"Faster G0W0 calculations using Lanczos algorithm and Sternheimer equation"

Laflamme Janssen, Jonathan ; Bérubé, Nicolas ; Geadah-Antonius, Gabriel ; Côté, Michel

Abstract

G0W0 corrections to DFT band structures are a popular way to go beyond the accuracy DFT is able to provide. However, the calculation of such corrections with the ABINIT code is currently prohibitive for systems with more than a few hundreds of electrons. What limits the calculations to this system size is the need in the current implementation to invert the dielectric matrix and to carry out some summation over conduction bands. This poster presents a strategy to avoid both of these limitations for the screened-exchange contribution to the self-energy. In ABINIT's implementation, the dielectric matrix is expressed in a plane wave basis, which needs to be relatively big to properly describe the matrix. This poster explains how a Lanczos basis can be generated to substantially reduce the size of the matrix. Also, the number of conduction bands needed to reach convergence in the summation is usually an order of magnitude bigger than the number of valence bands. Here, the calculation o...

Document type : *Communication à un colloque (Conference Paper)*

Référence bibliographique

Laflamme Janssen, Jonathan ; Bérubé, Nicolas ; Geadah-Antonius, Gabriel ; Côté, Michel. *Faster G0W0 calculations using Lanczos algorithm and Sternheimer equation.*5th International ABINIT Developer Workshop (Han-sur-lesse, Belgique, du 11/04/2011 au 14/04/2011).

Faster G_0W_0 calculations
using Lanczos algorithm and Sternheimer equation

Jonathan Laflamme Janssen, Gabriel Antonius, Nicolas Bérubé and Michel Côté

Université de Montréal

Regroupement Québécois pour les Matériaux de Pointe (RQMP)

Abstract

G_0W_0 corrections to DFT band structures are a popular way to go beyond the accuracy DFT is able to provide. However, the calculation of such corrections with the ABINIT code is currently prohibitive for systems with more than a few hundreds of electrons. What limits the calculations to this system size is the need in the current implementation to invert the dielectric matrix and to carry out some summations over conduction bands. This poster presents a strategy to avoid both of these limitations for the screened-exchange contribution to the self-energy (Σ^{SEX}). In ABINIT's implementation, the dielectric matrix is expressed in a plane wave basis, which needs to be relatively big to properly describe the matrix. This poster explains how a Lanczos basis can be generated to substantially reduce the size of the matrix. Also, the number of conduction bands needed to reach convergence in the summation is usually an order of magnitude bigger than the number of valence bands. Here, the calculation of all the conduction states is avoided by reformulating the summation into a linear equation problem (Sternheimer equation), which also substantially reduces the computation time.

Introduction

GW : real quasiparticles

$$(\hat{T} + \hat{V}_{ext} + \hat{V}_H)|\phi_n\rangle + \hat{\Sigma}(\varepsilon_n)|\phi_n\rangle = \varepsilon_n|\phi_n\rangle$$

$$\hat{G}(\omega) = \sum_n \frac{|\phi_n\rangle\langle\phi_n|}{\omega - \varepsilon_n + i\eta\,sgn(\varepsilon_n - \mu)}$$

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

$$\hat{W}(\omega) = \hat{v} + \hat{v}\hat{P}(\omega)\hat{W}(\omega)$$

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

G_0W_0 : First order corrections to ε_n

We approximate quasiparticles states φ_n and energies ε_n by Kohn-Sham states and energies :

$$(\hat{T} + \hat{V}_{ext} + \hat{V}_H)|\phi_n\rangle + \hat{V}_{xc}|\phi_n\rangle = \varepsilon_n|\phi_n\rangle$$

To obtain an estimate of the Self-Energy for a specific state :

$$\hat{P}(\omega) = ...$$
$$\hat{W}(\omega) = ...$$
$$\Sigma(\varepsilon_e) \equiv \langle\phi_e|\hat{\Sigma}(\varepsilon_e)|\phi_e\rangle$$
$$= \frac{i}{2\pi}\sum_n\int_{-\infty}^{+\infty}d\omega e^{i\eta\omega}\frac{\langle\phi_e\phi_n^*|\hat{W}(\omega)|\phi_n^*\phi_e\rangle}{\omega - (\varepsilon_n - \varepsilon_e - i\eta\,sgn(\varepsilon_n - \mu))}$$
$$= \Sigma^{SEX}(\varepsilon_e) + \Sigma^{COH}(\varepsilon_e)$$

Pôles of G(ω)	Pôles of W(ω)

To obtain a first order estimate of quasiparticle energies :

$$\varepsilon_n^{QP} \approx \varepsilon_n + \langle\phi_n|\hat{\Sigma}(\varepsilon_n^{QP}) - \boxed{\hat{V}_{xc}}|\phi_n\rangle$$

Screened exchange

In our work, only screened exchange is implemented (for now) :

$$\Sigma^{SEX}(\varepsilon_e) = -\sum_v\langle\phi_e\phi_v^*|\hat{W}(\varepsilon_v - \varepsilon_e)|\phi_v^*\phi_e\rangle$$
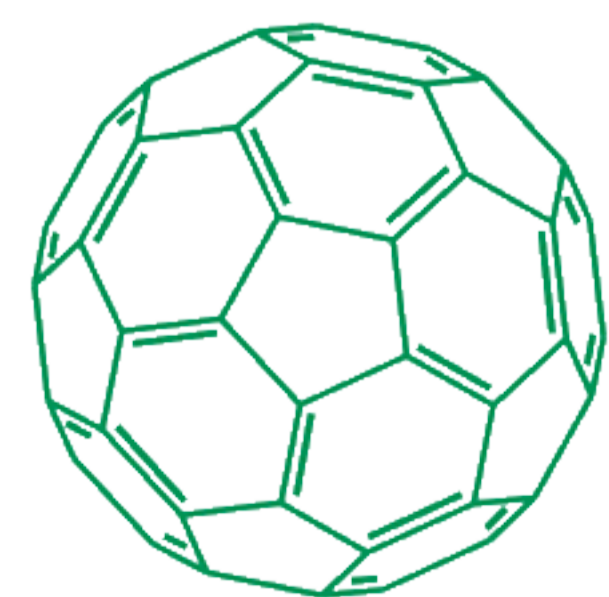
Method, part 1 :
Sternheimer equation

Bottleneck 1 : sum over conduction states

To obtain $\Sigma^{SEX}(\varepsilon_e)$, we need $P(\varepsilon_v - \varepsilon_e)$.

This requires a sum over conduction states :

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

Problematic example : C_{60}

$$N_c \approx 10N_v = 1200 \quad \& \quad ecut = 30\,Ha \quad \Rightarrow \quad mpw = 300\,000$$

$$\{\varphi_c, \varepsilon_c\} : 9\,Gb\,RAM\,\&\,days\,of\,CPU\,time$$

Solution to bottleneck 1

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

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

$$\equiv \sum_v|f_v^{+*}v\rangle + |f_v^{-*}v\rangle$$

$$\Rightarrow |f_v^{\pm}\rangle = -\sum_c\frac{|\phi_c\rangle\langle\phi_c|}{\varepsilon_c - \varepsilon_v \pm \omega}|\psi^*\phi_v\rangle$$

$$= -\sum_c\frac{|\phi_c\rangle\langle\phi_c|}{\hat{H} - \varepsilon_v \pm \omega}|\psi^*\phi_v\rangle$$

Eliminating $\sum_c$

$$= -\frac{\hat{P}_c}{\hat{H} - \varepsilon_v \pm \omega}|\psi^*\phi_v\rangle$$

$$\Rightarrow (\hat{H} - \varepsilon_v \pm \omega)|f_v^{\pm}\rangle = -\hat{P}_c|\psi^*\phi_v\rangle \quad \text{(Sternheimer's equation)}$$

Implementation of solution

1) H is sparse $\Rightarrow$ iterative method

2) H - $\varepsilon_v \pm \omega$ is indefinite $\Rightarrow$ SYMMLQ instead of CG

3) Convergence slow $\Rightarrow$ preconditionning

(6x faster for benzene, antracene, pentacene, heptacene and C_{60})

Method, part 2 :
Lanczos algorithm

Bottleneck 2 : dielectric matrix inversion

To obtain W(ω), one usually invert the dielectric matrix :

$$\hat{W}(\omega) = \hat{v} + \hat{v}\hat{P}(\omega)\hat{W}(\omega)$$

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

$$\Rightarrow \hat{W}(\omega) = \hat{\epsilon}^{-1}(\omega)\,\hat{v}$$

Problematic example : C_{60} again

ecuteps = 3 Ha $\Rightarrow$ 10 000 x 10 000 dielectric matrix

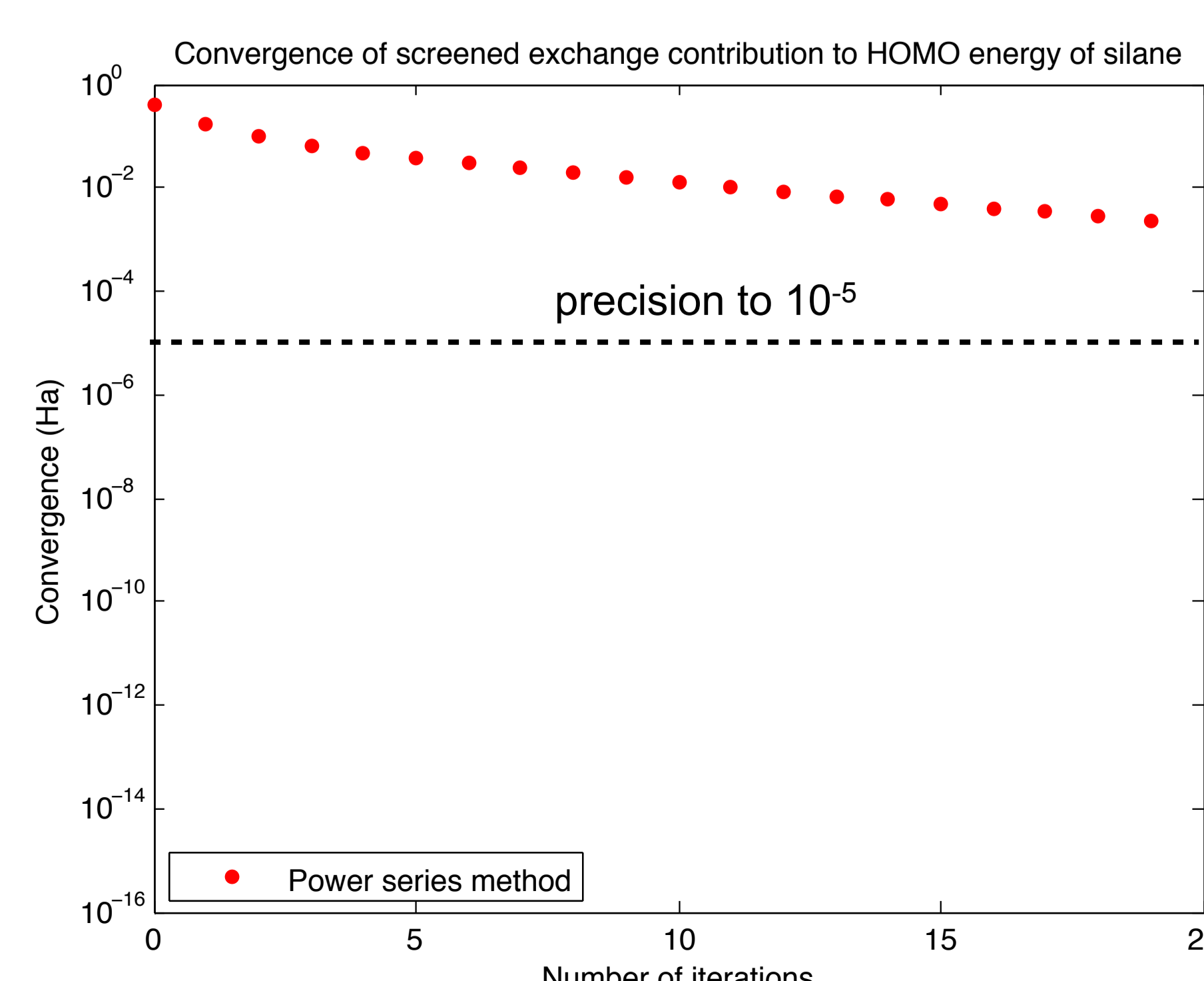ϵ^{-1} $\Rightarrow$ 1,5 Gb RAM & days CPU time

Solution to bottleneck 2

Idea : Use geometrical serie to express matrix inversion...

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

$$= \sum_{m=0}^{\infty}(\hat{v}\hat{P}(\omega))^m\hat{v}$$

$$\Rightarrow \Sigma^{SEX}(\varepsilon_e) = -\sum_v\sum_{m=0}^{\infty}\langle\phi_e\phi_v^*|\,(\hat{v}\hat{P}(\varepsilon_v - \varepsilon_e))^m\hat{v}\,|\phi_v^*\phi_e\rangle$$

Problem : convergence slow...

Convergence of screened exchange contribution to HOMO energy of silane. precision to 10^{-5}. Power series method.

Method, part 2 :
Lanczos algorithm

Better solution to bottleneck 2

Power series version of screened exchange :
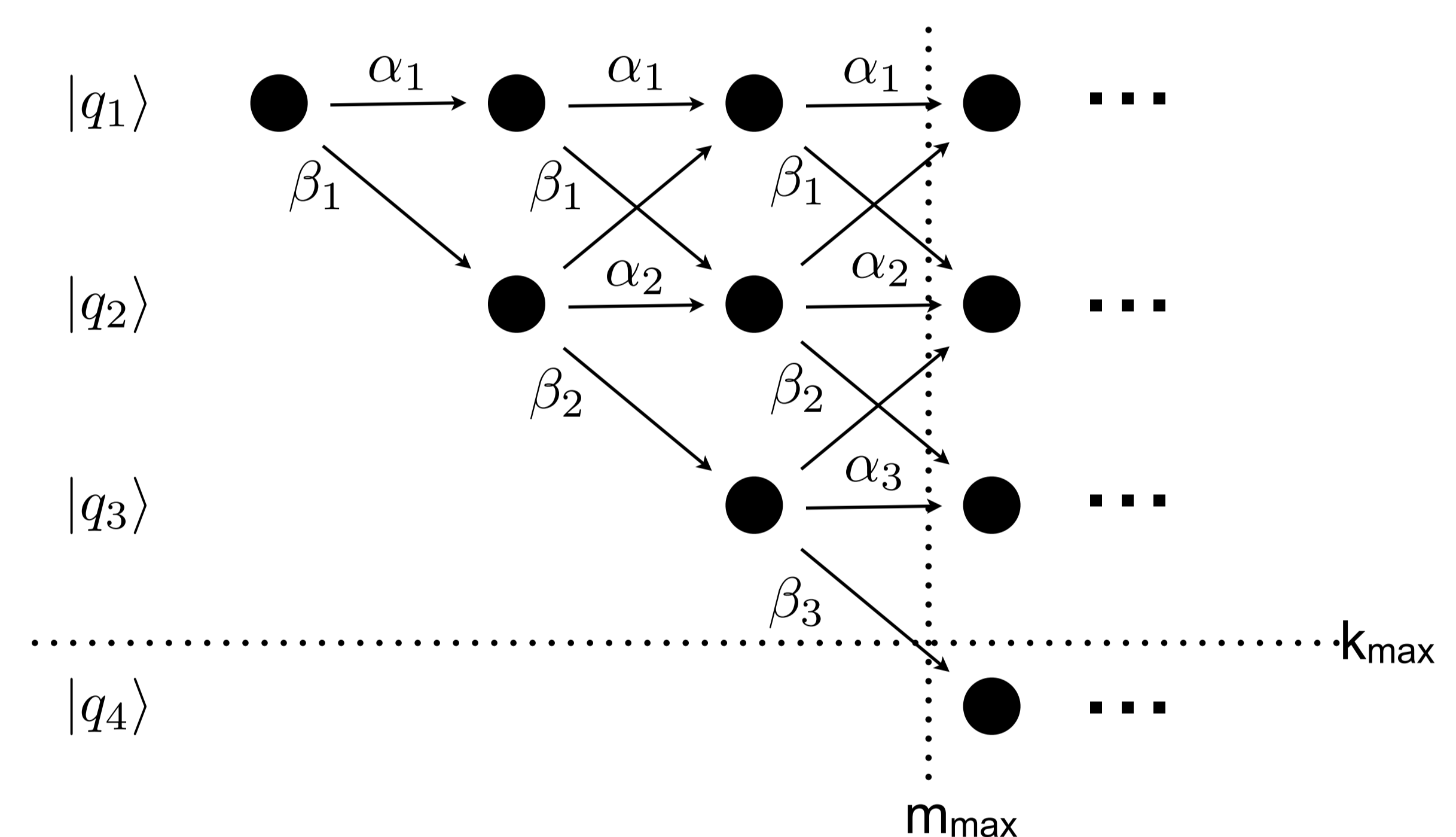
$$\Sigma^{SEX}(\varepsilon_e) = -\sum_{m=0}^{\infty}\sum_v\langle\phi_e\phi_v^*|\,\sqrt{\hat{v}}\left(\sqrt{\hat{v}}\hat{P}(\omega_{ve})\sqrt{\hat{v}}\right)^m\sqrt{\hat{v}}|\phi_v^*\phi_e\rangle$$

Idea : Tridiagonalize the operator in (), taking the vector it acts on as the first vector of tridiagonal basis $\{q_i\}$.

$$\Sigma^{SEX}(\varepsilon_e) = -\sum_{m=0}^{\infty}\sum_v\sum_{ij}\langle\phi_e\phi_v^*|\sqrt{\hat{v}}|q_i\rangle\,\hat{T}_{kmax}^m\langle q_j|\sqrt{\hat{v}}|\phi_v^*\phi_e\rangle$$

$$= -\sum_{m=0}^{\infty}\sum_v\begin{bmatrix}1 & 0 & \cdots & 0\end{bmatrix}\begin{bmatrix}\alpha_1 & \beta_1 & \cdots & 0 \\ \beta_1 & \alpha_2 & & \vdots \\ \vdots & & \ddots & \beta_{k_{max}-1} \\ 0 & \cdots & \beta_{k_{max}-1} & \alpha_{k_{max}}\end{bmatrix}^m\begin{bmatrix}1 \\ 0 \\ \vdots \\ 0\end{bmatrix}$$

Applying tridiagonal operator m times gives schematically :

Constructing a k_{max} x k_{max} tridiagonal matrix costs the same as iterating $m_{max} = k_{max} - 1$ times the power series.
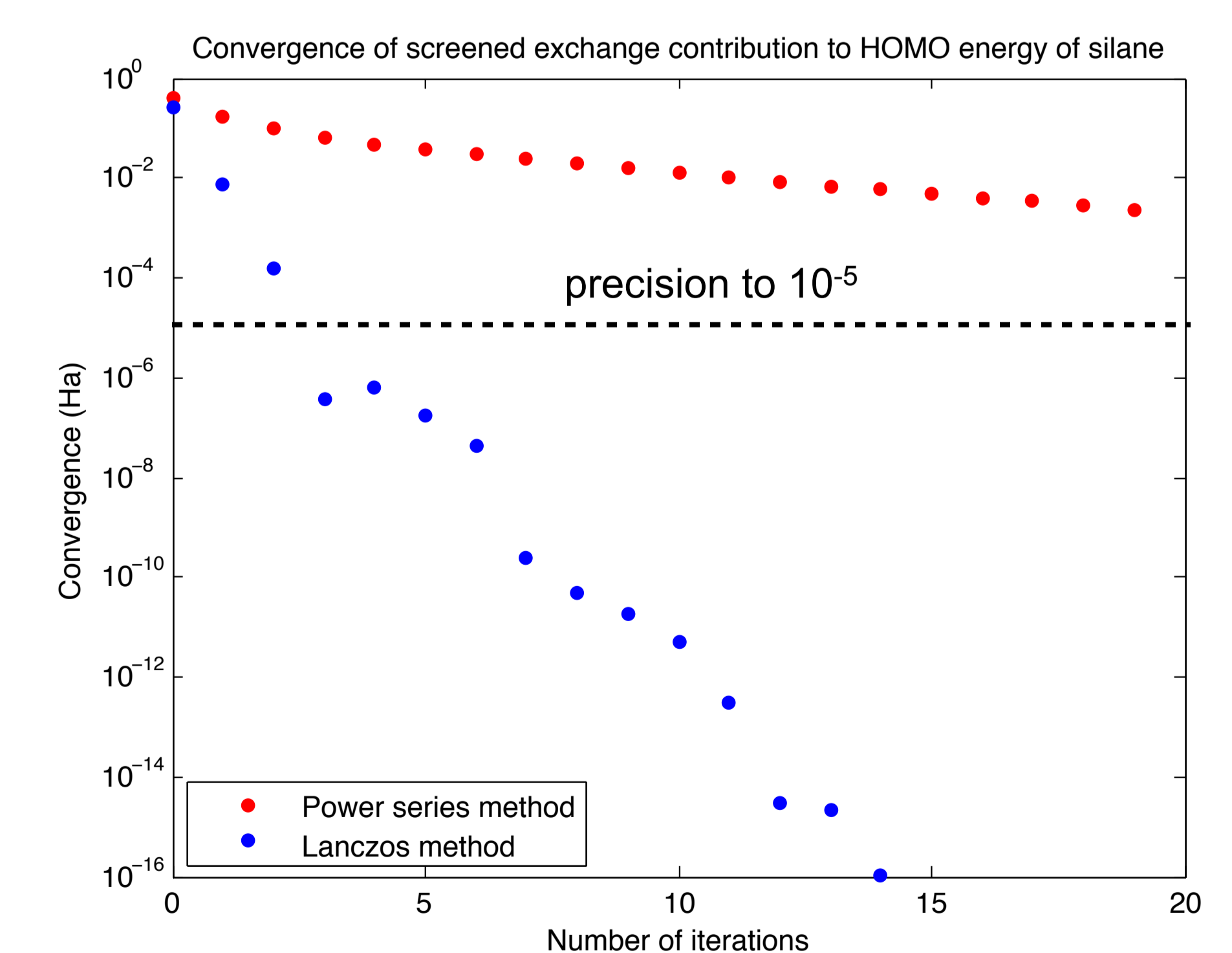
But a T matrix of k_{max} dimensions contains a more precise estimate of Σ^{SEX} than a power series with $m_{max} = k_{max} - 1$.

Keeping k_{max} finite but letting $m_{max} \to \infty$, we have :

$$\sum_{m=0}^{\infty}(\sqrt{\hat{v}}\hat{P}(\omega_{ve})\sqrt{\hat{v}})^m \approx \sum_{m=0}^{\infty}\hat{T}_{kmax}^m = (1 - \hat{T}_{kmax})^{-1}$$

$$\Sigma^{SEX}(\varepsilon_e) \approx -\sum_v\sum_{ij}\langle\phi_e\phi_v^*|\sqrt{\hat{v}}|q_i\rangle\,(1 - \hat{T}_{kmax})_{ij}^{-1}\langle q_j|\sqrt{\hat{v}}|\phi_v^*\phi_e\rangle$$

Which converges a lot faster :

Convergence of screened exchange contribution to HOMO energy of silane. precision to 10^{-5}. Power series method. Lanczos method.

Conclusion

- $\Sigma^{SEX}(\varepsilon_e)$ implemented without $\{\varphi_c, \varepsilon_c\}$ and without substantial time spent on matrix inversion.

- Preconditionning in SYMMLQ causes 6x increase of speed in organic systems.

- Lanczos method is dramatically faster that power method.