Practical calculations with the GW approximation and Bethe-Salpeter equation in BerkeleyGW

David A. Strubbe

Department of Physics,

University of California, Merced



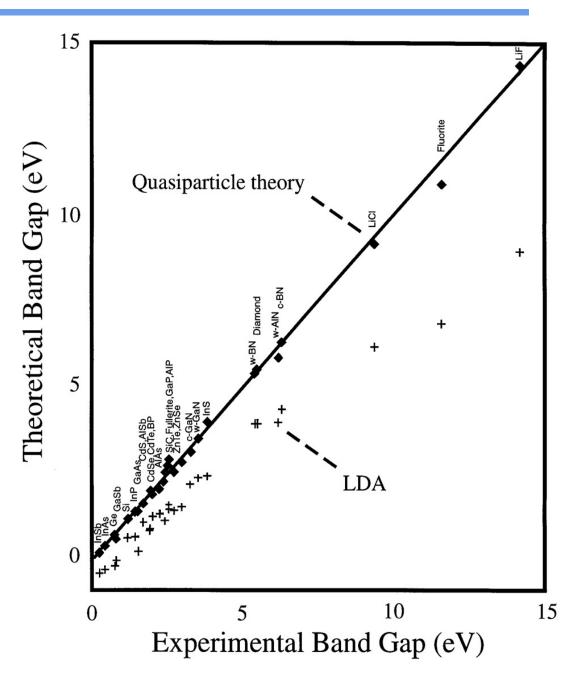




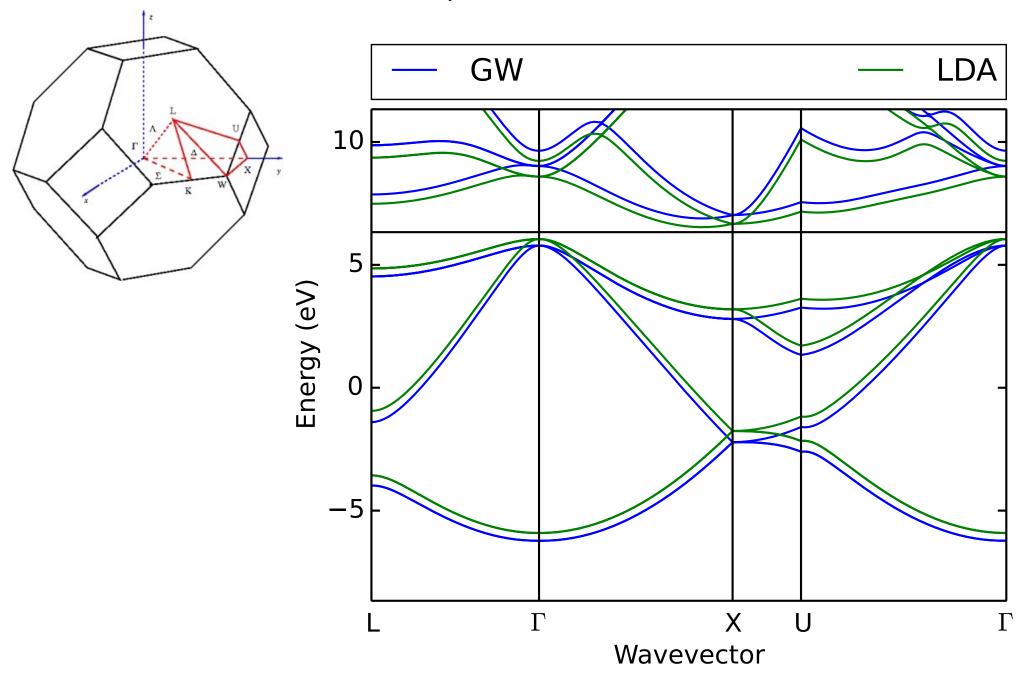
Benasque, Spain 24 October 2022

Band gaps: DFT-LDA and GW

Materials: InSb, InAs Ge GaSb Si InP GaAs CdS AISb, AIAs CdSe, CdTe BP SiC C_{60} GaP AIP ZnTe, ZnSe c-GaN, w-GaN InS w-BN, c-BN diamond w-AIN LiCI **Fluorite** LiF



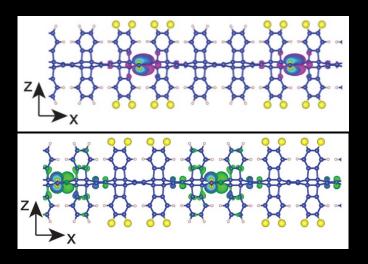
Silicon interpolated bandstructure

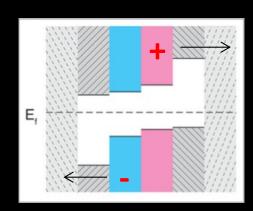


GW/Bethe-Salpeter for graphene-based PV

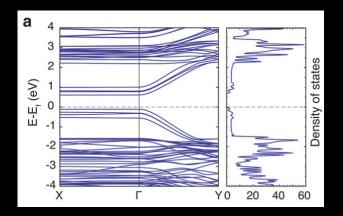
valence

conduction

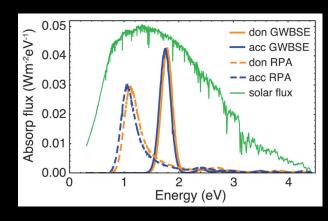


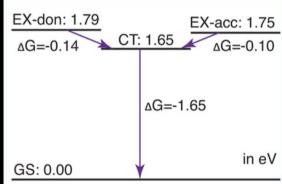


Band structure and density of states



Excitons (electron-hole pairs) from Bethe-Salpeter equation





H. Li, D. A. Strubbe, and Jeffrey C. Grossman, Adv. Funct. Mater. 25, 5199 (2015)



Preview: GW approximation/Bethe-Salpeter eqn

Start with wavefunctions and energies from DFT as mean field

Add perturbation: difference between Vxc and true exchange-correlation

GW self-energy: single-electron energy levels (band structure)

$$\left[-\frac{1}{2} \nabla^2 + V_{\text{ion}} + V_{\text{H}} + \Sigma (E_{n\mathbf{k}}^{\text{QP}}) \right] \psi_{n\mathbf{k}}^{\text{QP}} = E_{n\mathbf{k}}^{\text{QP}} \psi_{n\mathbf{k}}^{\text{QP}}$$

Bethe-Salpeter equation: electron-hole interaction for optical properties

$$\left(E_{c\mathbf{k}}^{\mathrm{QP}} - E_{v\mathbf{k}}^{\mathrm{QP}}\right)A_{vc\mathbf{k}}^{S} + \sum_{v'c'\mathbf{k}'}\left\langle vc\mathbf{k}|K^{\mathrm{eh}}|v'c'\mathbf{k}'\right\rangle = \Omega^{S}A_{vc\mathbf{k}}^{S} \quad \Psi(\mathbf{r}_{e},\mathbf{r}_{h}) = \sum_{\mathbf{k},c,v}A_{vc\mathbf{k}}^{S}\psi_{\mathbf{k},c}(\mathbf{r}_{e})\psi_{\mathbf{k},v}^{*}(\mathbf{r}_{h})$$



widely used massively parallel code www.berkeleygw.org

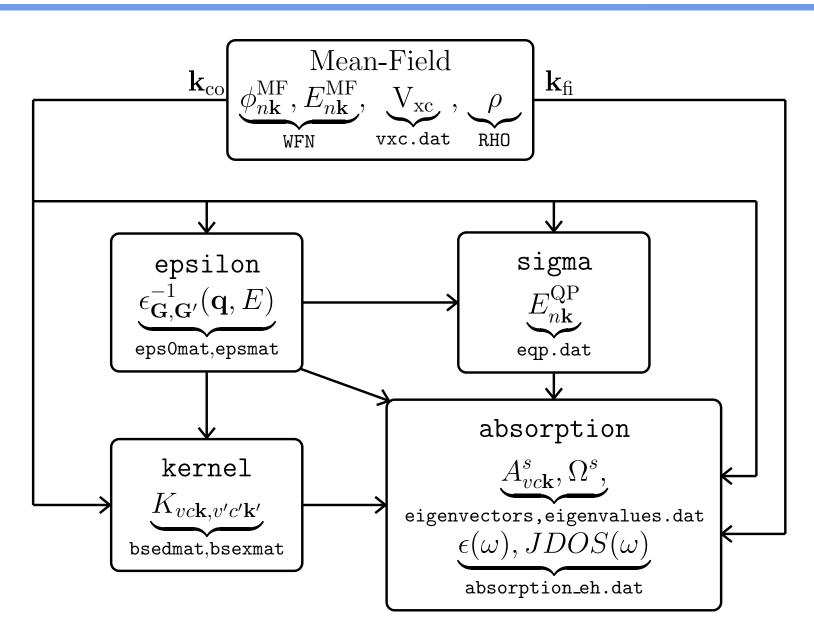
J. Deslippe, G. Samsonidze, D. A. Strubbe, M. Jain, M. L. Cohen, and S. G. Louie, *Comput. Phys.* Comm. 183, 1269 (2012)

Why Use BerkeleyGW



- Supports a large set of Mean-Field codes: PARATEC, Quantum ESPRESSO, ABINIT, Octopus, PARSEC, SIESTA, EPM (TBPW)
- Supports 3D, 2D, 1D and Molecular Systems. Coulomb Truncation
- Support for Semiconductor, Metallic and Semi-Metallic Systems
- Efficient Algorithms and Use of Libraries. (BLAS, FFTW3, LAPACK, SCALAPACK, ELPA, OpenMP, HDF5)
- Massively Parallel. Scales to 100,000 CPUs, distributed Memory.
- Efficient accurate solution to BSE via k-point Interpolation
- Support for LDA/GGA/Hybrid/HF/COHSEX starting points as well as offdiagonal Σ calculations





Full-Frequency vs. Generalized Plasmon Pole (GPP)



$$\langle n\mathbf{k}| \, \Sigma_{\text{CH}}(E) \, |n'\mathbf{k}\rangle = \frac{i}{2\pi} \sum_{n''} \sum_{\mathbf{q}\mathbf{G}\mathbf{G}'} M_{n''n}^*(\mathbf{k}, -\mathbf{q}, -\mathbf{G}) M_{n''n'}(\mathbf{k}, -\mathbf{q}, -\mathbf{G}') \quad (20)$$

Full-Frequency

$$\times \int_{0}^{\infty} dE' \frac{\left[\epsilon_{\mathbf{G}\mathbf{G}'}^{\mathrm{r}}\right]^{-1}(\mathbf{q}; E') - \left[\epsilon_{\mathbf{G}\mathbf{G}'}^{\mathrm{a}}\right]^{-1}(\mathbf{q}; E')}{E - E_{n''\mathbf{k} - \mathbf{q}} - E' + i\delta} v(\mathbf{q} + \mathbf{G}')$$

$$\langle n\mathbf{k} | \Sigma_{\text{CH}}(E) | n'\mathbf{k} \rangle = \frac{1}{2} \sum_{n''} \sum_{\mathbf{q}\mathbf{G}\mathbf{G}'} M_{n''n}^*(\mathbf{k}, -\mathbf{q}, -\mathbf{G}) M_{n''n'}(\mathbf{k}, -\mathbf{q}, -\mathbf{G}')$$
(22)
$$\times \frac{\Omega_{\mathbf{G}\mathbf{G}'}^2(\mathbf{q}) (1 - i \tan \phi_{\mathbf{G}\mathbf{G}'}(\mathbf{q}))}{\tilde{\omega}_{\mathbf{G}\mathbf{G}'}(\mathbf{q}) (E - E_{n''\mathbf{k} - \mathbf{q}} - \tilde{\omega}_{\mathbf{G}\mathbf{G}'}(\mathbf{q}))} v(\mathbf{q} + \mathbf{G}')$$

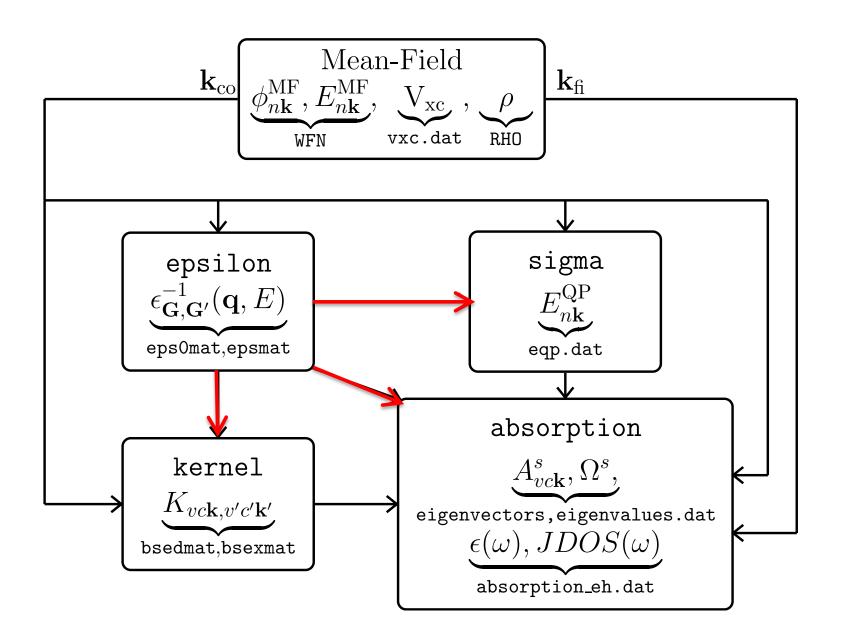
GPP is significantly faster. The integral over frequencies can be performed analytically if assume the dielectric response is dominated by a single plasmon pole.

BerkeleyGW supports both. With full-frequency you can compute spectral functions, lifetimes and weights. By contrast, GPP is (incorrectly) purely real.

Practical issues for *GW*

- 1. Screening models for Epsilon
- 2. Construction of k-grids
- 3. Symmetry and degeneracy
- 4. Real and complex version
- 5. Solving Dyson's equation
- 6. Convergence
- 7. Use of Octopus

Screening models: How do we use ε ?



Screening models: How do we use ε ?

Sigma integrates over **q** with $\varepsilon^{-1}(\mathbf{q})$

$$\langle n\mathbf{k} | \Sigma(E) | n'\mathbf{k} \rangle = \frac{i}{2\pi} \sum_{n''} \sum_{\mathbf{q} \mathbf{G} \mathbf{G}'} M_{n''n}^*(\mathbf{k}, -\mathbf{q}, -\mathbf{G}) M_{n''n'}(\mathbf{k}, -\mathbf{q}, -\mathbf{G}')$$

$$\times \int_{-\infty}^{\infty} dE' e^{-i\delta E'} \frac{\left[\epsilon_{\mathbf{G} \mathbf{G}'}\right]^{-1}(\mathbf{q}; E')}{E - E_{n''\mathbf{k} - \mathbf{q}} - E' - i\delta_{n''\mathbf{k} - \mathbf{q}}} v(\mathbf{q} + \mathbf{G}')$$

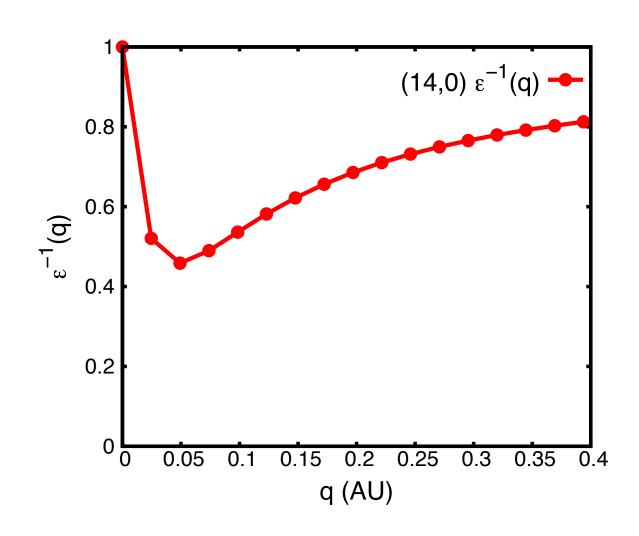
Absorption interpolates kernel over **q** with $W(\mathbf{q}) = \varepsilon^{-1}(\mathbf{q}) v(\mathbf{q})$

$$\langle vc\mathbf{k}|K^{\mathrm{d}}|v'c'\mathbf{k}'\rangle = \sum_{\mathbf{GG}'} M_{c'c}^{*}(\mathbf{k},\mathbf{q},\mathbf{G})W_{\mathbf{GG}'}(\mathbf{q};0)M_{v'v}(\mathbf{k},\mathbf{q},\mathbf{G}')$$

Problem 1: Non-smooth behavior

(14, 0) carbon nanotube wire truncation

General for truncation: see BN tutorial



Problem 2: Divergent behavior

$$\chi_{\mathbf{G}\mathbf{G}'}(\mathbf{q};0) = \sum_{n}^{\mathrm{occ}} \sum_{\mathbf{k}}^{\mathrm{emp}} \sum_{\mathbf{k}} M_{nn'}^*(\mathbf{k},\mathbf{q},\mathbf{G}) M_{nn'}(\mathbf{k},\mathbf{q},\mathbf{G}') \frac{1}{E_{n\mathbf{k}+\mathbf{q}} - E_{n'\mathbf{k}}}.$$

Head: G = 0, G' = 0

Wing: **G** = 0, **G'** \neq 0

Wing': $G \neq 0$, G' = 0

Body: **G** \neq 0, **G'** \neq 0

Cannot calculate at **q** = 0!

$$q^2/q^2$$

$\epsilon_{\mathbf{G}\mathbf{G}'}^{-1}$	head	wing	wing'	body
Semiconductor	const	q	\mathbf{q}/q^2	const
Metal	q^2	q^2	const	const
$W_{\mathbf{G}\mathbf{G}'}$	head	wing	wing'	body
Semiconductor	$1/q^2$	${f q}/q^2$	${f q}/q^2$	const
Metal	const	const	const	const

diverges

Solution: Screening models

Calculate at $\mathbf{q_0} \approx 0.001$ in periodic direction use to parametrize screening model

Sigma: Integrate over region around $\mathbf{q} = 0$

Kernel: Interpolate in parts

$$\left\langle vc\mathbf{k}\left|K\right|v'c'\mathbf{k'}\right\rangle = \frac{a_{vc\mathbf{k}v'c'\mathbf{k'}}}{A\left(\mathbf{q}\right)} + \frac{b_{vc\mathbf{k}v'c'\mathbf{k'}}}{B\left(\mathbf{q}\right)} + \frac{c_{vc\mathbf{k}v'c'\mathbf{k'}}}{C\left(\mathbf{q}\right)}$$

head

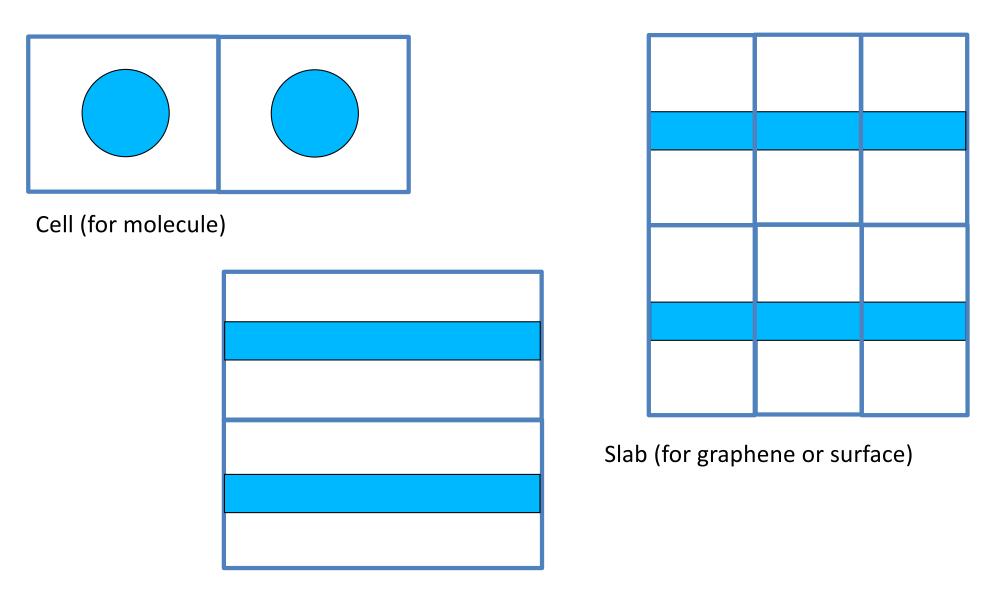
$\epsilon_{\mathbf{G}\mathbf{G}'}^{-1}$	head	wing	wing'	body
Semiconductor	const	q	${f q}/q^2$	const
Metal	q^2	q^2	const	const
$W_{\mathbf{G}\mathbf{G}'}$	head	wing	wing'	body
Semiconductor	$1/q^2$	${f q}/q^2$	${f q}/q^2$	const
Metal	const	const	const	const

wing, wing'

body

Truncation for non- or partially periodic systems

Periodicity in 0, 1, 2, or 3 dimensions. Eliminate spurious image interactions.



Wire (for nanotube or nanowire)

Truncation of Coulomb potential

GW and BSE utilize the Coulomb and screened Coulomb interaction

$$W = \varepsilon^{-1} V_c$$

- Long-range interactions make it computationally infeasible to increase lattice vectors until periodic images do not interact.
- Rule of thumb: non-periodic direction should include 99% of density

Truncation Schemes within BerkeleyGW

Cell box: 0D

Cell wire: 1D

Cell slab: 2D

Spherical: Define radius of truncation

$$v_{
m t}({f r}) = rac{\Theta(f({f r}))}{r}$$

- Cell truncation: at half lattice vector length
 - Analytical form for Coulomb potential in k-space
- Spherical truncation: convenient, available in many packages

Regular k-grids

Epsilon

$$\chi_{\mathbf{G}\mathbf{G}'}(\mathbf{q};0) = \sum_{n}^{\text{occ}} \sum_{\mathbf{k}}^{\text{emp}} \sum_{\mathbf{k}} M_{nn'}^*(\mathbf{k}, \mathbf{q}, \mathbf{G}) M_{nn'}(\mathbf{k}, \mathbf{q}, \mathbf{G}') \frac{1}{E_{n\mathbf{k}+\mathbf{q}} - E_{n'\mathbf{k}}}.$$

where

$$M_{nn'}(\mathbf{k}, \mathbf{q}, \mathbf{G}) = \langle n\mathbf{k} + \mathbf{q} | e^{i(\mathbf{q} + \mathbf{G}) \cdot \mathbf{r}} | n'\mathbf{k} \rangle$$

Sigma

$$\langle n\mathbf{k} | \Sigma_{\text{SX}}(E) | n'\mathbf{k} \rangle = -\sum_{n''}^{\text{occ}} \sum_{\mathbf{q}\mathbf{G}\mathbf{G}'} M_{n''n}^*(\mathbf{k}, -\mathbf{q}, -\mathbf{G}) M_{n''n'}(\mathbf{k}, -\mathbf{q}, -\mathbf{G}')$$

$$\times \left[\epsilon_{\mathbf{G}\mathbf{G}'} \right]^{-1} (\mathbf{q}; E - E_{n''\mathbf{k} - \mathbf{q}}) v(\mathbf{q} + \mathbf{G}')$$

Kernel

$$\langle vc\mathbf{k}|K^{\mathrm{d}}|v'c'\mathbf{k}'\rangle = \sum_{\mathbf{G}\mathbf{G}'} M_{c'c}^{*}(\mathbf{k}, \mathbf{q}, \mathbf{G})W_{\mathbf{G}\mathbf{G}'}(\mathbf{q}; 0)M_{v'v}(\mathbf{k}, \mathbf{q}, \mathbf{G}')$$

$$\epsilon^{-1}(\mathbf{q}) \text{ for } \mathbf{q} = \mathbf{k} - \mathbf{k}'.$$

k-grids and bands

recommended approach

	k-grid	# bands	Comments
SCF	Uniform, 0.5 shift	occupied	as usual in DFT
WFN	Uniform, 0.5 shift	many	
WFNq	WFN + q -shift	occupied	
epsilon.inp q -points	WFN but no shift, q ₀	many	bands to sum over
WFN_inner	WFN but no shift	many	bands to sum over
sigma.inp k -points	subset of WFN_inner	few	can choose to calculate Sigma just for bands of interest
WFN_co	WFN_inner	few	
WFN_fi (absorption)	Uniform, random shift	few	
WFNq_fi	WFN_fi + q -shift	occupied	
WFN_fi (inteqp)	anything	few	whatever is of interest

epsilon.inp

Semiconductors

end

begin qpoints 0.000000 1.0 0.000000 0.005000 0.000000 0.000000 0.062500 1.0 0 1.0 0.000000 0.000000 0.125000 0.000000 0.000000 0.187500 1.0

eps0mat:

$$\epsilon_{\mathbf{G}\mathbf{G}'}^{-1}\left(\mathbf{q}_{0}\right)$$

epsmat:

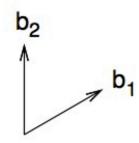
$$\epsilon_{\mathbf{G}\mathbf{G}'}^{-1} \left(\mathbf{q} \neq \mathbf{q}_0 \right)$$

k-grid construction: 4×4 grid for graphene

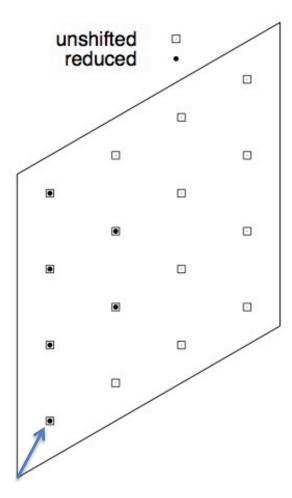
(0.5, 0.5) Monkhorst-Pack shift

kgrid.x

Uniform -> unfold -> shift with **q** -> reduce

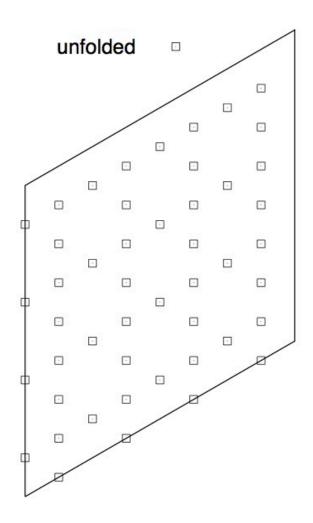


Unfolding gives more points!



(0.5, 0.5)

Main grid (WFN) 16 in full BZ Reduced to 6

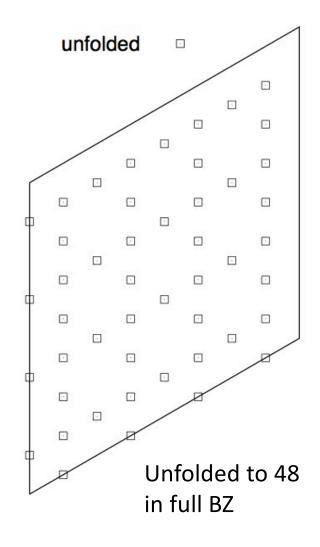


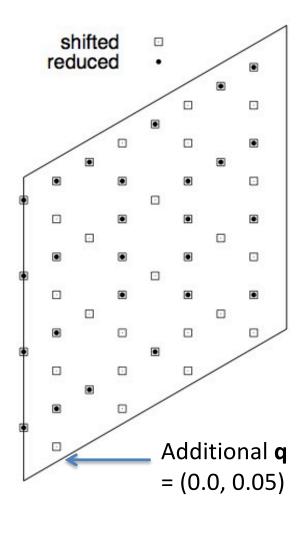
Unfolded to 48 in full BZ

k-grid construction: 4×4 grid for graphene

kgrid.x

Uniform -> unfold -> shift with **q** -> reduce





Unfolding and breaking symmetry gives more points!

Shifted grid (WFNq) 48 in full BZ Reduced to 26

Degeneracy

Epsilon, Sigma: symmetry of Hamiltonian

$$\langle n\mathbf{k} | \Sigma_{\text{SX}}(E) | n'\mathbf{k} \rangle = -\sum_{n''}^{\text{occ}} \sum_{\mathbf{q}\mathbf{G}\mathbf{G}'} M_{n''n}^*(\mathbf{k}, -\mathbf{q}, -\mathbf{G}) M_{n''n'}(\mathbf{k}, -\mathbf{q}, -\mathbf{G}')$$
$$\times \left[\epsilon_{\mathbf{G}\mathbf{G}'} \right]^{-1} (\mathbf{q}; E - E_{n''\mathbf{k} - \mathbf{q}}) v(\mathbf{q} + \mathbf{G}')$$

Absorption: symmetry of e-h basis

$$\left(E_{c\mathbf{k}}^{\mathrm{QP}} - E_{v\mathbf{k}}^{\mathrm{QP}}\right) A_{vc\mathbf{k}}^{S} + \sum_{v'c'\mathbf{k}'} \left\langle vc\mathbf{k} | K^{\mathrm{eh}} | v'c'\mathbf{k}' \right\rangle = \Omega^{S} A_{vc\mathbf{k}}^{S}$$

Summing over only some of a degenerate space will break symmetry.

Degeneracy in mean-field => broken in *GW*!

Results depends on arbitrary linear combinations in mean-field. Not reproducible! Incorrect oscillator strengths in absorption!

Degeneracy check utility

```
$ degeneracy_check.x WFN
Reading eigenvalues from file WFN
Number of spins:
Number of bands:
                              35
Number of k-points:
== Degeneracy-allowed numbers of bands (for epsilon and sigma) ==
           14
           18
           20
           32
Note: cannot assess whether or not highest band 35 is degenerate.
```

So, use number bands 32 in Epsilon.

Real or complex flavor?

e.g. bin/epsilon.real.x, bin/epsilon.cplx.x

Complex is general, but real is faster, uses less memory and disk space

Real: only with inversion symmetry about the origin $u(-\mathbf{r}) = au(\mathbf{r})$

and time-reversal symmetry $u^*(\mathbf{r}) = bu(\mathbf{r})$

a, b each equal to ± 1

What breaks time-reversal? Magnetic fields, spin-polarization, spinors Plane-wave codes generally just use complex wavefunctions. Conditions for reality depends on the basis! Real-space: k = 0, time-reversal.

Real output not implemented in Octopus yet.

Solving Dyson's equation in Sigma

$$E_{n\mathbf{k}}^{\mathrm{QP}} = E_{n\mathbf{k}}^{\mathrm{MF}} + \langle \psi_{n\mathbf{k}} | \Sigma(E_{n\mathbf{k}}^{\mathrm{QP}}) - \Sigma^{\mathrm{MF}} | \psi_{n\mathbf{k}} \rangle$$

How can we solve when we don't know E^{QP} yet?

(1) eqp0: evaluate at E^{MF} .

$$E_{n\mathbf{k}}^{\text{QPO}} = E_{n\mathbf{k}}^{\text{MF}} + \langle \psi_{n\mathbf{k}} | \Sigma(E_{n\mathbf{k}}^{\text{MF}}) - \Sigma^{\text{MF}} | \psi_{n\mathbf{k}} \rangle$$

(2) eqp1: solve linearized approximation (Newton's Method)

$$E_{n\mathbf{k}}^{\text{QP1}} = E_{n\mathbf{k}}^{\text{QP0}} + \frac{d\Sigma_{n\mathbf{k}}/dE}{1 - d\Sigma_{n\mathbf{k}}/dE} \left(E_{n\mathbf{k}}^{\text{QP0}} - E_{n\mathbf{k}}^{\text{MF}} \right)$$

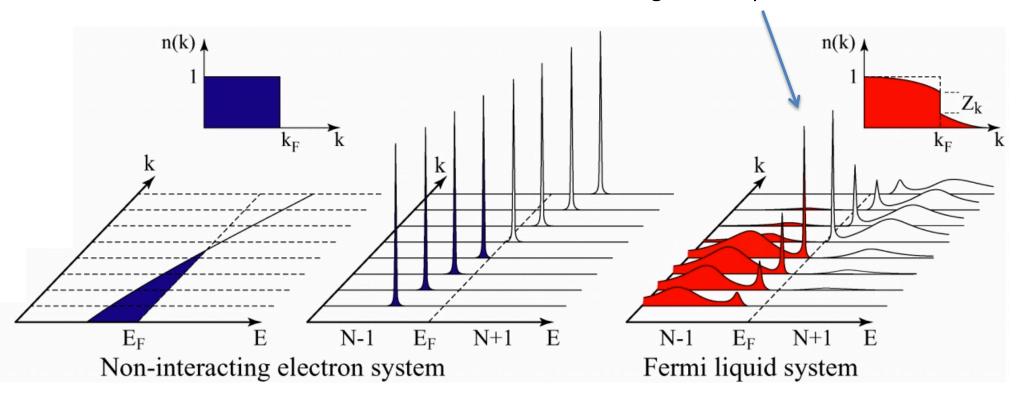
Available as columns in sigma_hp.log, and eqp0.dat and eqp1.dat files

Quasiparticle renormalization factor Z

$$E_{n\mathbf{k}}^{\text{QP1}} = E_{n\mathbf{k}}^{\text{QP0}} + (Z_{n\mathbf{k}} - 1) \left(E_{n\mathbf{k}}^{\text{QP0}} - E_{n\mathbf{k}}^{\text{MF}} \right)$$

$$Z_{n\mathbf{k}} = \frac{1}{1 - d\Sigma_{n\mathbf{k}}/dE}$$

Between 0 and 1 Weight in QP peak



A. Damascelli, Z. Hussain, Z.-X Shen, Rev. Mod. Phys. 75, 473 (2003)

There are many convergence parameters in a GW calculations: convergence with each must be checked

Screened cutoff

Empty bands (dielectric matrix)

$$\chi_{\mathbf{GG'}}(\mathbf{q};0) = \sum_{n}^{\mathrm{occ}} \sum_{\mathbf{k}}^{\mathrm{emp}} \sum_{\mathbf{k}} M_{nn'}^{*}(\mathbf{k},\mathbf{q},\mathbf{G}) M_{nn'}(\mathbf{k},\mathbf{q},\mathbf{G}') \frac{1}{E_{n\mathbf{k}+\mathbf{q}}-E_{n'\mathbf{k}}}$$
 q-grid

Bands in CH summation (sigma)

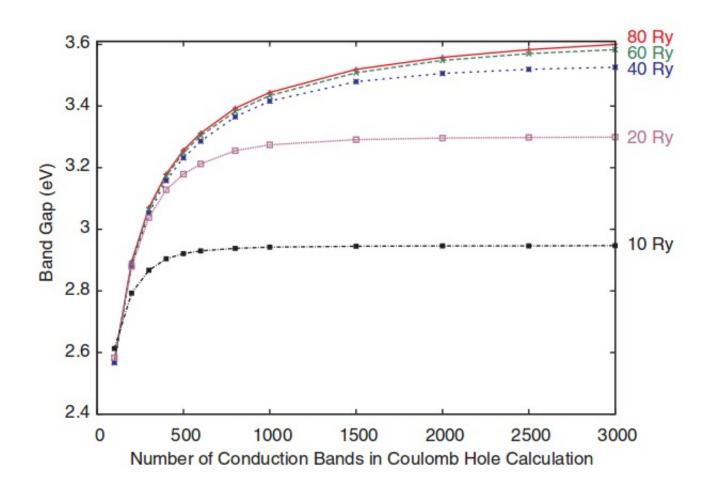
$$\langle n\mathbf{k} | \Sigma_{\text{CH}}(E) | n'\mathbf{k} \rangle = \frac{1}{2} \sum_{n''} \sum_{\mathbf{q}\mathbf{G}\mathbf{G}'} M_{n''n}^*(\mathbf{k}, -\mathbf{q}, -\mathbf{G}) M_{n''n'}(\mathbf{k}, -\mathbf{q}, -\mathbf{G}')$$

$$\times \frac{\Omega_{\mathbf{G}\mathbf{G}'}^2(\mathbf{q}) (1 - i \tan \phi_{\mathbf{G}\mathbf{G}'}(\mathbf{q}))}{\tilde{\omega}_{\mathbf{G}\mathbf{G}'}(\mathbf{q}) (E - E_{n''\mathbf{k} - \mathbf{q}} - \tilde{\omega}_{\mathbf{G}\mathbf{G}'}(\mathbf{q}))} v(\mathbf{q} + \mathbf{G}')$$

$$M_{nn'}(\mathbf{k}, \mathbf{q}, \mathbf{G}) = \langle n\mathbf{k} + \mathbf{q} | e^{i(\mathbf{q} + \mathbf{G}) \cdot \mathbf{r}} | n'\mathbf{k} \rangle$$

Wavefunction cutoff (matrix elements)

Coupled convergence parameters



ZnO: B. Shih et al., Phys. Rev. Lett. 105, 146401 (2010)

See convergence and "When things go wrong" slides on BerkeleyGW 2022 tutorial page!

Octopus interface to BerkeleyGW

Real space transformed to plane-waves for GW.

Can only produce complex wavefunctions currently.

Good for:

- very large systems
- finite or perhaps partially periodic systems (molecules, nanowires, 2D sheets)
- charged systems
- model systems
- interfacing with special features of Octopus

Domain parallelization for real-space scales better than plane waves.

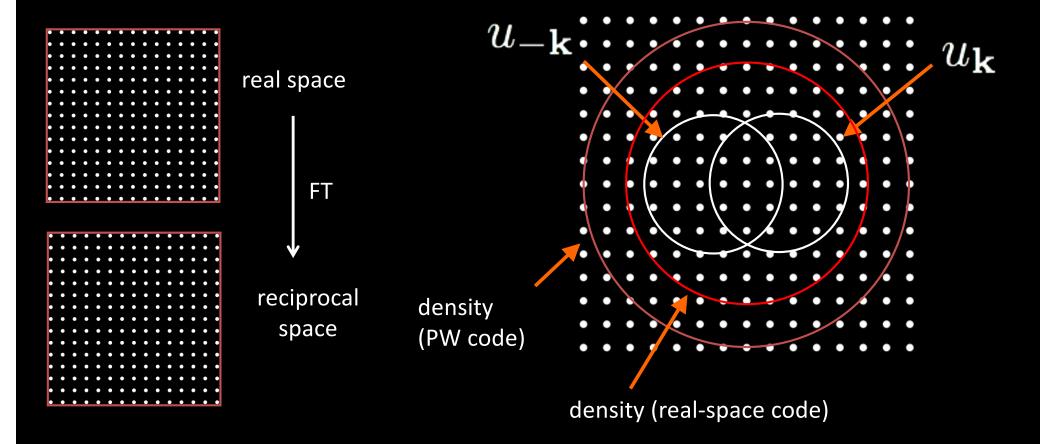
Recent application for spin-flip Bethe-Salpeter equation: https://arxiv.org/abs/2207.04549 (2022)

Fourier transform to plane-wave form

$$\psi_{n\mathbf{k}}(\mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{r}}u_{n\mathbf{k}}(\mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{r}}\sum_{\mathbf{G}}u_{n\mathbf{k}}(\mathbf{G})$$

real space: cube for all quantities

$$|\mathbf{k} + \mathbf{G}|^2 < E_{\text{cutoff}}$$



Must check norm, renormalize.

Finite vs. periodic boundary conditions for vacuum

Are plane waves more natural basis for unbound vacuum states?

"Particle in a box" spectrum

$$x = 0$$
 $x = L$

$$\psi_n(x) \sim e^{in\pi x/I}$$

$$\psi_n(x) \sim e^{in\pi x/L}$$

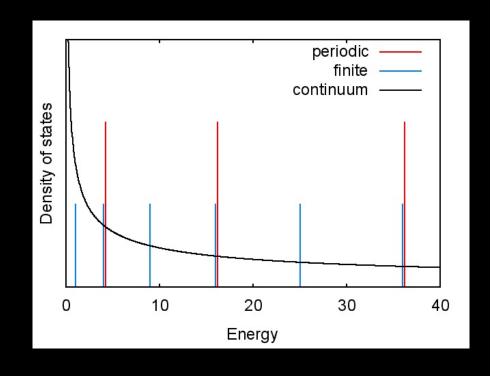
$$E_n = \frac{n^2 \pi^2 \hbar^2}{2mL^2}$$

periodic (plane waves)
$$\psi'(0)=\psi'(L)=0$$

$$n = \pm 2, \pm 4, \pm 6, \dots$$

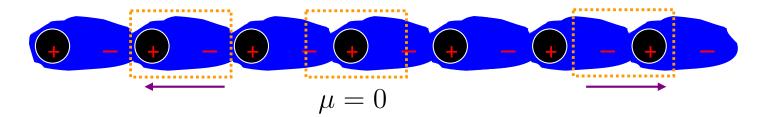
finite (real space)
$$\psi(0)=\psi(L)=0$$

$$n = 1, 2, 3, \dots$$



An opportunity: real-space dipole matrix elements

Bethe-Salpeter optical spectrum velocity gauge
$$\epsilon_2(\omega) = \frac{16\pi^2 e^2}{\omega^2} \sum_S \left| \mathbf{e} \cdot \langle 0 | \mathbf{v} | S \rangle \right|^2 \delta \left(\omega - \Omega^S \right)$$



Plane waves (inherently periodic) need the auxiliary shifted grid (WFNq fi).

length gauge
$$\langle v\mathbf{k} | \mathbf{r} | c\mathbf{k} \rangle = \lim_{\mathbf{q} \to 0} \frac{\langle v\mathbf{k} + \mathbf{q} | e^{i\mathbf{q} \cdot \mathbf{r}} - 1 | c\mathbf{k} \rangle}{iq}$$

$$= -i \lim_{\mathbf{q} \to 0} \frac{\langle v\mathbf{k} + \mathbf{q} | e^{i\mathbf{q} \cdot \mathbf{r}} | c\mathbf{k} \rangle}{q}$$

In finite direction, Octopus can calculate directly. Only one grid. Simpler, less computation, no concern of finite differences or consistent phases

In principle, can use k.p perturbation theory for derivatives in periodic directions.

The tutorial

Three examples:

- (1) benzene molecule, with Octopus
- (2) silicon, with plane waves
- (3) LiCl exciton visualization

Instructions at

http://sail.ucmerced.edu/Benasque_BerkeleyGW.html

Calculations will be performed on Cori supercomputer at NERSC.

