High-fidelity Green's functions in Correlated systems

Mark van Schilfgaarde, National Renewable Energy Lab

Why use Green's functions as the fundamental variable ?

Wave-function (ψ) methods are the king for high-fidelity

Density-functional (ρ) methods are very efficient

Goldilocks principle: Green's function (G) methods straddle the ρ- and ψ- methods, intermediate in both accuracy and efficiency.

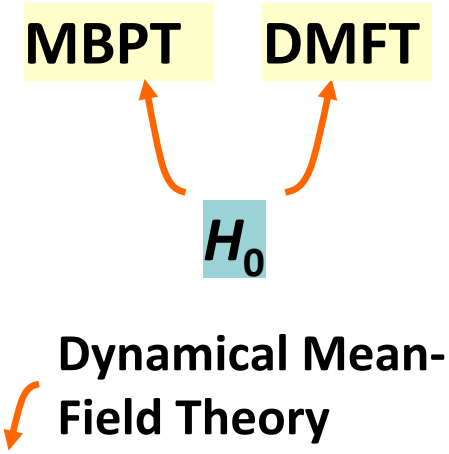
Possible to *systematically improve* fidelity higher order diagrams and self consistency.

Very often interest lies in excitations & response functions G-methods are natural for both --- intrinsic to the theory

What about fidelity ?

Primary tracks for implementation of G methods

Two traditional routes to extend one-body descriptions. Both traditionally start from independent-particle H_0 (e.g. DFT.)



1 Many body perturbation theory

Usually H_0 constructed from DFT

Add low order diagrams (usu. GW)

Implies correlations weak or moderate

2 Assume strong correlations mostly local (DMFT)

Partition hamiltonian into strongly correlated, local sector, and a weakly correlated nonlocal sector

Solve Anderson impurity problem to all orders, embedded in a noninteracting bath H_0 (*usually* DFT)

Why GW is so effective

Hartree Fock ≠ DFT: robust, sharply-defined, norm conserving.

But … HF is terrible in extended systems !

Screening is the 800 pound gorilla among many body effects (see Richard Martin's book): essentially what GW captures.

Write Fock exchange in Green's function terminology:

$$\Sigma_x = V_x(\mathbf{r}) = i\int G(\mathbf{r},\mathbf{r}')\frac{1}{|\mathbf{r}-\mathbf{r}'|}d^3\mathbf{r}' = iGv$$

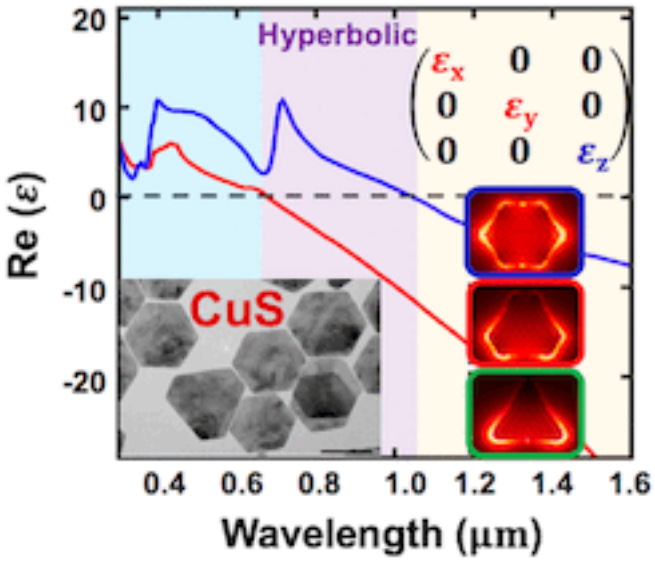
Replace bare coulomb $v \rightarrow$ dynamically screened W :

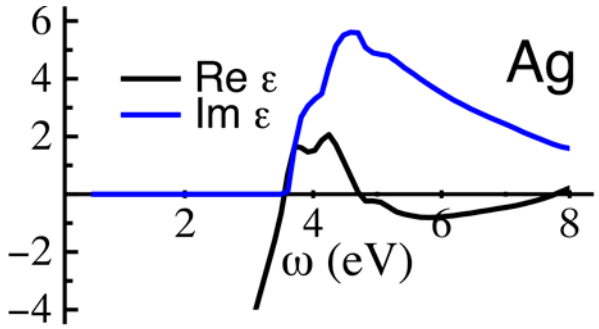
$$V_{\text{bare}}(\mathbf{r},\mathbf{r}') = \frac{1}{|\mathbf{r}-\mathbf{r}'|} \rightarrow W(\mathbf{r},\mathbf{r}',\omega) = \varepsilon^{-1}v; \quad \Sigma = iGW$$

GW is vastly better than HF. Add ladders (GW → GŴ) improves screening. Then low-order MBPT can describe charge excitations very well, even in correlated insulators (see later)

Energy scales of spin and charge excitations

Plasmon ω_p typically $\gtrsim$5 eV, e.g. Ag





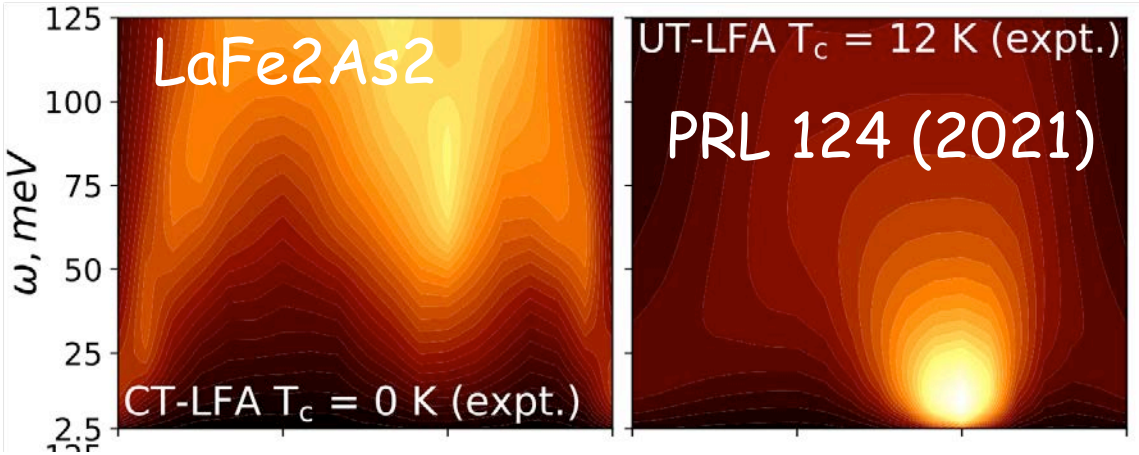
Special case:

CuS with $\omega_p<1$

Aside: Hyperbolic dispersion (region of ω with $\varepsilon_x<0$ and $\varepsilon_z>0$) key for plasmonic devices (**ACS Nano 13, 6550 (2019)**)

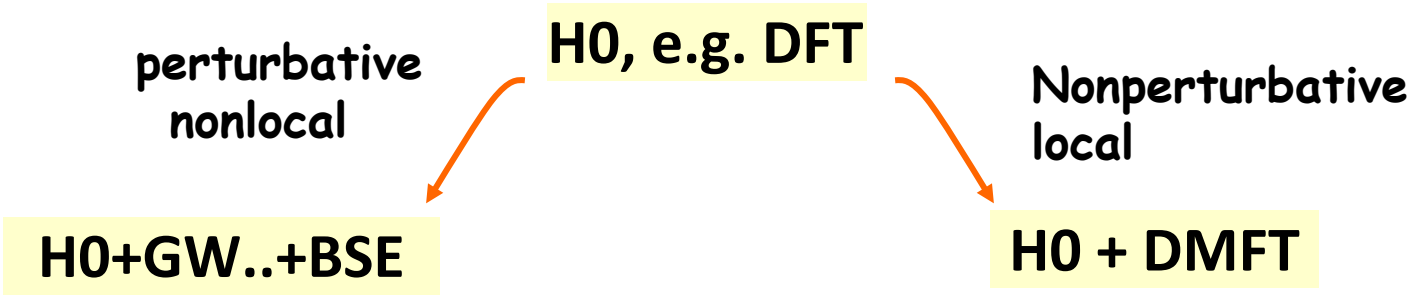
Magnetic excitations:

peaks in the range $\omega_m \sim$ 5-500 meV



The structure of spin χ^{+-} explains a great deal about the origins of unconventional superconductivity

Two tracks of G originate from different energy scales



Weak spin fluctuations

Spin is adequately described by the Fock diagram Screening of spin is weak)

Charge fluctuations are high energy, and long range.

Low order MBPT will adequately capture electronic structure

Strong spin fluctuations

Low energy ⇒ many competing processes. Rich phase diagrams.

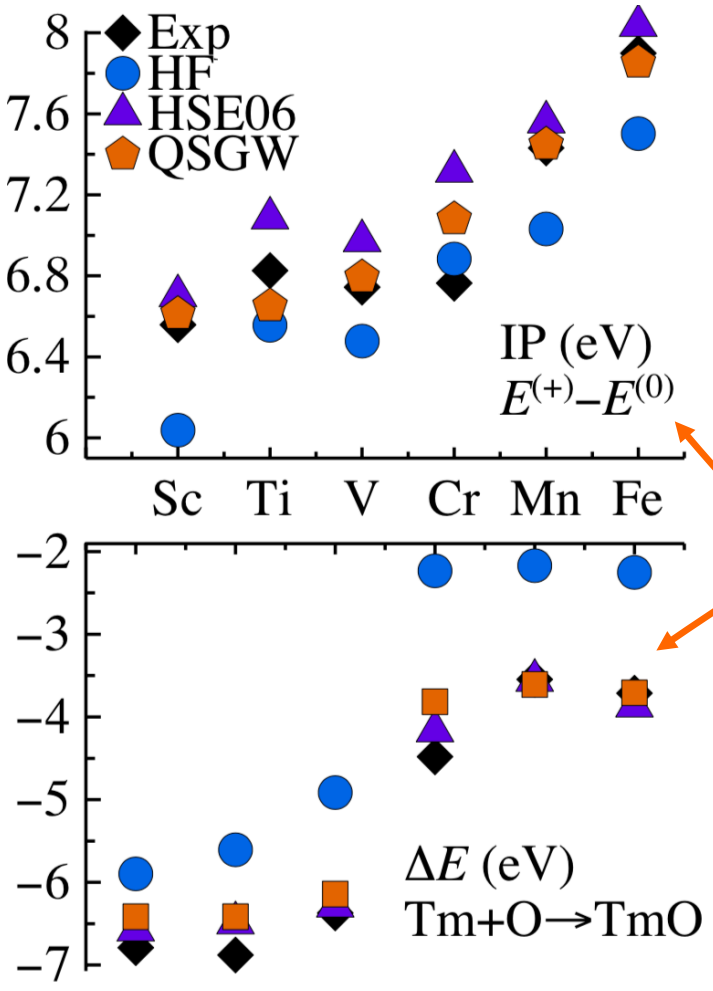
(Unconventional superconductors are usually bad metals)

✕ nonperturbative : low order diagrams not sufficient!

✔ The effective interaction is mostly site-local (DMFT)

Ambiguities in GW from starting point

GW is true *ab initio* (unlike many extensions to the LDA),



but *GW* is perturbation around H_0

Ambiguities through choice of H_0
⇒ not really *ab initio* any more.

Example: TM & TM-O dimatomic molec.

From RPA total energy calculate:

Ionization potential

Tm-O heat of reaction

Compare three choices for starting H_0:

• Hartree Fock • HSE06 • QSGW

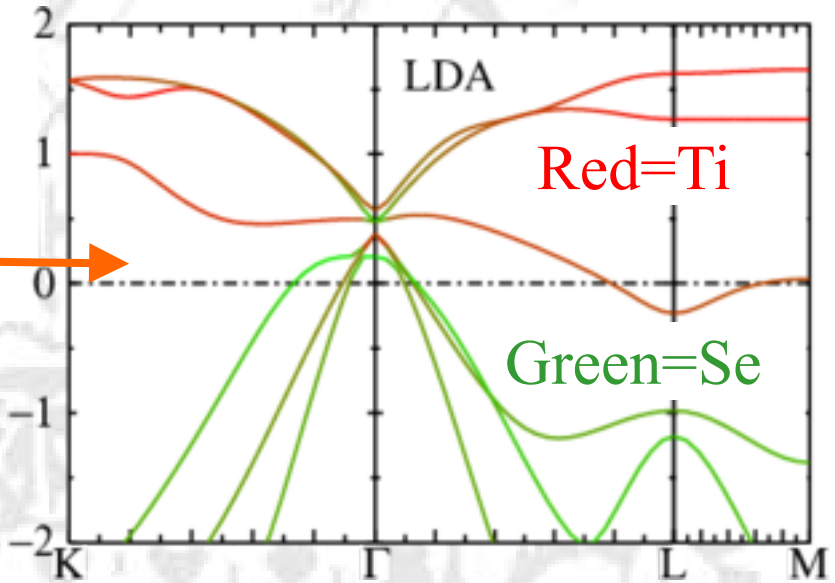
Arbitrariness gives freedom to improve result, but not universal or predictive

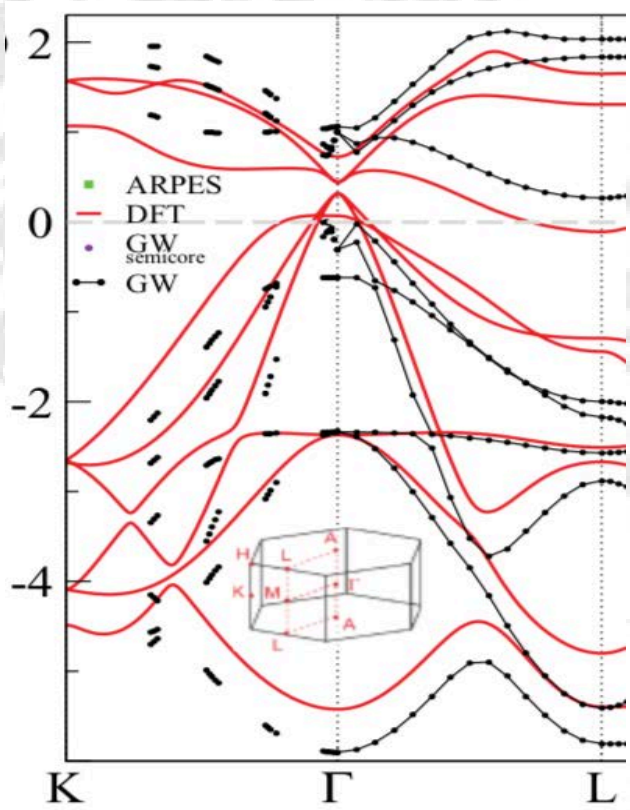
Why Self-Consistency Matters: $TiSe_2$ Case Study

At RT, $TiSe_2$ has a simple unit cell.

Band gap is not reliably known, but thought to be < 0.05 eV.

LDA predicts inverted gap





Cazzaniga et al PRB 85 '12 added *GW* to LDA ($G^{LDA}W^{LDA}$). Result : an insulator with a gap ~0.5 eV … suggests usual problem with LDA

QP Renormalization by density

Turns out that the positive gap is an artifact of $G^{LDA}W^{LDA}$!

LDA eigenfunctions ψ should be different from *GW*.

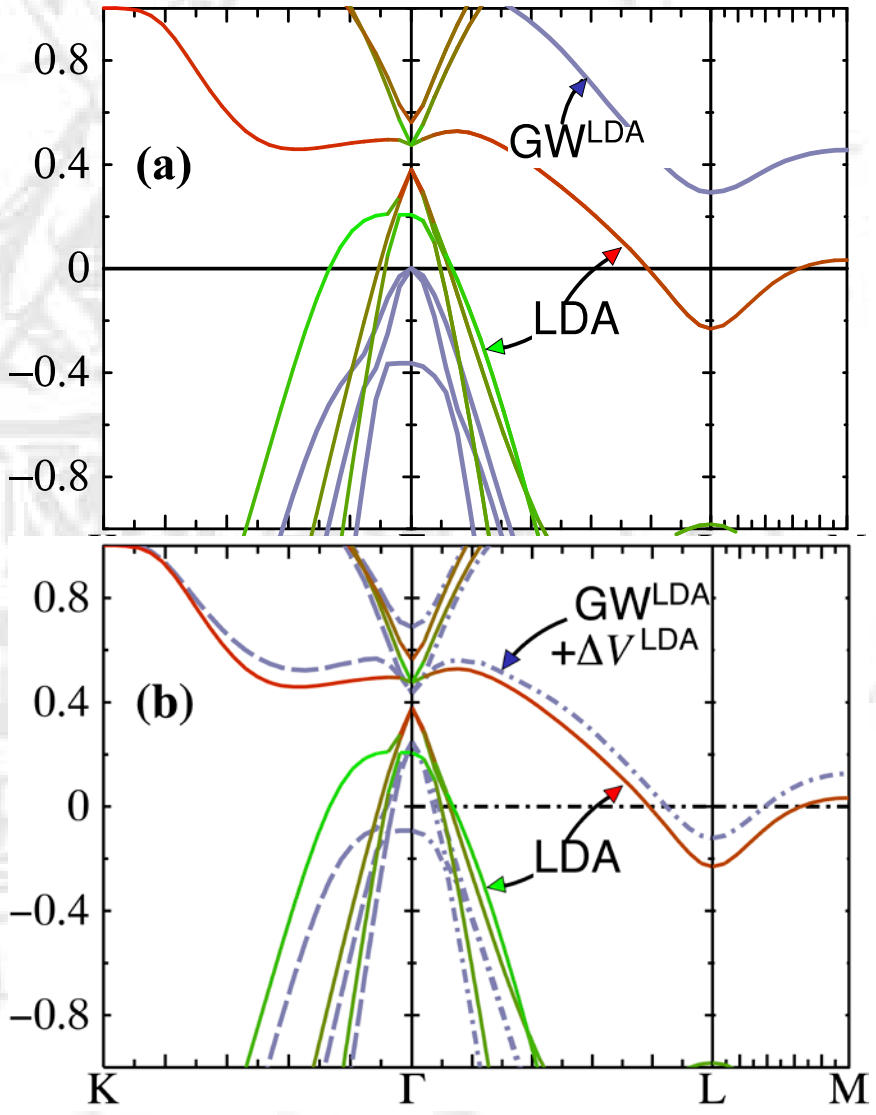
Off-diagonal self-energy $\Sigma^{nn'}$ modifies ψ and density $n(\mathbf{r})$ and V (requires full matrix Σ)

Simple ansatz: assume LDA adequately yields $\delta V/\delta n$. The potential becomes

$$\Sigma - V_{\mathrm{Hxc}}{}^{\mathrm{LDA}}[n^{\mathrm{LDA}}] + V_{\mathrm{Hxc}}{}^{\mathrm{LDA}}[n^{\mathrm{GW}}]$$

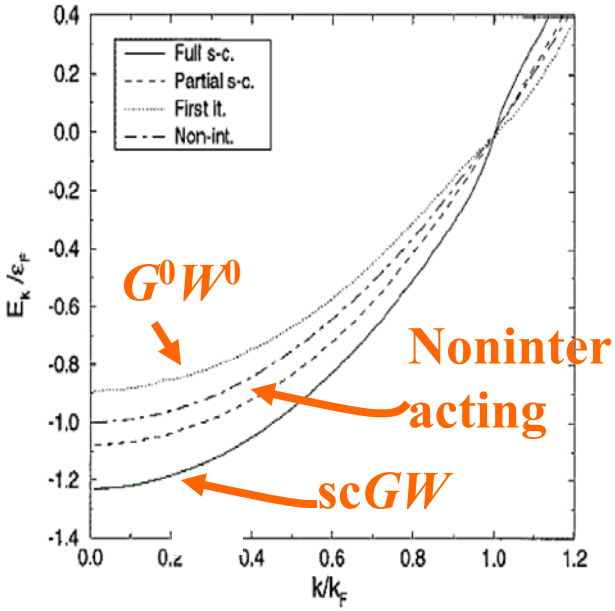
Iterate to self-consistency.

Gap becomes negative again!



True self-consistent GW

$$G \Rightarrow P = -iGG \Rightarrow W = \varepsilon^{-1} v \Rightarrow \Sigma = iGW \Rightarrow G = \frac{1}{\omega - (T + V^H + V^{ext} + \Sigma_{xc})}$$



Starting-point dependence can be surmounted by making *G* self-consistent

→ Based on Luttinger-Ward functional.

→ Keeps symmetry for *G*

→ Conserving approximation

But ...

ε strongly violates *f* sum rule [Tamme, PRL '99]

P loses its usual meaning as derivative $\delta n/\delta V$

... And it is poor in practice, even for the electron gas

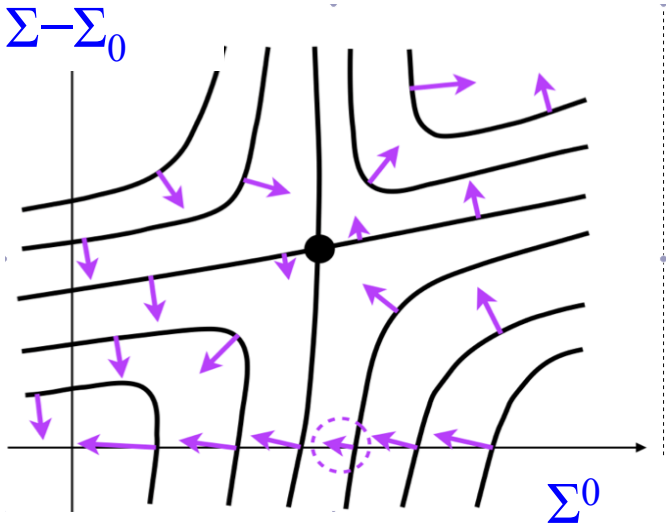
B. Holm and U. von Barth, PRB57, 2108 (1998). The *scGW* bandwidth *widens* by ~20% when it should narrow by 10% (30% error)

Quasiparticle self-consistecy

Better to perform GW (GŴ) around some noninteracting H_0.

How to find the best possible H_0 ?
Requires a prescription for minimizing the difference ΔV between H_0 and the full H :

$$\Delta V = H - H_0 = G^{-1} - G_0{}^{-1}$$



Quasiparticle Self-Consistent *GW* : a self-consistent perturbation theory where self-consistency determines the best H_0 within the GW approximation (or within GŴ)

- Surmounts starting point dependence
- Optimal construction for a given level of theory
- Uniform discrepancies w/ expt, their origin transparent

Prescription for Optimal G_0

Start with some trial V_{xc} (e.g. from LDA, or …). Defines G_0 :

$$H_0 = \frac{-1}{2m}\nabla^2 + V^{\text{ext}}(\mathbf{r}) + V^{\text{H}}(\mathbf{r}) + \Sigma_0(\mathbf{r},\mathbf{r}')$$

Analog of LDA V^{xc}

$$H_0\psi_i = E_i\psi_i \longrightarrow G_0(\mathbf{r},\mathbf{r}',\omega) = \sum_i \frac{\psi_i(\mathbf{r})\psi_i^*(\mathbf{r}')}{\omega - E_i}$$

GWA determines ΔV and thus H :

$$G_0 \xrightarrow{RPA} \varepsilon(iG_0G_0) \xrightarrow{GWA} \Sigma(\mathbf{r},\mathbf{r}',\omega) = iG_0W; \qquad \Delta V = \Sigma - \Sigma_0$$

Find a new Σ_0 that minimizes norm N, a measure of $\Delta V\, G_0$.

$$\Sigma_0 = \frac{1}{2}\sum_{ij}\langle\psi_i|\text{Re}\Big(\Sigma(E_i) + \Sigma(E_j)\Big)|\psi_j\rangle$$

(approximate) result of min N

Iterate to self-consistency.

At self-consistency, E_i of G matches E_i of G_0 (real part).

How to assess importance of what is left out?

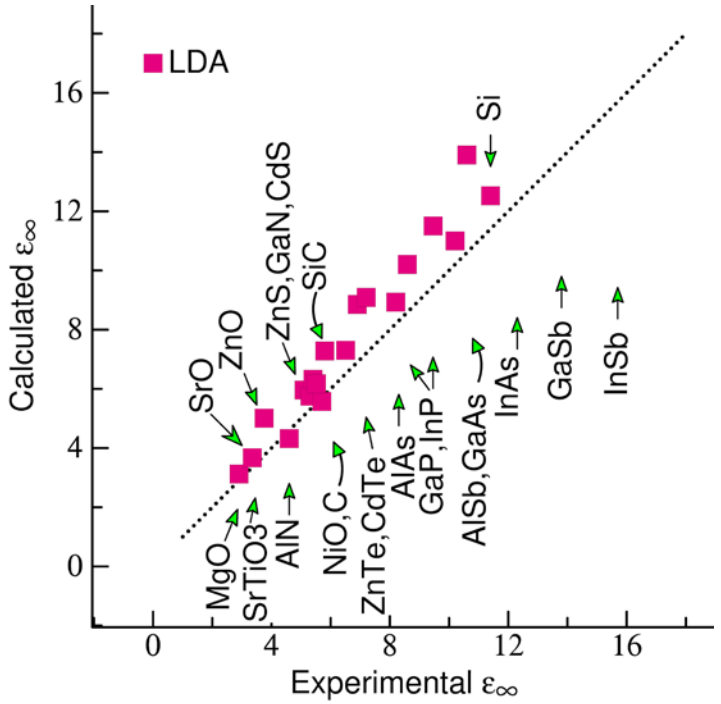
Ambiguities in H_0 means cannot *systematically* determine effects of omitted diagrams in *GW*

Classic example: RPA description of ε_∞ when $H_0 = H_{\mathrm{LDA}}$

H_{LDA} *underestimates* bandgaps ⇒ ε_∞ should be *too large*

RPA misses electron-hole attraction ⇒ ε_∞ should be *too small*

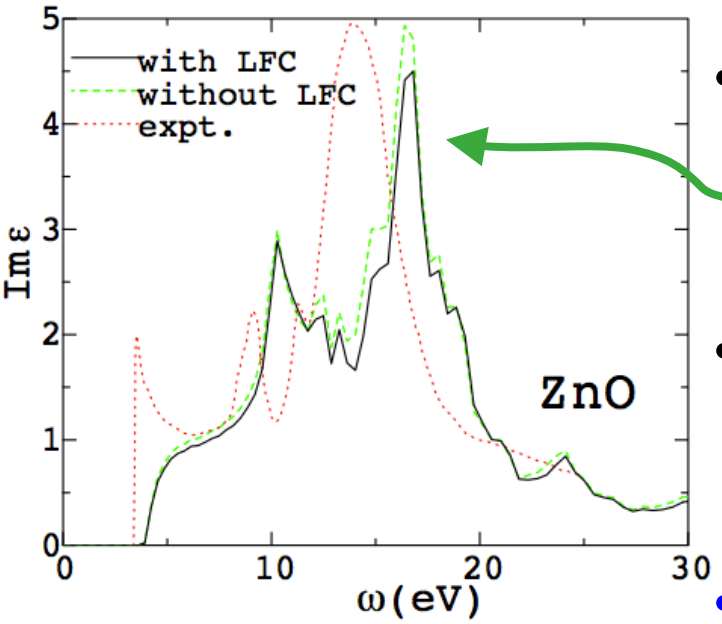
Errors approximately cancel in simple semiconductors ⇒ ε_∞ is fortuitously well described



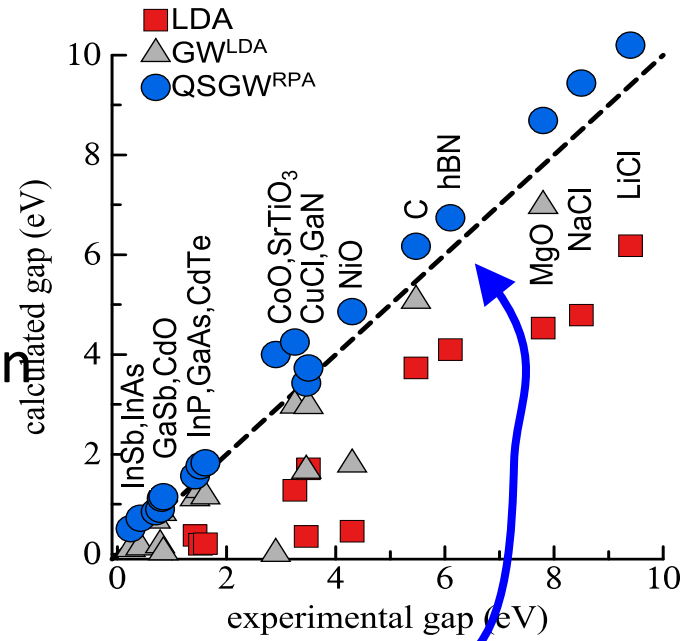
sp

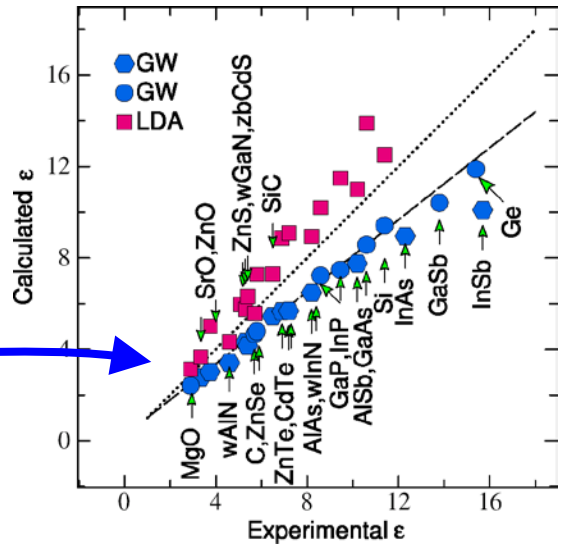
Electronic structure, QSGW

Absent significant spin fluctuations, how well does QSGW predict spectral properties ?



- Plasmon peaks in Imε(ω) are blue shifted
- Bandgaps are systematically too large
- ε_∞ is universally underestimated by a factor 0.8





Missing diagrams in W

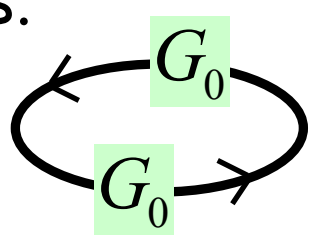
Kramer's Kronig relates real and imaginary parts of ε :

$$\Delta \,\mathrm{Re}\,\chi_1(0) = \frac{1}{\pi}\int_{-\infty}^{\infty}\left[\frac{\delta(\omega'-\omega_{\mathrm{th}})}{\omega'} - \frac{\delta(\omega'-\omega_{\mathrm{exp}})}{\omega'}\right]d\omega' = \frac{1}{\pi}\left(\frac{1}{\omega_{\mathrm{th}}} - \frac{1}{\omega_{\mathrm{exp}}}\right) < 0$$

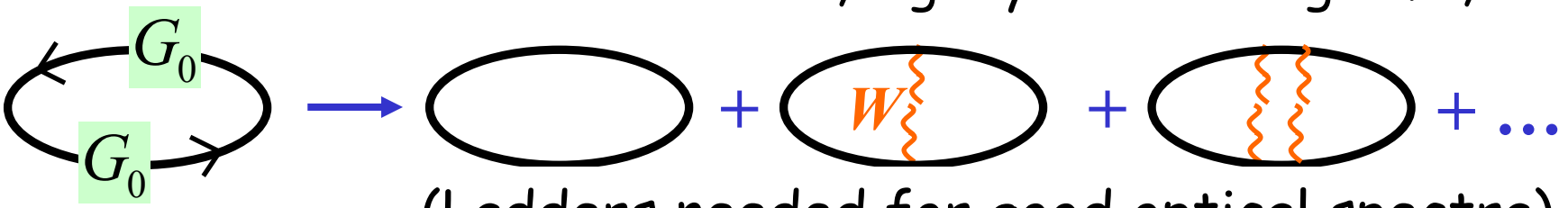
ε_∞ too small because of *blue shifts* in plasmon peaks.

GW uses *RPA approximation* for the polarizability $\Pi = iG_0 \times G_0$, and

$$W = \left(1 - \Pi v\right)^{-1} v = \varepsilon^{-1} v$$



But e^- and h^+ *are attracted* via W, e.g. by ladder diagrams,



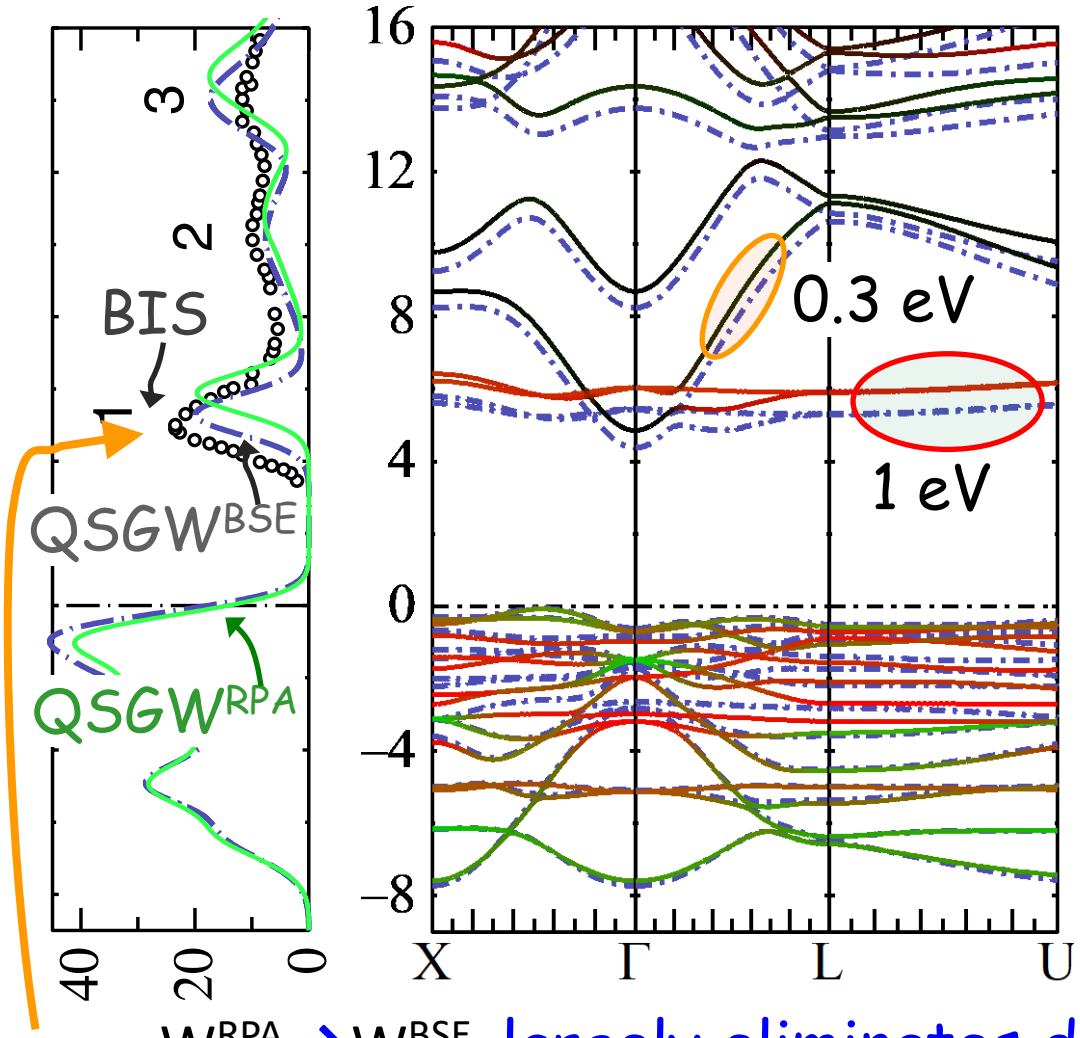
(Ladders needed for good optical spectra)

Conclusion: W calculated via RPA is *too large*, by 25% at ω=0.

Also if: GW is *too large* ⇒ bandgaps overestimated.

Compare QSGW^{RPA} ,QSGW^{BSE} bands to BIS in NiO

Brian Cunningham, M. Gruening added ladders to improve W.



NiO has both dispersive *sp* bands

peak +0.3 eV too high

… and a flat *d* band

1 eV too high

Effect on dispersive *sp* bands

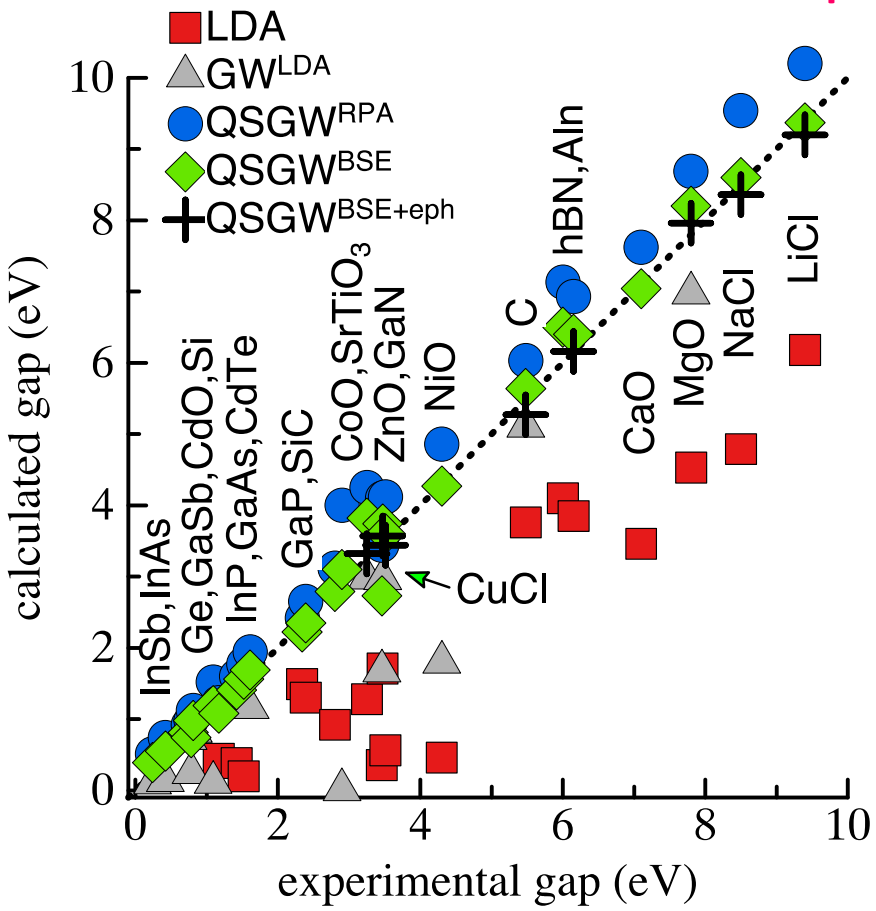
W^{RPA} → W^{BSE} ⇒ -0.3 eV shift

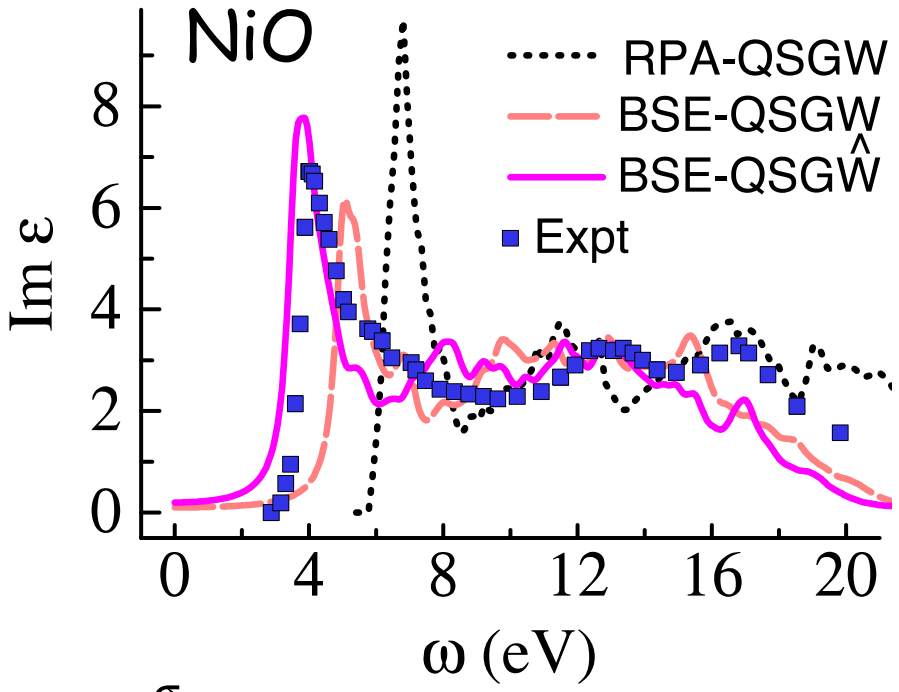
… *d* band W^{RPA} → W^{BSE} ⇒

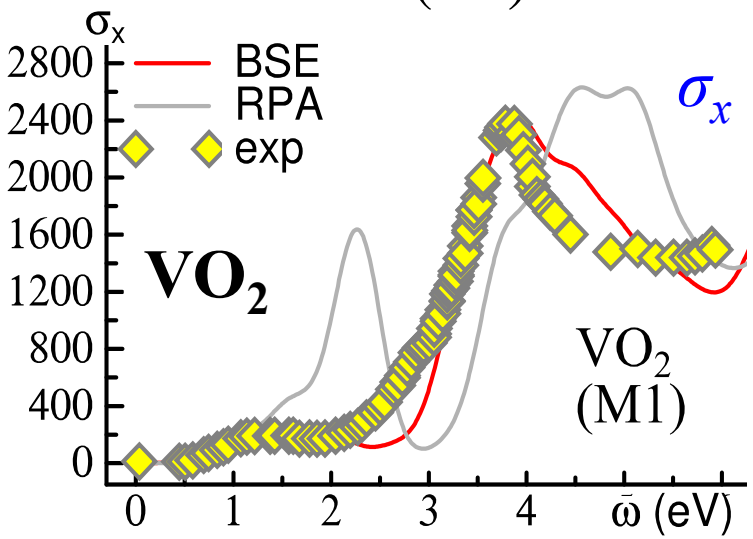
-1 eV shift

… W^{RPA} → W^{BSE} largely eliminates discrepancies in BIS

QSGW^{BSE} : Optical properties





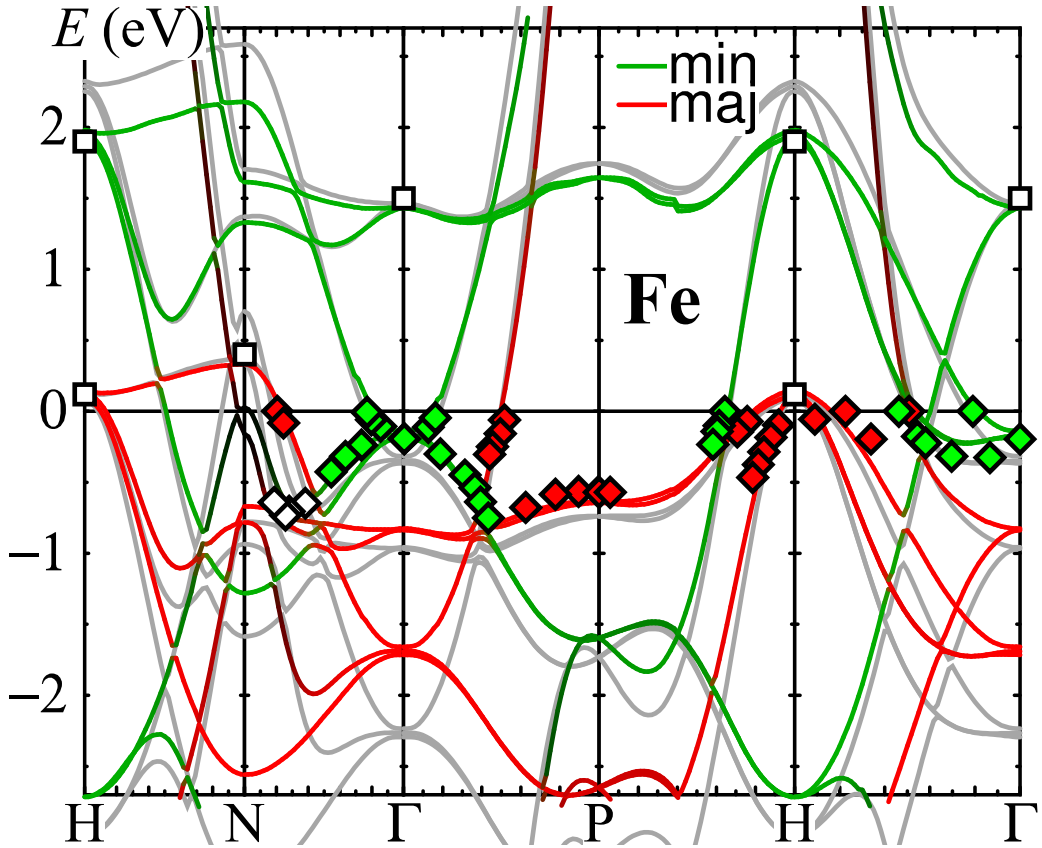


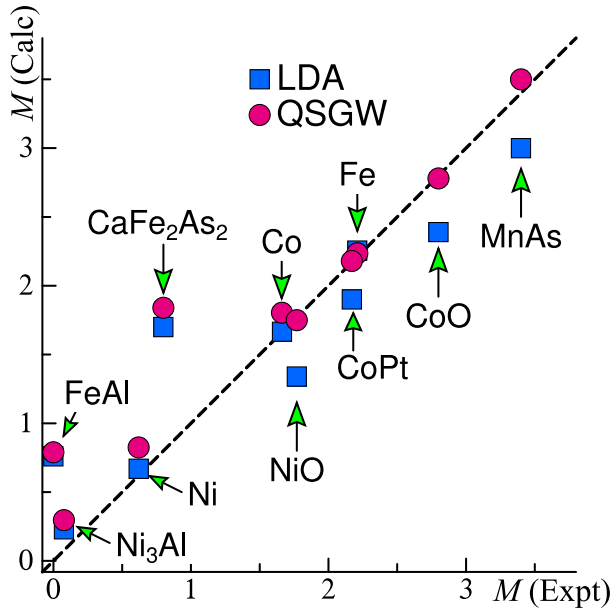
$W^{RPA} \rightarrow W^{BSE} = \hat{W}$ reduces systematic error in gaps and e(w)

$\Rightarrow$ Charge channel dominated by low-order diagrams

Fe: archetypal local-moment system

Weak spin fluctuations in local-moment magnetic systems, e.g. Fe, MnAs.





QSGW describes Fermi-liquid region of ARPES to within ~0.02 eV (including final-state and e-ph corrections)

Local moment system well described in Fe, but …

Phys. Rev. B 95, 041112(R)

Where QSGW breaks down

QSGW breaks down when dynamical spin fluctuations matter.

Archetypal example: compare M1 and M2 phases of VO_2 .

1. Phase transition monoclinic→rutile at 67C

2. Rutile: every V has 2 equal NN @2.89Å

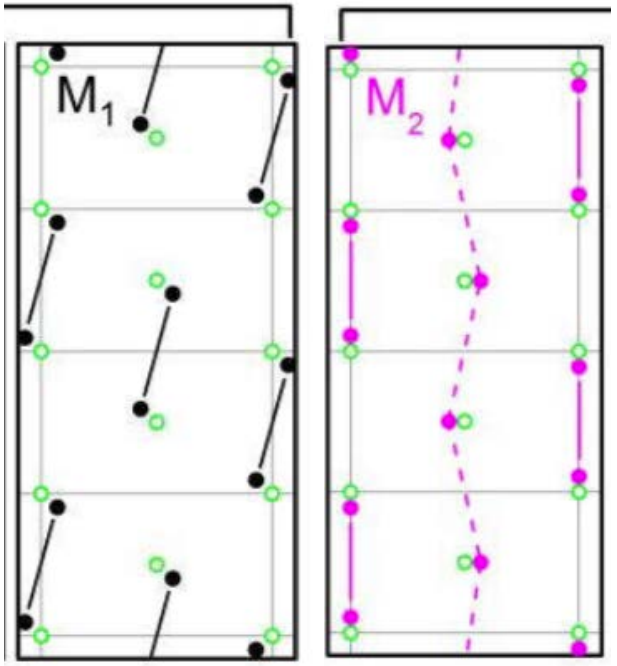
3. Monoclinic: unit cell (V_4O_8) is doubled

   The 4 V atoms pair into two dimers:

   The NN bonds split into @2.65Å, 3.13Å

4. Rutile is widely thought to be more or less a weakly correlated metal.

5. There is a *second* monoclininic phase M2 (metastable). One V-V pair dimerizes like M1; the other does not, similar to rutile.

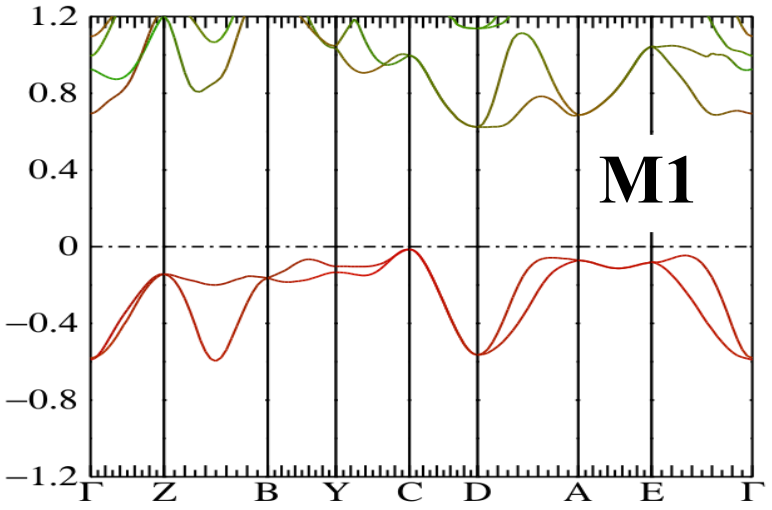


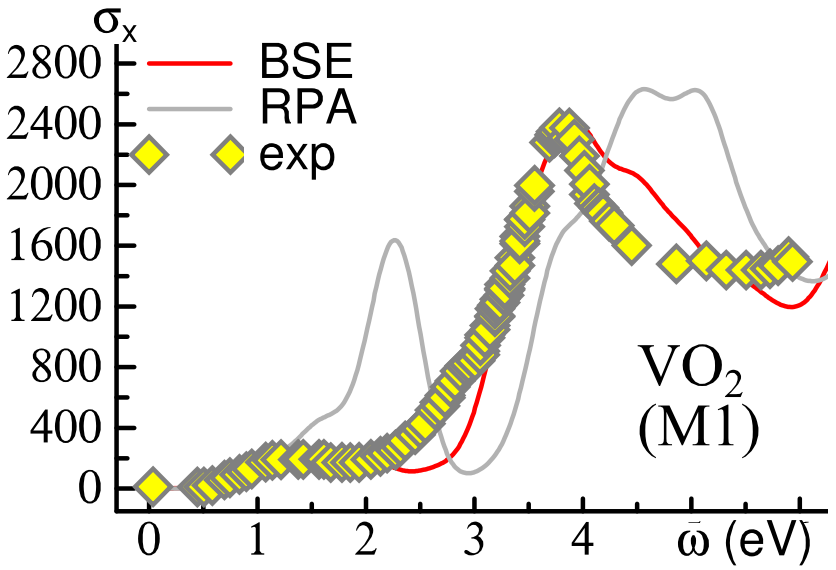
QSGW of VO_2 in M1 phase

QS*GW* first performed by Gatti et al (PRL 99, 266402 '07)

Nonmagnetic calculation yielded gap ~0.7 eV, similar to expt.

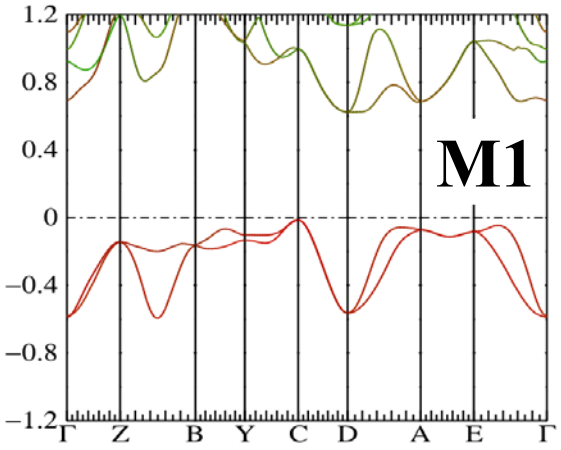
We find a similar gap, and also σ_x in excellent agreement with σ_x measured on single crystals---provided ladders are included.

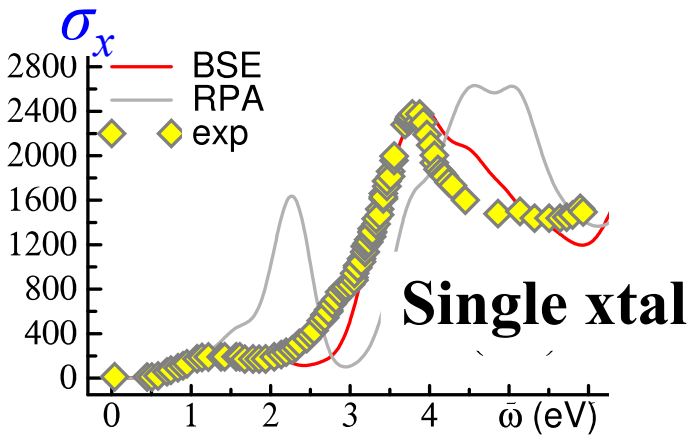


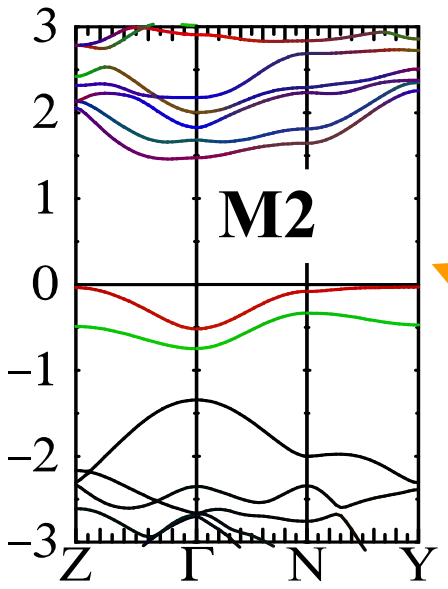


Mott gap would be associated with strong spin fluctuations --- not present here. Strong support for Peierls picture, not Mott

The M2 phase







Nonmagnetic QS*GW* predicts M2 to be metal, contrary to expt.

NonMagnetic ≠ Paramagnetic ~ antiferromagnetic. Mag QS*GW*:

Dimerized V pair: moment vanishes (like M1) but …

Undimerized V pair orders antiferromagnetically.

AFM causes a gap to form, slightly larger than M1 gap

(probably slightly overestimated, as no spin fluctuations.)

Paramagnetic like antiferromagnetic, but spins disordered.

DMFT can well describe the (disordered) PM state

Why not DFT+DMFT?

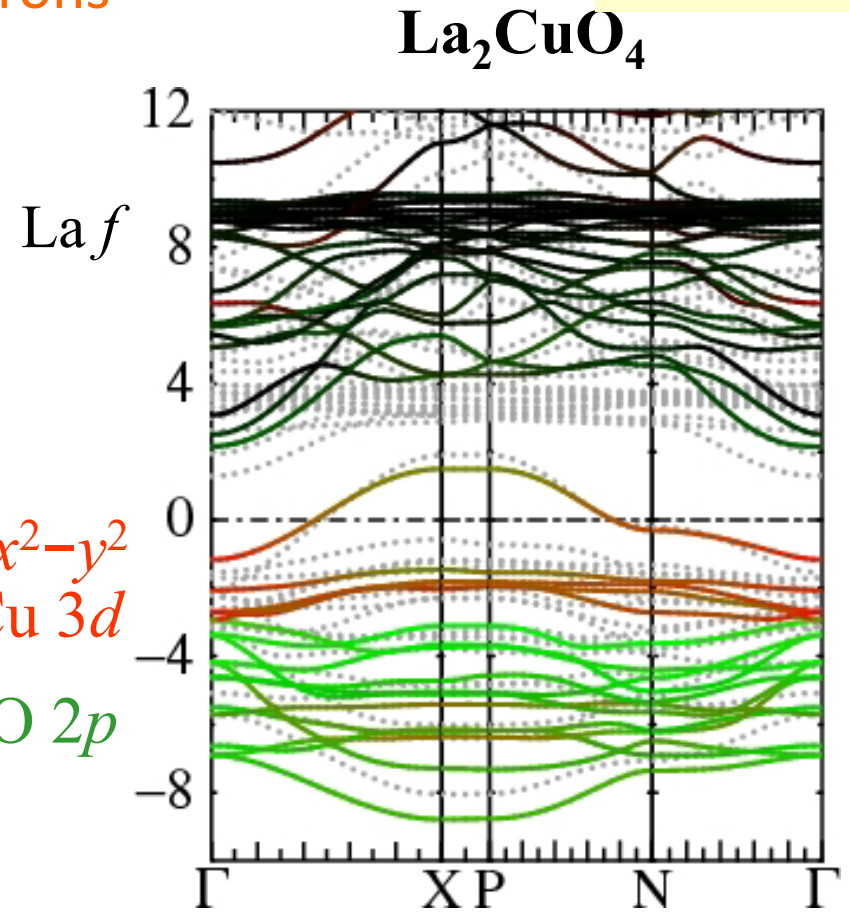
DFT → DFT+DMFT

✕ DFT is a poor reference for DMFT!

$V^{\mathrm{eff}}[n] \Rightarrow V^{\mathrm{eff}}$ is the same for all electrons

A zoo of patches (LDA+U, hybrids, van der Waals, …) but starting ansatz makes systematic improvement problematic.
Errors in

1-body description are propagated to DMFT.



✕ Most many-body phenomena are inherently nonlocal! χ^S, χ^C, superconductivity, quantum criticality, pseudogap …

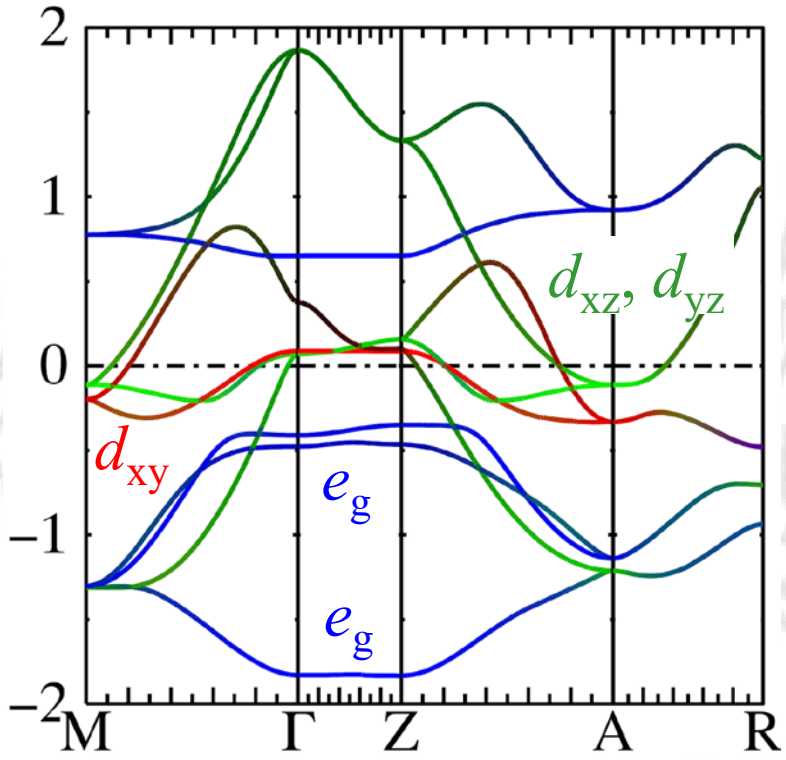
Nonlocalities in χ add a contribution to $\Sigma(\mathbf{k})$.

FeSe: Paramagnet with strong spin fluctuations

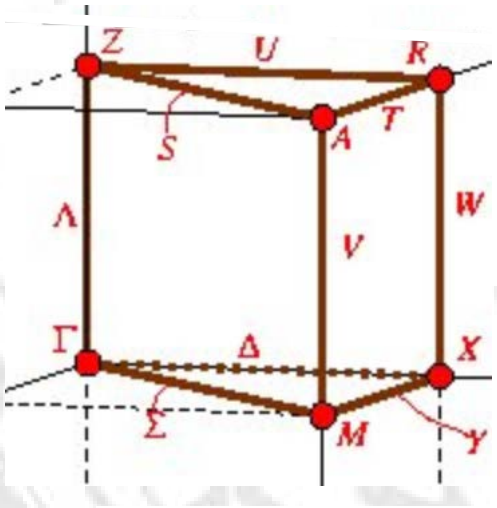
Experimental facts:

e^- pockets at M & A (d_{xz} and d_{yz})

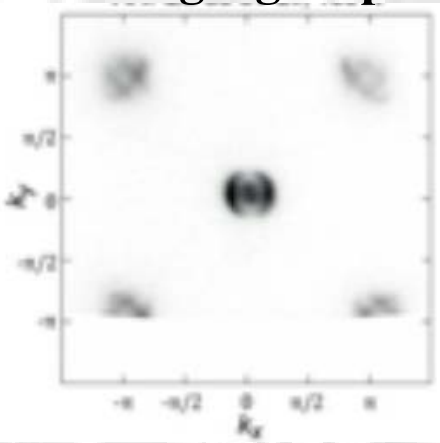
h^+ pockets at Γ & Z : (d_{xy} + d_{xz} + d_{yz})

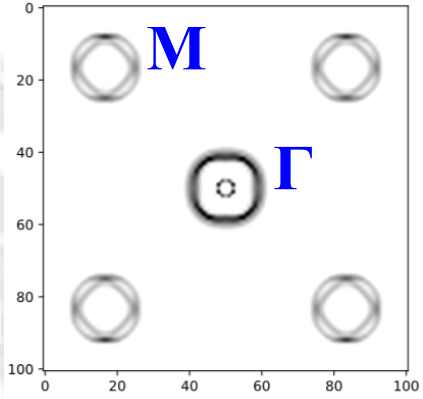


What does spin do?



ARPES from Eshrig's group





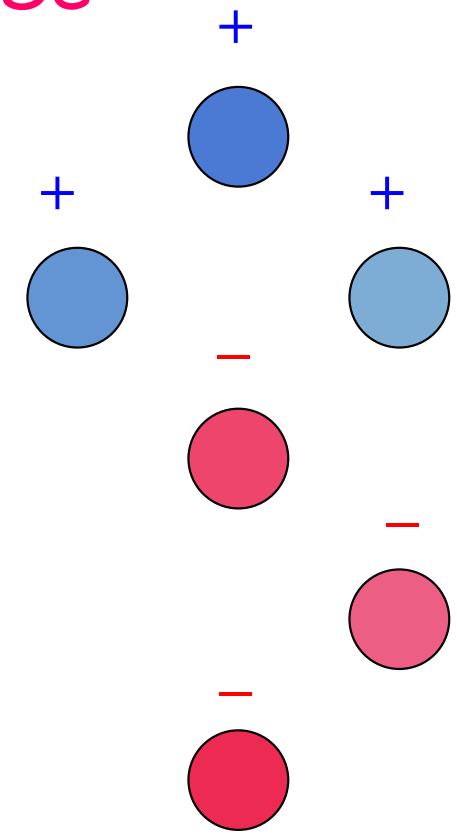
Nonmagnetic QSGW predicts:

- Pockets are too large
- ~100 meV discrepancy with ARPES in FL regime (later)

QSGW and Spin Fluctuations in FeSe

Simulate paramagnetism with SQS6 structure

A low moment QSGW solution can be stabilized with $\langle M \rangle = 0.2 \pm 0.15\ \mu_B$.



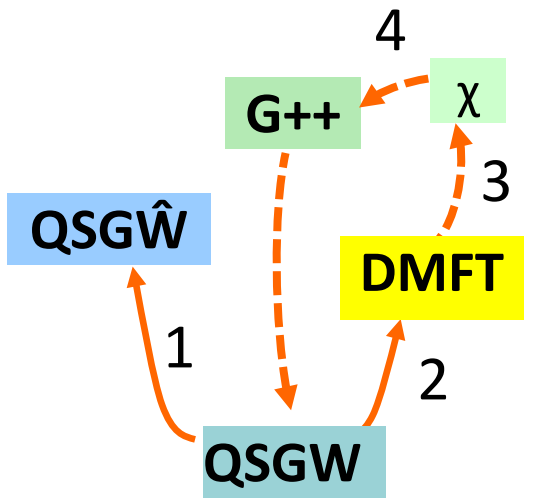
QP Levels shift towards ARPES data, but still a significant discrepancy.

	Γ		M		Z		A	
LDA,nm	+109	+113	−204	−337	+254	+141	−208	−582
QSGW,nm	+41	+44	−107	−202	+131	+56	−113	−334
SQS6	+45	(60)	−52	−70	+31	+68	−59	−72
ARPES	+9	−18	−22	−42	+7	+34	−16	−25

Shows spin fluctuations are important but (QS)*GW* does not adequately capture them

QSGW+DMFT+BSE + …

Path to tractable *ab initio* framework for strong correlations



Partition problem :

1. Charge fluctuations governed by long-range interactions … but they can be treated accurately with low-order perturbation theory (QSGW)

2. Spin-spin bare vertex U mostly between orbitals on one site (See C. Friedrich, Ni).

Solve local impurity problem with DMFT embedded in QSGW ⇒ G^{loc}.

Embed G^{loc} $[\Sigma^{DMFT}(\omega)]$ into bath ⇒ $G^{crys}(\mathbf{k};\omega)$. Iterate⇒self-consistent

3. $G^{loc} \Rightarrow G_2{}^{loc}$

$\Rightarrow \Gamma^{loc,q} + \Gamma^{loc,m}$

$\Gamma^{loc} + G(\mathbf{k};\omega) * G(\mathbf{k};\omega)$

⇒ susceptibilities

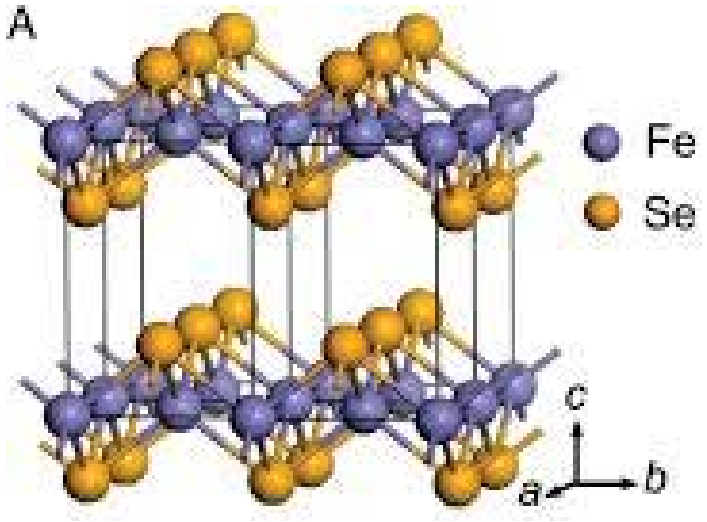
spin $\chi^s(k,\omega)$, and

charge $\chi^c(k,\omega)$

4. New step (for future!)

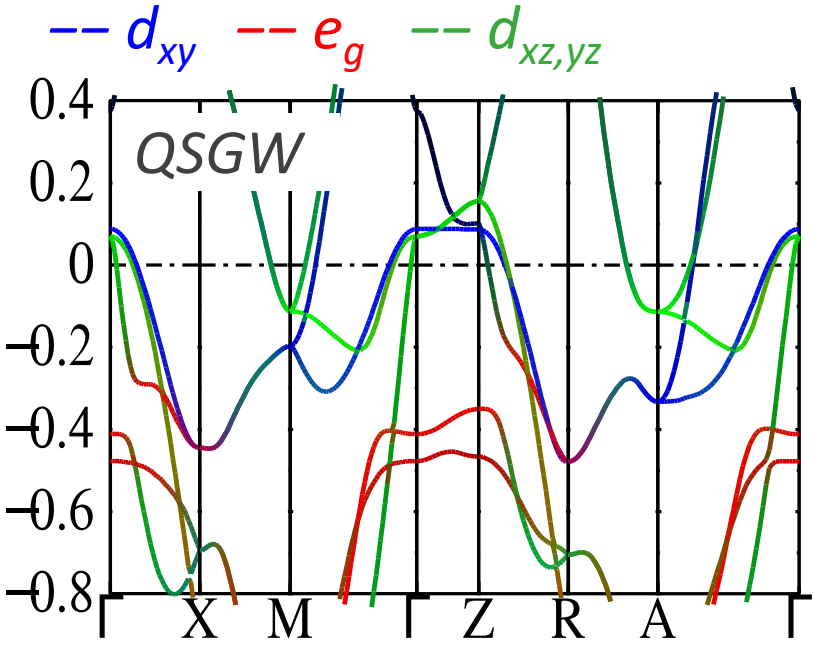
$\Gamma^{loc}\chi^s\Gamma^{loc} \Rightarrow \Sigma(k,\omega)$ (G^{++})

FeSe : spin fluctuations and superconductivity



Tetragonal: a=b≠c

States near E_F mostly of Fe *d* character. Largely 2D FeSe sheets weakly coupled along *z*



M: 2 e^- pockets made of $d_{xz,yz}$

Γ: 2 h^+ pockets made of $d_{xz,yz}$

1 h^+ pocket made of d_{xy}.

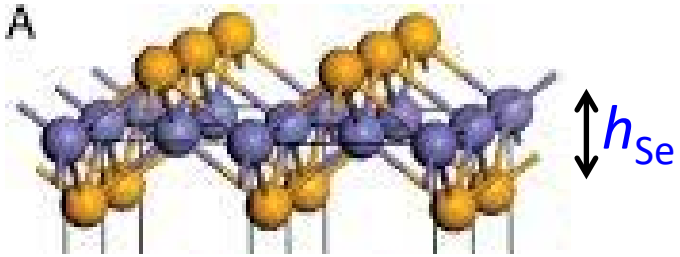
d_{xy} *is* the crucial ~~band~~ to watch out for! Very sensitive to h_{Se}.

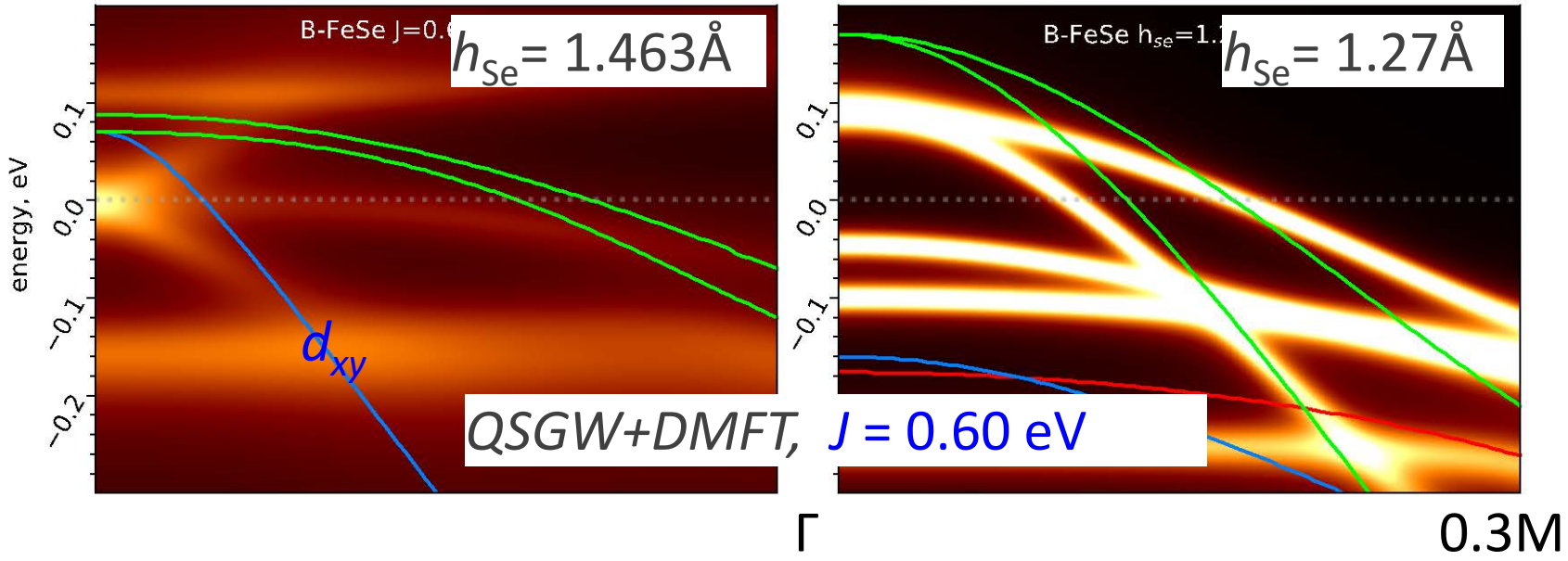
QSGW pockets are a little larger than experiment.

d_{xy} probably should not cross E_F

Proximity d_{xy} to E_F

Proximity of d_{xy} to E_F is very sensitive to geometry, e.g. the height h_{Se} above the Fe plane





When d_{xy} is very near E_F ⇒ "bad metal," lots of incoherence.

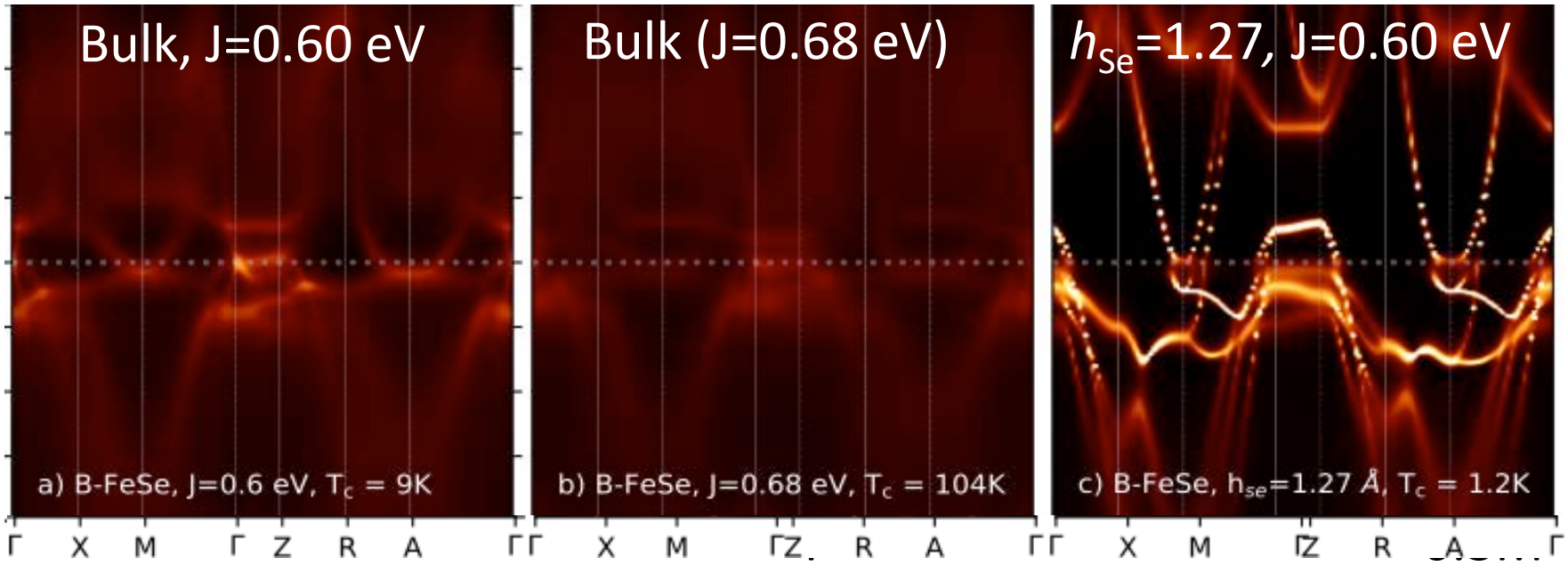
Reduce h_{Se} ⇒ d_{xy} is pushed below E_F ⇒ more coherent

T_c predicted by the theory drops from 9K to 1.2K

Role of Hundness

DMFT requires Hubbard *U* and *J* for the effective interaction

U, *J* = HF "bare" direct and exchange, but screened by the bath.
Calculated in RPA : yields *U*=3.5 eV, *J*=0.60 eV.

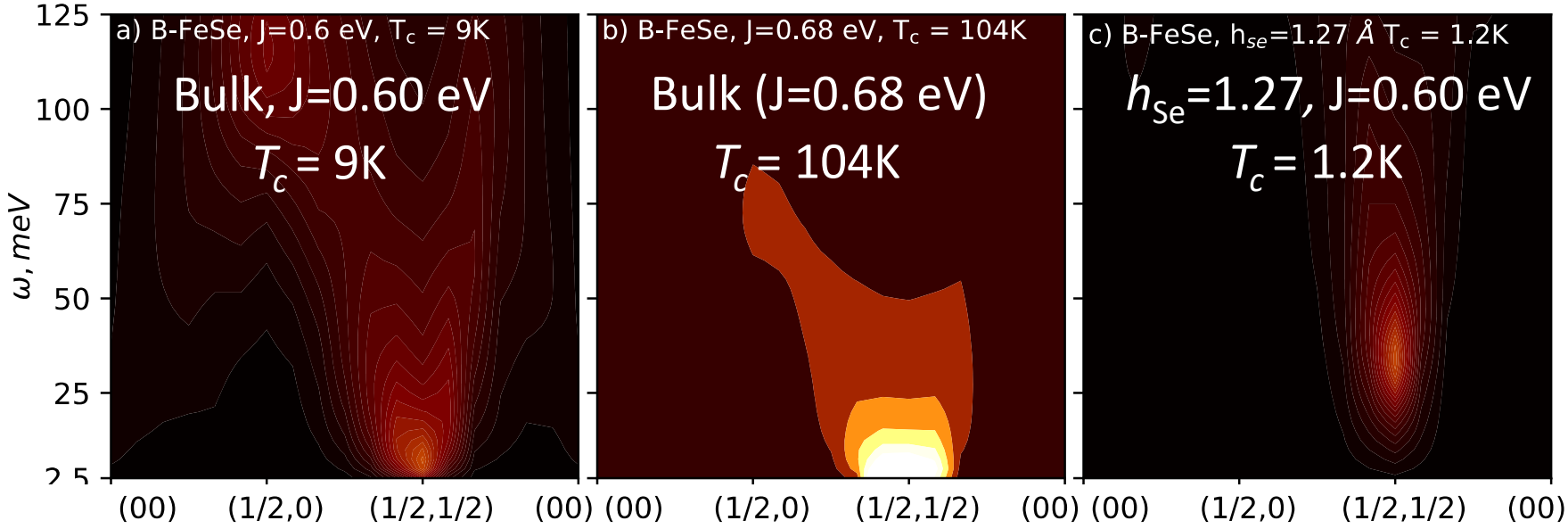


Treat J as parameter. Modest, gradual evolution for 0 < J < 0.6

For J > 0.6, incoherence is hugely sensitive to *J* !

Hundness and Spin susceptibility

Incoherence has a huge effect on spin susceptibility $\chi^s(\boldsymbol{q},\omega)$



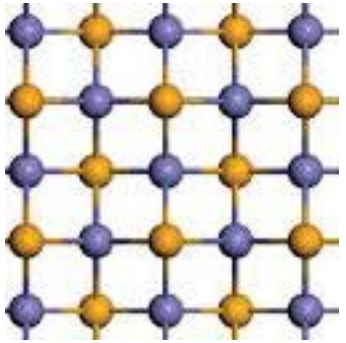
$\chi^s(\boldsymbol{q},\omega)$ is the glue (boson) that drives superconductivity.

Peak @5 meV near the AM **q** vector (1/2,1/2)

Becomes very intense J = 0.6 → 0.66

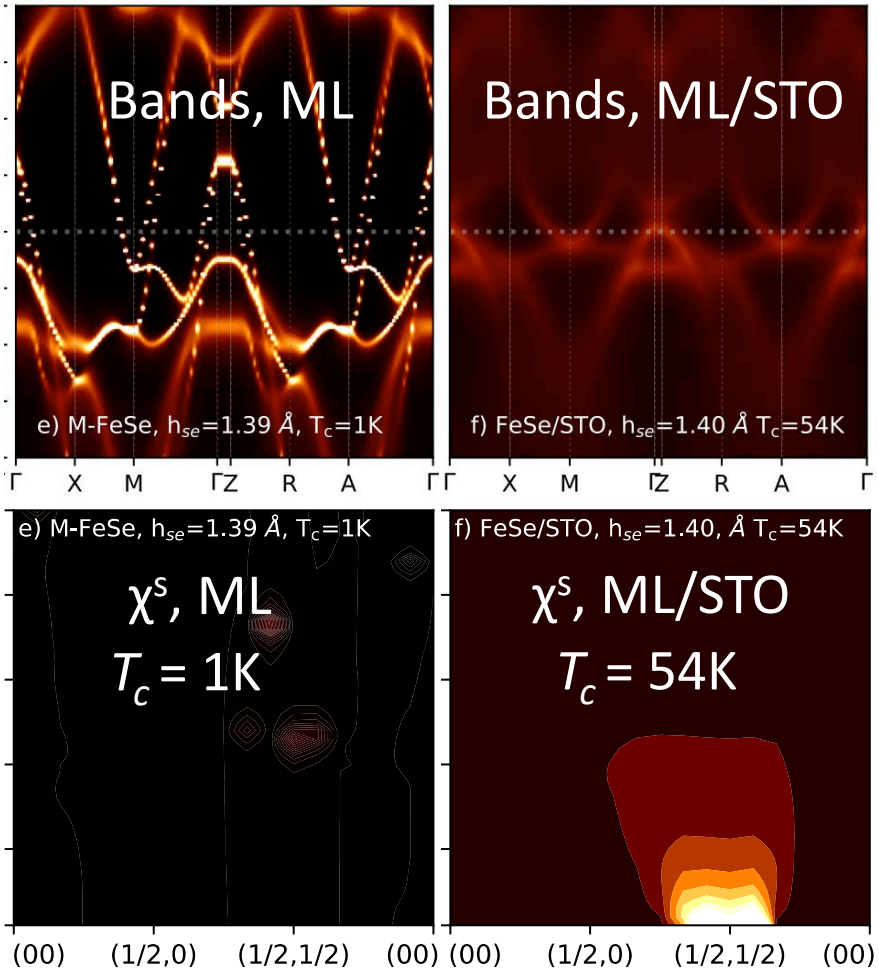
Opposite trend: with h_{Se} = 1.463 → 1.27

peak weakens, shifts to 30 meV,



Fe on $SrTiO_3$

Big story: as a single monolayer on $SrTiO_3$, FeSe has Tc≈80K!



Q: Why?

A: screening reduced: CRPA predicts *J* to increase from 0.60 eV to the magic 0.66 eV

Yet… d_{xy} is pushed down in free-standing ML. Weak Spin fluctuations ⇒ low T_c.

But then … the $SrTiO_3$ substrate pushes d_{xy} near E_F once more. Incoherence restored in d_{xy} ⇒ intense spin fluctuations. T_c rises dramatically.

Conclusions

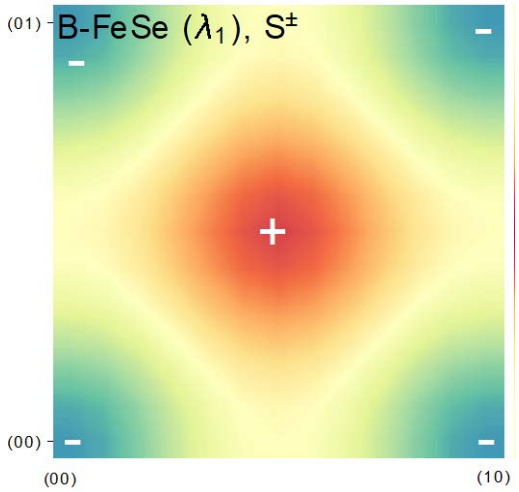
1. Many-Body Perturbation Theory (GW++)

Low-order, but no partitioning, real axis

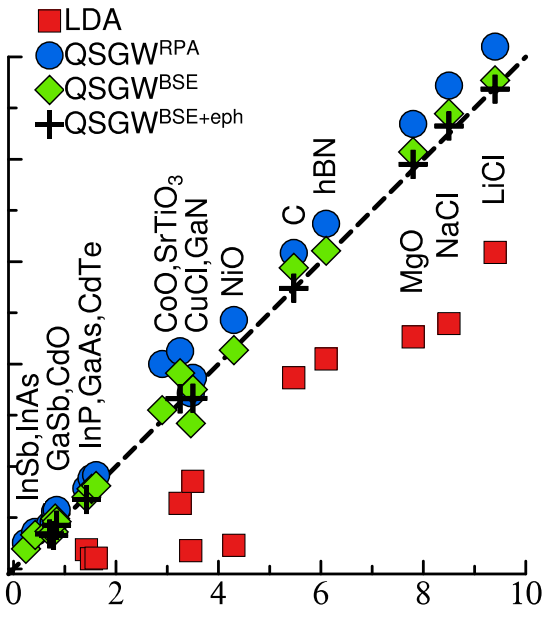
QS*GW* resolves starting point ambiguity; vastly improves consistency, reliability.

QS*GW*→QS*GŴ* removes systematic errors

Excellent for charge fluctuations, not spin



Explains why T_c changes FeSe(s) → FeSe/SrTiO3



2. Dynamical Mean Field Theory for spin fluctuations. QSGW+local diagrams seems able to describe 1-particle quantities remarkably well;

QSGW+DMFT+BSE seems to provide solid path to predict properties of strongly correlated systems ab initio

Contributors to this work

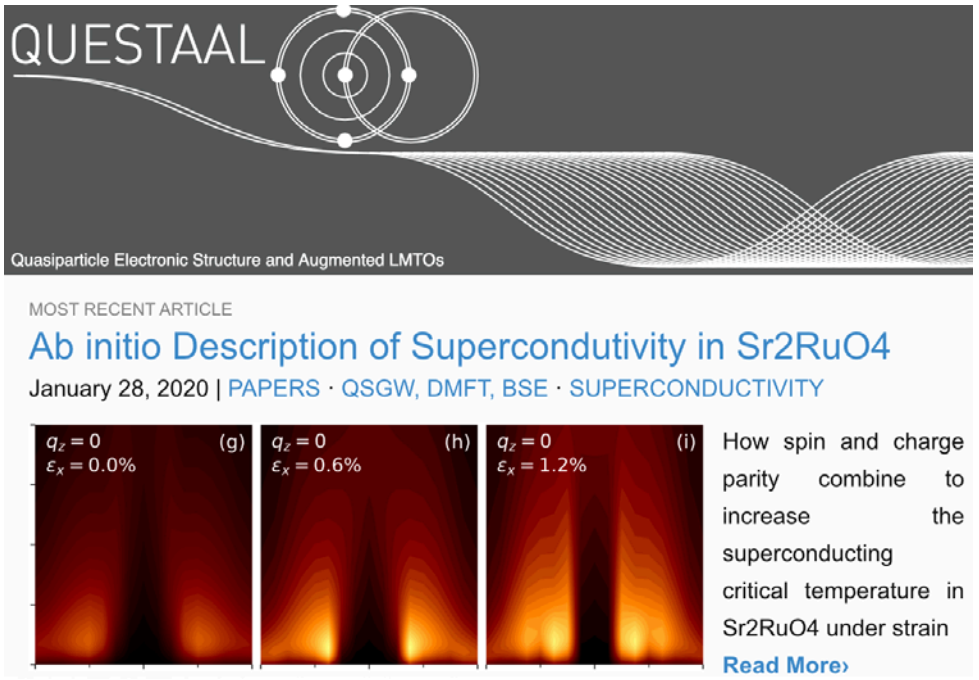
Dimitar Pashov

Myrta Gruening

Brian Cunningham

Francois Jamet

Mainly the work of Swagata Acharya



NREL/PR-5900-82029

Code is free to anyone!.

https://www.questaal.org/

Looking for 2 postdocs at to work at NREL … please contact me!