{r}, \mathbf{r}', i\omega) = \int d\tau\, e^{i\omega\tau}\, \Sigma(\mathbf{r}, \mathbf{r}', i\tau)$$

$$\varepsilon^{G_0W_0}_{n\mathbf{k}} = \varepsilon^{\text{DFT}}_{n\mathbf{k}} + \Sigma_{n\mathbf{k}}(\varepsilon^{G_0W_0}_{n\mathbf{k}}) - v^{\text{xc}}_{n\mathbf{k}}$$

Rojas, Godby, Needs, Phys. Rev. Lett. **74**, 1827 (1995)
Liu, Kaltak, Klimes, Kresse, Phys. Rev. B **94**, 165109 (2016)

Numerical precision of low-scaling *GW* for molecules



[J. Wilhelm, P. Seewald, D. Golze, J. Chem. Theory Comput. **17**, 1662-1677 (2021)]

- *GW*100 benchmark set [J. Chem. Theory Comput. **11**, 5665-5687 (2015)]
- 5 excluded HOMOs: BN, O_3, MgO, BeO, CuCN (128 Padé parameters necessary)
- Similar accuracy in Fiesta [I. Duchemin, X. Blase, J. Chem. Theory Comput. **17**, 2383 (2021)]

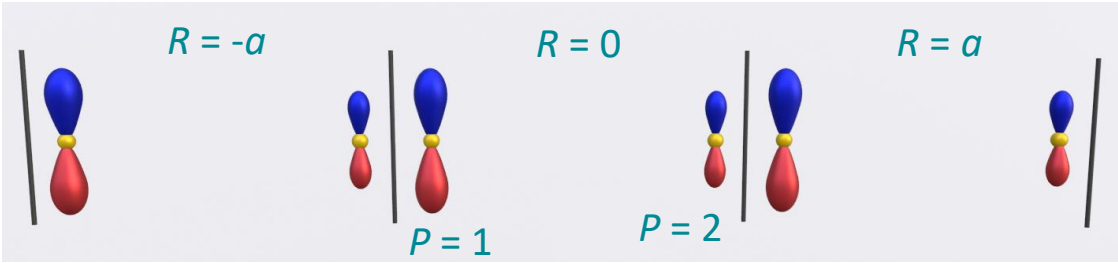
k-points in density response

Full k-points:

$$\chi_{PQ}(i\omega,\mathbf{q}) = \frac{2}{N_\mathbf{k}}\sum_\mathbf{k}(i\mathbf{k}\,a\mathbf{k}{-}\mathbf{q}\,P\mathbf{q})\,\frac{\varepsilon_{i\mathbf{k}}-\varepsilon_{a\mathbf{k}-\mathbf{q}}}{\omega^2+(\varepsilon_{i\mathbf{k}}-\varepsilon_{a\mathbf{k}-\mathbf{q}})^2}\,(i\mathbf{k}\,a\mathbf{k}{-}\mathbf{q}\,Q\mathbf{q})$$

Γ-point only:

$$\chi_{PQ}(i\omega,\mathbf{0}) = 2\,(i\mathbf{0}\,a\mathbf{0}\,P\mathbf{0})\,\frac{\varepsilon_{i\mathbf{0}}-\varepsilon_{a\mathbf{0}}}{\omega^2+(\varepsilon_{i\mathbf{0}}-\varepsilon_{a\mathbf{0}})^2}\,(i\mathbf{0}\,a\mathbf{0}\,Q\mathbf{0})$$



Γ-point and real space:

$$\chi_{PQ}(i\omega,\mathbf{0}) = \begin{pmatrix}\alpha & \gamma\\ \gamma & \beta\end{pmatrix} = \sum_\mathbf{R}\chi^\mathbf{R}_{PQ}(i\omega) = \chi^{-a}_{PQ}(i\omega)+\chi^{0}_{PQ}(i\omega)+\chi^{a}_{PQ}(i\omega) \equiv \begin{pmatrix}0 & \gamma\\ 0 & 0\end{pmatrix}+\begin{pmatrix}\alpha & 0\\ 0 & \beta\end{pmatrix}+\begin{pmatrix}0 & 0\\ \gamma & 0\end{pmatrix}$$

large cell

$$\chi^\mathbf{R}_{PQ} := \langle\varphi^\mathbf{0}_P|\chi|\varphi^\mathbf{R}_Q\rangle = \int d\mathbf{r}d\mathbf{r}'\,\varphi^\mathbf{0}_P(\mathbf{r})\,\chi(\mathbf{r},\mathbf{r}')\,\varphi^\mathbf{R}_Q(\mathbf{r}')$$

Γ-point only ➡ real space for large cells:

$$\chi^\mathbf{R}_{PQ}(i\omega) = \begin{cases}\chi_{PQ}(\mathbf{0},i\omega) & \text{if } \varphi^\mathbf{0}_P,\,\varphi^\mathbf{R}_Q \text{ closest},\\ 0 & \text{else}.\end{cases}$$

Real space ➡ k-points:

$$\chi_{PQ}(i\omega,\mathbf{q}) = \sum_\mathbf{R}\chi^\mathbf{R}_{PQ}(i\omega)\,e^{i\mathbf{q}\cdot\mathbf{R}} \equiv \begin{pmatrix}\alpha & \gamma e^{-iq_xa}\\ \gamma e^{iq_xa} & \beta\end{pmatrix}$$

GW space-time method with PBC in a Gaussian basis

$$\hat{H}_{\mathrm{KS}}\psi_{n\mathbf{k}}(\mathbf{r}) = \varepsilon_{n\mathbf{k}}\psi_{n\mathbf{k}}(\mathbf{r})$$

$$\psi_{n\mathbf{k}}(\mathbf{r}) = \sum_{\mu} C_{n\mu}(\mathbf{k}) \sum_{\mathbf{R}} e^{i\mathbf{k}\cdot\mathbf{R}}\, \phi_{\mu}^{\mathbf{R}}(\mathbf{r})$$

$$G_{\mu\nu}(\mathbf{0}, i\tau) = \theta(\tau) \sum_{a}^{\mathrm{virt}} C_{a\mu}(\mathbf{0})\, C_{a\nu}(\mathbf{0}) \exp(-(\varepsilon_a - \varepsilon_{\mathrm{F}})\tau)$$
$$-\theta(-\tau) \sum_{i}^{\mathrm{occ}} C_{i\mu}(\mathbf{0})\, C_{i\nu}(\mathbf{0}) \exp(-(\varepsilon_i - \varepsilon_{\mathrm{F}})\tau)$$

$$\chi_{PQ}(\mathbf{0}, i\tau) = \sum_{\lambda\nu\mu\sigma} (\mu\nu|P)\, G_{\mu\lambda}(\mathbf{0}, -i\tau)\, (\lambda\sigma|Q)\, G_{\nu\sigma}(\mathbf{0}, i\tau)$$

$$(\mu\nu|P) = \sum_{\mathbf{R}_1\mathbf{R}_2} \int d\mathbf{r}\, d\mathbf{r}'\, \phi_{\mu}^{\mathbf{R}_1}(\mathbf{r})\, \phi_{\nu}^{\mathbf{R}_2}(\mathbf{r})\, V_{r_c}(\mathbf{r}, \mathbf{r}')\, \varphi_P^{\mathbf{0}}(\mathbf{r}')$$

$$V_{r_c}(\mathbf{r}, \mathbf{r}') = \begin{cases} 1/|\mathbf{r} - \mathbf{r}'| & \text{if } |\mathbf{r} - \mathbf{r}'| \le r_{\mathrm{c}}\,, \\ 0 & \text{else}\,, \end{cases}$$

$$\chi_{PQ}^{\mathbf{R}}(i\tau) = \begin{cases} \chi_{PQ}(\mathbf{0}, i\tau) & \text{if } \varphi_P^{\mathbf{0}}, \varphi_Q^{\mathbf{R}} \text{ closest}\,, \\ 0 & \text{else}\,. \end{cases}$$

$$\chi_{PQ}(\mathbf{k}, i\omega) = \int d\tau\, \cos(\omega\tau) \sum_{\mathbf{R}} e^{-i\mathbf{k}\cdot\mathbf{R}}\, \chi_{PQ}^{\mathbf{R}}(i\tau)\,.$$

$$\epsilon(\mathbf{k}, i\omega) = \mathbf{1} - \mathbf{L}^{\dagger}(\mathbf{k})\mathbf{S}^{-1}(\mathbf{k})\chi(\mathbf{k}, i\omega)\mathbf{S}^{-1}(\mathbf{k})\mathbf{L}(\mathbf{k})$$

$$M_{PQ}(\mathbf{k}) = \sum_{\mathbf{R}} e^{-i\mathbf{k}\cdot\mathbf{R}} \int d\mathbf{r}\, d\mathbf{r}'\, \varphi_P^{\mathbf{0}}(\mathbf{r})\, V_{r_c}(\mathbf{r}, \mathbf{r}')\, \varphi_Q^{\mathbf{R}}(\mathbf{r}')$$

$$V_{PQ}(\mathbf{k}) = \sum_{\mathbf{R}} e^{-i\mathbf{k}\cdot\mathbf{R}} \int d\mathbf{r}\, d\mathbf{r}'\, \varphi_P^{\mathbf{0}}(\mathbf{r})\, \frac{1}{|\mathbf{r} - \mathbf{r}'|}\, \varphi_Q^{\mathbf{R}}(\mathbf{r}')$$

$$\mathbf{W}(\mathbf{k}, i\omega) = \mathbf{V}_{r_c}(\mathbf{k}) + \mathbf{V}^{0.5}(\mathbf{k}) \left(\boldsymbol{\epsilon}^{-1}(\mathbf{k}, i\omega) - \mathbf{Id}\right) \mathbf{V}^{0.5}(\mathbf{k})$$

$$W_{PQ}^{\mathbf{R}}(i\omega) = \frac{1}{\Omega_{\mathrm{BZ}}} \int d\mathbf{k}\, e^{i\mathbf{k}\cdot\mathbf{R}}\, W_{PQ}(\mathbf{k}, i\omega)$$

$$W_{PQ}^{\mathrm{MIC}}(i\omega) = W_{PQ}^{\mathbf{R}_{PQ}^{\min}}(i\omega)$$

$$\tilde{\mathbf{W}}(i\omega) = \mathbf{M}^{-1}(\mathbf{0})\, \mathbf{W}^{\mathrm{MIC}}(i\omega)\, \mathbf{M}^{-1}(\mathbf{0})$$

$$\Sigma_{\lambda\sigma}(\mathbf{0}, i\tau) = i \sum_{\nu\mu PQ} (\lambda\mu|Q)\, G_{\mu\nu}(\mathbf{0}, i\tau)\, (\nu\sigma|P)\, \tilde{W}_{PQ}(i\tau)\,.$$

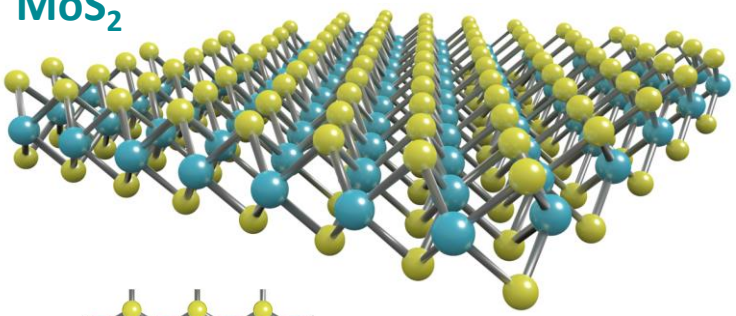
$$\Sigma_{\lambda\sigma}(\mathbf{k}, i\tau) = \sum_{\mathbf{R}} e^{i\mathbf{k}\cdot\mathbf{R}} \begin{cases} \Sigma_{\lambda\sigma}(\mathbf{0}, i\tau) & \text{if } \phi_{\lambda}^{\mathbf{0}}, \phi_{\sigma}^{\mathbf{R}} \text{ closest}\,, \\ 0 & \text{else}\,. \end{cases}$$

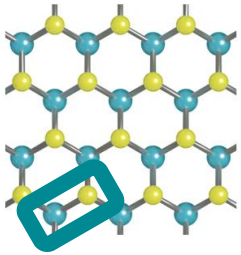
$$\Sigma_{n\mathbf{k}}(i\tau) = \sum_{\lambda\sigma} C_{n\lambda}^{*}(\mathbf{k})\, \Sigma_{\lambda\sigma}(\mathbf{k}, i\tau)\, C_{n\sigma}(\mathbf{k})\,.$$

$$\varepsilon_{n\mathbf{k}}^{G_0W_0} = \varepsilon_{n\mathbf{k}} - v_{n\mathbf{k}}^{\mathrm{xc}} + \mathrm{Re}\, \Sigma_{n\mathbf{k}}(\varepsilon_{n\mathbf{k}}^{G_0W_0})$$

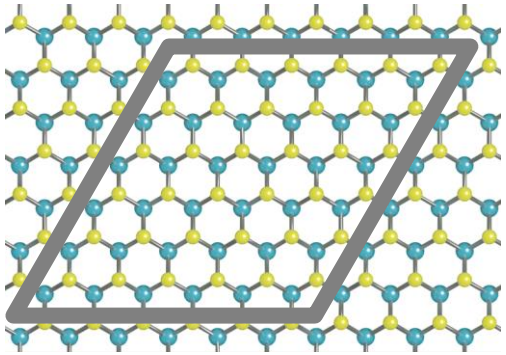
Benchmark calculations on MoS_2

MoS_2

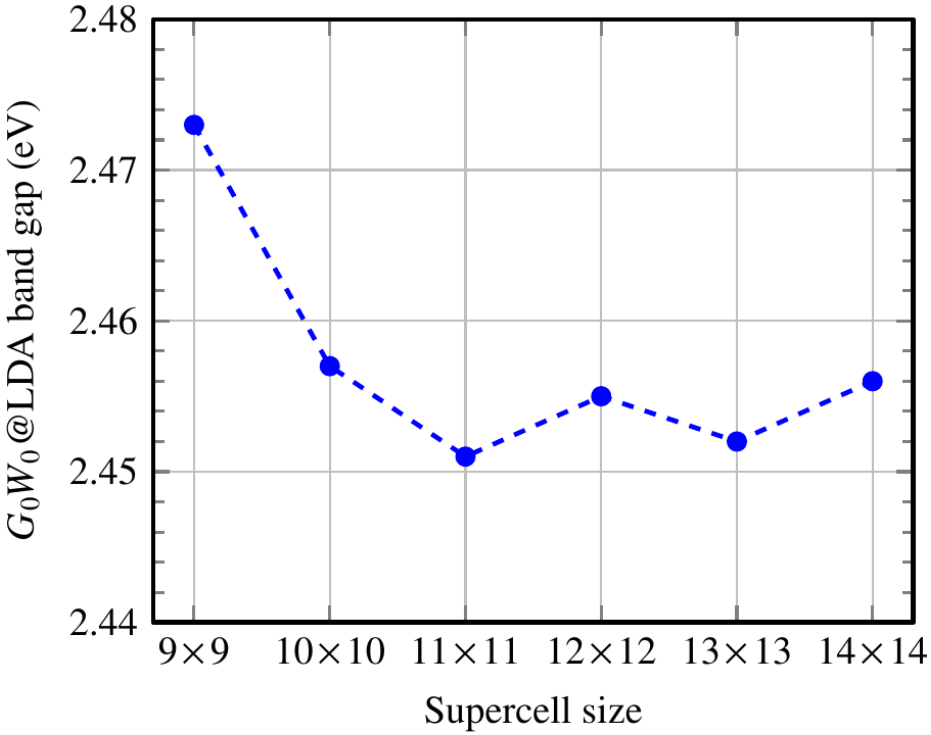




unit cell



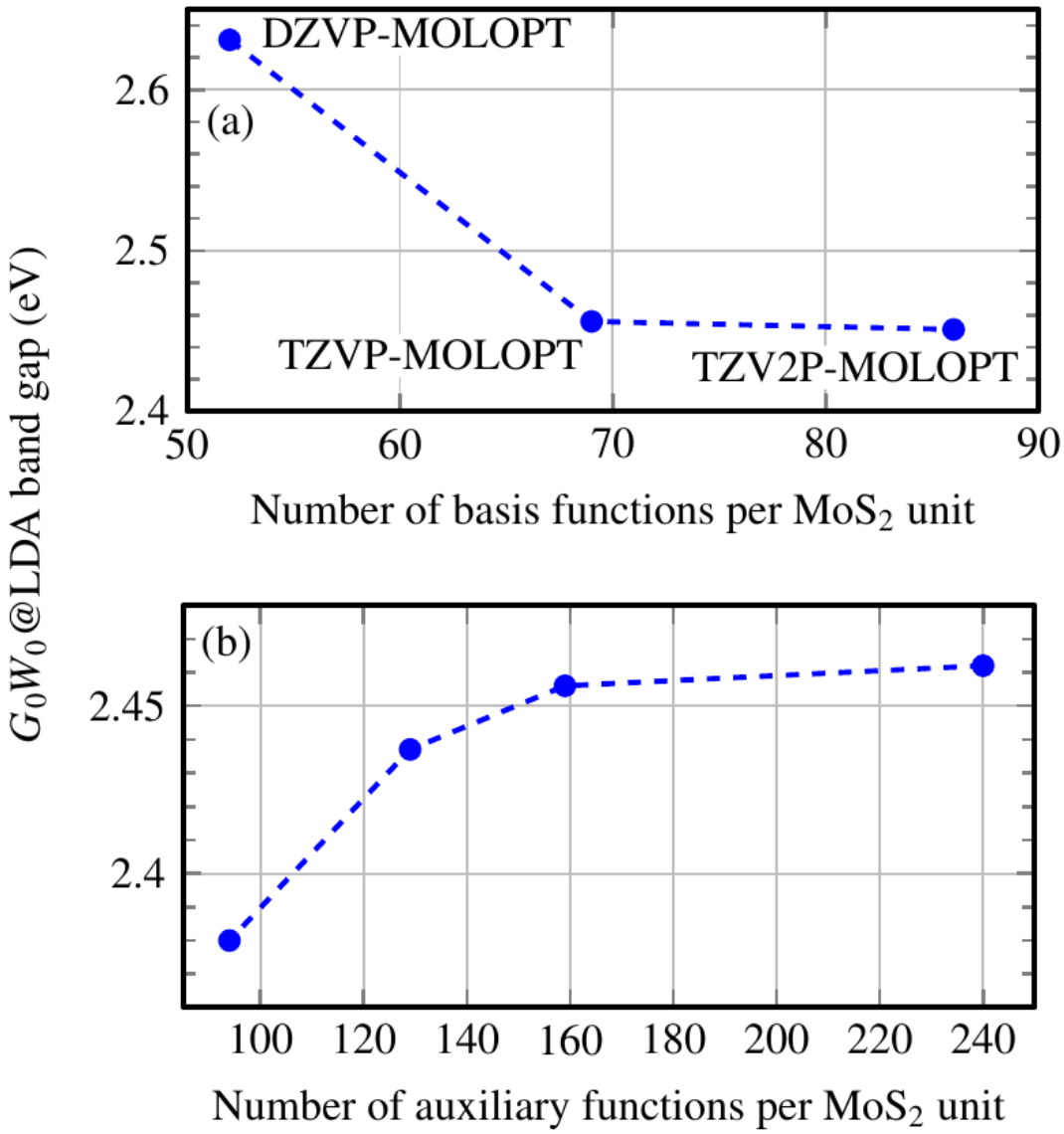
6x6 supercell



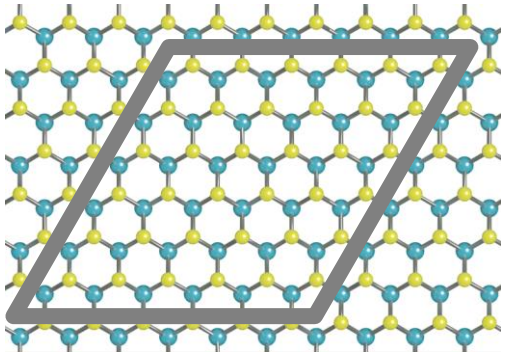
Reference value: G_0W_0@LDA: 2.45 eV

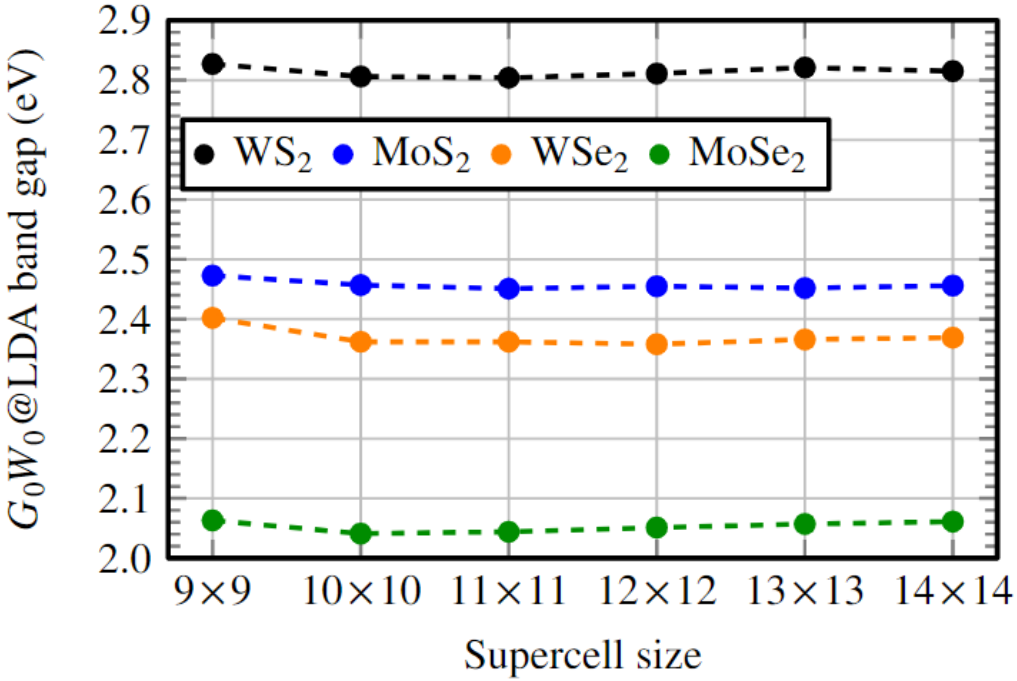
Qiu, da Jornada, Louie, PRB 93, 235435 (2016)

Basis set convergence (MoS_2)



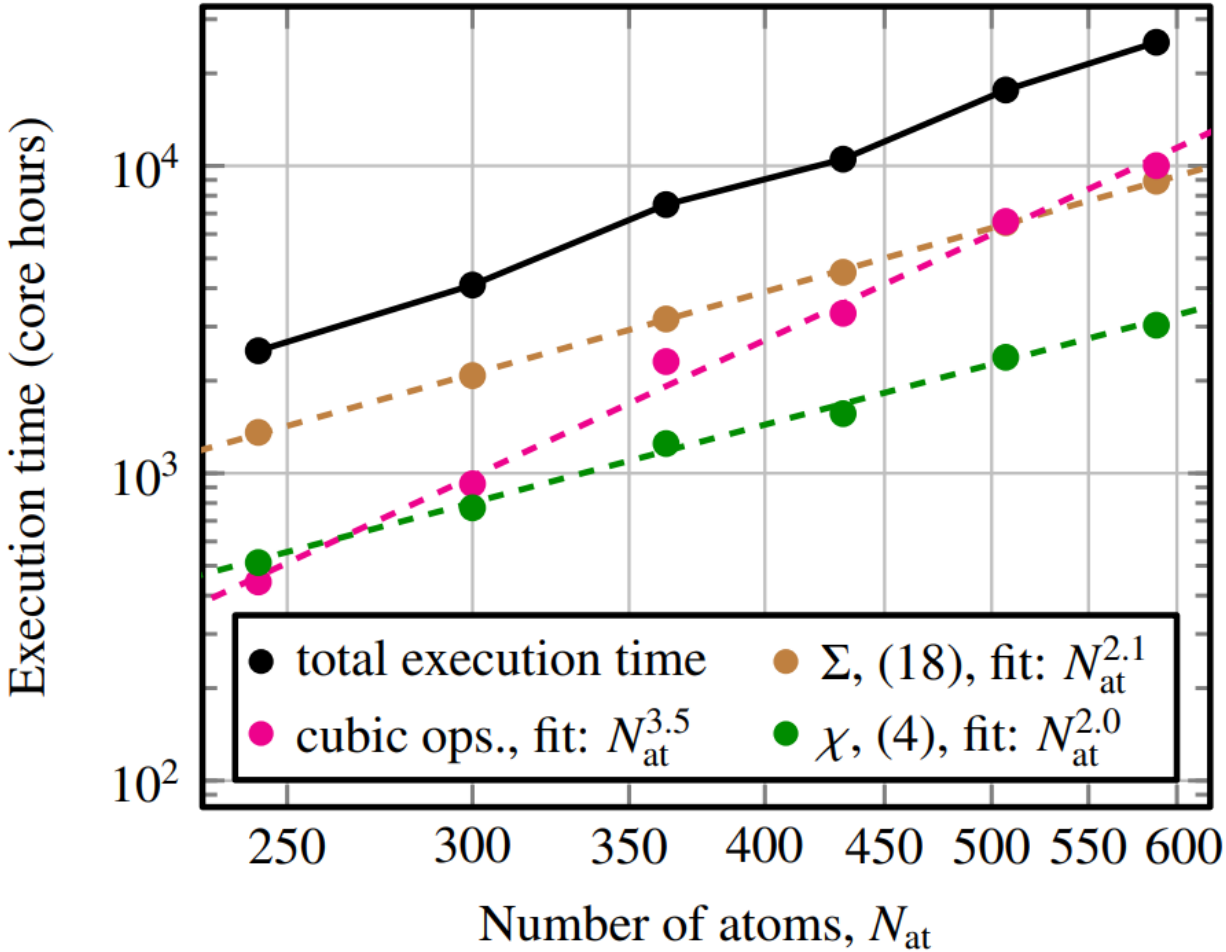
Benchmark calculations on MoS_2, $MoSe_2$, WS_2, WSe_2



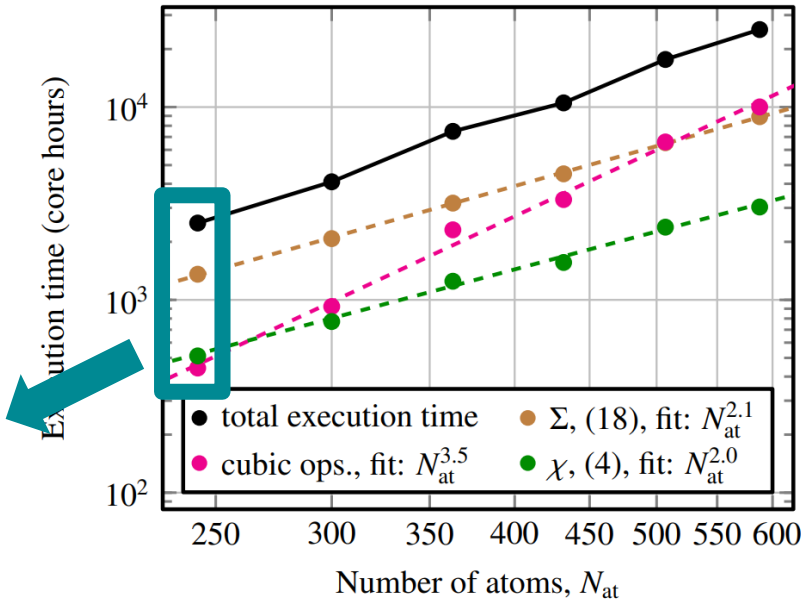


Software package	MoS_2	$MoSe_2$	WS_2	WSe_2
This work, (CP2K [51, 52])	2.47	2.07	2.81	2.37
GPAW [43] (SOC removed)	2.53	2.12	2.75	2.30
BerkeleyGW [40] (details in SI)	2.45	2.09	2.61	2.34
VASP [44]	2.50	2.06	2.70	2.34

Scaling


Execution time (core hours)
10^4
10^3
10^2
250
300
350
400
450
500
550
600
Number of atoms, N_{at}
total execution time
Σ, (18), fit: $N_{at}^{2.1}$
cubic ops., fit: $N_{at}^{3.5}$
χ, (4), fit: $N_{at}^{2.0}$

Scaling



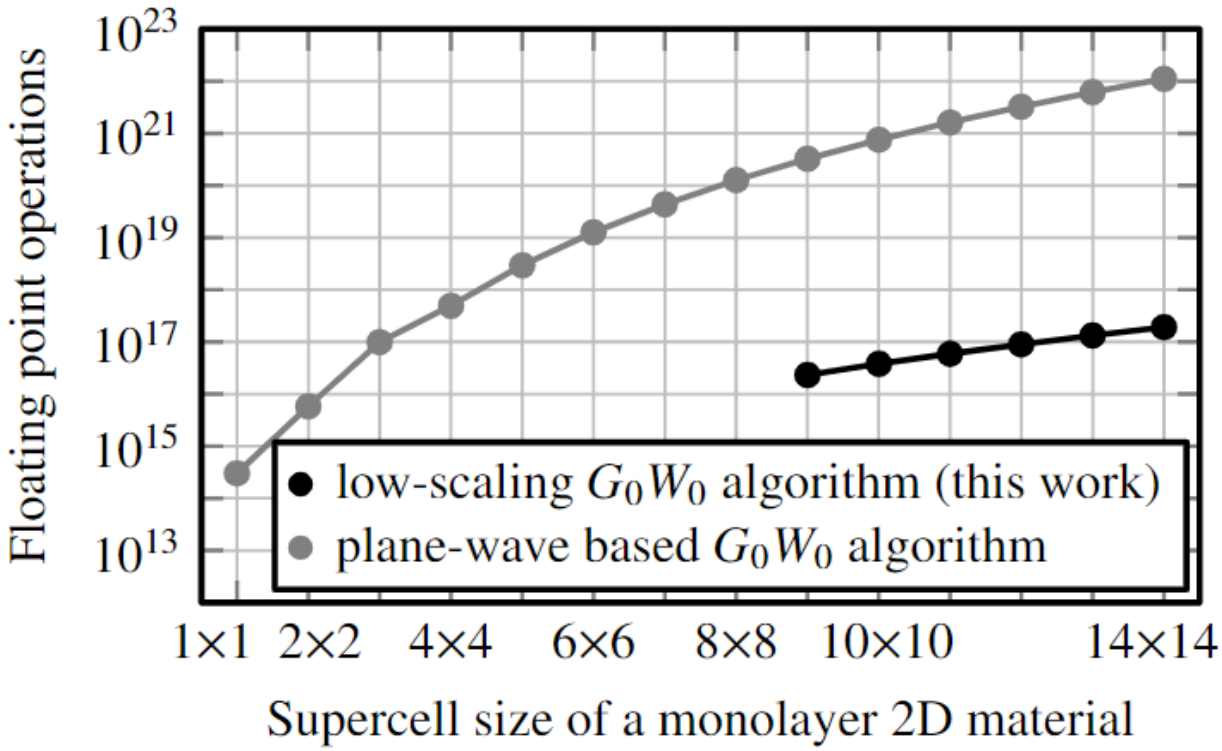
6 nodes (= 288 cores)
for 9 hours

$$\text{Comp. time} = \left(\frac{N_{at}}{243}\right)^2 T_{\text{quadr., 243 atoms}} + \left(\frac{N_{at}}{243}\right)^3 T_{\text{cubic, 243 atoms}}$$

	Number of atoms in a G_0W_0 calculation
100 000 core hours	1200
Maximum job size on Supermuc-NG:	4500

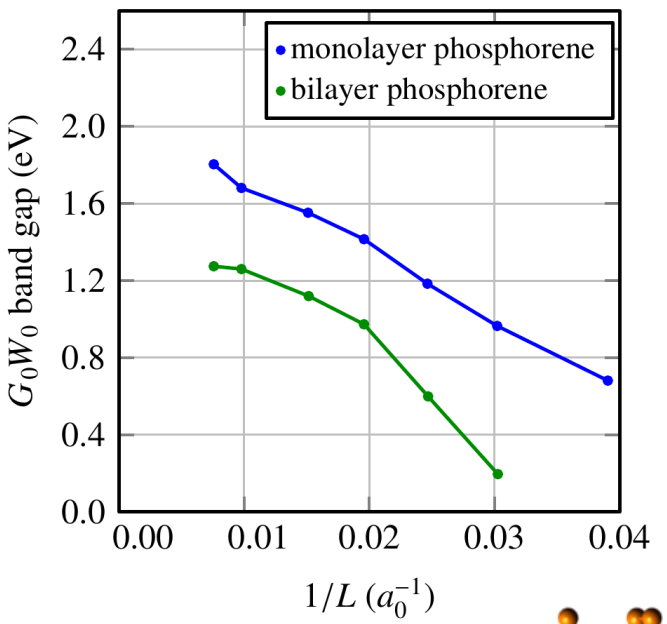
152000 cores for 24 hours = 3.6 Mio core hours

Computation time versus brute-force N^4 plane-wave *GW* algorithm

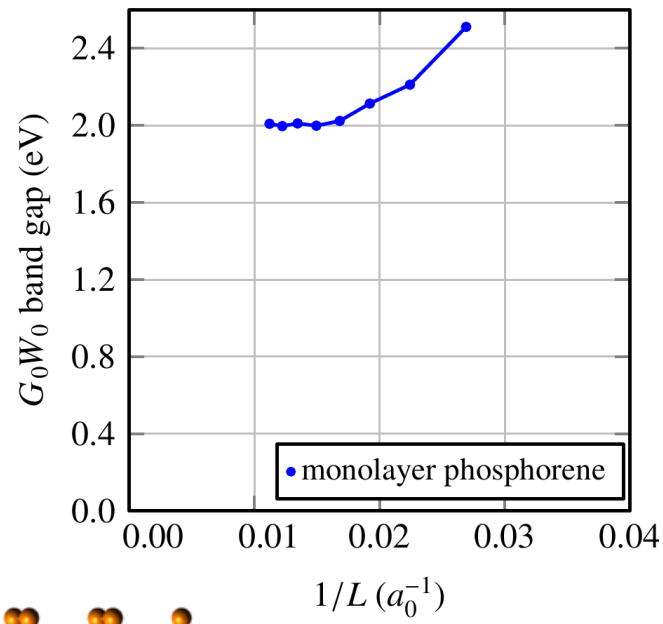


Size extrapolation of stochastic *GW*

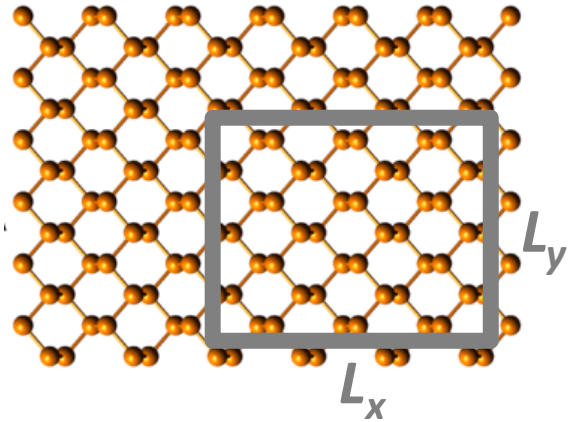
Stochastic *GW*



GW space-time with Gaussians



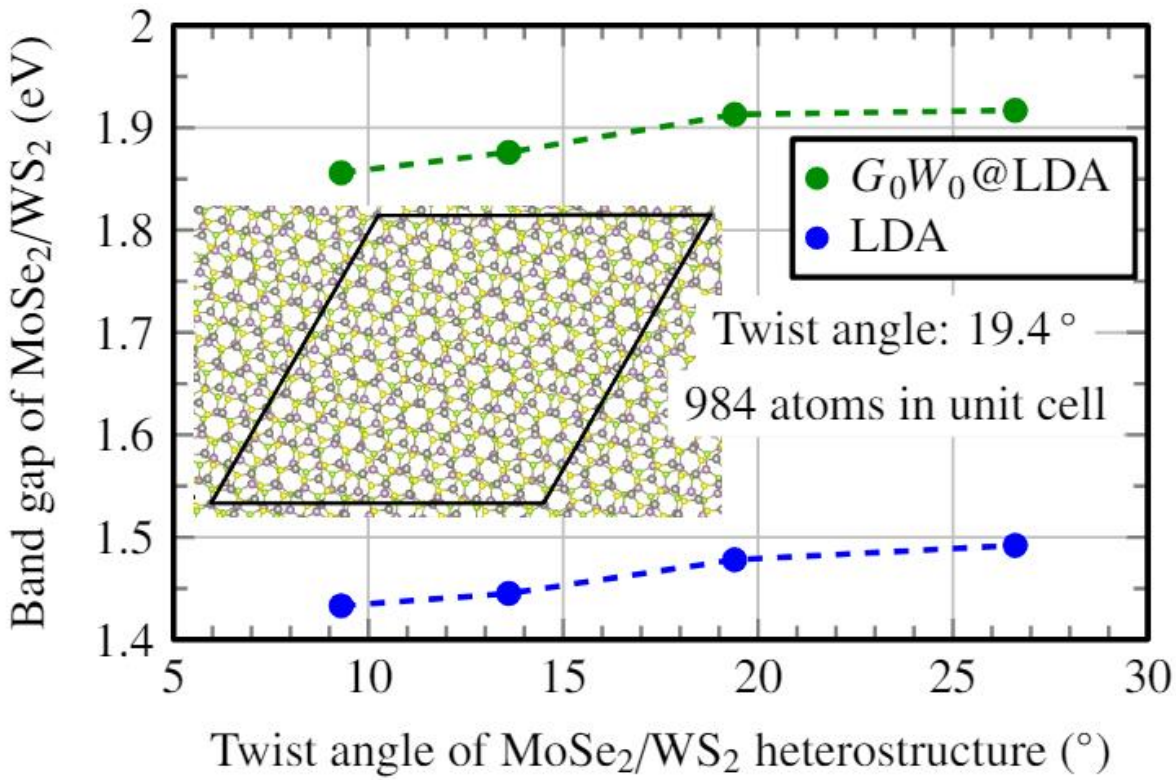
Brooks, Wenig, Taylor, Vlcek,
J. Phys. Condens. Matter **32**, 234001 (2020)



Graml, Zollner, Hernangomez, Faria, Wilhelm, arXiv:2306.16066 (2023)

$$L = (L_x L_y)^2$$

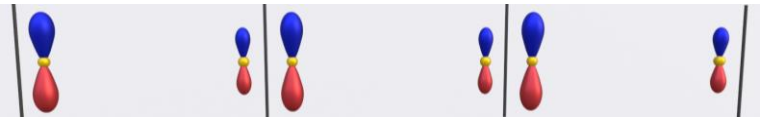
$MoSe_2$ / WS_2 heterostructure



984 atoms, 128 nodes (12 cores per node) for 42 hours

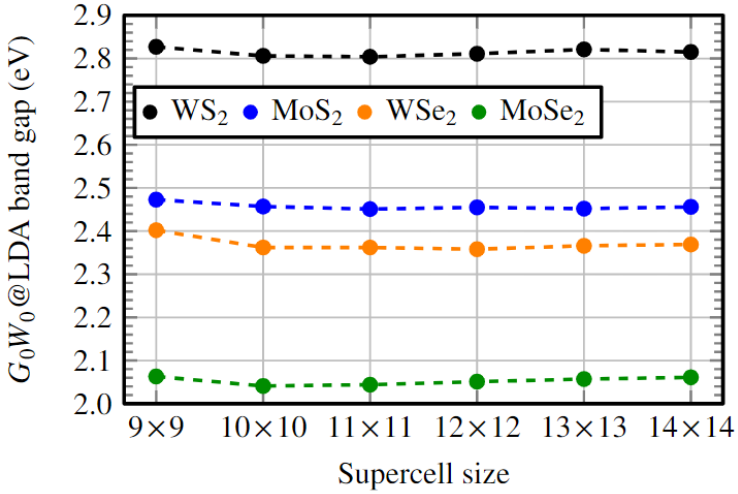
Summary

- **Real space ➡ k-points:** $\chi_{PQ}(i\omega, \mathbf{0}) = \begin{pmatrix} \alpha & \gamma \\ \gamma & \beta \end{pmatrix}$ ➡ $\chi_{PQ}(i\omega, \mathbf{q}) = \begin{pmatrix} \alpha & \gamma e^{-iq_x a} \\ \gamma e^{iq_x a} & \beta \end{pmatrix}$



- **Resulting *GW* algorithm converges in cell size & is numerically precise**

Software package	MoS_2	$MoSe_2$	WS_2	WSe_2
This work, (CP2K [51, 52])	2.47	2.07	2.81	2.37
GPAW [43] (SOC removed)	2.53	2.12	2.75	2.30
BerkeleyGW [40] (details in SI)	2.45	2.09	2.61	2.34
VASP [44]	2.50	2.06	2.70	2.34



- **Low computational cost, 9x9 cell (243 atoms) for 9 hours on 6 nodes**

M. Graml, K. Zollner, D. Hernangomez, P. E. Faria, J. Wilhelm, *arXiv:2306.16066 (2023)*

Emmy Noether-Programm
DFG Deutsche Forschungsgemeinschaft