

Stochastic resolution of identity second-order Matsubara Green's function theory

Cite as: J. Chem. Phys. 151, 044114 (2019); doi: 10.1063/1.5108840

Submitted: 2 May 2019 • Accepted: 2 July 2019 •

Published Online: 30 July 2019



View Online



Export Citation



CrossMark

Tyler Y. Takeshita,^{1,2,a)} Wenjie Dou,^{2,b)} Daniel G. A. Smith,^{3,c)} Wibe A. de Jong,^{4,d)} Roi Baer,^{5,e)} Daniel Neuhauser,^{6,f)} and Eran Rabani^{1,2,7,8,g)}

AFFILIATIONS

¹Mercedes-Benz Research and Development North America, Sunnyvale, California 94085, USA

²Department of Chemistry, University of California Berkeley, Berkeley, California 94720, USA

³The Molecular Sciences Software Institute, Blacksburg, Virginia 24060, USA

⁴Computational Research Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA

⁵Fritz Haber Center for Molecular Dynamics, Institute of Chemistry, The Hebrew University of Jerusalem, Jerusalem 91904, Israel

⁶Department of Chemistry and Biochemistry, University of California, Los Angeles, California 90095, USA

⁷Materials Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA

⁸The Raymond and Beverly Sackler Center of Computational Molecular and Materials Science, Tel Aviv University, Tel Aviv 69978, Israel

^{a)}Electronic mail: tyler.takeshita@daimler.com

^{b)}Electronic mail: douw@berkeley.edu

^{c)}Electronic mail: dgasmith@vt.edu

^{d)}Electronic mail: wadejong@lbl.gov

^{e)}Electronic mail: roi.baer@huji.ac.il

^{f)}Electronic mail: dxn@chem.ucla.edu

^{g)}Electronic mail: eran.rabani@berkeley.edu

ABSTRACT

We develop a stochastic resolution of identity representation to the second-order Matsubara Green's function (sRI-GF2) theory. Using a stochastic resolution of the Coulomb integrals, the second order Born self-energy in GF2 is decoupled and reduced to matrix products/contractions, which reduces the computational cost from $O(N^5)$ to $O(N^3)$ (with N being the number of atomic orbitals). The current approach can be viewed as an extension to our previous work on stochastic resolution of identity second order Møller-Plesset perturbation theory [T. Y. Takeshita *et al.*, J. Chem. Theory Comput. **13**, 4605 (2017)] and offers an alternative to previous stochastic GF2 formulations [D. Neuhauser *et al.*, J. Chem. Theory Comput. **13**, 5396 (2017)]. We show that sRI-GF2 recovers the deterministic GF2 results for small systems, is computationally faster than deterministic GF2 for $N > 80$, and is a practical approach to describe weak correlations in systems with 10^3 electrons and more.

Published under license by AIP Publishing. <https://doi.org/10.1063/1.5108840>

I. INTRODUCTION

To understand the fundamental chemistry and physics that drive emergent phenomena in complex chemical systems under real-world conditions requires an accurate description of the ground and excited electronic states. However, such a task poses unique

challenges. First, quantitative electronic structure calculations for realistic chemical systems can be prohibitively expensive due to the system size and poor scaling of current methods. Second, complex chemical systems with multiple components lead to situations where electron correlation and excited states must be accurately described throughout multiple phases and at their interfaces.

Finally, it is insufficient to study these systems at equilibrium as a great deal of important chemistry emerges under nonequilibrium conditions.

For weakly correlated systems, the many-body perturbation theory (MBPT) within the Green's function (GF) method provides an accurate framework to describe ground state properties beyond mean field density functional theory (DFT)^{1–5} and Hartree-Fock (HF)^{6–8} and has been proven extremely fruitful.^{9–21} These post DFT/HF techniques allow for the inclusion of electron correlation in a controlled and systematic fashion through approximations of the self-energy and have been used to accurately describe total ground state energies, quasiparticle excitations, lifetimes, etc.^{22–30} However, their steep computational scaling and costs restrict their use to small molecular systems or to bulk materials with a small unit cell.

A particularly interesting implementation is the self-consistent GF approach based on the second-order approximation to the electron self-energy (GF2), which is experiencing a renaissance.^{30–36} Modern implementations of the finite temperature GF2 (FT-GF2) method have calculated accurate ionization potentials^{35–38} and total energies for organic molecules as large as C₆H₆ as well as transition metal containing triatomics.^{39,40} Recently, GF2 has been applied to capture static correlation, producing potential energy curves of a symmetrically stretched H₆ ring that are in qualitative agreement with truncated configuration interaction results.⁴¹

Although the results of recent studies are extremely promising, the GF2 approach suffers from a prohibitively high computational cost, scaling as $O(N^5)$ (N is the number of orbitals used). Thus, GF2 methods have yet to be applied to molecular systems of experimentally relevant sizes. It would be extremely advantageous if methods with inherent and systematic structure, e.g., GF methods, were to become a viable option and can be performed routinely for complex chemical systems. This is especially true of the largely unexplored set of chemical systems that lie in the intermediate regime between small molecules and the bulk limit.

Recently, a stochastic method was developed that reduces the scaling of the ground state and thermal GF2 using a stochastic decomposition of the imaginary time Green's function. The method was tested successfully on long linear hydrogen chains.⁴²

In this work, we introduce an alternative stochastic orbital approach to reduce the scaling of the GF2 methods that would lend itself very simply to higher-order approximations to the self-energy and for going beyond limited to the ground state and thermal calculations. Unlike the stochastic GF2 method developed in Ref. 42 which uses random variables to represent the GF, the current approach builds upon our recently developed stochastic orbital electron repulsion integral (ERI) method⁴³ that decouples the 4-index ERIs.

Decoupling the indices of a “core element” in electronic structure theory enables the derivation of new systems of equations that can be solved at a fraction of the computational cost of the deterministic methods. As a result, large and accurate basis sets, traditionally considered impractical, can be used to perform electronic structure calculations within the framework of GF2 at a similar formal scaling to DFT and HF for systems that are impractical in deterministic calculations by introducing a controlled error.

The manuscript is structured as follows: Sec. II describes the GF2 approach and provides two alternative implementations, one

based on going between imaginary time and frequency and the other solely in the imaginary time domain. We also review the stochastic representation of the resolution of identity and apply the formalism to the GF2 self-energy. In Sec. III, we compare the frequency and imaginary time domain approaches for a hydrogen chain as well as provide results for the total correlation energy and the scaling of the stochastic vs deterministic approaches. Section IV concludes.

II. THEORY

Matsubara, or finite-temperature imaginary time, Green's function methods are well suited for the study of complex chemical systems as they enable us to evaluate equilibrium thermodynamic properties. In practice, the temperature is selected such that the Matsubara GF technique is better thought of as a ground state electronic structure method. In the following subsections, we outline the Matsubara GF method and the corresponding stochastic orbital implementation.

A. Matsubara Green's functions

The Matsubara GF method provides a self-consistent solution to the Dyson equation in imaginary frequency (we use atomic units where $\hbar = 1$ throughout the paper, unless otherwise noted),

$$\tilde{\mathbf{G}}(i\omega_n) = \frac{1}{[\tilde{\mathbf{G}}_0(i\omega_n)]^{-1} - \tilde{\mathbf{\Sigma}}(i\omega_n)}, \quad (1)$$

where $i\omega_n = i(2n + 1)\frac{\pi}{\beta}$ is the Matsubara frequency (β the inverse temperature), $\tilde{\mathbf{G}}(i\omega_n)$, $\tilde{\mathbf{G}}_0(i\omega_n)$, and $\tilde{\mathbf{\Sigma}}(i\omega_n)$ are the single-particle Green's function of the interacting system, the single-particle Green's function of the Hartree-Fock reference system, and the self-energy, respectively. In the second order approximation, i.e., GF2, the self-energy is given by

$$\Sigma_{ij}(\tau) = \sum_{klmnpq} G_{kl}(\tau) G_{mn}(\tau) G_{pq}(\beta - \tau) v_{imqk} (2v_{lpnj} - v_{nplj}). \quad (2)$$

Here, $\mathbf{\Sigma}(\tau)$ and $\mathbf{G}(\tau)$ are the imaginary time self-energy and Green's function, respectively. Here, we have restricted ourselves to the closed shell case. v_{ijmn} is the 4-index electron repulsion integrals (ERIs),

$$v_{ijmn} = (ij|mn) = \iint dr_1 dr_2 \frac{\chi_i(r_1)\chi_j(r_1)\chi_m(r_2)\chi_n(r_2)}{r_{12}}, \quad (3)$$

where χ_i is the atomic orbital. Equations (1) and (2) need to be solved self-consistently since the self-energy depends on Green's function.

We present two methods to obtain the self-consistent solution to Eq. (1), the first follows the procedure of Ref. 42 which builds on the techniques developed in Refs. 34, 40, and 39, and the second, developed in this work, is based on an imaginary time propagation technique. The self-consistent procedure of Ref. 42 may be summarized in the following steps:

1. Initialize $\tilde{\mathbf{G}}(i\omega_n) = \tilde{\mathbf{G}}_0(i\omega_n) = [(\mu + i\omega_n)\mathbf{S} - \mathbf{F}]^{-1}$, where \mathbf{S} is the overlap matrix and \mathbf{F} is the Fock matrix from a converged Hartree-Fock calculation, and μ is the chemical potential, initialized at the middle point between the highest occupied molecular orbital (HOMO) and the lowest unoccupied molecular orbital (LUMO).
2. Perform a discrete Fourier transform (FT) to obtain Green's function in imaginary time, $\mathbf{G}(\tau) = \frac{1}{\beta} \sum_n e^{-i\omega_n\tau} \tilde{\mathbf{G}}(i\omega_n)$. To reduce the number of time and frequency points, we use the approach developed in Ref. 42.
3. Construct $\mathbf{\Sigma}(\tau)$ from $\mathbf{G}(\tau)$, and then perform a FT to the imaginary frequency space to obtain $\mathbf{\Sigma}(i\omega_n)$.
4. Solve the Dyson equation $\tilde{\mathbf{G}}(i\omega_n) = [(\tilde{\mathbf{G}}_0(i\omega_n))^{-1} - \mathbf{\Sigma}(i\omega_n)]^{-1}$.
5. Update the Fock matrix, $F_{ij} = h_{ij} + \sum_{ab} P_{ab} (v_{ijab} - \frac{1}{2} v_{ibaj})$, using the relation $\mathbf{G}(\tau = \beta) = -\frac{1}{2} \mathbf{P}$, where \mathbf{P} is the density matrix.
6. Tune the chemical potential, μ , so that the total number of electrons, $N_e = \text{Tr}[\mathbf{PS}]$, is conserved.
7. Repeat steps 2–6 until self-consistency is reached.

Step 3 in the above algorithm is the computational bottleneck, i.e., the evaluation of $\mathbf{\Sigma}(\tau)$. The computational cost of the second-order Born approximation to the self-energy, $\mathbf{\Sigma}^{(2)}(\tau)$ [see definition in Eq. (2)], scales as $O(N^5)$ (N is the number of basis functions) and must be done for all N_τ imaginary time points.

Alternatively, one can solve the Dyson equations in the time domain, rather than going back and forth between imaginary time and Matsubara frequency domains. In this case, the complication associated with the FT is replaced with an imaginary time integration and imposing the boundary condition such that $\mathbf{P} = -2\mathbf{G}(\tau = \beta)$:

1. Initialize $\mathbf{G}(\tau) = \mathbf{G}_0(\tau) = \mathbf{X} e^{-\tau(\tilde{\mathbf{F}} - \mu I)} \left[\frac{\theta(-\tau)}{1 + e^{\beta(\tilde{\mathbf{F}} - \mu I)}} - \frac{\theta(\tau)}{1 + e^{-\beta(\tilde{\mathbf{F}} - \mu I)}} \right] \mathbf{X}^T$, where \mathbf{X} satisfies $\mathbf{X}\mathbf{X}^T = \mathbf{S}^{-1}$ and $\tilde{\mathbf{F}} = \mathbf{X}^T \mathbf{F} \mathbf{X}$ is the Fock matrix in orthogonal basis. Again, \mathbf{F} is initialized from a converged Fockian, and μ is initialized at the middle point of the HOMO-LUMO gap.
2. Construct $\mathbf{\Sigma}(\tau)$ from $\mathbf{G}(\tau)$ using Eq. (2).
3. Solve the Dyson equation in the imaginary time to obtain new $\mathbf{G}_{new}(\tau)$,

$$\mathbf{G}_{new}(\tau) = \mathbf{G}_0(\tau) + \int d\tau' d\tau'' \mathbf{G}_0(\tau - \tau') \mathbf{\Sigma}(\tau' - \tau'') \mathbf{G}(\tau'').$$

Imaginary time integral gets solved using a proper quadrature.³⁰ To be specific, we use 1000 Legendre polynomials. The Legendre quadrature points are cubic splined from 256 Chebyshev grid points of imaginary time. To reduce the computational cost, the double integral is decoupled into two separated integrals. Furthermore, to make the most use of the spline, the integral is evaluated for the spline coefficient, such that we only do the quadrature integral once.

4. Update the Fock matrix $F_{ij} = h_{ij} + \sum_{mn} P_{mn} (v_{ijmn} - \frac{1}{2} v_{inmj})$ using the new Green's function, $\mathbf{G}_{new}(\tau = \beta) = -\frac{1}{2} \mathbf{P}$.
5. Tune the chemical potential, μ , so that the total number of electrons, $N_e = \text{Tr}[\mathbf{PS}]$, is conserved.
6. Use the new $\mathbf{G}_{new}(\tau)$ to calculate the self-energy and repeat steps 2–5 until self-consistency is reached.

Again, the computational bottleneck is in Step 2 of calculating the self-energy. In Sec. III, we will compare the two algorithms, i.e., imaginary time and imaginary frequency Green's function

methods and discuss the computational complexity of each approach. Before this is done, we first turn to describe a stochastic formulation for the self-energy that reduces the scaling of obtaining the self-energy to $O(N^3)$.

B. Stochastic resolution of identity

To address the high computational cost of Matsubara GF2 calculations, we begin by noting that a significant reduction in the computational prefactor may be achieved by approximating the 4-index electron repulsion integrals (ERIs), v_{ijmn} , in Eq. (3) in terms of the 2- and 3-index integrals. This is the resolution of the identity (RI) or density fitting approximation.⁴³ The 3- and 2-index ERIs are defined as

$$(\alpha\beta|A) = \int \int dr_1 dr_2 \frac{\chi_\alpha(r_1) \chi_\beta(r_1) \chi_A(r_2)}{r_{12}} \quad (4)$$

and

$$V_{AB} = \int \int dr_1 dr_2 \frac{\chi_A(r_1) \chi_B(r_2)}{r_{12}}. \quad (5)$$

In the above equations, we use the usual notation: the AO Gaussian basis functions are represented by $\chi_\alpha(r)$ and Greek indices $\alpha, \beta, \gamma, \delta, \dots$, while the auxiliary basis functions are represented by the indices A, B, \dots . Finally, the total number of AO basis functions, auxiliary basis functions, are N_{AO} and N_{aux} , respectively. Furthermore, both N_{aux} and N_{AO} are proportional to the system size with N_{aux} typically 3–6 times N_{AO} .

The approximate 4-index RI-ERIs are then expressed symmetrically in terms of the lower-rank integrals according to

$$\begin{aligned} (\alpha\beta|\gamma\delta) &\approx \sum_{AB}^{N_{aux}} (\alpha\beta|A) [V^{-1}]_{AB} (B|\gamma\delta) \\ &= \sum_Q^{N_{aux}} \left[\sum_A^{N_{aux}} (\alpha\beta|A) [V^{-\frac{1}{2}}]_{AQ} \right] \left[\sum_B^{N_{aux}} [V^{-\frac{1}{2}}]_{QB} (B|\gamma\delta) \right]. \end{aligned} \quad (6)$$

Defining

$$K_{\alpha\beta}^Q \equiv \sum_A^{N_{aux}} (\alpha\beta|A) V_{AQ}^{-\frac{1}{2}}, \quad (7)$$

yields

$$(\alpha\beta|\gamma\delta) \approx \sum_Q^{N_{aux}} K_{\alpha\beta}^Q K_{\gamma\delta}^Q. \quad (8)$$

The stochastic orbital implementation of the RI approximation⁴³ utilizes the same set of 2- and 3-index ERIs while introducing an additional set of N_s stochastic orbitals, $\{\theta^\xi\}$, $\xi = 1, 2, \dots, N_s$. Stochastic orbitals are arrays of length N_{aux} with randomly selected elements $\theta_A^\xi = \pm 1$ and have the following property:⁴³

$$\langle \theta \otimes \theta^T \rangle_\xi = I, \quad (9)$$

where we have denoted the stochastic average over N_s stochastic orbitals by $\langle \rangle_\xi$. Inserting the stochastic identity matrix into the deterministic RI-ERIs in Eq. (6), we obtain the expression for the sRI-ERIs,

$$\begin{aligned}
 (\alpha\beta|\gamma\delta) &\approx \sum_{PQ} \sum_{AB}^{N_{aux}} (\alpha\beta|A) V_{AP}^{-\frac{1}{2}} I_{PQ} V_{QB}^{-\frac{1}{2}} (B|\gamma\delta) \\
 &= \sum_{PQ} \sum_{AB}^{N_{aux}} (\alpha\beta|A) V_{AP}^{-\frac{1}{2}} \left(\left(\theta \otimes \theta^T \right)_{\xi} \right)_{PQ} V_{QB}^{-\frac{1}{2}} (B|\gamma\delta) \\
 &= \left\langle \left[\sum_A^{N_{aux}} (\alpha\beta|A) \sum_P^{N_{aux}} V_{AP}^{-\frac{1}{2}} \theta_P \right] \left[\sum_B^{N_{aux}} (B|\gamma\delta) \sum_Q^{N_{aux}} \theta_Q^T V_{QB}^{-\frac{1}{2}} \right] \right\rangle_{\xi}, \quad (10)
 \end{aligned}$$

where $\left(\left(\theta \otimes \theta^T \right)_{\xi} \right)_{PQ}$ is the PQ^{th} element of the stochastic identity matrix. We now define the ξ^{th} elements of the stochastic average as

$$R_{\alpha\beta}^{\xi} = \sum_A^{N_{aux}} (\alpha\beta|A) \left[\sum_P^{N_{aux}} \left[V_{AP}^{-\frac{1}{2}} \theta_P^{\xi} \right] \right] \equiv \sum_A^{N_{aux}} (\alpha\beta|A) L_A^{\xi}. \quad (11)$$

With this definition, the ERI in the AO basis [Eq. (10)] is now given by a stochastic average, an $O(N_s N_{AO}^4)$ step,

$$(\alpha\beta|\gamma\delta) \approx \frac{1}{N_s} \sum_{\xi} R_{\alpha\beta}^{\xi} R_{\gamma\delta}^{\xi} \equiv \langle R_{\alpha\beta} R_{\gamma\delta} \rangle_{\xi}. \quad (12)$$

The introduction of the stochastic identity matrix decouples the index Q in Eq. (6). Previous work has shown that the trade-off between *decoupling* this index and the introduction of stochastic noise allows for a reduced overall scaling and computational time.⁴³ It was also shown that the stochastic error of the elements of the identity matrix and therefore the error of the ERIs is governed by the number of stochastic orbitals, N_s . This is due to the fact that it is the *length* of stochastic arrays, N_{aux} , that increases with the system size, not the number of stochastic orbitals.

C. Stochastic resolution of identity second-order Born self-energy

In this section, we demonstrate how the sRI-ERIs may be applied to the second-order approximation to the self-energy to reduce the overall scaling of computing $\Sigma(\tau)$ [cf., Eq. (2)]. Using the stochastic resolution of identity of ERIs [Eq. (12)], the second-order Born self-energy in Eq. (3) can be represented by

$$\begin{aligned}
 \Sigma(\tau)_{ij} &= \left\langle \sum_{klmnpq} -G_{kl}(\tau) G_{mn}(\tau) G_{pq}(\beta - \tau) R_{ik} R_{mq} (2R'_{ij} R'_{pn} - R'_{nj} R'_{ip}) \right\rangle_{\zeta, \zeta'} \\
 &= \left\langle \sum_{klmnpq} -2G_{kl}(\tau) G_{mn}(\tau) G_{pq}(\beta - \tau) R_{ik} R_{mq} R'_{ij} R'_{pn} \right. \\
 &\quad \left. + G_{kl}(\tau) G_{mn}(\tau) G_{pq}(\beta - \tau) R_{ik} R_{mq} R'_{nj} R'_{ip} \right\rangle_{\zeta, \zeta'} \\
 &= \left\langle \sum_{klmnpq} -2\Sigma^{dir}(\tau) + \Sigma^{ex}(\tau) \right\rangle_{\zeta, \zeta'}. \quad (13)
 \end{aligned}$$

We note that, in the above equations, R and R' matrices use different sets of stochastic orbitals. In the above equations, we have also defined

$$\begin{aligned}
 \Sigma_{ij}^{dir}(\tau) &= \left\langle \sum_{klmnpq} -2G_{kl}(\tau) G_{mn}(\tau) G_{pq}(\beta - \tau) R_{ik} R_{mq} R'_{ij} R'_{pn} \right\rangle_{\zeta, \zeta'}, \quad (14) \\
 \Sigma_{ij}^{ex}(\tau) &= \left\langle \sum_{klmnpq} G_{kl}(\tau) G_{mn}(\tau) G_{pq}(\beta - \tau) R_{ik} R_{mq} R'_{nj} R'_{ip} \right\rangle_{\zeta, \zeta'}.
 \end{aligned}$$

In the above equations, the evaluation of $\Sigma_{ij}^{dir}(\tau)$ and $\Sigma_{ij}^{ex}(\tau)$ is done now simply by matrix products (or matrix contractions), which scale as $O(N^3)$, as long as the number of stochastic orbitals does not change with the size of the system. It should also be noted that sRI-ERIs can also be used to reduce the scaling to $O(N^3)$ when evaluating the third order or higher self-energy for all contractions in self-energy are now matrix products. That being said, for now, we limit ourselves to the second order Born self-energy.

III. RESULTS AND DISCUSSION

To study the observed scaling, stochastic errors, and the impact of the prefactors, N_s and N_t , on the sRI-GF2 method, we test the approach on a hydrogen dimer chain H_n with sto-3g basis, with n being the number of H atoms. In such a hydrogen dimer chain, the distance between two hydrogen atoms within a dimer is 0.74 Å, and the distance between two H atoms over the nearest hydrogen dimers is 1.26 Å. For such a system, we use cc-pvdz as a fitting basis set (we use cc-pvdz-jkfit for Fock Matrix elements and cc-pvdz-ri for self-energy). In addition, we restrict ourselves to closed shell calculations. This system served as a test bed for other GF2 implementations as well as for the stochastic approach of Ref. 42.

We do not take advantage of the locality of the atomic basis and the resulting sparsity of the ERIs which makes the current applications more realistic for 3D structures, where sparsity often occurs at much larger system sizes. We note in passing that using the sparsity of the 4-index ERIs would further reduce the cost of computing the self-energy.

A. Comparison of GF2 in frequency and imaginary time domain

We first compare the imaginary time and imaginary frequency sRI-GF2 methods. One of the most important properties we can gain from Matsubara GF is the correlation energy, which is given by

$$E_{corr} = \int_0^{\beta} d\tau \text{Tr}(\Sigma(\tau) \mathbf{G}(\beta - \tau)). \quad (15)$$

We generally used $N_{\tau} = 256$ imaginary time points for the integrals. In Table I, we compare correlation energy from imaginary time and imaginary frequency sRI-GF2 methods on H_2 , H_{10} , H_{20} , and H_{40} . We find excellent agreement between the two sRI-GF2 methods, up to an error of 6 meV for the total correlation energy (about 0.2 meV for correlation energy per electron). The computational cost of the two methods is comparable and is mainly determined by the evaluation of the self-energy. The results shown in the remainder of the paper were obtained with the frequency domain calculation, but the analysis applies to both methods.

B. Correlation energy

In Fig. 1, we plot the correlation energy for the hydrogen dimer chain as a function of number of hydrogen atoms. We use

TABLE I. Correlation energy (in eV) from time sRI-GF2 and frequency sRI-GF2, and the CPU time ratio for sRI-GF2 between frequency and time domain formulations.

	E_{corr} from time sRI-GF2	E_{corr} frequency sRI-GF2	CPU time ratio (%)
H_2	-0.6546	-0.6548	47.9
H_{10}	-3.7541	-3.7565	37.7
H_{20}	-6.3811	-6.3835	63.1
H_{40}	-13.2038	-13.2092	61.1

$N_s = 800$ stochastic orbitals for each sRI-GF2 calculation. The final results are then averaged over 10 independent runs. We find that sRI-GF2 results provide a very good agreement with the deterministic GF2 results. Note also that the correlation energy grows nearly linearly with the system size, which signifies weak correlations in the hydrogen dimer chain. We also note that the difference in correlation energy per electron [Fig. 1(b)] between deterministic and stochastic GF2 is within the statistical error.

Note also that from Fig. 1(b), the statistical error bar for energy per electron does not increase with the system size. In other words, to obtain constant accuracy for energy per electron for different system sizes, one does not need to increase the number of stochastic orbitals; therefore, our sRI-GF2 approach scales as N^3 with the system size or number of electrons. However, if one does want to obtain the same accuracy for the total energy (instead of energy per electron), since the stochastic error is proportional to the inverse square root of the sampling size (central limit theorem), our sRI-GF2 approach will scale as N^5 with system size, which is the same scaling as the deterministic implementation of GF2. That being said, for extended systems, e.g., bulk materials, the meaningful quantities are energy per particle or per unit cell instead of the total energy. We expect our sRI-GF2 method will be very useful for such extended systems.

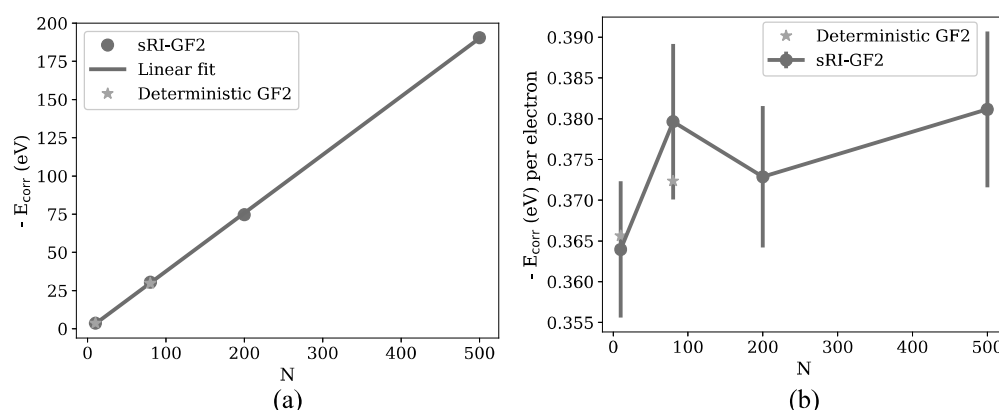


FIG. 1. (a) Correlation energy as a function of the number of H atoms ($N = 10, 80, 200, 500$). We have averaged over 10 independent sRF-GF2 runs. sRI-GF2 results agree well with deterministic GF2 well. For a weakly correlated system such as H dimer chain, the correlation energy grows almost linearly with the number of H atoms. (b) Correlation energy per electron as a function of the number of H atoms. The stochastic error is estimated using 10 independent sRI-GF2 runs, $\frac{\sigma}{\sqrt{10}}$. Here, σ is the standard deviation of 10 independent sRI-GF2 runs with different sampling seeds. The difference in correlation energy per electron between deterministic and stochastic GF2 are within the statistical error. In both cases, we have used $N_s = 800$ stochastic orbitals in our sRI-GF2 calculations.

C. Error estimation

In Fig. 2, we plot the average correlation energy per electron as a function of the number of stochastic orbitals. The stochastic error is the standard deviation of the 10 independent sRI-GF2 runs, which is roughly 0.02 eV. Clearly, the error bars decrease with the increasing number of stochastic orbitals and do not depend on the system size. Note also that the average correlation energy does not show a monotonic behavior with the number of stochastic orbitals. In Ref. 42, the authors report a linear dependence between the average correlation energy and the number of stochastic orbitals due to a systematic bias which is observable due to the small statistical errors. Here, we do not observe a clean linear dependence, suggesting that our bias, if there is any, is much smaller than our statistical error. This also implies that in order to converge the results to a given statistical error, the number of stochastic orbitals does not need to increase with the system size (at least for the system size we are testing here, under 1000 electrons).

D. Timing

In Fig. 3, we show the overall computational scaling of our sRI-GF2 method and compare it to the deterministic version. In all cases, we used, as mentioned, $N_r = 256$, and roughly 20 SCF cycles were required for convergence. Furthermore, $N_\omega = 20000$ frequency points (cubic splined from 200 frequency points⁴²) were used. All calculations used the 32-core Intel-Xeon processor E5-2698 v3 (“Haswell”) at 2.3 GHz on a single node. The wall time for sRI-GF2 is tested for a single run with 800 stochastic orbitals.

In our calculations, we do not rely on the sparsity of the 4-index ERIs nor on the sparsity of the overlap matrix or the Fockian. Using the sparsity of the 4-index ERIs would further reduce the computational cost for both the deterministic and stochastic GF2 implementations. For larger systems, we find that the computational cost scales as N^3 , with N being the number of hydrogen atoms. As

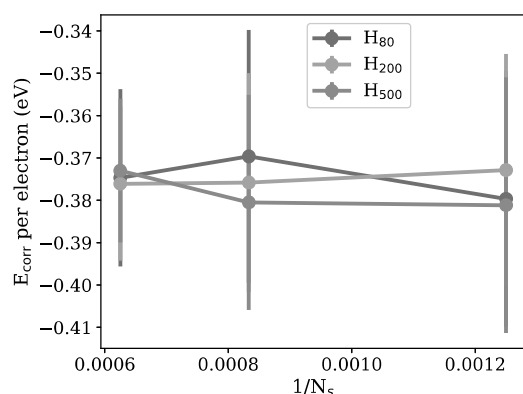


FIG. 2. Correlation energy per electron as a function of the number of stochastic orbitals N_s , for H_{80} , H_{200} , and H_{500} , respectively. We run 10 independent sRI-GF2 calculations, and the stochastic error is the standard deviation of the 10 correlation energy per electrons. In Ref. 42, systemic bias is reported where the correlation energy decreases almost linearly with the increase in stochastic orbitals. Here, there is no clear linear dependence on N_s since the statistical error is larger compared to the bias.

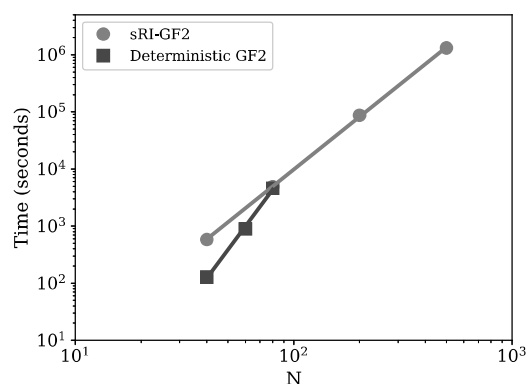


FIG. 3. Wall time as a function of number of electrons on a single node (32-core) modern Xeon CPU. We see that the computational cost scales as $N^{3.1}$ (from fitting), with N being the number of H atoms. The original GF2, without such stochastic resolution of identity, however, scales as $N^{5.1}$ (from fitting) with the number of H atoms N .

stated above, the most time consuming part is calculating the self-energy. Using the stochastic resolution of identity, calculating the self-energy is simply matrix multiplication, which scales N^3 . The deterministic GF2 scales as N^5 . We find that the crossover between the deterministic and stochastic GF2 is around 80 hydrogen atoms for the accuracy of 0.01 eV in correlation energy per particle (as estimated by the standard error of the mean value over 10 independent runs, $\frac{\sigma}{\sqrt{10}}$) given by 800 stochastic orbitals.

IV. CONCLUSIONS

We have presented two stochastic versions of the Matsubara Green's function method to calculate ground states properties of

molecular systems. Using the resolution of identity, the second order Born self-energy is represented in a matrix product form, which reduces the scaling from the original N^5 to N^3 . Our sRI-GF2 can be seen as an extension of the stochastic MP2 results as presented in Ref. 43 and offers an alternative to the one as presented in Ref. 42. We have shown that our sRI-GF2 reproduces deterministic GF2 up to the stochastic error. In addition, we do not see systemic bias as we increase the number of stochastic orbitals. Future work will investigate how our stochastic resolution of identity can be implemented in the real time GF method. Such work is ongoing.

ACKNOWLEDGMENTS

The authors would like to thank Dr. Ming Chen for fruitful discussions. T.Y.T. is supported by a fellowship from the Molecular Sciences Software Institute (MolSSI) under NSF Grant No. ACI-1547580. D.N. and E.R. are grateful for support by the Center for Computational Study of Excited State Phenomena in Energy Materials (C2SEPEM) at the Lawrence Berkeley National Laboratory, which is funded by the U.S. Department of Energy, Office of Science, Basic Energy Sciences, Materials Sciences and Engineering Division under Contract No. DEAC02-05CH11231 as part of the Computational Materials Sciences Program. R.B. is grateful for support by Binational US-Israel Science Foundation Grant No. BSF 2015687. Resources of the National Energy Research Scientific Computing Center (NERSC), a U.S. Department of Energy Office of Science User Facility operated under Contract No. DE-AC02-05CH11231, are greatly acknowledged.

REFERENCES

- P. Hohenberg and W. Kohn, *Phys. Rev.* **136**, B864 (1964).
- W. Kohn and L. J. Sham, *Phys. Rev.* **140**, A1133 (1965).
- S. Baroni, S. de Gironcoli, A. Dal Corso, and P. Giannozzi, *Rev. Mod. Phys.* **73**, 515 (2001).
- C. Freysoldt, B. Grabowski, T. Hickel, J. Neugebauer, G. Kresse, A. Janotti, and C. G. van de Walle, *Rev. Mod. Phys.* **86**, 253 (2014).
- T. Frauenheim, G. Seifert, M. Elstner, T. Niehaus, C. Köhler, M. Amkreutz, M. Sternberg, Z. Hajnal, A. Di Carlo, and S. Suhai, *J. Phys.: Condens. Matter* **14**, 3015 (2002).
- P. Fulde, *Electron Correlations in Molecules and Solids* (Springer Science & Business Media, 2012), Vol. 100.
- R. M. Martin, *Electronic Structure: Basic Theory and Practical Methods* (Cambridge University Press, 2004).
- A. L. Fetter and J. D. Walecka, *Quantum Theory of Many-Particle Systems* (Courier Corporation, 2012).
- M. Häser and J. Almlöf, *J. Chem. Phys.* **96**, 489 (1992).
- L. Hedin, *Phys. Rev.* **139**, A796 (1965).
- D. Bohm and D. Pines, *Phys. Rev.* **82**, 625 (1951).
- M. Gell-Mann and K. A. Brueckner, *Phys. Rev.* **106**, 364 (1957).
- J. Schirmer, *Phys. Rev. A* **26**, 2395 (1982).
- M. S. Hybertsen and S. G. Louie, *Phys. Rev. Lett.* **55**, 1418 (1985).
- M. Shishkin and G. Kresse, *Phys. Rev. B* **75**, 235102 (2007).
- N. Marom, F. Caruso, X. Ren, O. T. Hofmann, T. Körzdörfer, J. R. Chelikowsky, A. Rubio, M. Scheffler, and P. Rinke, *Phys. Rev. B* **86**, 245127 (2012).
- F. Caruso, P. Rinke, X. Ren, A. Rubio, and M. Scheffler, *Phys. Rev. B* **88**, 075105 (2013).
- M. J. van Setten, F. Weigend, and F. Evers, *J. Chem. Theory Comput.* **9**, 232 (2012).

- ¹⁹S. Baroni, P. Giannozzi, and A. Testa, *Phys. Rev. Lett.* **58**, 1861 (1987).
- ²⁰G. Onida, L. Reining, and A. Rubio, *Rev. Mod. Phys.* **74**, 601 (2002).
- ²¹J. Zaanen, G. A. Sawatzky, and J. W. Allen, *Phys. Rev. Lett.* **55**, 418 (1985).
- ²²M. F. Herman, K. F. Freed, and D. Yeager, *Adv. Chem. Phys.* **48**, 1 (1981).
- ²³W. von Niessen, G. H. F. Diercksen, and L. S. Cederbaum, *J. Chem. Phys.* **67**, 4124 (1977).
- ²⁴M. Rohlfing and S. G. Louie, *Phys. Rev. B* **62**, 4927 (2000).
- ²⁵S. Refaely-Abramson, S. Sharifzadeh, N. Govind, J. Autschbach, J. B. Neaton, R. Baer, and L. Kronik, *Phys. Rev. Lett.* **109**, 226405 (2012).
- ²⁶W. G. Aulbur, M. Städele, and A. Görling, *Phys. Rev. B* **62**, 7121 (2000).
- ²⁷P. Koval, D. Foerster, and D. Sánchez-Portal, *Phys. Rev. B* **89**, 155417 (2014).
- ²⁸L. Cederbaum and W. Domcke, *Adv. Chem. Phys.* **36**, 205 (1977).
- ²⁹V. Vlcek, E. Rabani, D. Neuhauser, and R. Baer, *J. Chem. Theory Comput.* **13**, 4997 (2017).
- ³⁰G. Stefanucci and R. van Leeuwen, *Nonequilibrium Many-Body Theory of Quantum Systems: A Modern Introduction* (Cambridge University Press, 2013).
- ³¹L. Cederbaum, *J. Phys. B: At. Mol. Phys.* **8**, 290 (1975).
- ³²L. Holleboom and J. Snijders, *J. Chem. Phys.* **93**, 5826 (1990).
- ³³N. E. Dahlen, R. van Leeuwen, and U. von Barth, *Int. J. Quantum Chem.* **101**, 512 (2005).
- ³⁴J. J. Phillips and D. Zgid, *J. Chem. Phys.* **140**, 241101 (2014).
- ³⁵F. Pavošević, C. Peng, J. Ortiz, and E. F. Valeev, *J. Chem. Phys.* **147**, 121101 (2017).
- ³⁶Y.-y. Ohnishi and S. Ten-no, *J. Comput. Chem.* **37**, 2447 (2016).
- ³⁷A. R. Welden, J. J. Phillips, and D. Zgid, e-print arXiv:1505.05575 (2015).
- ³⁸N. E. Dahlen and R. van Leeuwen, *J. Chem. Phys.* **122**, 164102 (2005).
- ³⁹A. A. Kananenka, J. J. Phillips, and D. Zgid, *J. Chem. Theory Comput.* **12**, 564 (2016).
- ⁴⁰A. A. Rusakov and D. Zgid, *J. Chem. Phys.* **144**, 054106 (2016).
- ⁴¹A. A. Rusakov, J. J. Phillips, and D. Zgid, *J. Chem. Phys.* **141**, 194105 (2014).
- ⁴²D. Neuhauser, R. Baer, and D. Zgid, *J. Chem. Theory Comput.* **13**, 5396 (2017).
- ⁴³T. Y. Takeshita, W. A. de Jong, D. Neuhauser, R. Baer, and E. Rabani, *J. Chem. Theory Comput.* **13**, 4605 (2017).