Faster G_0W_0 implementation
for more accurate material design

Jonathan Laflamme Janssen[1,2], Bruno Rousseau[1], and Michel Côté[1]
[1]*Université de Montréal, Canada* and *Regroupement Québécois pour les Matériaux de Pointe (RQMP)*
[2]*Université Catholique de Louvain, Louvain-la-Neuve, Belgium* and *European Theoretical Spectroscopy Facility (ETSF)*

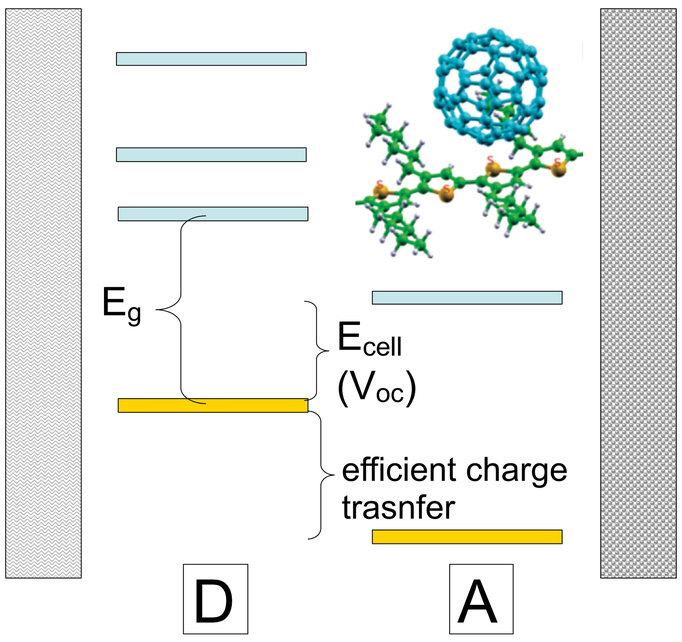
Abstract

Density-functional theory (DFT) is currently the *ab initio* method most widely used to predict electronic energy levels of new molecules. However, approximations intrinsic to the theory limit the accuracy of calculated energy levels to about ±0.5 eV. More efficient theoretical design of molecules and materials could be achieved if more precise methods were available. The G_0W_0 approach is an *ab initio* method that provides such an enhanced precision, with predicted energy levels accurate to about ±0.05 eV. However, such calculations are currently prohibitive for systems with more than a few tens of electrons, thus limiting their use in many applications. What limits calculations to this system size in traditional plane waves (PW) implementations of G_0W_0 is the need to invert the dielectric matrix in a large PW basis and the need to carry out summations over a large number of conduction states. This poster presents a strategy to avoid both of these bottlenecks.

Material design : organic photovoltaïcs

Criteria :
- Gap : optimal for solar spectrum
- Ionisation Energy (IE) :
  - lower than C_{60} (for good charge transfer)
  - high enough for good V_{oc} and air stability

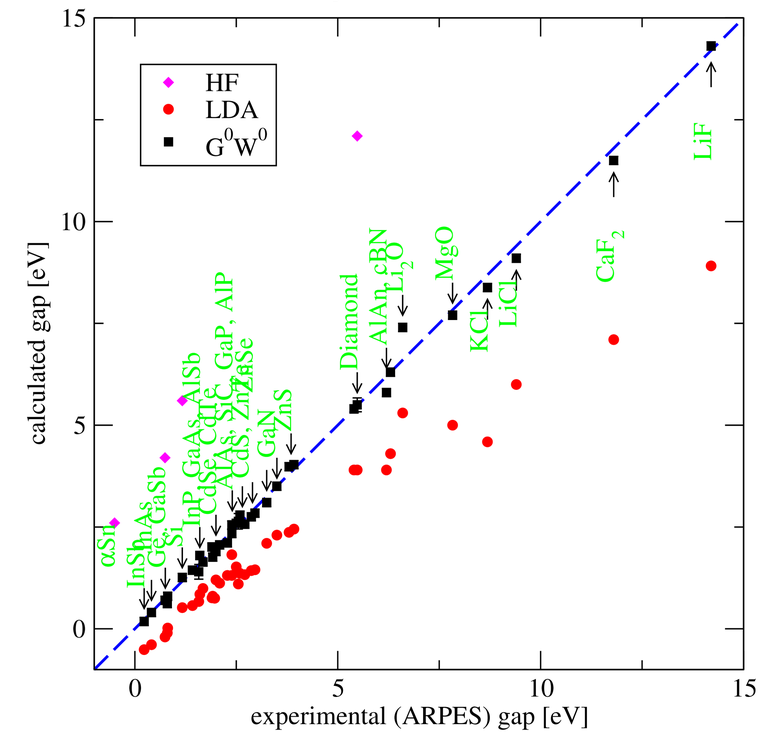
Morale: Precise electron removal and addition energies are crucial in some applications



The G_0W_0 method

DFT levels offer quick sorting of candidate materials.
Further refinement is still desirable given the ressources and time required to synthesize a new material.
G_0W_0 can provide such results

Accuracy of DFT and G_0W_0



In G_0W_0, DFT eigenstates and eigenvalues are used as a starting point :

$$\left(\hat{T}+\hat{V}_{ext}+\hat{V}_{xc}\right)|n\rangle=\varepsilon_n|n\rangle$$

They are then corrected to first order in perturbation theory using the GW exchange-correlation operator $\Sigma(\omega)$:

$$\Delta\varepsilon_n \approx \varepsilon_n+\langle n|\hat{\Sigma}(\varepsilon_n+\Delta\varepsilon_n)-\hat{V}_{xc}|n\rangle$$

The expectation value of $\Sigma(\omega)$ being :

$$\langle n|\hat{\Sigma}(\omega)|n\rangle=\frac{i}{2\pi}\sum_{n'}\int_{-\infty}^{+\infty}d\omega\,\frac{\langle n|\langle n'^{*}|\hat{\epsilon}^{-1}(\omega)\hat{v}|n'^{*}\rangle|n\rangle}{\omega-(\varepsilon_{n'}-\varepsilon_n)}$$

where $\hat{v}$ is the coulomb operator and where :

$$\hat{\epsilon}(\omega)=\hat{1}-\hat{v}\hat{P}(\omega)$$

$$\hat{P}(\omega)=\sum_{c,v}|c^{*}\rangle|v\rangle\left[\frac{1}{\omega-(\varepsilon_c-\varepsilon_v)}-\frac{1}{\omega+(\varepsilon_c-\varepsilon_v)}\right]\langle v|\langle c^{*}|$$

- O : bottleneck #1 : sum over conduction states
- O : bottleneck #2 : inversion of the dielectric matrix

Solution to bottleneck 1
Sternheimer equation

Idea : transform $\sum_c$ into a linear equation problem...

$$\hat{P}(\omega)|\psi\rangle=\sum_{c,v}|c^{*}\rangle|v\rangle\left[\frac{1}{\omega-(\varepsilon_c-\varepsilon_v)}-\frac{1}{\omega+(\varepsilon_c-\varepsilon_v)}\right]\langle v|\langle c^{*}||\psi\rangle$$
$$\equiv\sum_v|f_v^{+*}v\rangle+|f_v^{-*}v\rangle$$

$$\Rightarrow|f^{\pm}{}_v\rangle=-\sum_c\frac{|c\rangle\langle c|}{\varepsilon_c-\varepsilon_v\pm\omega}|\psi^{*}\rangle|v\rangle$$
$$=-\sum_c\frac{|c\rangle\langle c|}{\hat{H}-\varepsilon_v\pm\omega}|\psi^{*}\rangle|v\rangle$$
$$=-\frac{\hat{P}_c}{\hat{H}-\varepsilon_v\pm\omega}|\psi^{*}\rangle|v\rangle$$

Eliminating $\sum_c$

$$\Rightarrow\left(\hat{H}-\varepsilon_v\pm\omega\right)|f^{\pm}{}_v\rangle=-\hat{P}_c|\psi^{*}\rangle|v\rangle \quad \text{(Sternheimer's equation)}$$

Implementation of solution

1) H + FFTs = N logN op. ⇒ iterative method

2) H - $\varepsilon_v \pm \omega$ can be singular ⇒ SQMR instead of CG

Solution to bottleneck 2
Lanczos algorithm

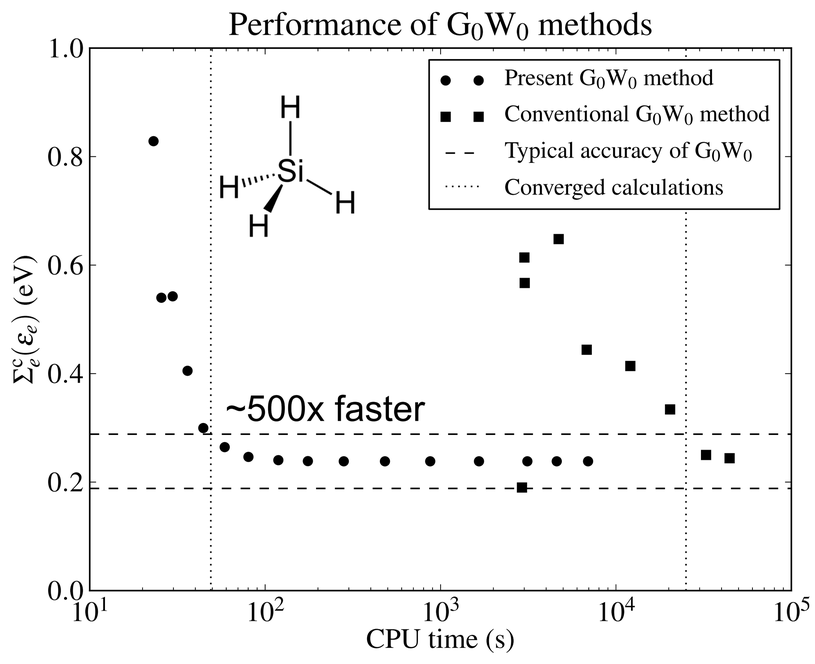
Usually, $\hat{\epsilon}(\omega)$ is expressed in a planewave basis.
Here, we decrease the size of the matrix by constructing a basis that automatically focuses on the relevant subspace :

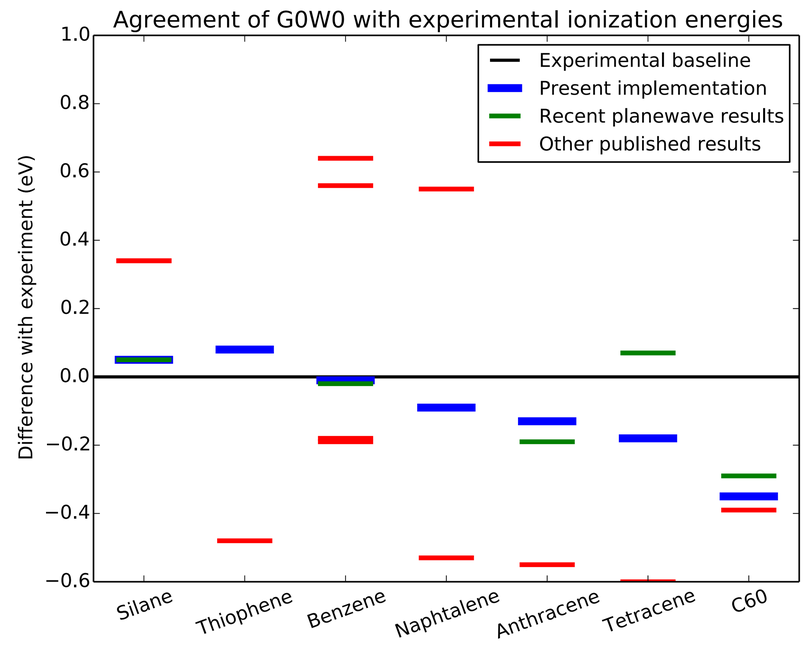
$$\left\{|\phi_0\rangle,\hat{\epsilon}(\omega)|\phi_0\rangle,\hat{\epsilon}^2(\omega)|\phi_0\rangle,\ldots,\hat{\epsilon}^N(\omega)|\phi_0\rangle\right\}$$

and then orthonormalize it to obtain the Lanczos basis :

$$\left\{|l_0(\omega)\rangle,|l_1(\omega)\rangle,|l_2(\omega)\rangle,\ldots,|l_N(\omega)\rangle\right\}$$

which is substantially smaller than a planewave basis of equivalent accuracy.





Conclusion

- 2 bottlenecks in G_0W_0 implementations :
  1. The sum over conduction states
  2. The inversion of the dielectric matrix
- We assess these bottlenecks using :
  3. Sternheimer's equation
  4. Lanczos algorithm
- and obtain a 500-fold increase in speed
- without any further approximation

See arXiv:1408.3036 for more details

Université de Montréal · RQMP · Calcul Québec · compute • calcul CANADA · Fonds de recherche Nature et technologies Québec · NSERC CRSNG · fnrs FREEDOM TO RESEARCH · ETSF · UCL Université catholique de Louvain