University at Buffalo
The State University of New York

Quasiparticle and electron-hole excitations in solids: theory and computation (Part 2)

Peihong Zhang

Department of Physic, University at Buffalo

Outline

- Quasiparticle concept, quasiparticle band structure and band gap
- Quasiparticle equation, self-energy, and the GW approximation
- First-principles GW method
- Success and challenges of the GW method
- Acceleration techniques: Energy-integration
- GW calculations of 2D materials
- Electron-hole excitations and the Bethe-Salpeter equation

GW calculations for large systems? High-throughput GW calculations?

- (Forget about ZnO or CuCl) For a 2-atom MgO, we need about 1,000 conduction bands to converge the result

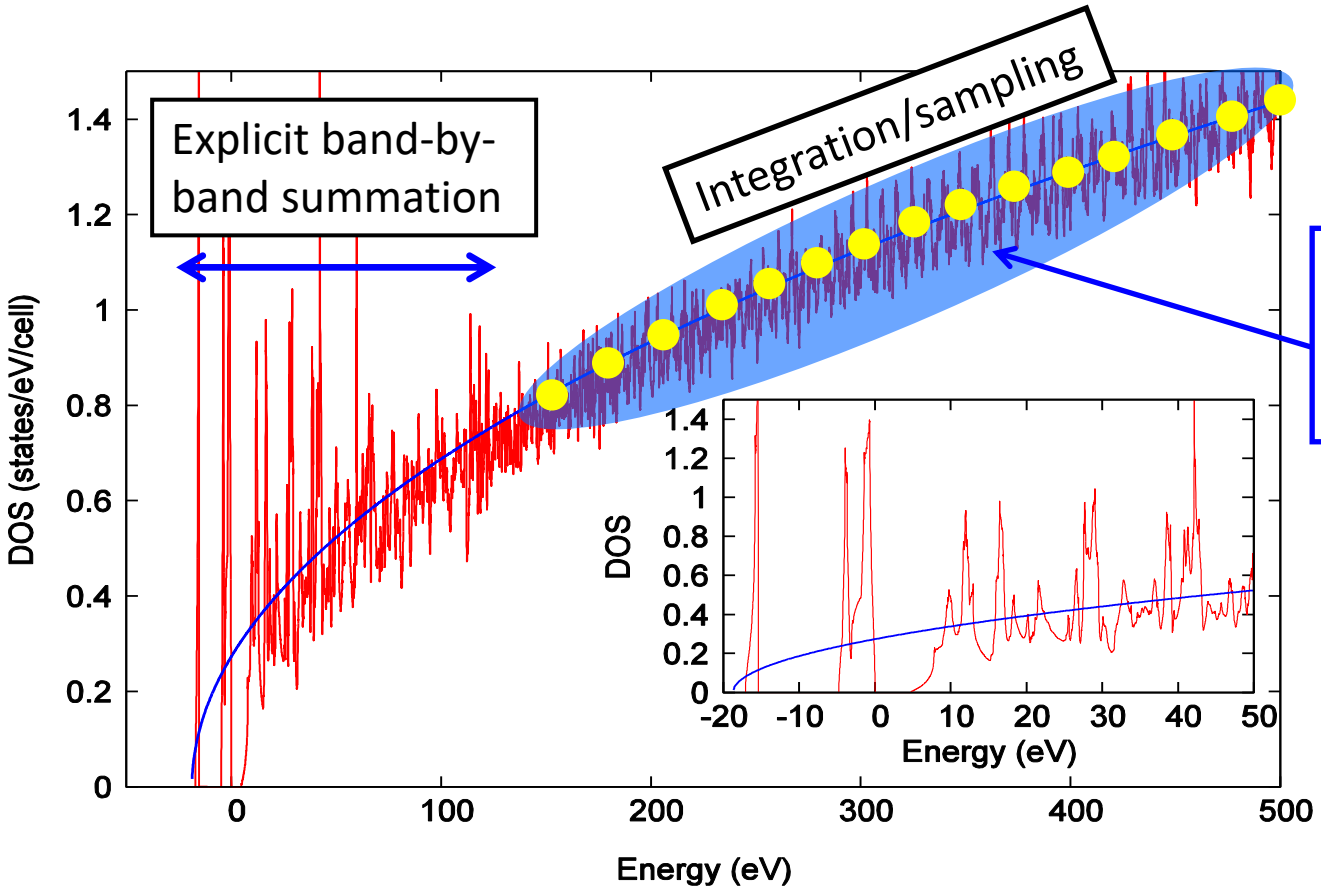
- Suppose we are interest in a system containing 200 atoms (e.g, a supercell containing a defect), we will need

$$1000 \times (200/2) = 100{,}000 \ \text{conduction bands}$$

to achieve the same level of convergence

- Not only calculating the wave functions is hard (if possible), but storing these wave functions is extremely problematic, not to mention the subsequent GW calculations

Speed up GW calculations for large systems?



DOS of MgO

Do we really need an explicit band-by-band summation?

Scientific Reports **6**, 36849 (2016).

Speed up GW calculation for large systems

- Contributions from high energy states are calculated by an energy integration:

Low-energy states

$$\chi^0_{\mathbf{G},\mathbf{G}'}(\mathbf{q},\omega=0) \approx \sum_{c}^{N_0}\left[\sum_{v\mathbf{k}} \frac{M_{vc}(\mathbf{k},\mathbf{q},\mathbf{G})M^*_{vc}(\mathbf{k},\mathbf{q},\mathbf{G}')}{E_{v\mathbf{k}+\mathbf{q}}-E_{c\mathbf{k}}}\right]$$

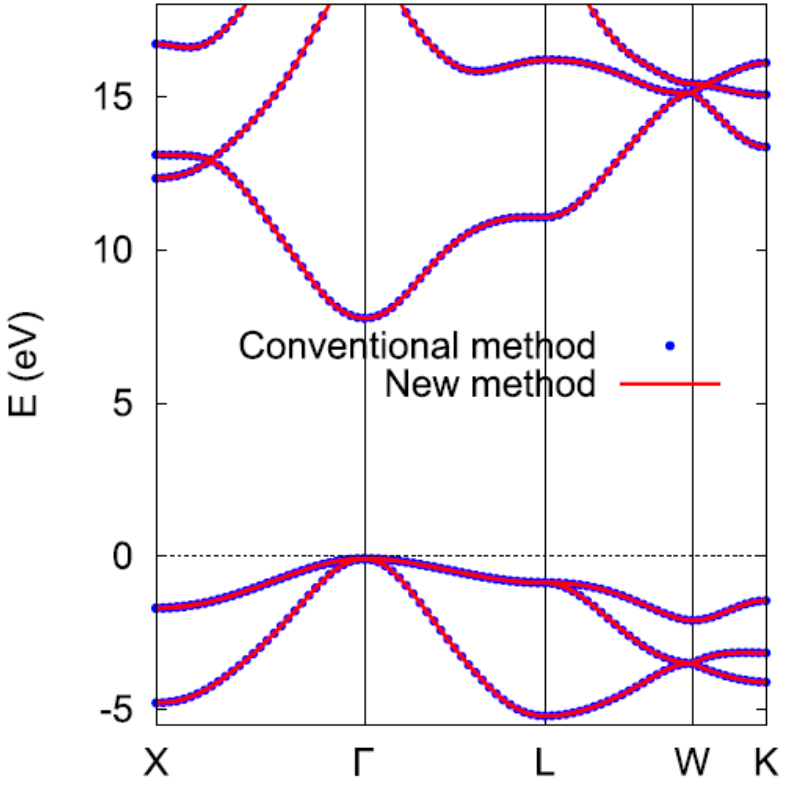
$$+\int_{E_0}\sum_{v\mathbf{k}} \frac{M_{vE'}(\mathbf{k},\mathbf{q},\mathbf{G})M^*_{vE'}(\mathbf{k},\mathbf{q},\mathbf{G}')}{E_{v\mathbf{k}+\mathbf{q}}-E'} g(E')\mathrm{d}E',$$

High-energy states

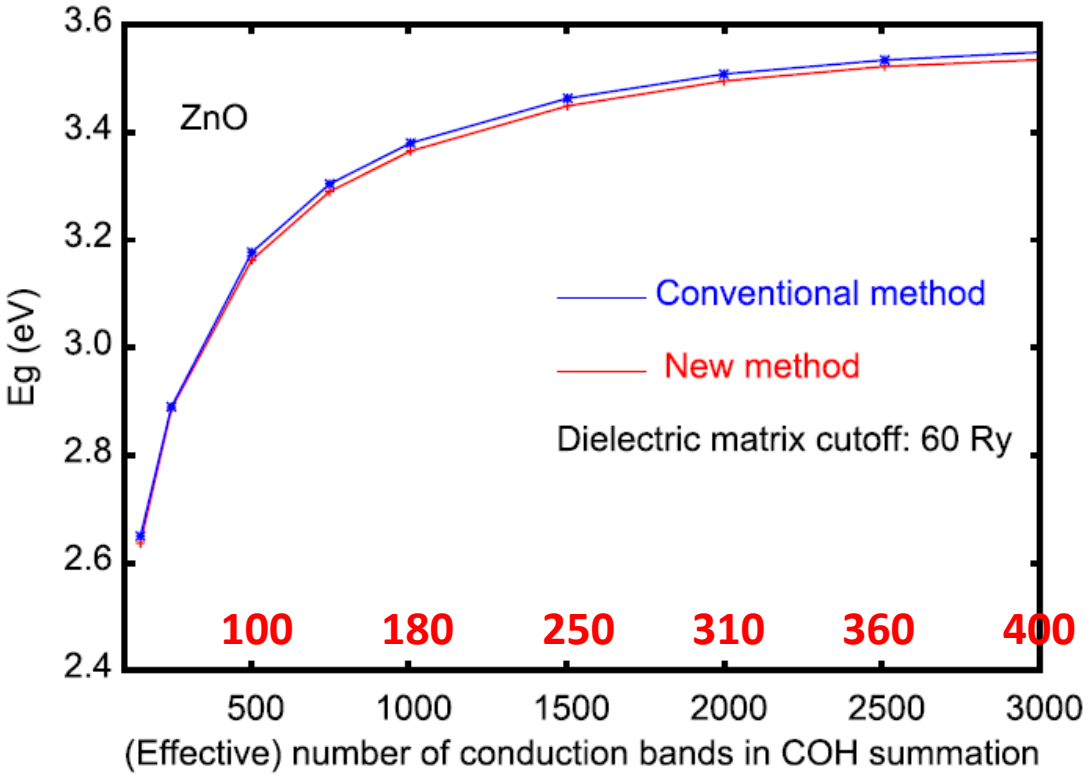
$g(E)=\frac{\Omega}{\pi^2}\sqrt{2(E-V_{xc}(0))}$: free-electron-like DOS

Scientific Reports **6**, 36849 (2016).

Accuracy/performance of the new method



QP band structure of MgO



QP band gap of ZnO

Scientific Reports **6**, 36849 (2016).

Large scale GW calculations

- **Band gap of MgO supercells**

# of atoms	New method		Conventional method		Speed-up factor	ΔE_g
	$N_0 + N_E$	E_g	N_c	E_g		
2	170	7.86	1,000		5.9	0.00
16	320	7.84	8,000		25.0	−0.02
64	920	7.89	32,000	7.86	34.8	+0.03
128	1060	7.83	64,000		60.4	−0.03
256	1580	7.86	128,000		81.0	0.00

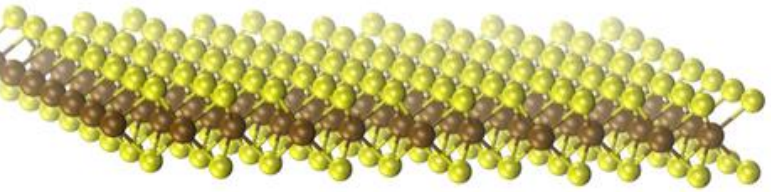
- A speed-up factor of nearly two orders of magnitude is achieved
- Numerical error: less than $\pm$0.05 eV

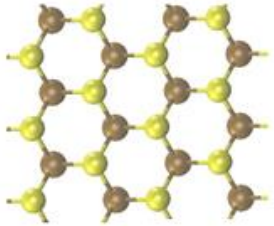
Scientific Reports **6**, 36849 (2016).

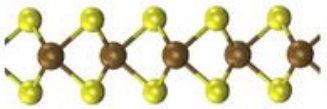
Outline

- Quasiparticle concept, quasiparticle band structure and band gap
- Quasiparticle equation, self-energy, and the GW approximation
- First-principles GW method
- Success and challenges of the GW method
- Acceleration techniques
- GW calculations of 2D materials
- Electron-hole excitations and the Bethe-Salpeter equation

What about 2D materials?







MoS_2

- Reported GW band gap of MoS_2:
  2.41 ~ 2.84 eV
  (without including spin-orbit interactions)

- Methods: G^0W^0, G^1W^0, self-consistent GW, etc

- Parameters used:
  - Number of conduction bands (3-atom unit cell): 96 to 10,000
  - k-point sampling: 6x6x1 to 24x24x1
  - Who do you trust?

PRB 85, 205302 (2012); PRB 86, 115409 (2012);
PRB 87, 155304 (2013); PRB 88, 045412 (2013);
PRL 115, 119901 (2013)

GW calculations for 2D materials: Challenges

- Need to avoid fictitious interlayer interaction

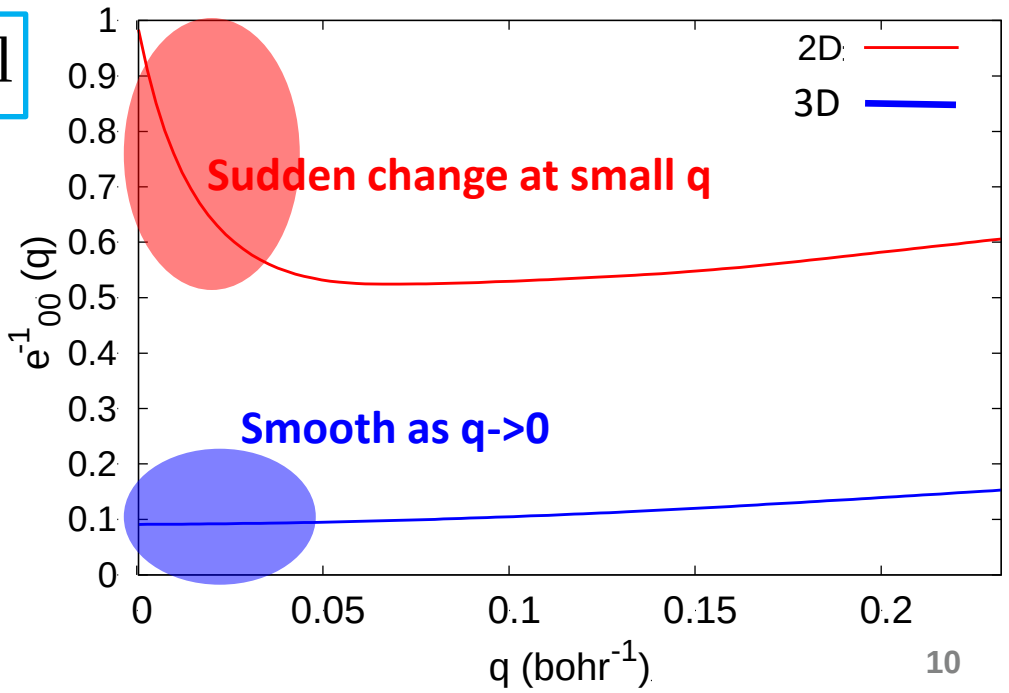
  - Long-range Coulomb interaction means slow convergence with respect to the interlayer distance; a large vacuum layer is needed even if truncated Coulomb interaction is used

  - A large number of conduction bands (Nc) is needed

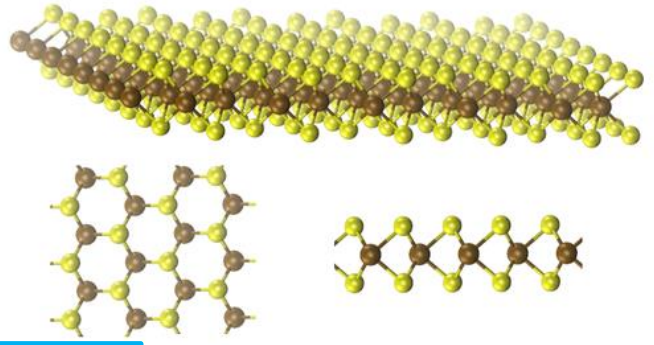
$$N_c \propto \text{volume of the unit cell}$$

- Asymptotic behavior of the dielectric function at small q

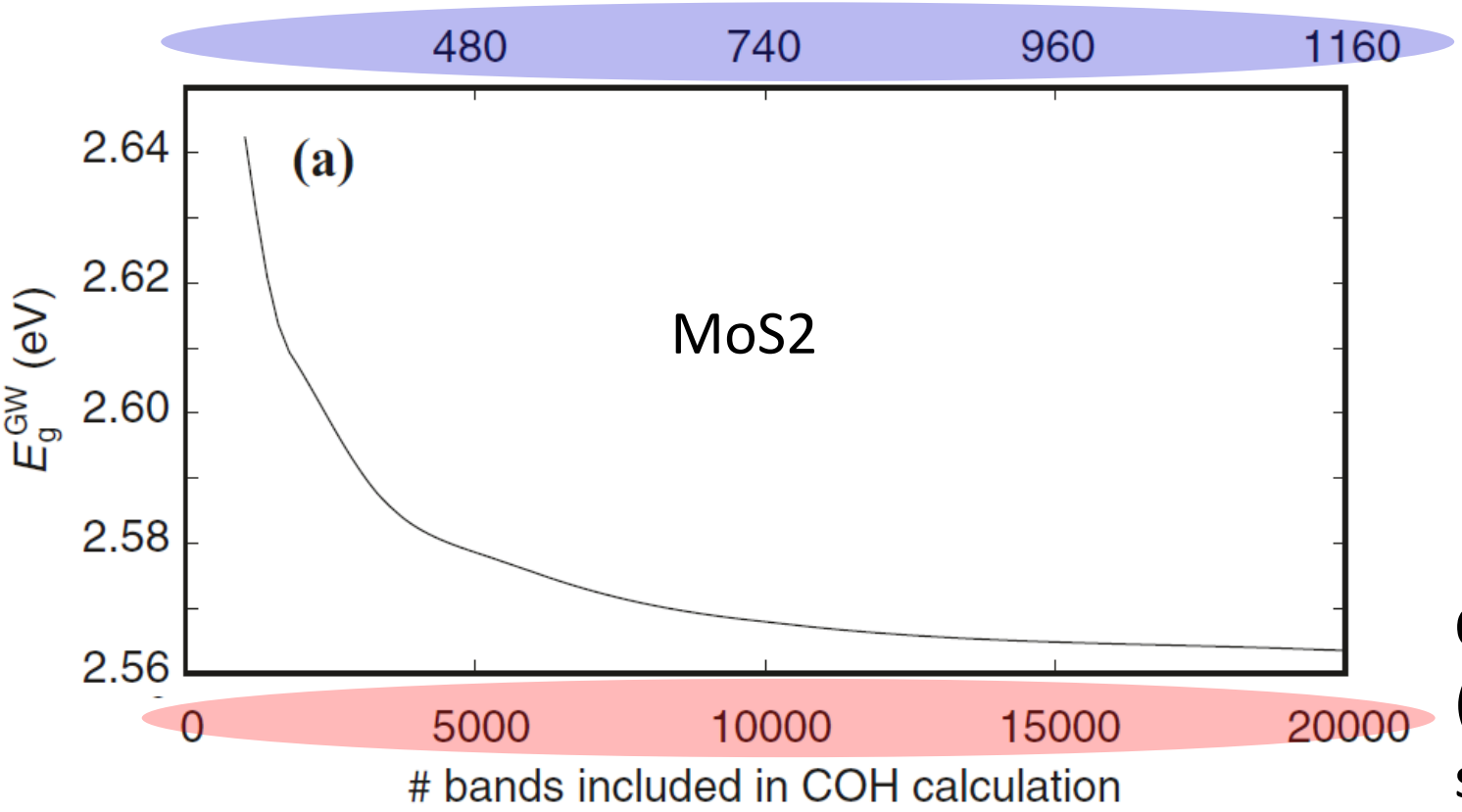
- An extremely dense k-grid is needed for 2D GW calculations



2D materials



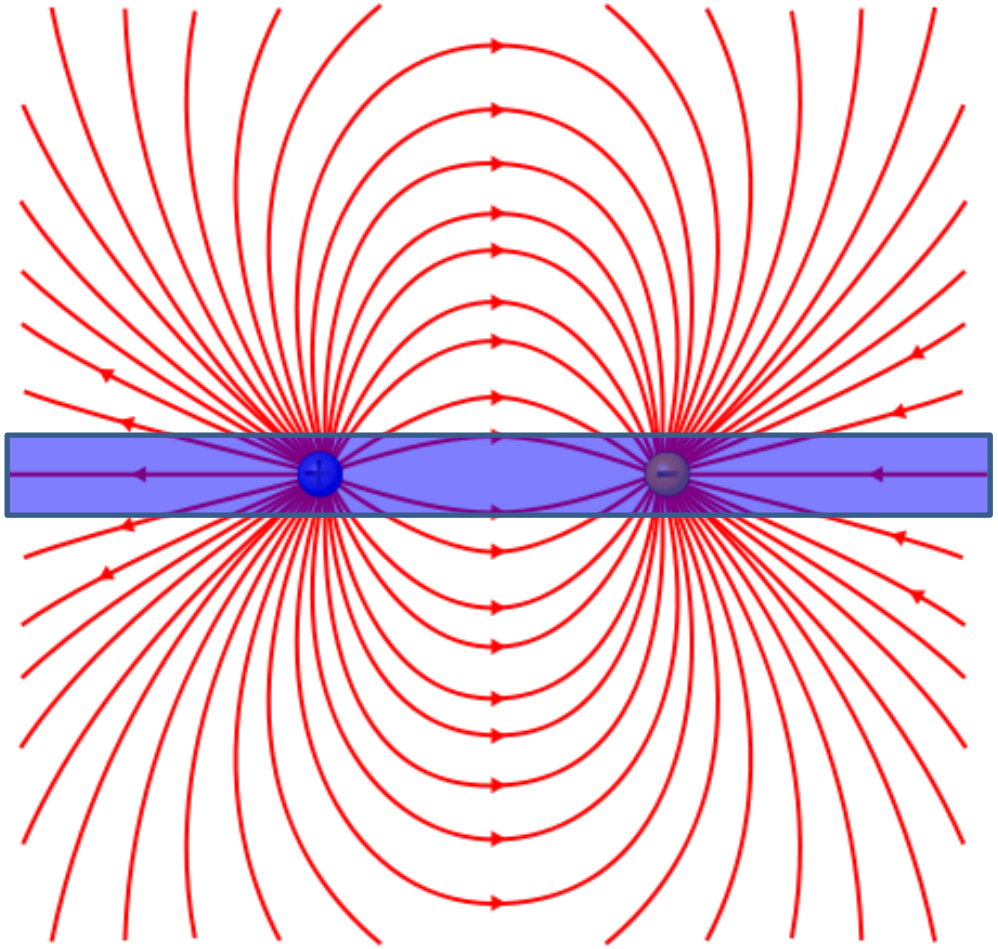
GW band gap of MoS2 as a function of number of bands included in the calculation



New method (Energy integration)

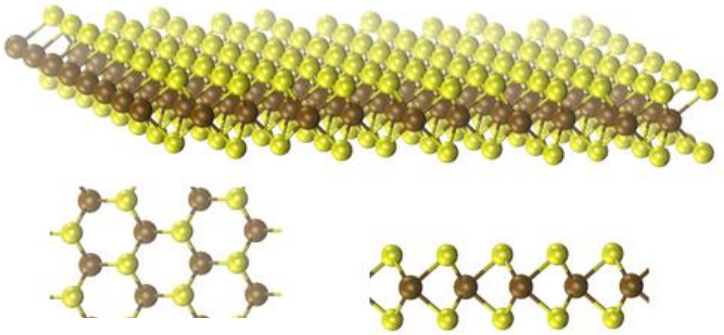
Old method (band-by-band summation)

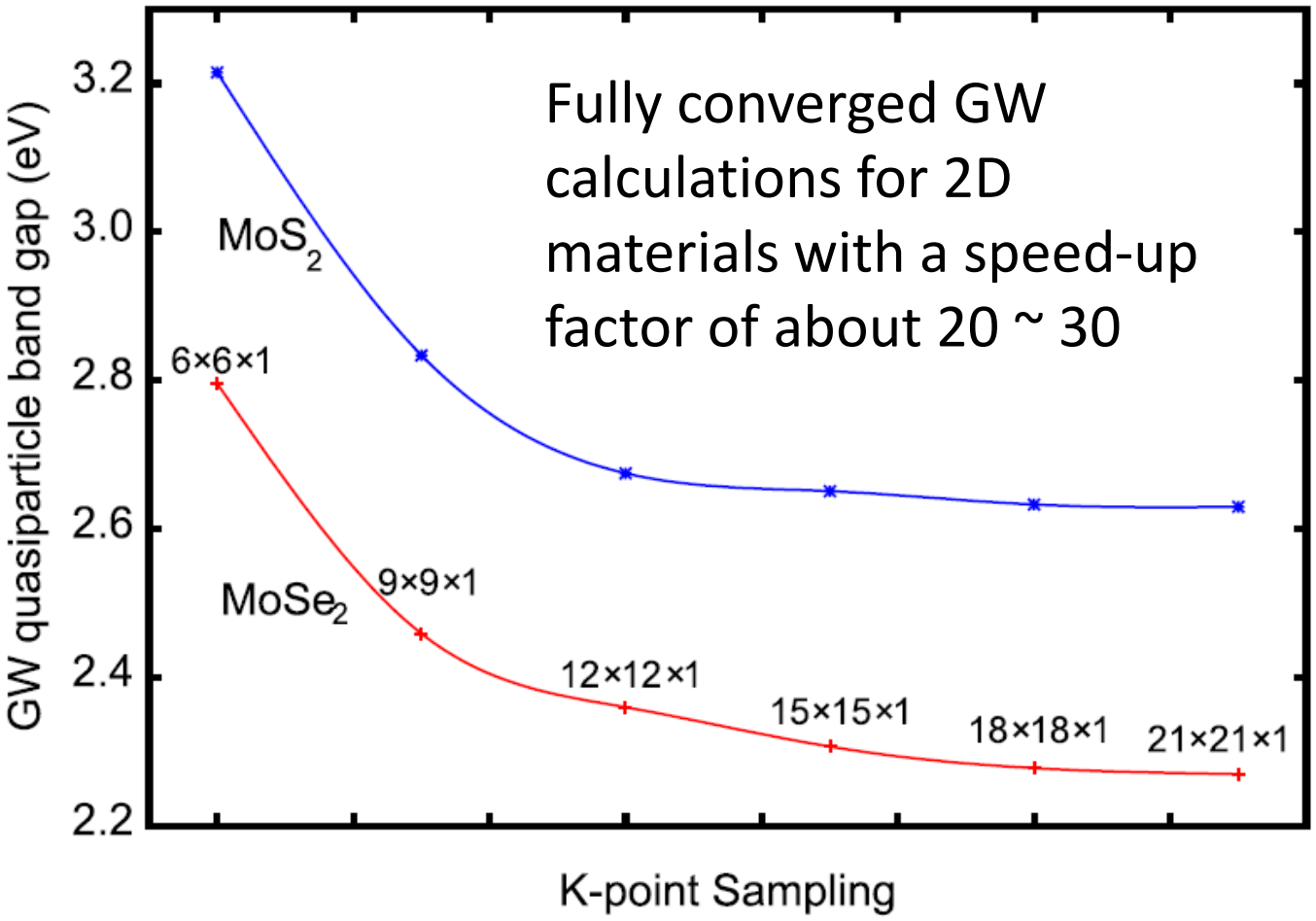
2D dielectric screening



Ineffective long wave length screening

2D materials





The GW band gap converges extremely slowly! Why?

GW calculations for 2D materials: Challenges

- For simple 2D materials such as MoS_2 and $MoSe_2$, we need **~10,000** bands and at least a **24x24x1** k-grid to properly converge the quasiparticle properties using conventional GW methods
- Scaling of the computational cost of GW calculations with respect to number of k points: $O(N_k^2)$

  - Compared with a calculation using a **6x6x1** k-grid, a calculation using a **24x24x1** k-grid is

$$24^4 / 6^4 = 256 \text{ times more expensive}$$

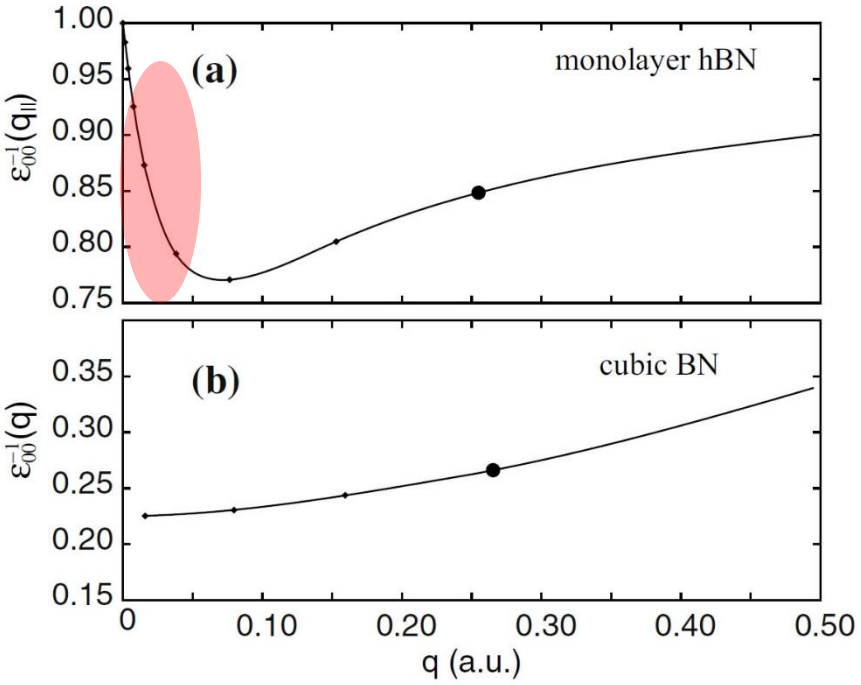
- Also, the number of bands will scale linearly with the system size, and the computation cost scales as $O(N_{atom}^4)$
- It is nearly impossible to carry out fully converged GW calculations for complex 2D materials using current methods

GW calculations for 2D materials: Challenges

- In the conventional GW approach, the integration of the self-energy within the Brillouin zone (BZ) is carried out by a summation on a uniform k-grid:

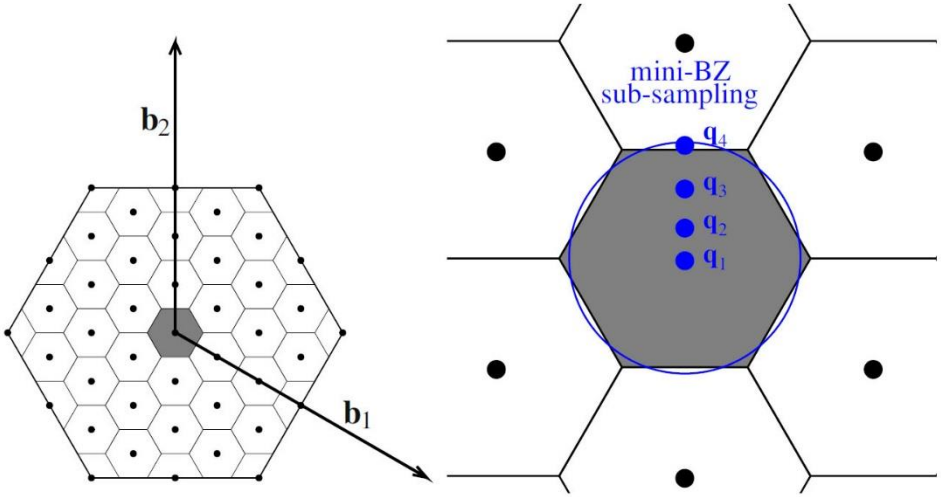
$$\Sigma_{n\vec{k}}^{QP} = \frac{1}{\Omega_{BZ}} \int_{\Omega_{BZ}} \Sigma_{n\vec{k}}^{QP}(\vec{q}) d\vec{q} = \sum_{q \in BZ} W_{\vec{q}} \Sigma_{n\vec{k}}^{QP}(\vec{q})$$

- This approach is highly inefficient for 2D materials because of the sharp change in the dielectric function near q = 0.



Mini-BZ sub-sampling fitting and analytical integration

- Our approach: Sub-sampling the mini-BZ C_q near q = 0

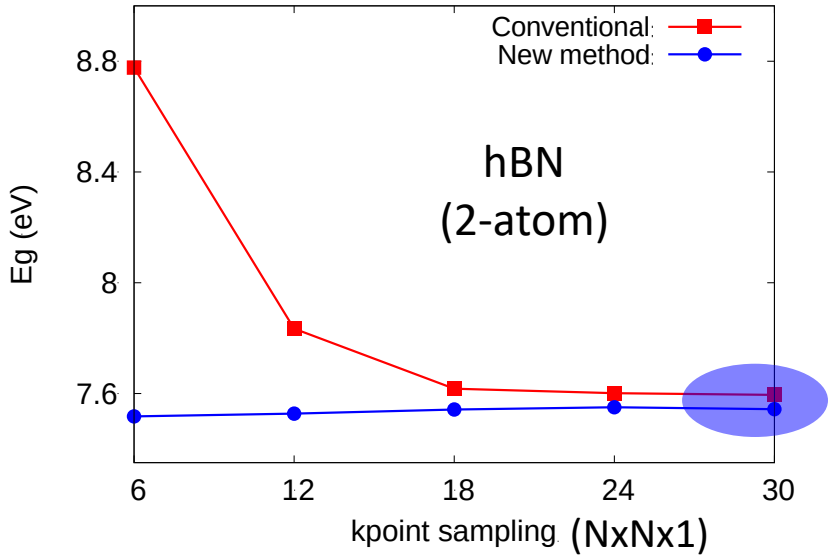


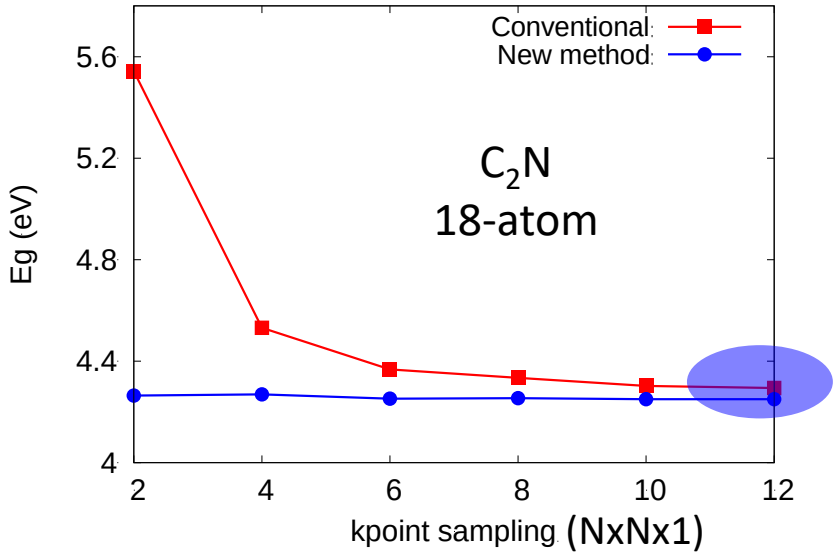
- The BZ summation of the electron self-energy is separated into two parts: conventional summation of all k points other the Γ point, and an analytical integration in the mini-BZ

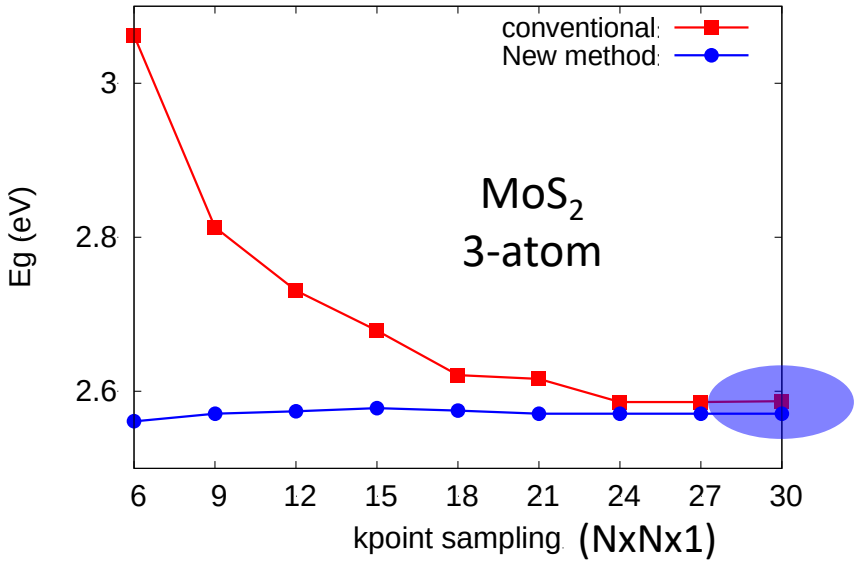
$$\langle nk|\Sigma(\omega)|nk\rangle \approx \frac{1}{A}\int_{mBZ} \boldsymbol{\Sigma}(\boldsymbol{q},\omega)dA + \sum_{\boldsymbol{q}\neq\boldsymbol{0}} \boldsymbol{\Sigma}(\boldsymbol{q},\omega)$$

2D Materials **6**, 015018 (2019). Nanoscale, 11, 3993 (2019).
NPJ Comput. Mater. 6, 118 (2020).

Performance of the new Method

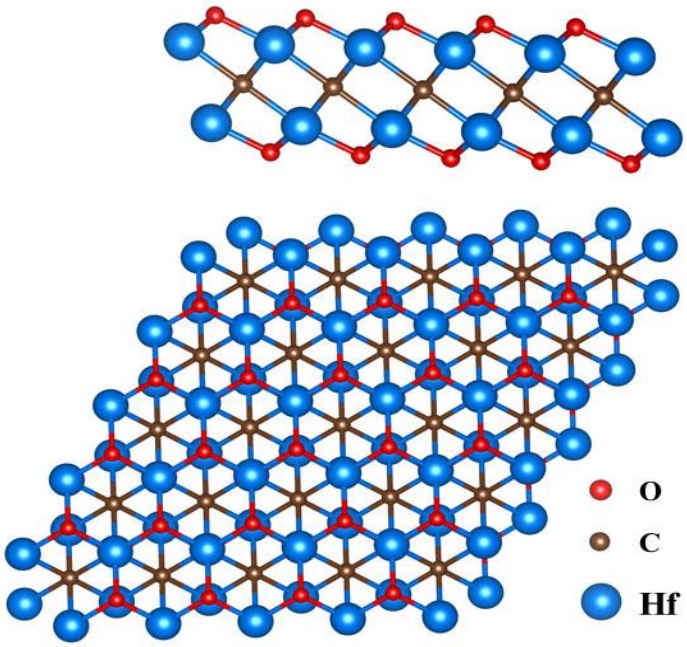




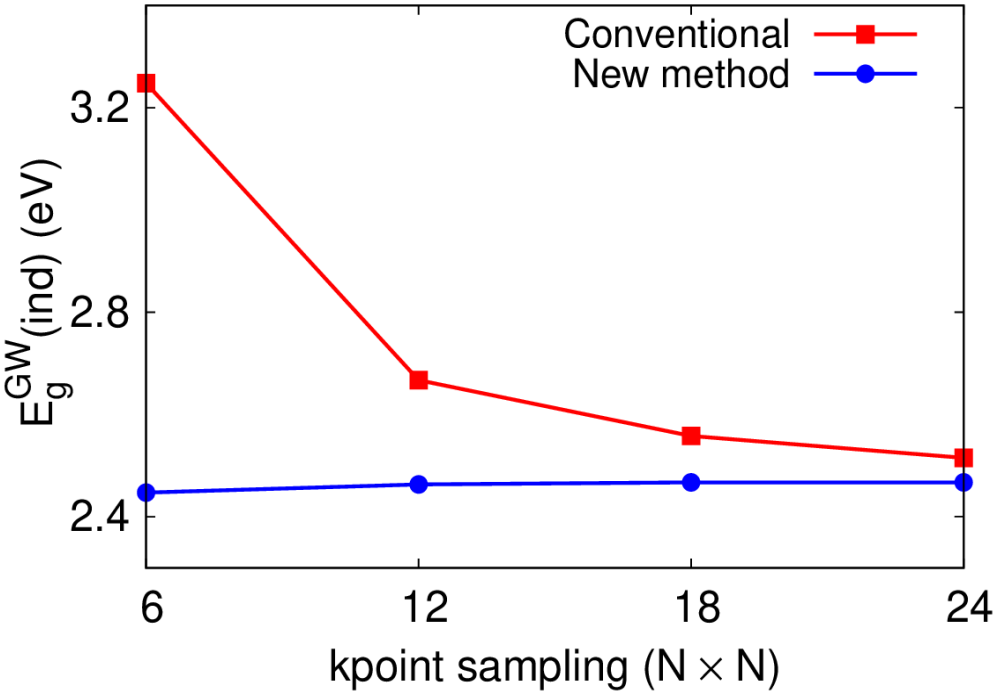


Note that the results calculated using the conventional k-point sampling approach do not seem to converge even at extremely high k-density

Performance of the new method



Nanoscale, **11**, 3993 (2019).



GW band gap of single-layer Hf_2CO_2 MXene as a function of k-point sampling density.

Our new methods result in a combined speed-up factor of ~ 1,000 times for GW calculations of complex 2D materials.

Outline

- Quasiparticle concept, quasiparticle band structure and band gap
- Quasiparticle equation, self-energy, and GW approximation
- First-principle GW method
- Success and challenges of the GW method
- Acceleration techniques
- GW calculations of 2D materials
- Electron-hole excitations and the Bethe-Salpeter equation

Optical absorption

- Optical absorption spectrum of solids is typically described by the imaginary part of the (frequency-dependent) macroscopic dielectric function, i.e., $\varepsilon_2(\omega)$
- Various other optical properties can be calculated using both the real and the imaginary parts of the dielectric function. For example, the absorption coefficient α:

$$\alpha(\omega) = \frac{\sqrt{2}\omega}{c}[\sqrt{\varepsilon_1^2(\omega) + \varepsilon_2^2(\omega)} - \varepsilon_1(\omega)]^{1/2}$$

- On a single particle level, one have

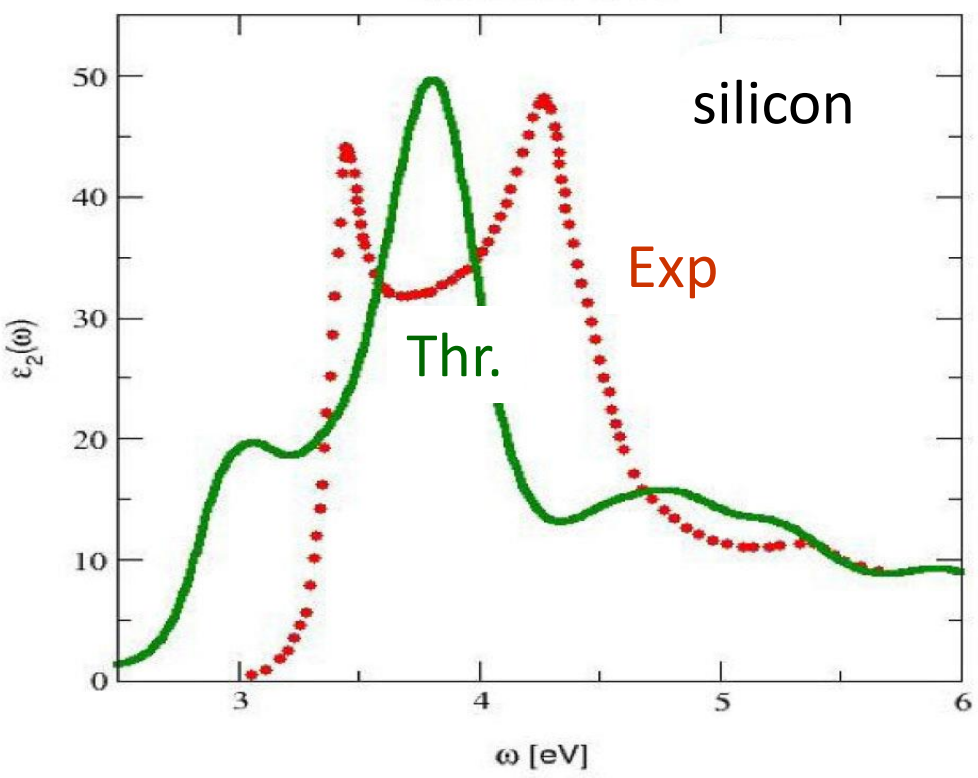
$$\varepsilon_2(\omega) = \frac{16\pi^2 e^2}{\omega^2} \sum_{c,v,\vec{k}} |\vec{\lambda} \cdot < v\vec{k} \,|\, \vec{v} \,|\, c\vec{k} >|^2 \; \delta(\omega - (\varepsilon_{c\vec{k}} - \varepsilon_{v\vec{k}}))$$

where $\vec{\lambda}$ is the polarization vector of light, and $\vec{v}$ is the velocity operator, and $|\, v\vec{k} >, \text{and} \,|\, c\vec{k} >$ are the valence and conduction states

Optical absorption: single particle calculation

$$\varepsilon_2(\omega) = \frac{16\pi^2 e^2}{\omega^2} \sum_{c,v,\vec{k}} |\, \vec{\lambda} \cdot < v\vec{k} \,|\, \vec{v} \,|\, c\vec{k} > |^2 \; \delta(\omega - (\varepsilon_{c\vec{k}} - \varepsilon_{v\vec{k}}))$$

- How well does theory work?

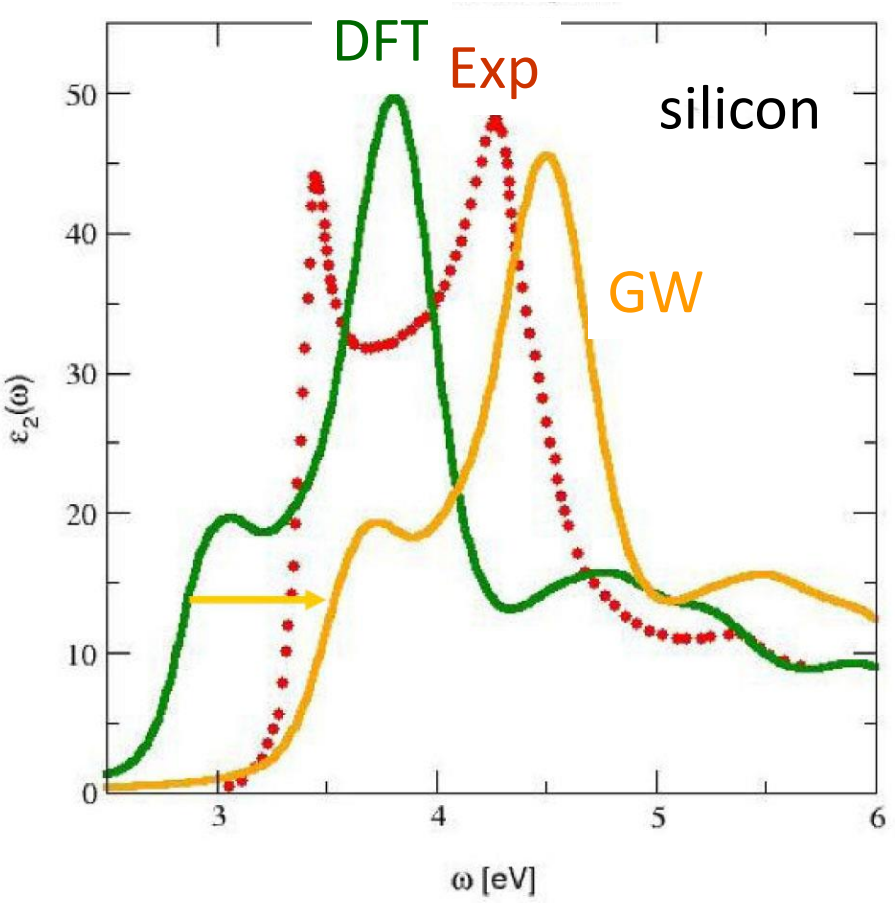


Pretty bad!

Optical absorption: single particle calculation

$$\varepsilon_2(\omega) = \frac{16\pi^2 e^2}{\omega^2} \sum_{c,v,\vec{k}} |\vec{\lambda} \cdot < v\vec{k} \,|\, \vec{v} \,|\, c\vec{k} >|^2 \; \delta(\omega - (\varepsilon_{c\vec{k}} - \varepsilon_{v\vec{k}}))$$

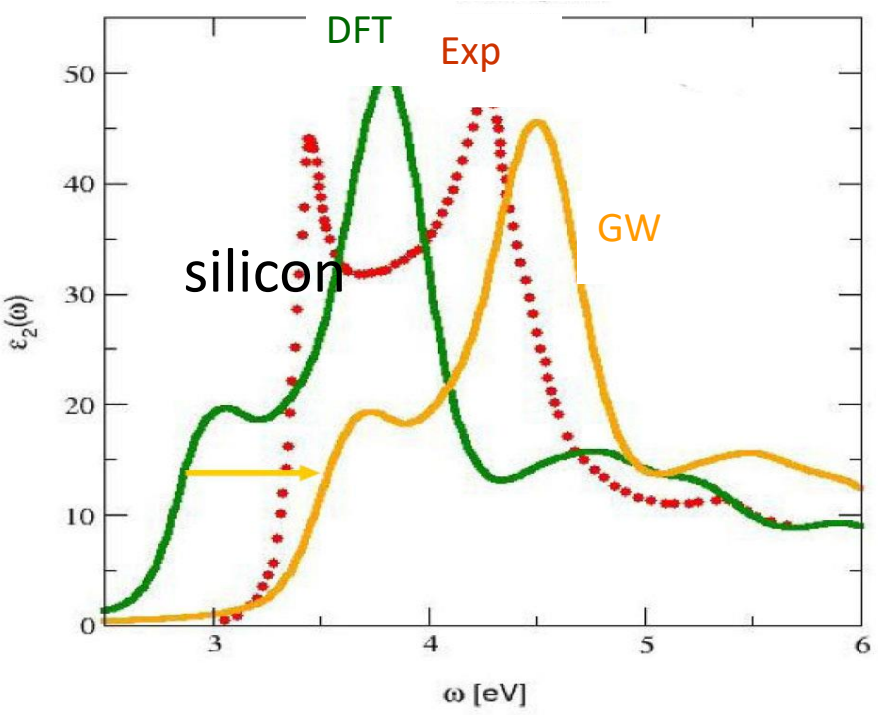
- What goes wrong?
- If calculated using DFT results, the KS band gap could be way off. This can be resolved by including the GW self-energy correction.
- Still NOT right!



Optical absorption: single particle calculation

$$\varepsilon_2(\omega) = \frac{16\pi^2 e^2}{\omega^2} \sum_{c,v,\vec{k}} |\vec{\lambda} \cdot < v\vec{k} \,|\, \vec{v} \,|\, c\vec{k} >|^2 \; \delta(\omega - (\varepsilon_{c\vec{k}} - \varepsilon_{v\vec{k}}))$$

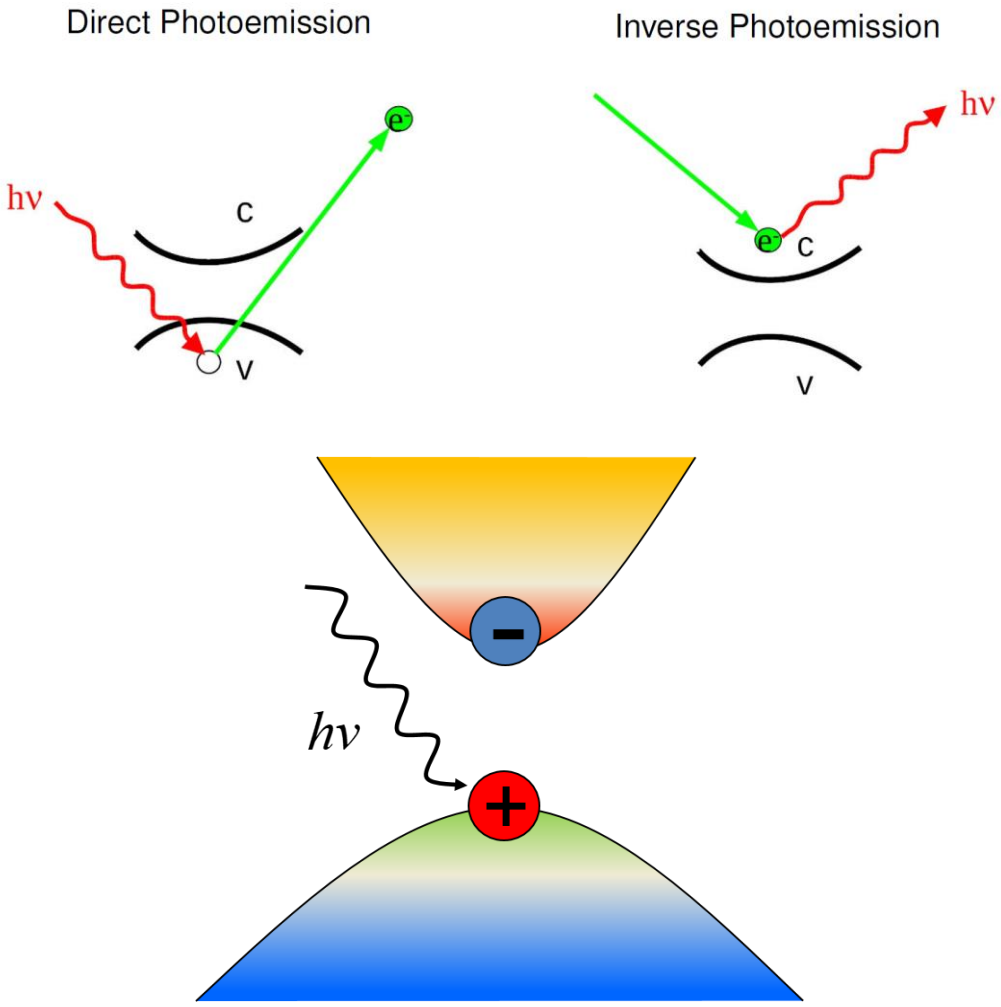
- What is still missing?



- The excited electron and the hole it leaves behind have interaction!

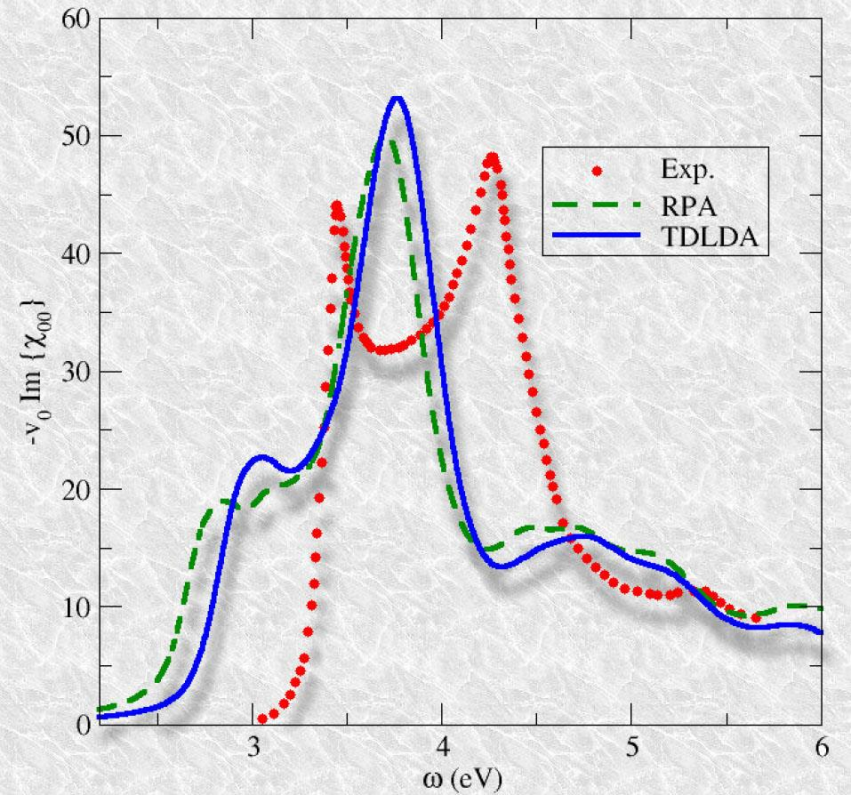
Electron-hole excitations

- The GW theory can describe ONE-particle excitations well
- However, optical excitations actually involve TWO particles: an electron and a hole
- Interaction between the electron and hole must be included in the calculation



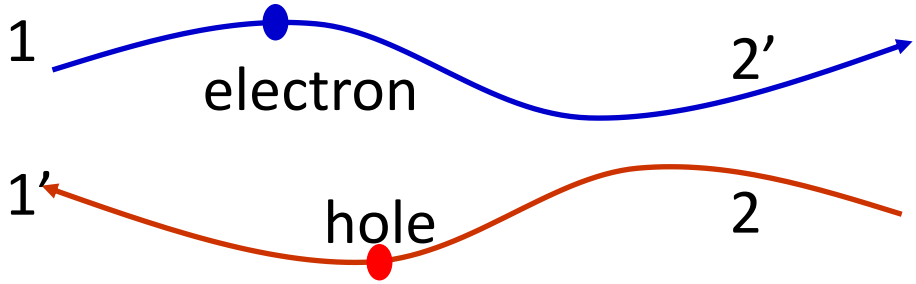
What about TDDFT?

- TDDFT seems to work better for localized/isolated systems with highly non-uniform charge densities, e.g., atoms and molecules
- Unless specialized functionals and kernels are used, TDDFT does not give accurate optical absorption spectra for solids

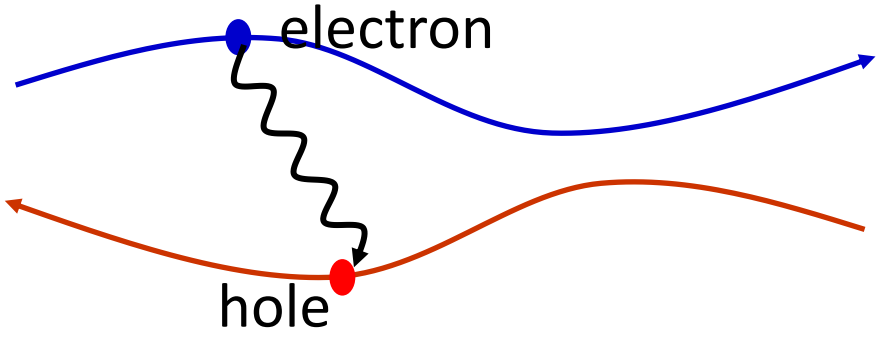


Electron-hole excitations and the pair correlation function

- The propagation of an electron-hole pair is described by the two-particle correlation function L



- Non-interacting pair correlation function: $L_0(12;1'2') = iG(12')G(21')$ (indices 1,2,3,4 stand for space, time, and spin variables)
- With interaction,



$L(12;1'2')$

Explicit expression for the pair correlation function

- The pair correlation function can be written explicitly if we know the (interacting or non-interacting) pair states:

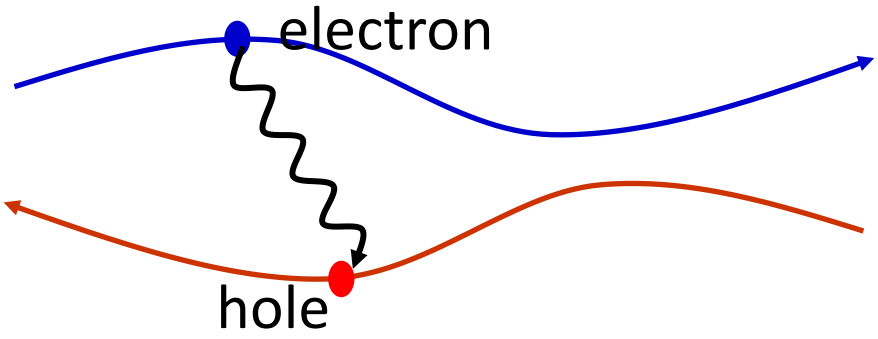
$$L_0(12;1'2') = i\sum_{vc} \frac{\varphi_c(\vec{r}_1)\varphi_v^*(\vec{r}_1')\varphi_v(\vec{r}_2)\varphi_c^*(\vec{r}_2')}{\omega-(\varepsilon_c-\varepsilon_v)} - \frac{\varphi_v(\vec{r}_1)\varphi_c^*(\vec{r}_1')\varphi_c(\vec{r}_2)\varphi_v^*(\vec{r}_2')}{\omega+(\varepsilon_c-\varepsilon_v)}$$

$$L(12;1'2') = i\sum_{S} \frac{\chi_S(\vec{r}_1,\vec{r}_1')\chi_S^*(\vec{r}_2',\vec{r}_2)}{\omega-\Omega_S} - \frac{\chi_S(\vec{r}_2,\vec{r}_2')\chi_S^*(\vec{r}_1',\vec{r}_1)}{\omega+\Omega_S}$$

where φ_c, φ_v are the non-interacting conduction and valence states and $\chi_S(\vec{r}_1,\vec{r}_1')$ are the interacting pair states with excitation energy Ω_S.

Bethe-Salpeter equation for the pair correlation function

- Formally, the interacting pair correlation function and the non-interacting one is related by the Dyson equation



- In this case, it is also known as the Bethe-Salpeter equation (BSE)

$$L(12;1'2') = L_0(12;1'2') + \int d(3456) L_0(14;1'3) K(35;46) L(62;52')$$

where K is called the electron-hole kernel

$$K(35;46) = \frac{\delta[V_{\text{coul}}(3)\delta(3,4) + \Sigma(3,4)]}{\delta G(6,5)} = K^x(35;46) + K^d(35;46)$$

$$\begin{cases} K^x(35;46) : \text{The exchange term} \\ K^d(35;46) : \text{The direct (screened) Coulomb term} \end{cases}$$

Solving the Bethe-Salpeter equation

- The main idea of using many-body perturbation theory to solve the BSE is that one can start with mean-field KS solutions to obtain quasiparticle properties within the GW approximation. The electron-hole interaction is then
- In the BSE, the excitations and de-excitations are coupled, making the calculations rather difficult.
- Often one use the so-called Tamm-Dancoff approximation (TDA) to decouple the excitation and de-excitations, leading to a much simplified BSE

The TDA and the Bethe-Salpeter equation

- Using the TDA, the BSE can be casted into a simplified eigenvalue problem:

$$(E_{c\vec{k}} - E_{v\vec{k}})A^S_{vc\vec{k}} + \sum_{v',c',\vec{k}'} K_{vck,v'c'\vec{k}'} A^S_{v'c'\vec{k}'} = \Omega_S A^S_{vc\vec{k}}$$

where Ω_S is the e-h excitation energy and the eigen vector $A^S_{vc\vec{k}}$ can be used to constructor the e-h pair (excitonic) states:

$$\Psi^{eh}_S(\vec{r}_e, \vec{r}_h) = \sum_{v,c,\vec{k}} A^S_{vc\vec{k}} \psi_{c\vec{k}}(\vec{r}_e) \psi^*_{v\vec{k}}(\vec{r}_h)$$

- At this point, we need to examine more carefully the spin aspect of the e-h pair. For a given e-h pair, we have the following spin states:

$$|v\uparrow c\uparrow\rangle;\ |v\uparrow c\downarrow\rangle;\ |v\downarrow c\uparrow\rangle;\ |v\downarrow c\downarrow\rangle$$

which give rise to spin singlet and triplet solutions.

The TDA and the Bethe-Salpeter equation

- The electron-hole Hamiltonian matrix looks like this

$$H^{e-h} = \begin{pmatrix} \omega_{cv}^{QP} + K^d + K^x & 0 & 0 & K^x \\ 0 & \omega_{cv}^{QP} + K^d & 0 & 0 \\ 0 & 0 & \omega_{cv}^{QP} + K^d & 0 \\ K^x & 0 & 0 & \omega_{cv}^{QP} + K^d + K^x \end{pmatrix} \begin{matrix} k,v\uparrow c\uparrow \\ k,v\uparrow c\downarrow \\ k,v\downarrow c\uparrow \\ k,v\downarrow c\downarrow \end{matrix}$$

$$\begin{matrix} k',v'\uparrow c'\uparrow & k',v'\uparrow c'\downarrow & k',v'\downarrow c'\uparrow & k',v'\downarrow c'\downarrow \end{matrix}$$

- Within the singlet subspace, $\frac{1}{\sqrt{2}}(|v\uparrow c\uparrow> - |v\downarrow c\downarrow>)$, the Hamiltonian is simplified

$$H^{e-h} = \omega_{cv}^{QP} + K^d + 2K^x \qquad (\omega_{cv}^{QP} = E_c^{QP} - E_v^{QP})$$

- Within the triplet space, $H^{e-h} = \omega_{cv}^{QP} + K^d$

The TDA and the Bethe-Salpeter equation

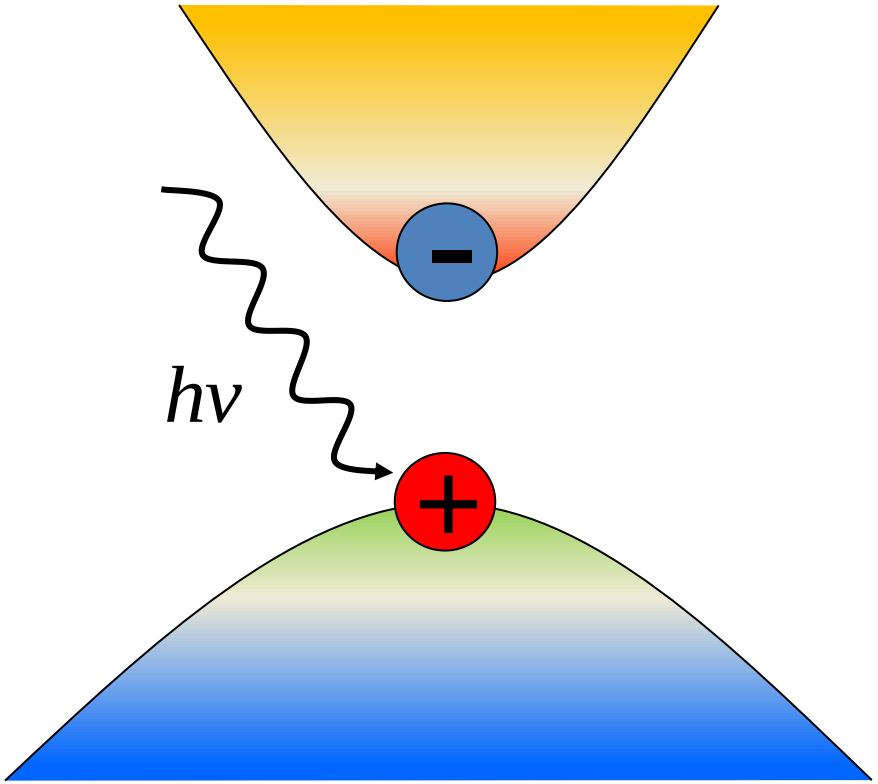
- The imaginary part of the dielectric function is then

$$\varepsilon_2(\omega) = \frac{16\pi^2 e^2}{\omega^2} \sum_S |\, \vec{\lambda} \cdot < 0 \,|\, \vec{v} \,|\, \Psi_S^{eh} >|^2 \; \delta(\omega - \Omega_S)$$

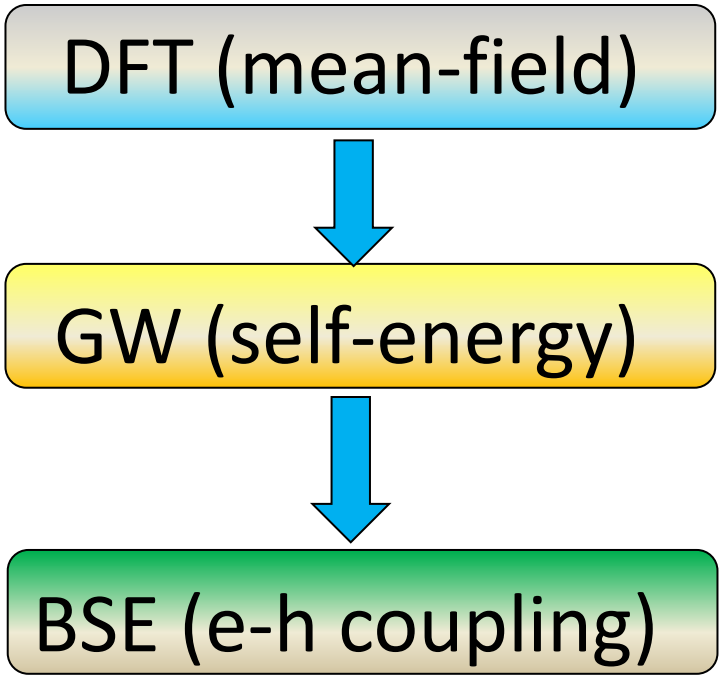
- Or

$$\varepsilon_2(\omega) = \frac{16\pi^2 e^2}{\omega^2} \sum_S |\sum_{c,v,\vec{k}} \vec{\lambda} \cdot < v\vec{k} \,|\, \vec{v} \,|\, c\vec{k} > A^S_{cv\vec{k}} \,|^2 \; \delta(\omega - \Omega_S)$$

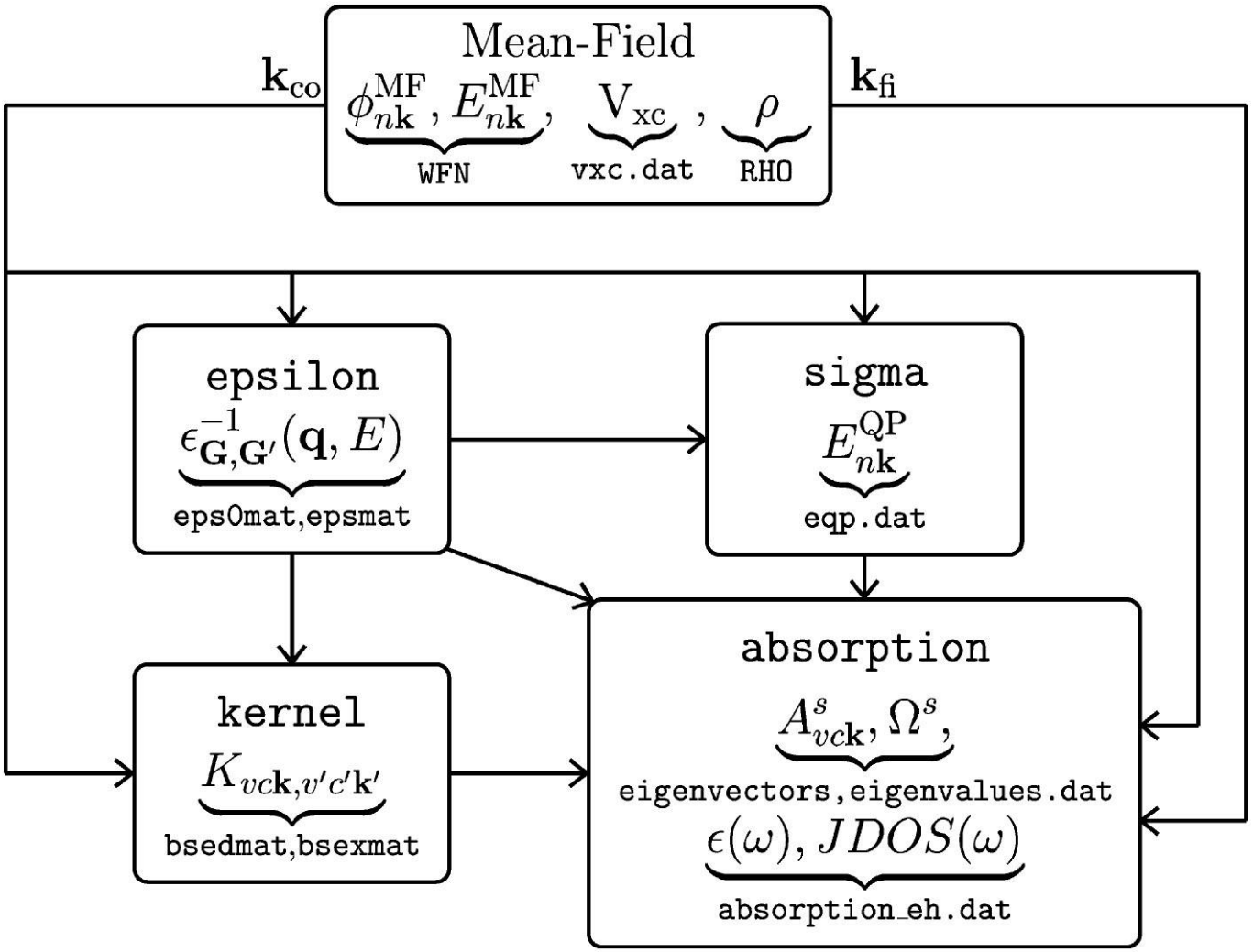
Modern approach to electronic excitations in solids



Electron-hole excitations

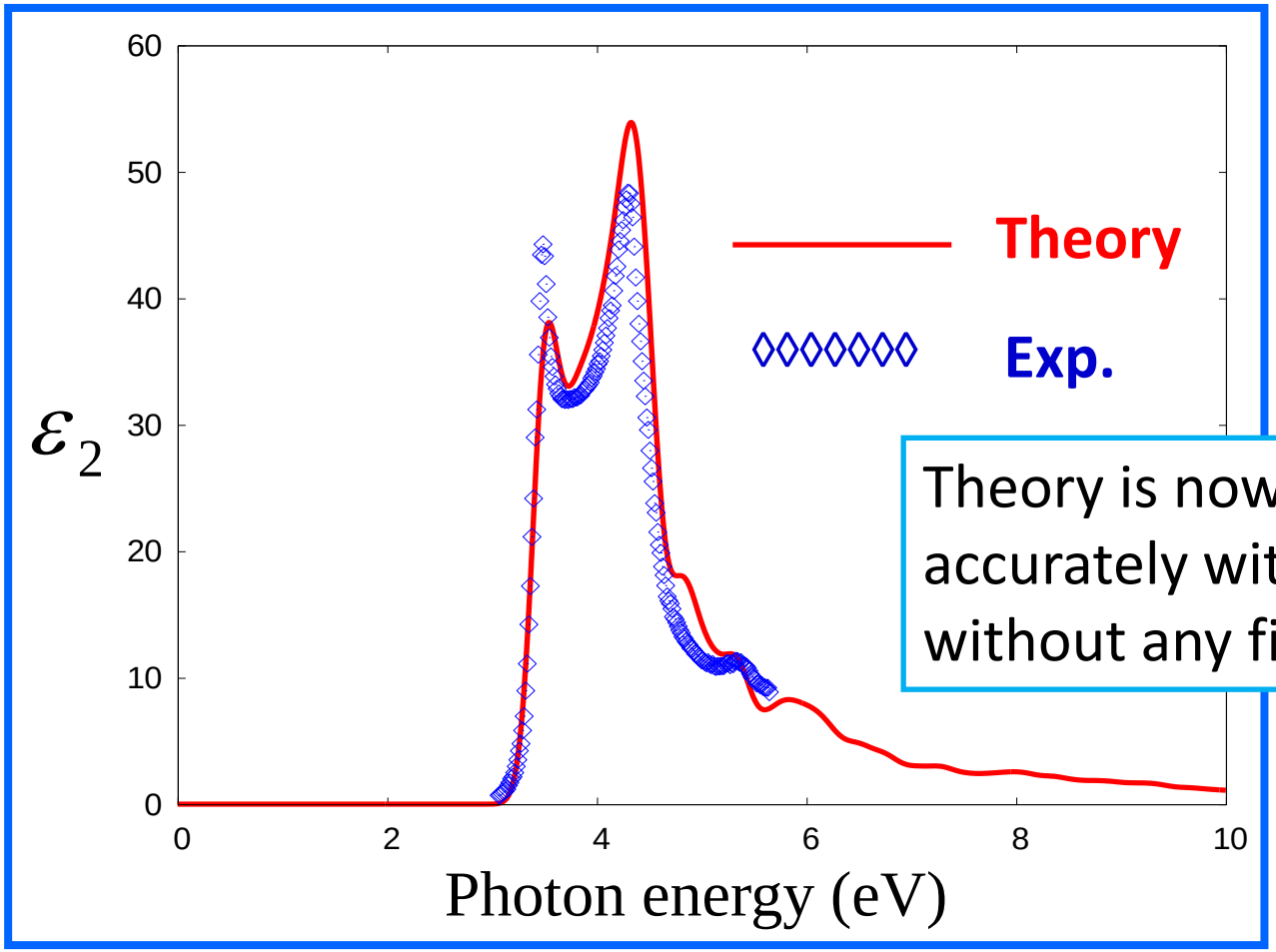


The BerkeleyGW package (https://berkeleygw.org/)



Computer Physics Communications 183, 1269 (2012)

Optical absorption of semiconductors from first-principles: Silicon



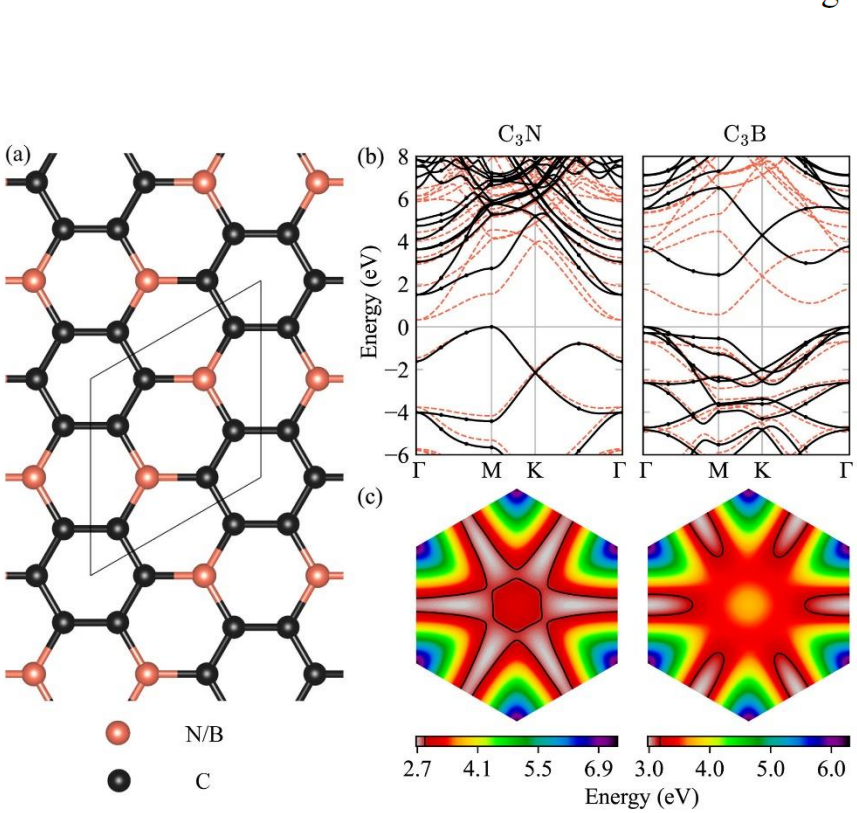
Theory is now able to compare accurately with experiment without any fitting parameters!

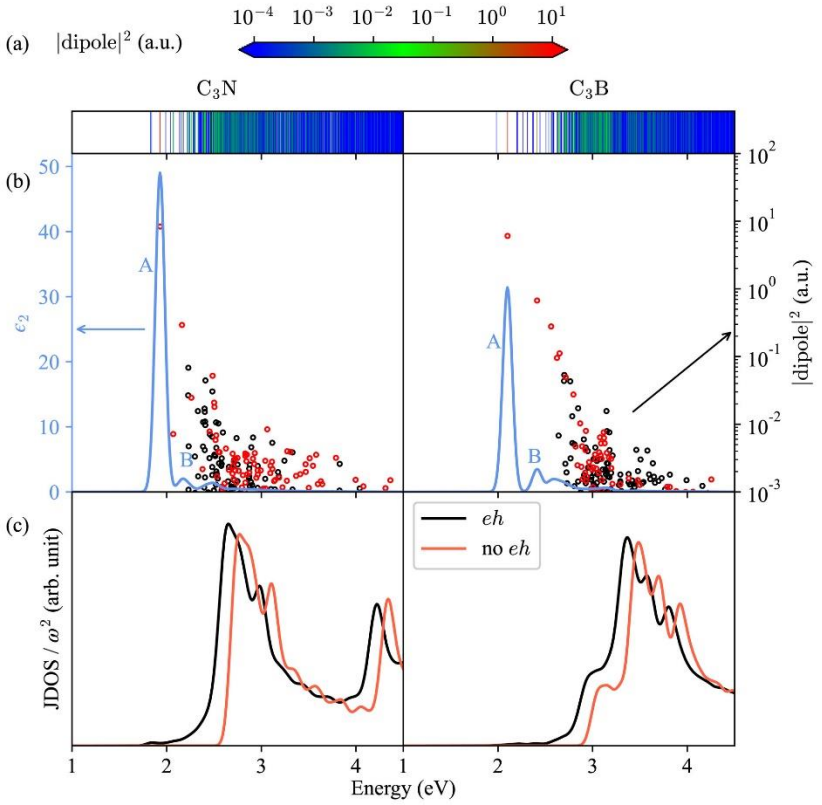
Our recent work

PHYSICAL REVIEW APPLIED **17,** 034068 (2022)

Giant Narrow-Band Optical Absorption and Distinctive Excitonic Structures of Monolayer C_3N and C_3B

Zhao Tang,[1] Greis J. Cruz,[1] Yabei Wu,[2] Weiyi Xia,[3] Fanhao Jia,[1,4] Wenqing Zhang,[2] and Peihong Zhang[1,*]





Our recent work

npj computational materials

www.nature.com/npjcompumats

ARTICLE OPEN

Check for updates

Prediction of protected band edge states and dielectric tunable quasiparticle and excitonic properties of monolayer $MoSi_2N_4$

Yabei Wu[1,2,3,7], Zhao Tang[4,7], Weiyi Xia[4], Weiwei Gao[5], Fanhao Jia[4,6], Yubo Zhang[1,2], Wenguang Zhu[3], Wenqing Zhang[1,2]✉ and Peihong Zhang[4]✉