Supporting information for "Breaking the theoretical scaling limit for predicting quasi-particle energies: The stochastic GW approach"

Daniel Neuhauser, Yi Gao, Christopher Arntsen, and Cyrus Karshenas

Department of Chemistry, University of California at Los Angeles, CA-90095 USA

Eran Rabani

School of Chemistry, The Sackler Faculty of Exact Sciences,

Tel Aviv University, Tel Aviv 69978, Israel

Roi Baer

Fritz Haber Center for Molecular Dynamics, Institute of Chemistry,

The Hebrew University of Jerusalem, Jerusalem 91904, Israel

I. SOME TRICKS OF THE TRADE CONCERNING RANDOM ORBITALS

A. Real random orbitals

A real random orbital $\zeta(\mathbf{r})$ is a set of random numbers associated with each grid point and can take the values $\pm h^{-3/2}$ with equal probability, where h is the grid spacing. In the algorithm below by "draw a random orbital" we mean that one has to produce by a pseudorandom number generator a sample of $\zeta(\mathbf{r})$.

The mean of the projection $\langle |\zeta\rangle\langle\zeta| \rangle_\zeta$ is the unit operator on the grid: $\langle |\zeta\rangle\langle\zeta| \rangle_\zeta = \hat{\mathbf{1}}$. For an operator $\hat{A}$ the stochastic trace formula[1] is $tr\left[\hat{A}\right] = \left\langle \left\langle \zeta \left| \hat{A} \right| \zeta \right\rangle \right\rangle_\zeta$.

B. Chebyshev evaluation

The algorithm below assumes we need to compute the self-energy $\tilde{\Sigma}(\omega;\varepsilon)$ for several Kohn-Sham energies ε (say, ε_n, $n = 1, \ldots, N_\varepsilon$). A filtered real random orbital is associated with each energy ε using an appropriate filtered $\phi_\varepsilon = f_\sigma\left(\hat{h}_{KS} - \varepsilon\right)\phi$ where ϕ is a random orbital.

The filter function $f_\sigma\left(\hat{h}_{KS} - \varepsilon\right)$ is applied onto ϕ using Chebyshev expansion techniques [2] by which $\phi_\varepsilon = \sum_m c_m(\varepsilon)\phi_m$ where $\phi_0 = \phi$, $\phi_1 = \hat{h}\phi$ and the following iteration is used for the other terms: $\phi_{m+1} = 2\hat{h}\phi_m - \phi_{m-1}$. Note that only 3 auxiliary wave functions are needed for applying this series. Here, $\hat{h} = \frac{\hat{h}_{KS} - \bar{h}}{\Delta h}$ is a shifted-scaled Hamiltonian, such that the eigenvalues of $\hat{h}$ are contained in the interval $[-1, 1]$ and the expansion coefficients are obtained from $c_m(\varepsilon) = \int_0^\pi f_\sigma\left(\Delta h \cos\theta + \bar{h} - \varepsilon\right)\cos m\theta d\theta$.

When ϕ_ε for several values of the parameters ε are needed one can save numerical effort by exploiting the fact that ϕ_m are not ε dependent and thus summing $\phi_\varepsilon = \sum_m c_m(\varepsilon)\phi_m$ simultaneously for all values of ε.

C. The use of $\bar{\phi}$

For evaluating the expectation value $a_\varepsilon = Q(\varepsilon)^{-1} tr\left[f_\sigma\left(\hat{h}_{KS} - \varepsilon\right)^2 \hat{A}\right]$, where $Q(\varepsilon) = tr\left[f_\sigma\left(\hat{h}_{KS} - \varepsilon\right)^2\right]$, one can use the stochastic trace formula $a_\varepsilon = Q(\varepsilon)^{-1}\left\langle \left\langle \phi_\varepsilon \left| \hat{A} \right| \phi_\varepsilon \right\rangle \right\rangle_\phi$, $Q(\varepsilon) = \langle\langle \phi_\varepsilon | \phi_\varepsilon \rangle\rangle_\phi$. However this is expensive numerically, since one needs to operate with

$\hat{A}$ on ϕ N_ε times, for each desired value of ε. Instead, one can define $\bar{\phi} = \sum_\varepsilon \phi_\varepsilon$ and then $a_\varepsilon = Q(\varepsilon)^{-1} \left\langle \left\langle \phi_\varepsilon \left| \hat{A} \right| \bar{\phi} \right\rangle \right\rangle_\phi$,where now $Q(\varepsilon) = \left\langle \left\langle \phi_\varepsilon | \bar{\phi} \right\rangle \right\rangle_\phi$. Since $\bar{\phi}$ does not depend on ε, one has to act with $\hat{A}$ on $\bar{\phi}$ only once and then for each value of ε preform the stochastic average to obtain a_ε.

D. Matrix/Operator Compression and decoupling

Given an operator $\hat{A}$ we draw a real random orbital $\zeta(\mathbf{r})$ and $\hat{A} = \left\langle \hat{A} |\zeta\rangle \langle\zeta| \right\rangle_\zeta = \left\langle |\zeta_A\rangle \langle\zeta| \right\rangle_\zeta$; therefore, as a grid-matrix: $\left\langle \mathbf{r}_1 \left| \hat{A} \right| \mathbf{r}_2 \right\rangle = \left\langle \zeta_A(\mathbf{r}_1) \zeta(\mathbf{r_2})^* \right\rangle_\zeta$ and one achieves a compact and "decoupled" stochastic representation of $\hat{A}$.

E. Compression of the density matrix and the exchange self-energy

The exchange self-energy at orbital energy ε can be obtained by one of two choices. The first uses $\bar{\phi}$ and ϕ_ε as defined above to calculate the exchange self-energy for all sampled energies simultaneously:

$$\Sigma^X(\varepsilon) = -\frac{1}{\left\langle \left\langle \phi_\varepsilon | \bar{\phi} \right\rangle \right\rangle_\phi} \left\langle \iint \phi_\varepsilon(\mathbf{r}_1) \frac{\rho(\mathbf{r}_1, \mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|} \bar{\phi}(\mathbf{r}_2) d^3r_1 d^3r_2 \right\rangle_\phi, \tag{1}$$

For large enough β ($\beta \gg \frac{1}{E_g}$, where E_g is the QP gap), the density matrix (DM), $\rho(\mathbf{r}_1, \mathbf{r}_2)$, can be expressed as the matrix elements of the operator

$$\hat{\rho} \equiv \theta_\beta\left(\mu - \hat{h}_{KS}\right) = \frac{1}{2}\left\{1 + \mathrm{erf}\left(\beta\left(\mu - \hat{h}_{KS}\right)\right)\right\}, \tag{2}$$

where $\hat{h}_{KS}$ is the KS Hamiltonian and μ the chemical potential. Since $\hat{\rho} = \hat{\rho}^2$, one can define a compact representation of $\hat{\rho}$ in terms of a new set of stochastic orbitals, φ, as $\left\langle \varphi_\mu(\mathbf{r}_1) \varphi_\mu(\mathbf{r_2})^* \right\rangle_\zeta$, where $\varphi_\mu = \frac{1}{2}\left[1 + \mathrm{erf}\left(\beta\left(\mu - \hat{h}_{KS}\right)\right)\right]\varphi$. Note that these are exactly the same orbitals used for the stochastic time propagation. Using this representation of $\hat{\rho}$, we rewrite the exchange self-energy as:

$$\Sigma^X(\varepsilon) = -\frac{1}{\left\langle \left\langle \phi_\varepsilon | \bar{\phi} \right\rangle \right\rangle_\phi} \left\langle \int\int d^3r_1 d^3r_2 \phi_\varepsilon(\mathbf{r}_1) \varphi_\mu(\mathbf{r}_1) \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} \varphi_\mu(\mathbf{r}_2) \bar{\phi}(\mathbf{r}_2) \right\rangle_{\varphi\phi}. \tag{3}$$

For every pair of random orbitals φ and ϕ we need to compute φ_μ, ϕ_ε and $\bar{\phi} = \sum_\varepsilon \phi_\varepsilon$, and then through a Fourier transform perform the convolution to obtain the exchange energy.

The second variant uses the fact that most of the numerical effort in the exchange energy goes towards the determination of the orbitals φ_μ rather than the actual calculation of the exchange integral in Eq. (3). Therefore, a better numerical convergence with very little additional cost is achieved if the exchange energy is calculated for each energy separately, replace $\bar{\phi}$ by ϕ_ε:

$$\Sigma^X(\varepsilon) = -\frac{1}{\langle\langle\phi_\varepsilon|\phi_\varepsilon\rangle\rangle_\phi}\left\langle \int\int d^3r_1 d^3r_2 \phi_\varepsilon(\mathbf{r}_1)\varphi_\mu(\mathbf{r}_1)\frac{1}{|\mathbf{r}_1-\mathbf{r}_2|}\varphi_\mu(\mathbf{r}_2)\phi_\varepsilon(\mathbf{r}_2)\right\rangle_{\varphi\phi}.$$

This is the variant we used in the computations reported in the paper. We find that for a fixed statistical error, the number of φ's and $\phi's$ is independent of system size, which implies that the calculations of the exchange self-energy is scaling linearly.

II. ALGORITHM

The steps below are done I_{sGW} times. In each stochastic sampling we

1. Draw a real random orbital $\zeta(\mathbf{r})$ and a real random orbital $\phi(\mathbf{r})$ and by filtering produce $\phi_\varepsilon(\mathbf{r}) = \left\langle \mathbf{r}\left| f_\sigma\left(\hat{h}_{KS}-\varepsilon\right)\right|\phi\right\rangle$. Generate $\bar{\phi} = \sum_\varepsilon \phi_\varepsilon$. Typically we filtered two or four orbitals simultaneously, e.g., two near the HOMO and two near the LUMO. This is done with a single Chebyshev expansion, as explained above.

2. Add the contribution to the exchange correlation energy, $\int v_{XC}(\mathbf{r})|\phi_\varepsilon(\mathbf{r})|^2 d^3r$, and the contributions $\int|\phi_\varepsilon(\mathbf{r})|^2 d^3r$ and $\int\phi_\varepsilon(\mathbf{r})\bar{\phi}(\mathbf{r})dr$ to the denominators, $Q(\varepsilon) = \left\langle\left\langle\phi_\varepsilon|\bar{\phi}\right\rangle\right\rangle$ (used for the polarization) and $\langle\langle\phi_\varepsilon|\phi_\varepsilon\rangle\rangle$ (used for the exchange and the Kohn-Sham energy).

3. Draw N_ψ real random orbitals $\psi_\ell(\mathbf{r})$, $\ell = 1,\ldots,N_\psi$. These will be used to generate $W^{(r)}(\zeta\phi_\varepsilon,\psi_\ell,t_k)$ and $\left\langle\phi_\varepsilon\bar{\zeta}(t)^*\right|\psi_\ell\rangle$ in steps (5) and (8).

4. Calculate $\Delta(\mathbf{r}) = \int u_C(\mathbf{r}-\mathbf{r}')\zeta(\mathbf{r}')\bar{\phi}(\mathbf{r}')d^3r'$.

5. Draw N_{TDH} random orbitals $\varphi_m(\mathbf{r})$, $m = 1,\ldots,N_{TDH}$ and project each one on the occupied space: $\varphi_{m,\mu}(\mathbf{r}) = \left\langle\mathbf{r}|\theta_\beta\left(\mu-\hat{h}_{KS}\right)|\varphi_m\right\rangle$. These orbitals are then used for two purposes:

(a) Evaluate the Fock exchange self-energy for which we accumulate the contribution:

$$-\frac{1}{N_{TDH}}\sum_{m=1}^{N_{TDH}}\int\int \phi_{\varepsilon}(\mathbf{r})\varphi_{m,\mu}(\mathbf{r})u_C(|\mathbf{r}-\mathbf{r}'|)\phi_{\varepsilon}(\mathbf{r}')\varphi_{m,\mu}(\mathbf{r}')d^3r\,d^3r'. \tag{4}$$

(b) Propagate the orbitals using the stochastic time-dependent Hartree equations (only for positive times):

$$\begin{gathered} i\hbar\dot{\varphi}_{m,\mu}(\mathbf{r},t)=\left[\hat{h}_{KS}+\int\frac{\delta n(\mathbf{r}',t)}{|\mathbf{r}-\mathbf{r}'|}d^3r'\right]\varphi_{m,\mu}(\mathbf{r},t) \\ \delta n(\mathbf{r},t)=\frac{2}{N_{TDH}}\sum_{m=1}^{N_{TDH}}\left(|\varphi_{m,\mu}(\mathbf{r},t)|^2-|\varphi_{m,\mu}(\mathbf{r},0)|^2\right). \end{gathered} \tag{5}$$

This is done on a time grid $t=\tau\delta t$, $\delta t\equiv\frac{T}{N_T}$, $\tau=-N_T,\ldots,N_T$ containing $2N_T+1$ points for a maximum time $T\approx 100\hbar E_h^{-1}$ using a split operator technique with $\delta t=0.05$au. The propagation is done *twice*:

i. For an initial state $\varphi^{i}_{m,\mu}(\mathbf{r},t=0)=e^{-i\Delta(\mathbf{r})\eta}\varphi_{m,\mu}(\mathbf{r})$.

ii. For an initial state $\varphi^{ii}_{m,\mu}(\mathbf{r},t=0)=\varphi_{m,\mu}(\mathbf{r})$.

6. Use $\Delta n(\mathbf{r},\tau\delta t)=\frac{1}{\eta}\left[\delta n^{i}(\mathbf{r},\tau\delta t)-\delta n^{ii}(\mathbf{r},\tau\delta t)\right]$ to form N_ψ time-dependent retarded (r) overlaps:

$$W^{r}(\varepsilon,\ell,\tau)=\theta(\tau\delta t)\iint d^3r_1d^3r_2\psi_{\ell}(\mathbf{r}_1)\int u_C(\mathbf{r}_1-\mathbf{r}_2)\,\Delta n(\mathbf{r}_2,\tau\delta t). \tag{6}$$

7. Multiply the time-dependent overlaps $W^{r}(\varepsilon,\ell,\tau)$ by a regularization function $g(\tau\delta t)=\exp\left\{-\frac{(\Gamma\tau\delta t)^2}{2}\right\}$, where Γ is a damping parameter (typically we set $\Gamma=0.03\hbar^{-1}E_h$). Perform the transformation [3]

$$\tilde{W}(\varepsilon,\ell,\omega_\tau)\;=\;\tilde{W}^{r}(\varepsilon,\ell,\omega_\tau)\,\theta(\omega_\tau)+\tilde{W}^{r}(\varepsilon,\ell,\omega_\tau)^{*}\,\theta(-\omega_\tau) \tag{7}$$

where $\tilde{W}^{r}(\varepsilon,\ell,\omega_\tau)$ is the discrete Fourier transform of $W^{r}(\varepsilon,\ell,\tau)$. Transform $\tilde{W}(\varepsilon,\ell,\omega_\tau)$ back to the time domain to generate the time-ordered overlaps, $W(\varepsilon,\ell,\tau)$.

8. Filter the random orbital ζ to yield its occupied and unoccupied parts,

$$\begin{gathered} \zeta^{-}(\mathbf{r},t=0)=\langle\mathbf{r}|\theta_\beta\left(\mu-\hat{h}_{KS}\right)\zeta\rangle \\ \zeta^{+}(\mathbf{r},t=0)=\zeta(\mathbf{r})-\zeta^{-}(\mathbf{r},t=0) \end{gathered} \tag{8}$$

9. Propagate ζ^+ forward in time and ζ^- backward in time under $\hat{h}_{KS}$.

10. Calculate and store $G(\varepsilon, \ell, \tau) = \left\langle \phi_\varepsilon \bar{\zeta}(\tau\delta t)^* \right| \left. \psi_\ell \right\rangle$, where $\bar{\zeta}(t) = \theta(t)\zeta^+(t) - \theta(-t)\zeta^-(-t)$.

11. Average the polarization self-energy

$$\Sigma^P(\tau\delta t, \varepsilon) = \frac{g(\tau\delta t)}{N_\psi} \sum_{\ell=1}^{N_\psi} \langle G(\varepsilon, \ell\tau) W(\varepsilon, \ell, \tau) \rangle_{\phi\zeta} .$$

At the end of the stochastic iterations Eq. (1) in the paper is solved to yield the final quasi-particle energy.

---

[1] M. F. Hutchinson. A stochastic estimator of the trace o the influence matrix of laplacian smoothing splines. *Comm. Stat. Simul. Comp.*, 18:1059–1076, 1989.

[2] R. Kosloff. Time-dependent quantum-mechanical methods for molecular- dynamics. *J. Phys. Chem.*, 92(8):2087–2100, 1988.

[3] A. L. Fetter and J. D. Walecka. *Quantum Theory of Many-particle Systems*. McGraw-Hill, New York, 1971.