Stochastic resolution of identity to CC2 for large systems: excited state properties

Chongxiao Zhao,[1,2,3] Qi Ou,[4] Joonho Lee,[5,*] and Wenjie Dou[2,3,*]

[1]*Department of Chemistry, Zhejiang University, Hangzhou 310027, China*

[2]*Department of Chemistry, School of Science, Westlake University, Hangzhou, Zhejiang 310024, China*

[3]*Institute of Natural Sciences, Westlake Institute for Advanced Study, Hangzhou, Zhejiang 310024, China*

[4]*AI for Science Institute, Beijing 100080, China*

[5]*Department of Chemistry and Chemical Biology, Harvard University, Cambridge, MA 02138, USA*

We apply a stochastic resolution of identity approximation (sRI) to the CC2 method for excitation energy calculations. A set of stochastic orbitals are employed to decouple the crucial 4-index electron repulsion integrals and optimize the contraction steps in CC2 response theory. The CC2 response for excitations builds upon sRI-CC2 ground-state calculations, which scales as $O(N^3)$, where N is a measure for the system size. Overall, the current algorithm for excited states also allows a sharp scaling reduction from original $O(N^5)$ to $O(N^3)$. We test the sRI-CC2 for different molecular systems and basis sets, and we show our sRI-CC2 method can accurately reproduce the results of deterministic CC2 approach. Our sRI-CC2 exhibits an experimental scaling of $O(N^{2.88})$ for a hydrogen dimer chain, allowing us to calculate systems with nearly thousands of electrons.

1. INTRODUCTION

Calculating excited state properties for complex molecular systems and extended systems is one of the most challenging tasks in theoretical chemistry. TDDFT is probably one of the most useful means to obtain the excitation energy, due to its affordable computation cost. However, TDDFT often overestimates the energy gaps due to lack of exiton interactions. CC theory is more accurate in predicting excited state properties, however, the high computional cost of the CC method prevents its application to large systems. In all post-HF methods, the implementation of 4-index electron repulsion integrals (ERI) is one of the most challenging steps in electronic structure approaches. High demands for both disk space and computational time hamper their applications to larger systems with hundreds of atoms and have become an intractable bottleneck, e.g., for the coupled cluster (CC) theory.

1

Among a hierarchy of CC models, CCS, CC2, CCSD, CC3, CCSDT, [1-4] etc., the CC2 model is the most affordable one, which offers accurate ground state energies as well as excitation energies, both correct to second order in the electron fluctuation potential for single-excitation dominated transitions.[5] The CC2 model was first formulated and implemented by Christiansen and Koch *et al*.[6] as an approximation to full CCSD. Compared with conventional standard CCSD model, the CC2 model is capable of the calculation of excitation energies of modest accuracy (≈ 0.3 eV)[7] within the framework of CC response theory[8-11] and formally scales as $O(N^5)$ (with N being a measure of the system size).

Over the past 30 years, many attempts to reduce the computational cost associated with 4-index ERIs have been reported. One strategy is to introduce the resolution of identity (RI) approximation[12-17]. A set of pre-optimized auxiliary basis functions are utilized to expensive tensor contractions without sacrificing the accuracy. Another approach is Cholesky decomposition (CD)[18-21] , which rearranges the 2-electron ERIs into an iterative process controlled by a single parameter. In both approaches, high rank integrals are disassembled into lower ones, such that storage requirements and CPU time are significantly reduced. However, the overall computational scaling of CC2 remains unaltered as $O(N^5)$, which inevitably hinders application of CC2 for extended systems.

Recently, a stochastic approach to RI approximation, abbreviated as sRI approximation, serves an alternative to handle the costly 4-index ERIs. In the sRI approach, an additional set of random orbitals is employed to further decouple the 4-index ERI tensors. Remarkably, the number of the stochastic orbitals do not need to increase with the system size for intensive quantities, which allows for significant reduction of the computational scalings. Till now, the stochastic orbital appaoaches have been implemented in many quantum chemistry methods, including MP2[22-24], GF2[25-29] and DFT[30,31] *et al*.[32,33]

In a previous work, we have implemented sRI approach to CC2 for ground state calculations[34], where we demonstrate that sRI-CC2 allows for a remarkable scaling reduction from $O(N^5)$ to $O(N^3)$ without significantly affecting the accuracy. In this work, we further extend the sRI approach to CC2 response theory for excitation energy calculations. As tested on a variety of molecular system, we show that our sRI-CC2 response theory not only provides accurate results for excited state as compared against RI-CC2, but also exhibits an overall computational scaling of $O(N^3)$. We expect that the sRI-CC2 response theory will be very useful in describing excited state properties for large

systems with 1000 electrons or more.

The manuscript is organized as follows: In the section II, we present the sRI-CC2 theory and the algorithm to calculate excitation energies within the framework of response theory in detail. Section III, we demonstrate the performance of the sRI-CC2 for a variety of molecular systems. We benchmark the sRI-CC2 against results from RI-CC2 in the Q-Chem package[35]. In the section IV, we conclude.

2. THEORY

We use the conventional notations in Table 1 to represent the items in the following sections. Furthermore, the total number of atomic orbital (AO) basis functions, auxiliary basis functions, occupied molecular orbitals (MOs), and virtual MOs are denoted by N_{AO}, N_{aux}, N_{occ}, and N_{virt} respectively.

Table 1. Summary of notations in the following equations.

item	function or indice
AO Gaussian basis functions	$\chi_\alpha(r_1), \chi_\beta(r_1), \chi_\gamma(r_1), \chi_\delta(r_1), \ldots$
auxiliary basis functions	$P, Q, R, S, \ldots$
general sets of AOs	$\alpha, \beta, \gamma, \delta, \ldots$
general sets of MOs	$p, q, r, s, \ldots$
occupied (active) MOs	$i, j, k, l, \ldots$
unoccupied (virtual) MOs	$a, b, c, d, \ldots$

2.1. CC2 Excitation Energy Calculation

In CC2 model, the amplitudes equations for singles and doubles can be written as

$$\Omega_{\mu_1} = \langle \mu_1 | \widehat{H} + [\widehat{H}, T_2] | HF \rangle = 0 \tag{1}$$

$$\Omega_{\mu_2} = \langle \mu_2 | \widehat{H} + [F, T_2] | HF \rangle = 0 \tag{2}$$

where Ω_{μ_1} and Ω_{μ_2} are single and double excitation vector functions. μ_1 and μ_2 are single and double excitation manifolds. T_2 is the double excitation cluster operator. $|HF\rangle$ and F are SCF Hartree-Fock reference state and SCF Fock

3

operator. The Hamiltonian H undergoes a T_1-transformation with the exponential function of the single excitation cluster operator T_1:

$$\hat{H} = exp(-T_1) \, H \, exp(T_1) \tag{3}$$

Similarly, the T_1-transformed two-electron MO integrals are given by:

$$(pq\hat{|}rs) = \sum_{\alpha\beta\gamma\delta} \Lambda^p_{\alpha p} \Lambda^h_{\beta q} \Lambda^p_{\gamma r} \Lambda^h_{\delta s} (\alpha\beta|\gamma\delta) \tag{4}$$

$$\Lambda^p = C(I - t_1^T) \tag{5}$$

$$\Lambda^h = C(I + t_1) \tag{6}$$

Here Λ^p and Λ^h are transformation matrices particle and hole operators. t_1 is the auxiliary matrix comprised of singles cluster amplitudes $\{t_i^a\}$:

$$t_1 = \begin{pmatrix} 0 & 0 \\ \{t_i^a\} & 0 \end{pmatrix} \tag{7}$$

The double-excitation amplitudes $\{t_{ij}^{ab}\}$ with their similarity transformed forms $\{\hat{t}_{ij}^{ab}\}$ can be calculated explicitly with a given set of $\{t_i^a\}$:

$$t_{ij}^{ab} = \frac{(ai\hat{|}bj)}{\varepsilon_i - \varepsilon_a + \varepsilon_j - \varepsilon_b} \tag{8}$$

$$\hat{t}_{ij}^{ab} = (1 + \delta_{ij}\delta_{ab})(2\,t_{ij}^{ab} - t_{ij}^{ba}) \tag{9}$$

In the RI-CC2 implementation by C. Hättig and F. Weigend[36], explicit expressions for the Ω_{μ_1} and Ω_{μ_2} have been presented. In our previous work on CC2 ground state properties[34], we have employed the sRI approximation to into these two excitation vector functions, such that we can achieve $O(N^3)$ scaling with iterative solution to the single excitation amplitudes. To further obtain the excitation energies, we need to solve the eigen problem of the Jacobian matrix $A_{\mu_i \nu_j}$ in CC response theory, which is defined as the derivative of the vector function:

$$A_{\mu_i \nu_j} = \frac{\partial \Omega_{\mu_i}}{\partial t_{\nu_j}} = \begin{pmatrix} \langle \mu_1 | [\hat{H}, \tau_{\nu_1}] + [[\hat{H}, \tau_{\nu_1}], T_2] | HF \rangle & \langle \mu_1 | [\hat{H}, \tau_{\nu_2}] | HF \rangle \\ \langle \mu_2 | [\hat{H}, \tau_{\nu_1}] | HF \rangle & \delta_{\mu_2 \nu_2} \epsilon_{\mu_2} \end{pmatrix} \tag{10}$$

Here the $A_{\mu_2 \nu_2}$ block in this Jacobian matrix is diagonal with the matrix elements:

$$\epsilon_{\mu_2} = \epsilon_{aibj} = \epsilon_a - \epsilon_i + \epsilon_b - \epsilon_j \tag{11}$$

The eigen problem of the Jacobian matrix can be solved interatively as well, after we partition the CC2 eigenvalue problem into the following two equations:

$$\sigma_{\mu_1}(\omega, E_{\mu_1}) = \left[A_{\mu_1 v_1} - \frac{A_{\mu_1 \gamma_2} A_{\gamma_2 v_1}}{\epsilon_{\gamma_2} - \omega} \right] E_{v_1} = A_{\mu_1 v_1}^{eff}(\omega) \, E_{v_1} = \omega E_{\mu_1} \tag{12}$$

$$E_{\mu_2} = -\frac{A_{\mu_2 v_1} E_{v_1}}{\epsilon_{\mu_2} - \omega} \tag{13}$$

Eq. (13) shows that the eigenvector in the double excitation manifold E_{μ_2} can be obtained from the single one E_{v_1}. Such that we can solve the non-linear eigenvalue problem in Eq. (12), which only contains the single excitation manifold.

An initial guess for eigenvalue ω and for eigenvector b_{μ_1} can be obtained from models with lower scaling such as CIS or CCS. Similar to Ω_{μ_1}, the transformed vector $\sigma_{\mu_1}(\omega, b_1)$ is splitted into many terms:

$$\sigma_{ai}(\omega, b_1) = \sigma_{ai}^0 + \sigma_{ai}^G + \sigma_{ai}^H + \sigma_{ai}^I + \sigma_{ai}^J \tag{14}$$

$$\sigma_{ai}^0 = \sum_b E_{ab}^{(1)} b_{bi} - \sum_j b_{aj} E_{ji}^{(2)} \tag{15}$$

$$\sigma_{ai}^G = +\sum_{dlc} \hat{b}_{il}^{cd} (ld\hat{|}ac) \qquad\qquad \sigma_{ai}^H = -\sum_{dlk} \hat{b}_{kl}^{ad} (ld\hat{|}ki)$$

$$\sigma_{ai}^I = \sum_{dl} (\hat{b}_{il}^{ad} \hat{F}_{ld} + \hat{t}_{il}^{ad} \bar{F}_{ld}) \qquad\qquad \sigma_{ai}^J = \bar{F}_{ai}' \tag{16$\sim$19}$$

$$E_{ab}^{(1)} = \hat{F}_{ab} - \sum_{dlk} \hat{t}_{kl}^{ad} (ld\hat{|}kb) \qquad\qquad E_{ij}^{(2)} = \hat{F}_{ij} + \sum_{dlc} \hat{t}_{il}^{cd} (ld\hat{|}jc) \tag{20$\sim$21}$$

$$\bar{F}_{ld} = \sum_{ck} [2(ld\hat{|}kc) - (lc\hat{|}kd)] b_{ck} \qquad \bar{F}_{ai}' = \sum_{ck} [2(ai\hat{|}kc) - (ac\hat{|}ki)] b_{ck} \tag{22$\sim$23}$$

where $\hat{F}$ denotes the similarity transformed Fock matrix

$$\hat{F}_{pq} = \sum_{pq} \Lambda_{\mu p}^p \Lambda_{vq}^h F_{\mu v} + \sum_{ck} [2(pq\hat{|}ck) - (pk\hat{|}cq)] t_{ck} \tag{24}$$

The auxiliary vector $\hat{b}_{ij}^{ab}$ in the above equation is defined as

$$\hat{b}_{ij}^{ab} = \frac{2(ai\bar{|}bj) - (bi\bar{|}aj)}{\varepsilon_i - \varepsilon_a + \varepsilon_j - \varepsilon_b + \omega} \tag{25}$$

Here we have further defined the modified 4-index ERIs $(ai\bar{|}bj)$ as

$$(ai\bar{|}bj) = \hat{P}_{ij}^{ab} \sum_{\alpha\beta\gamma\delta} (\bar{\Lambda}_{\alpha a}^p \Lambda_{\beta i}^h + \Lambda_{\alpha a}^p \bar{\Lambda}_{\beta i}^h) \Lambda_{\gamma b}^p \Lambda_{\delta j}^h (\alpha\beta|\gamma\delta) \tag{26}$$

$$\bar{\Lambda}^p = -C b_1^T \qquad\qquad \bar{\Lambda}^h = C b_1 \tag{27$\sim$28}$$

$$b_1 = \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \{b_i^a\} & \mathbf{0} \end{pmatrix} \tag{29}$$

$\hat{P}_{ij}^{ab}$ in the above equation is an operator that symmetrizes pairs index ai and bj: $\hat{P}_{ij}^{ab} f_{ij}^{ab} = f_{ij}^{ab} + f_{ji}^{ba}$. This concludes

the CC2 response theory for excitation energies. The computational cost of the original CC2 and RI-CC2 methods scales as $O(N^5)$. Below, we introduce the algorithm for sRI-CC2 method to achieve scaling of $O(N^3)$.

2.2. Algorithm for sRI-CC2 Excitation Energy

With stochastic orbitals, the four index electron integrals are decoupled into product of the stochastic resolution of the identity:

$$(\alpha\beta|\gamma\delta) \approx \frac{1}{N_s}\sum_{\xi=1}^{N_s} R_{\alpha\beta}^{\xi} R_{\gamma\delta}^{\xi} \equiv \left\langle R_{\alpha\beta}^{\xi} R_{\gamma\delta}^{\xi} \right\rangle_{\xi} \tag{30}$$

Here $R_{\alpha\beta}^{\xi}$ is the stochastic resolution of the identity. ξ indicates the number of stochastic orbitals. The sRI approach builds on RI approach. Details of the formulation of RI and sRI approaches can be found in Appendices A and B.

The sRI-CC2 ground state calculation introduced previously is our starting point for the sRI-CC2 response theory. Note that the intermediate quantities such as $\hat{t}_{ij}^{ab}$ and $(ai|bj)$ are computed in the ground state calculations. We proceed to compute the transformed vector $\sigma_{\mu_1}(\omega, b_1)$ in CC2 response theory as follow.

1. Calculate the ground-state part with sRI and obtain cluster amplitudes t_{μ_1} and $\hat{t}_{\mu_2}$. Note that the overall scaling of the sRI-CC2 for ground state calculation is $O(N^3)$.

2. Compute $E_{ab}^{(1)}$ and $E_{ij}^{(2)}$ from Eq. (20)-(21). Initialize the eigenvalue ω and eigenvector b_{μ_1}.

3. Construct $\bar{\Lambda}^p$ and $\bar{\Lambda}^h$ from Eq. (27)-(28).

4. Respectively obtain each partitions of the transformed vector and eventually add up to $\sigma_{\mu_1}(\omega, b_1)$. Intermediates $(ai|bj)$ and $\hat{b}_{\mu_2}$ are utilized "on the fly" with no need to calculate and store. Take the construction of σ_{ai}^G as an example

$$\left(ai|bj\right) = \hat{P}_{ij}^{ab}\sum_{\alpha\beta\gamma\delta}\left(\bar{\Lambda}_{\alpha a}^p\Lambda_{\beta i}^h + \Lambda_{\alpha a}^p\bar{\Lambda}_{\beta i}^h\right)\Lambda_{\gamma b}^p\Lambda_{\delta j}^h \left\langle R_{\alpha\beta}^{\xi} R_{\gamma\delta}^{\xi}\right\rangle_{\xi}$$

$$= \hat{P}_{ij}^{ab}\left\langle \sum_{\alpha\beta}\left(\bar{\Lambda}_{\alpha a}^p\Lambda_{\beta i}^h + \Lambda_{\alpha a}^p\bar{\Lambda}_{\beta i}^h\right)R_{\alpha\beta}^{\xi}\hat{R}_{bj}^{\xi}\right\rangle_{\xi} = \hat{P}_{ij}^{ab}\left\langle \bar{R}_{ai}^{\xi}\hat{R}_{bj}^{\xi}\right\rangle_{\xi} = \left\langle \bar{R}_{ai}^{\xi}\hat{R}_{bj}^{\xi} + \bar{R}_{bj}^{\xi}\hat{R}_{ai}^{\xi}\right\rangle_{\xi} \tag{31}$$

$$\sigma_{\mu_1}^G = \sum_{abj}\hat{b}_{ij}^{ab}(jb|ca)$$

$$= \sum_{abj}\frac{2(ai|bj) - (bi|aj)}{\varepsilon_i - \varepsilon_a + \varepsilon_j - \varepsilon_b + \omega}\left\langle \hat{R}_{jb}^{\xi'}\hat{R}_{ca}^{\xi'}\right\rangle_{\xi'}$$

$$= \left\langle \sum_{abj}\frac{2(\bar{R}_{ai}^{\xi}\hat{R}_{bj}^{\xi} + \bar{R}_{bj}^{\xi}\hat{R}_{ai}^{\xi}) - (\bar{R}_{bi}^{\xi}\hat{R}_{aj}^{\xi} + \bar{R}_{aj}^{\xi}\hat{R}_{bi}^{\xi})}{\varepsilon_i - \varepsilon_a + \varepsilon_j - \varepsilon_b + \omega}\hat{R}_{jb}^{\xi'}\hat{R}_{ca}^{\xi'}\right\rangle_{\xi\xi'} \tag{32}$$

6

Each item in Eq. (32) can be transformed into the expression of time integrals

$$\frac{\bar{R}^{\xi}_{ai}\hat{R}^{\xi}_{bj}}{\varepsilon_i - \varepsilon_a + \varepsilon_j - \varepsilon_b + \omega}\hat{R}^{\xi\,\prime}_{jb}\hat{R}^{\xi\,\prime}_{ca} = -\int_0^{\infty}\left[\hat{R}^{\xi}_{bj}\mathrm{e}^{(\varepsilon_j-\varepsilon_b)t}\hat{R}^{\xi\,\prime}_{jb}\right]\bar{R}^{\xi}_{ai}\mathrm{e}^{(\varepsilon_i-\varepsilon_a)t}\hat{R}^{\xi\,\prime}_{ca}\mathrm{e}^{\omega t}\,dt \tag{33}$$

$$\frac{\bar{R}^{\xi}_{bi}\hat{R}^{\xi}_{aj}}{\varepsilon_i - \varepsilon_a + \varepsilon_j - \varepsilon_b + \omega}\hat{R}^{\xi\,\prime}_{jb}\hat{R}^{\xi\,\prime}_{ca} = -\int_0^{\infty}\left[\bar{R}^{\xi}_{bi}\mathrm{e}^{(\varepsilon_i+\varepsilon_j-\varepsilon_b)t}\hat{R}^{\xi\,\prime}_{jb}\right]\hat{R}^{\xi}_{aj}\mathrm{e}^{(-\varepsilon_a)t}\hat{R}^{\xi\,\prime}_{ca}\mathrm{e}^{\omega t}\,dt \tag{34}$$

Items in the brackets are calculated first to decouple the indices and these four items cost $O(N_s N_t N^3)$, where N_s

is the number of stochastic orbitals and N_t is the number of quadrature points, about 10 with the assistance of

Laplace transform. Since both N_s and N_t are small prefactors, the scaling is thus reduced to $O(N^3)$. All sub-items

are processed in this way to complete $\sigma_{\mu_1}(\omega, b_1)$.

5. Once the transformed vector $\sigma^{(i)}_{\mu_1} = \sigma_{\mu_1}(\omega^{(i)}, b_1^{(i)})$ is constructed in iteration i, an updated eigenvalue $\omega^{(i+1)}$

    and the residual are calculated according to the following equations

    $$\omega^{(i+1)} = \frac{\sigma^{(i)}_{\mu_1} b^{(i)}_{\mu_1}}{||b_1^{(i)}||^2} \tag{35}$$

    $$r^{(i)}_{\mu_1} = \frac{\sigma^{(i)}_{\mu_1} - \omega^{(i+1)} b^{(i)}_{\mu_1}}{||b_1^{(i)}||\, n_{nocc} n_{virt}} \tag{36}$$

    , where n_{occ} and n_{virt} respectively denote the number of occupied and virtual MOs.

6. If the Frobenius norm of the residual $||r^{(i)}_{\mu_1}||$ is larger than a pre-set threshold, e.g., 10^{-5}, a perturbational estimate

    $u^{(i)}_{\mu_1}$ is formed to modify the eigenvector $b^{(i)}_{\mu_1}$, else the iteration is stopped.

    $$u^{(i)}_{\mu_1} = \frac{r^{(i)}_{\mu_1}}{\varepsilon_i - \varepsilon_a} \tag{37}$$

    $$b^{(i+1)}_{\mu_1} = \frac{b^{(i)}_{\mu_1} + u^{(i)}_{\mu_1}}{||b_1^{(i)}||} \tag{38}$$

    and a DIIS algorithm for b_{μ_1} can accelerate this procedure.

7. Repeat steps 3-6 until self-consistency is reached.

    From the steps above, we can see that except steps 1 and 4, scaling as $O(N^3)$, other steps mainly cost $O(N^2)$. In

summary, the overall cost of sRI-CC2 excitation energy calculation is theoretically reduced to $O(N^3)$.

3. RESULTS AND DISCUSSION

To test the accuracy and efficiency of our sRI-CC2, we apply it to a variety of systems with different basis sets for the lowest singlet excitation energies. All the sRI-CC2 energies or CPU time are averaged over 10 different randomly selected sampling seeds. The standard deviations are estimated by the 10 runs (note that the error in the stardard deviation is 0.746σ for the 10 finite number of runs). Meanwhile, we utilize the RI-CC2 in the Q-Chem program package with same parameters for comparison.

3.1. Hydrogen Dimer Chains H_n

We first apply our method to a series of hydrogen dimer chains H_n with sto-3g basis and 1600 stochastic orbitals as test samples, with n being the number of hydrogen atoms ranging from 10 to 300. The distance between two hydrogen atoms in each dimer is 0.74 Å, and the distance between two hydrogen atoms of each adjacent dimer is 1.26 Å.

In Table 2, we list the lowest excitation energies from sRI-CC2 in comparison with RI-CC2 in Q-Chem. We have also list the errors (the difference between the sRI results and the RI results) and standard deviations from 10 runs. We first observe that the CC2 excitation energies slowly decrease with the number of H atoms, in both RI and sRI calculations. Notice also that the results from sRI-CC2 and RI-CC2 agree well with each other, where the error between sRI and RI is always less than 0.2 eV and this would further decrease if we utilize more stochastic orbitals. We note that the standard deviation does not increase with the system size for fixed number of stochastic orbitals. Since the exciation energies are intensive properties, we expect that we do not need to increase the number of stochastic orbitals when increasing the system size.

Table 2. Lowest excitation energies (in eV) for a series of hydrogen dimer chains with sto-3g basis. Here N_H is the number of hydrogen in the corresponding dimer chain, the same value as the number of correlated electrons N_e. The standard deviation σ is calculated from 10 independent samples and in eV.

N_H	RI	sRI	error	std deviation
H_{10}	16.7772	16.8304	0.0532	0.3290
H_{50}	14.9868	14.9688	0.0180	0.3134
H_{100}	14.8981	14.7801	0.1180	0.1620

H$_{200}$	14.8721	15.0026	0.1305	0.2724
H$_{300}$	14.8666	14.6736	0.1930	0.3493
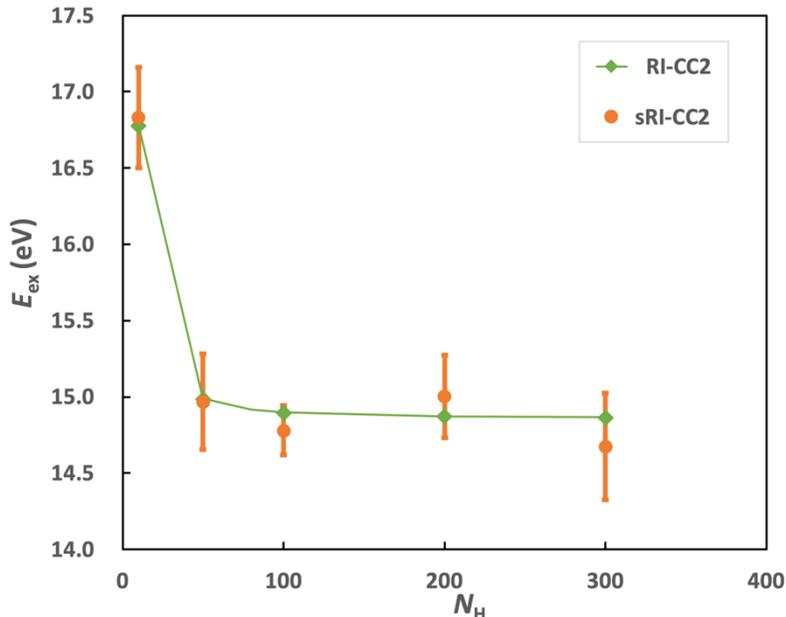
Figure. 1. Excitation energy for sRI-CC2 and RI-CC2 (obtained in Q-Chem) as a function of the number of hydrogen atoms with N_s = 1600. The error bar is from the standard deviation calculated from 10 different seeds.

A more direct comparison is shown in Figure 1. The error bar indicates the standard deviations of 10 independent runs. Results from our sRI-CC2 agree well with RI-CC2 method and all the values from RI-CC2 fall in the error bars, which indicates our sRI-CC2 provides a modest agreement with the standard results. Again, we observe that the statistical error bar does not increase with the system size of hydrogen dimer chains. This observation indicates that we can use fixed number of stochastic orbtials to achieve consistent accuracy for small and large systems. This claim has been made in stochastic MP2[22-24], GF2[25-29], DFT[30,31] and GW[33,37-39] methods as well.

We then test the computational scaling of our sRI-CC2 method. In Figure 2, we plot the CPU time for sRI-CC2 and RI-CC2 calculations as a function of the number of hydrogen atoms. All the calculations are implemented with the high-performance computing (HPC) with Intel CPU, utilizing a single compute node and 64 computational cores. We notice that we need roughly 20 cycles for the SCF interations for convergence in sRI-CC2 calculations, slightly larger than RI-CC2 number of iterations. The overall CPU time for sRI-CC2 is averaged over 10 sampling seeds. Notice that sRI-CC2 exhibits an experimental scaling of $O(N_H^{2.88})$, slightly better than the theoretical scaling of $O(N_H^3)$. RI-CC2 exhibits an experimental scaling of $O(N_H^{4.68})$, which is close to the theoretical scaling of $O(N_H^5)$. The crossover

9

happens at $N_H = 300$, meaning that for larger system, sRI-CC2 will exceed RI-CC2 perfermance. We are currently working on implementation of the sRI-CC2 in QChem package, such that we can further reduce the prefactor of the computational cost in sRI-CC2 calculations and apply the sRI-CC2 method to systems with 1000 electrons or more.
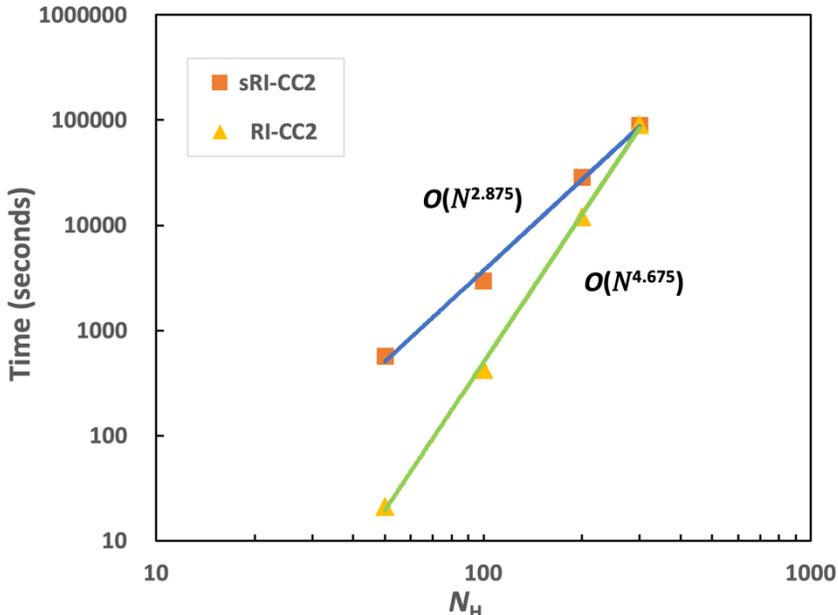
Figure. 2. CPU time as a function of the number of hydrogen atoms with $N_s = 1600$. The computational time of sRI-CC2 is averaged over 10 different runs.

3.2. (All-E)-Olefin Chains $C_{2n}H_{2n+2}$

To test our methods on more realistic molecular systems, we further apply sRI-CC2 to a series of (all-E)-alkenes $C_{2n}H_{2n+2}$, with n ranging from 1 to 8. Here, we have used cc-pVDZ and aug-cc-pVDZ as our basis sets. The experimental geometries[39] are provided in the Supporting Information. We have also used 800 stochastic orbitals for sRI calculations.

In Table 3, we list the lowest excitation energies from RI and sRI calculations. The error indicates the difference between RI results and sRI results. The standard deviations are estimated from 10 sRI runs. To better illustrate our results, we further plot the lowest excitation energies as a function of system size in Figure 3. The error bar indicates the standard deviation in sRI runs. We first notice that the excitation energies decrease with the system size, reaching to a plateau eventually. In addition, from the table and the figure, we observe that sRI results agree well with RI results, with all errors lying in the error bar (standard deviation). Notice also that neither the errors and the standard deviation increases with the system size. Finally, we observe that the stochastic error does not strongly depends on the basis set:

10

when we increase basis from cc-pvdz to aug-cc-pvdz, the stochastic error does not increase noticeably.

Table 3. Results for of (all-E)-$C_{2n}H_{2n+2}$ alkenes with n = 1~5. Here N_e is the number of correlated electrons. The excitation energies from RI and sRI, their errors, and the standard deviations of 10 sRI runs are in eV.

N_e	molecule	cc-pVDZ				aug-cc-pVDZ			
		RI	sRI	error	std deviation	RI	sRI	error	std deviation
16	C_2H_4	8.7502	8.5973	0.1529	0.4688	7.1618	7.2212	0.0594	0.1998
30	C_4H_6	6.6371	6.6706	0.0335	0.4164	6.1335	6.1205	0.0130	0.2352
44	C_6H_8	5.4894	5.6185	0.1291	0.2619	5.5858	5.2959	0.2899	0.3512
58	C_8H_{10}	4.7597	4.6807	0.0790	0.3017	5.2470	5.1938	0.0532	0.2745
72	$C_{10}H_{12}$	4.3393	4.3124	0.0269	0.2963	5.0650	5.1361	0.0711	0.3080
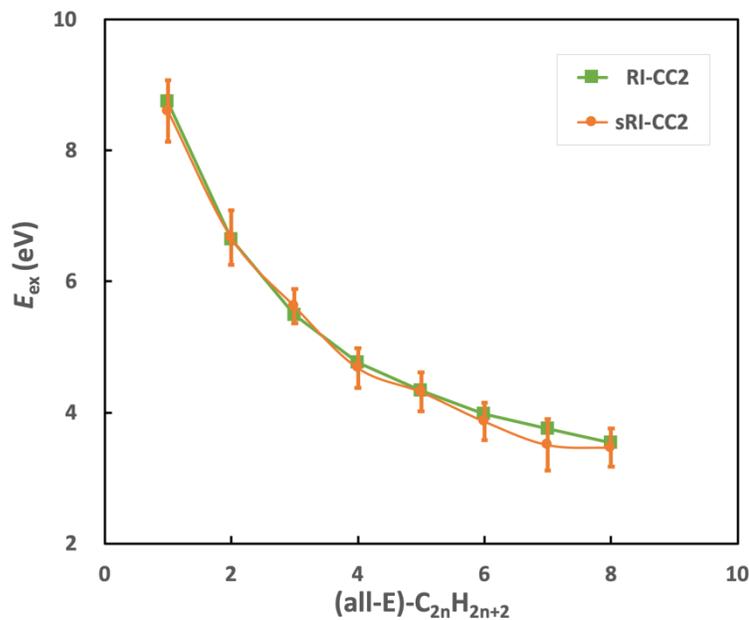
Figure. 3. Excitation energy for a series of (all-E)-$C_{2n}H_{2n+2}$ alkenes with cc-pVDZ basis set and $N_s = 800$. The error bar is from the standard deviation calculated from 10 different seeds.

In Figure 4, we plot the overall computational time of sRI-CC2 for $C_{2n}H_{2n+2}$ alkenes as a function of system size with different number of stochastic orbitals N_s = 800, 1600 and 3200. Note that regardless the number of stochastic orbitals, the overall scacling of the sRI-CC2 is around $O(N_e^{2.6})$ or $O(N_e^{2.8})$, slightly better than the theoretical scaling

$O(N_e^3)$. The RI-CC2 exhibits overall scaling of $O(N_e^{4.31})$, close to the theoretical scaling of $O(N_e^5)$. The crossover happens around $N_e = 60$, where the computational cost of RI-CC2 exceeds sRI-CC2 with 800 stochastic orbitals for larger system size.
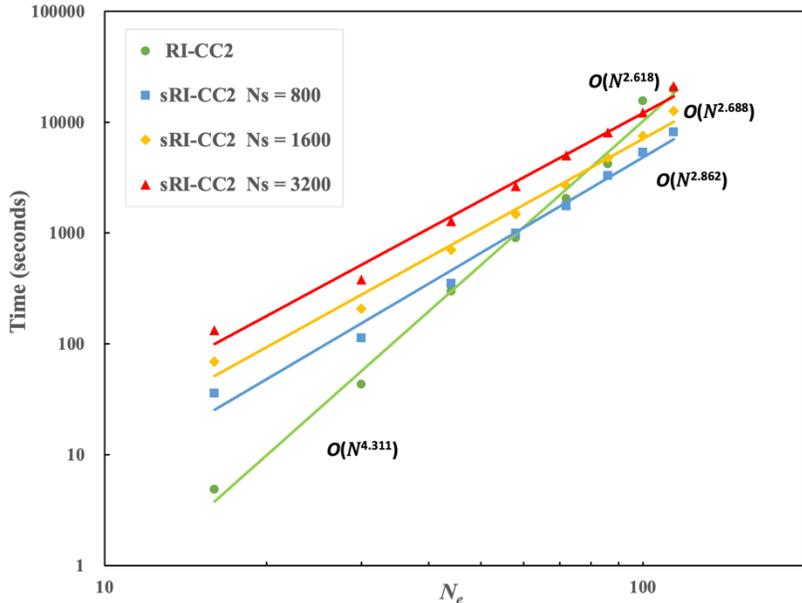
Figure. 4. CPU time as a function of the number of correlated electrons with $N_s = 800$, 1600 and 3200. The computational time of sRI-CC2 is averaged from 10 different runs.

3.3. Molecular Systems

Finally, we test the performance of sRI-CC2 method for a variety of molecules. The experimental geometries[39,40] of the molecules are provided in the Supporting Information. All results from sRI-CC2 are calculated with $N_s = 800$ stochastic orbtials. cc-pVDZ and aug-cc-pVDZ are chosen as our basis sets. The mean values, errors, and standard deviations are list in Table 4. Overall sRI-CC2 agrees with RI-CC2 well, with all errors lying in the standard deviations. Again, the stochastic error does not increase noticeably with the basis set. This concludes our claims that our sRI-CC2 will be useful for a variety of molecular systems.

Table 4. Results for of 11 different molecules. Here N_e is the number of correlated electrons. The excitation energies from Q-Chem and sRI, their errors, and the standard deviations of 10 sRI runs are in eV.

N_e	molecule	cc-pVDZ				aug-cc-pVDZ			
		RI	sRI	error	std deviation	RI	sRI	error	std deviation
2	H_2	14.0308	13.9505	0.0803	0.4111	12.7056	12.6157	0.0899	0.2679

10	H_2O	8.1101	8.1451	0.0350	0.5343	7.1020	7.0888	0.0132	0.4497
2	He	52.5900	52.3018	0.2882	0.9645	21.8743	21.6330	0.2413	0.3555
4	Be	5.4347	5.2817	0.1530	0.2755	5.2685	5.2349	0.0336	0.2285
10	Ne	50.2279	49.8211	0.4068	0.7243	18.7441	18.2055	0.5386	0.5642
10	CH_4	12.3851	12.2510	0.1341	0.4079	10.4742	10.4117	0.0625	0.3125
4	LiH	3.7399	3.8086	0.0687	0.4123	3.7841	3.8484	0.0643	0.4187
12	LiF	5.1234	5.1297	0.0063	0.4329	5.4413	5.2285	0.2128	0.3668
10	HF	10.5782	10.4202	0.1580	0.4481	9.8279	9.7500	0.0779	0.4019
22	$c\text{-}C_3H_4$	7.1215	7.2194	0.0979	0.6223	6.6596	6.5410	0.1186	0.3419
40	$EtCONH_2$	5.9118	5.9113	0.0005	0.4992	5.7510	5.5906	0.1604	0.5020

4. CONCLUSIONS

We develop a stochastic implementation of the RI-CC2 method for excitation energy calculations, namely sRI-CC2 theory. With the introduction of stochastic orbitals, the expensive 4-index ERIs are decoupled into lower-rank integrals, allowing for a reduction of computational scaling from the original $O(N^5)$ to $O(N^3)$. This work extends our previous implementation of sRI-CC2 for ground state to excited state, achieving efficient and accurate results for excitation energies within the CC2 model. As tested on a variety of molecular systems, sRI-CC2 provides exitaiton energies agree well with RI-CC2 method, demonstrating wide application of the sRI-CC2 method. We expect that the sRI-CC2 will be very useful in predicting excited states for very large systems.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at {molecule_geometries.txt}.

Information as mentioned in the text. (TXT)

AUTHOR INFORMATION

Corresponding Authors

*E-mail: joonholee@g.harvard.ed (J.L.).

*E-mail: douwenjie@westlake.edu.cn (W.D.).

ORCID

Chongxiao Zhao: 0009-0009-8627-8911

Joonho Lee: 0000-0002-9667-1081

Wenjie Dou: 0000-0001-5410-6183

Notes

The authors declare no competing financial interest.

APPENDIX

A. Deterministic Resolution of Identity

The 4-, 3- and 2-index ERIs are defined as

$$(\alpha\beta|\gamma\delta) \ = \ \iint dr_1 \, dr_2 \frac{\chi_\alpha(r_1)\chi_\beta(r_1)\chi_\gamma(r_2)\chi_\delta(r_2)}{r_{12}} \tag{39}$$

$$(\alpha\beta|P) \ = \ \iint dr_1 \, dr_2 \frac{\chi_\alpha(r_1) \, \chi_\beta(r_1) \, \chi_P(r_2)}{r_{12}} \tag{40}$$

$$V_{PQ} = (P|Q) \ = \ \iint dr_1 \, dr_2 \frac{\chi_P(r_1) \, \chi_Q(r_2)}{r_{12}} \tag{41}$$

The RI approximation enables the 4-index ERIs to be expanded into lower-rank integrals with auxiliary basis $\{P\}$:

$$(\alpha\beta|\gamma\delta) \approx \sum_{PR}^{N_{aux}} (\alpha\beta|P)[V^{-1}]_{PR} \, (R|\gamma\delta)$$

$$= \sum_{Q}^{N_{aux}} [\sum_{P}^{N_{aux}} [(\alpha\beta|P)[V^{-1/2}]_{PQ}][\sum_{R}^{N_{aux}} [V^{-1/2}]_{QR} \, (R|\gamma\delta)] \tag{42}$$

Defining

14

$$K_{\alpha\beta}^Q \equiv \sum_P^{N_{aux}} (\alpha\beta|P)\, V_{PQ}^{-1/2} \tag{43}$$

which, scaling as $O(N_{AO}^2 N_{aux}^2)$. Then the expression in Eq. (42) can be simplified as

$$(\alpha\beta|\gamma\delta) \approx \sum_Q^{N_{aux}} K_{\alpha\beta}^Q K_{\gamma\delta}^Q \tag{44}$$

Besides, since ERIs are commonly utilized in MO basis, their transformation from AO basis to MO basis can be completed in two steps, both costing $O(N_{AO}^3 N_{aux})$,

$$K_{p\beta}^Q = \sum_\alpha^{N_{AO}} C_\alpha^p\, K_{\alpha\beta}^Q \tag{45}$$

$$K_{pq}^Q = \sum_\beta^{N_{AO}} C_\beta^q\, K_{p\beta}^Q \tag{46}$$

where C_α^p and C_β^q is the usual SCF MO coefficient matrix.

Both N_{AO} and N_{aux} are dependent on the system size. Therefore, according to Eqs. (42)-(46), the overall cost of RI-ERIs scales as $O(N_{AO}^4 N_{aux})$, which we usually label as $O(N^5)$.

B. Stochastic Resolution of Identity

The stochastic optimization of the RI approximation adopts another set of stochastic orbitals $\{\theta^\xi\}$, $\xi = 1, 2, \ldots,$ N_s. All these stochastic orbitals are column arrays of length N_{aux} with each random element $\theta_A^\xi = \pm 1$. They satisfy the following equations:

$$\theta_A^\xi \theta_A^\xi = 1 \qquad\qquad \theta_A^\xi \theta_B^\xi = \pm 1 \tag{47$\sim$48}$$

$$\langle \theta \otimes \theta \rangle_\xi = \frac{1}{N_s} \sum_{\xi=1}^{N_s} \theta^\xi \otimes (\theta^\xi)^T = \begin{pmatrix} \langle \theta_1 \theta_1 \rangle_\xi & \langle \theta_2 \theta_1 \rangle_\xi & \cdots & \langle \theta_1 \theta_{N_{aux}} \rangle_\xi \\ \langle \theta_2 \theta_1 \rangle_\xi & \langle \theta_2 \theta_2 \rangle_\xi & & \langle \theta_2 \theta_{N_{aux}} \rangle_\xi \\ \vdots & & \ddots & \vdots \\ \langle \theta_{N_{aux}} \theta_1 \rangle_\xi & \langle \theta_{N_{aux}} \theta_2 \rangle_\xi & \cdots & \langle \theta_{N_{aux}} \theta_{N_{aux}} \rangle_\xi \end{pmatrix} \approx I \tag{49}$$

Accordingly, we insert this approximate identity matrix in Eq. (49) into the 4-index ERI and obtain its stochastic form:

$$(\alpha\beta|\gamma\delta) \approx \sum_{QS}^{N_{aux}} \sum_{PR}^{N_{aux}} (\alpha\beta|P)\, V_{PQ}^{-1/2}\, I_{QS}\, V_{SR}^{-1/2}\, (R|\gamma\delta)$$

$$= \sum_{QS}^{N_{aux}} \sum_{PR}^{N_{aux}} (\alpha\beta|P) \, V_{PQ}^{-1/2} \left(\langle \theta \otimes \theta^T \rangle_\xi \right)_{QS} (R|\gamma\delta) \, V_{SR}^{-1/2}$$

$$= \left\langle \left[\sum_{P}^{N_{aux}} (\alpha\beta|P) \sum_{Q}^{N_{aux}} \left(V_{PQ}^{-1/2} \theta_Q \right) \right] \left[\sum_{R}^{N_{aux}} (R|\gamma\delta) \sum_{S}^{N_{aux}} \left(V_{SR}^{-1/2} \theta_S^T \right) \right] \right\rangle_\xi \tag{50}$$

Similarly, we define the ξth element of the stochastic average as:

$$R_{\alpha\beta}^{\xi} = \sum_{P}^{N_{aux}} (\alpha\beta|P) \left[\sum_{Q}^{N_{aux}} \left(V_{PQ}^{-1/2} \theta_Q \right) \right] \equiv \sum_{P}^{N_{aux}} (\alpha\beta|P) \, L_P^{\xi} \tag{51}$$

where the construction of L_P^{ξ} costs $O(N_s N_{aux}^2)$ and $R_{\alpha\beta}^{\xi}$ costs $O(N_s N_{AO}^2 N_{aux})$. Thus the 4-index ERI can be rewritten as a stochastic average and scales as $O(N_s N_{AO}^4)$:

$$(\alpha\beta|\gamma\delta) \approx \frac{1}{N_s} \sum_{\xi=1}^{N_s} R_{\alpha\beta}^{\xi} R_{\gamma\delta}^{\xi} \equiv \left\langle R_{\alpha\beta}^{\xi} R_{\gamma\delta}^{\xi} \right\rangle_\xi \tag{52}$$

The ERI in the AO basis can be transformed to MO basis in the same way as RI, scaling as $O(N_s N_{AO}^3)$,

$$R_{p\beta}^{\xi} = \sum_{\alpha}^{N_{AO}} C_{\alpha}^{p} \, R_{\alpha\beta}^{\xi} \tag{53}$$

$$R_{pq}^{\xi} = \sum_{\beta}^{N_{AO}} C_{\beta}^{q} \, R_{p\beta}^{\xi} \tag{54}$$

The overall calculation of sRI-ERIs approximately scales as $O(N^4)$ for large-size systems since with the expansion of the system size, the N_s is almost unchanged and relatively small compared with N_{AO} and N_{aux}. Furthermore, since we only employ 4-index ERIs as intermediates and contract them with 3-index ERIs "on the fly", the actual cost can be further reduced in our sRI-CC2 algorithm.

DATA AVAILABILITY

The data that support the findings of this study are available within the article and supplementary material.

REFERENCES

(1) Christiansen, O.; Koch, H.; Jørgensen, P. The second-order approximate coupled cluster singles and doubles model CC2. *Chem. Phys. Lett.* **1995**, *243*, 409-418.

(2) Fedotov, D. A.; Coriani, S.; Hättig, C. Damped (linear) response theory within the resolution-of-identity coupled cluster singles and approximate doubles (RI-CC2) method. *J. Chem. Phys.* **2021**, *154*, 124110.

(3) Christiansen, O.; Stanton, J. F.; Gauss, J. A coupled cluster study of the $1\,^1A_{1g}$ and $1\,^1B_{2u}$ states of benzene. *J. Chem. Phys.* **1998**, *108*, 3987-4001.

(4) Hättig, C.; Köhn, A. Transition moments and excited-state first-order properties in the coupled-cluster model CC2 using the resolution-of-the-identity approximation. *J. Chem. Phys.* **2002**, *117*, 6939-6951.

(5) Christiansen, O.; Koch, H.; Halkier, A; Jørgensen, P.; Helgaker, T.; Sánchez de Merás, A. M. Large-scale calculations of excitation energies in coupled cluster theory: The singlet excited states of benzene. *J. Chem. Phys.* **1996**, *105*, 6921-6939.

(6) Purvis III, G. D.; Bartlett, R. J. A full coupled-cluster singles and doubles model: The inclusion of disconnected triples. *J. Chem. Phys.* **1982**, *76*, 1910-1918.

(7) Hättig, C.; Hald, K. Implementation of RI-CC2 triplet excitation energies with an application to trans-azobenzene. *Phys. Chem. Chem. Phys.* **2002**, *4*, 2111-2118.

(8) Sekino, H.; Bartlett, R. J. A linear response, coupled-cluster theory for excitation energy. *Int. J. Quantum Chem.* **1984**, *26*, 255-265.

(9) Geertsen, J.; Rittby, M.; Bartlett, R. J. The equation-of-motion coupled-cluster method: Excitation energies of Be and CO. *Chem. Phys. Lett.* **1989**, *164*, 57-62.

(10) Koch, H.; Jensen, H. J. A.; Jørgensen, P.; Helgaker, T. Excitation energies from the coupled cluster singles and doubles linear response function (CCSDLR). Applications to Be, CH^+, CO, and H_2O. *J. Chem. Phys.* **1990**, *93*, 3345-3350.

(11) Koch, H.; Jørgensen, P. Coupled cluster response functions. *J. Chem. Phys.* **1990**, *93*, 3333.

(12) Vahtras, O.; Almlöf, J.; Feyereisen, M. W. Integral approximations for LCAO-SCF calculations. *Chem. Phys. Lett.* **1993**, *213*, 514-518.

(13) Weigend, F.; Häser, M. RI-MP2: first derivatives and global consistency. *Theor. Chem. Acc.* **1997**, *97*, 331-340.

(14) Köhn, A.; Hättig, C. Comment on 'Efficient calculation of canonical MP2 energies' [P. Pulay, S. Saebø, K. Wolinski, Chem. Phys. Lett. 344 (2001) 543–552]. *Chem. Phys. Lett.* **2002**, *358*, 350-353.

(15) Hättig, C.; Köhn, A.; Hald, K. First-order properties for triplet excited states in the approximated coupled cluster model CC2 using an explicitly spin coupled basis. *J. Chem. Phys.* **2002**, *116*, 5401-5410.

(16) Hald, K.; Halkier, A.; Jørgensen, P.; Coriani, S.; Hättig, C.; Helgaker, T. A Lagrangian, integral-density direct formulation and implementation of the analytic CCSD and CCSD (T) gradients. *J. Chem. Phys.* **2003**, *118*, 2985-2998.

(17) Krause, K.; Klopper, W. Description of spin–orbit coupling in excited states with two-component methods based on approximate coupled-cluster theory. *J. Chem. Phys.* **2015**, *142*, 104109.

(18) Koch, H.; Sánchez de Merás, A. M.; Pedersen, T. B. Reduced scaling in electronic structure calculations using Cholesky decompositions. *J. Chem. Phys.* **2003**, *118*, 9481.

(19) Pedersen, T. B.; Sánchez de Merás, A. M.; Koch, H. Polarizability and optical rotation calculated from the approximate coupled cluster singles and doubles CC2 linear response theory using Cholesky decompositions. *J. Chem. Phys.* **2004**, *120*, 8887-8897.

(20) Baudin, P.; Marín, J. S.; Cuesta, I. G.; Sánchez de Merás, A. M. Calculation of excitation energies from the CC2 linear response theory using Cholesky decomposition. *J. Chem. Phys.* **2014**, *140*, 104111.

(21) Folkestad, S. D.; Kjønstad, E. F.; Koch, H. An efficient algorithm for Cholesky decomposition of electron repulsion integrals. *J. Chem. Phys.* **2019**, *150*, 194112.

(22) Neuhauser, D.; Rabani, E.; Baer, R. Expeditious stochastic approach for MP2 energies in large electronic systems. *J. Chem. Theory Comput.* **2013**, *9*, 24-27.

(23) Ge, Q.; Gao, Y.; Baer, R.; Rabani, E.; Neuhauser, D. A guided stochastic energy-domain formulation of the second

order Møller–Plesset perturbation theory. *J. Phys. Chem. Lett.* **2014**, *5*, 185-189.

(24) Takeshita, T. Y.; de Jong, W. A.; Neuhauser, D.; Baer, R.; Rabani, E. Stochastic formulation of the resolution of identity: Application to second order Møller–Plesset perturbation theory. *J. Chem. Theory Comput.* **2017**, *13*, 4605-4610.

(25) Neuhauser, D.; Baer, R.; Zgid, D. Stochastic self-consistent second-order Green's function method for correlation energies of large electronic systems. *J. Chem. Theory Comput.* **2017**, *13*, 5396-5403.

(26) Dou, W.; Takeshita, T. Y.; Chen, M.; Baer, R.; Neuhauser, D.; Rabani, E. Stochastic resolution of identity for real-time second-order Green's function: Ionization potential and quasi-particle spectrum. *J. Chem. Theory Comput.* **2019**, *15*, 6703-6711.

(27) Takeshita, T. Y.; Dou, W.; Smith, D. G.; de Jong, W. A.; Baer, R.; Neuhauser, D.; Rabani, E. Stochastic resolution of identity second-order Matsubara Green's function theory. *J. Chem. Phys.* **2019**, *151*, 044114.

(28) Dou, W.; Chen, M.; Takeshita, T. Y.; Baer, R.; Neuhauser, D.; Rabani, E. Range-separated stochastic resolution of identity: Formulation and application to second-order Green's function theory. *J. Chem. Phys.* **2020**, *153*, 074113.

(29) Neuhauser, D.; Baer, R.; Zgid, D. Stochastic self-consistent second-order Green's function method for correlation energies of large electronic systems. *J. Chem. Theory Comput.* **2017**, *13*, 5396-5403.

(30) Baer, R.; Neuhauser, D.; Rabani, E. Self-averaging stochastic Kohn-Sham density-functional theory. *Phys. Rev. Lett.* **2013**, *111*, 106402.

(31) Neuhauser, D.; Rabani, E.; Cytter, Y.; Baer, R. Stochastic optimally tuned range-separated hybrid density functional theory. *J. Phys. Chem. A* **2016**, *120*, 3071-3078.

(32) Lee, J.; Reichman, D. R. Stochastic resolution-of-the-identity auxiliary-field quantum Monte Carlo: Scaling reduction without overhead. *J. Chem. Phys.* **2020**, *153*, 044131.

(33) Rabani, E.; Baer, R.; Neuhauser, D. Time-dependent stochastic Bethe-Salpeter approach. *Phys. Rev. B* **2015**, *91*, 235302.

(34) Zhao, C.; Dou, W. Stochastic resolution of identity to CC2 for large systems: ground-state properties. 2022, arXiv: 2210.12897. arXiv.org ePrint archive. https://arxiv.org/abs/2210.12897.

(35) Shao, Y.; Gan, Z.; Epifanovsky, E.; Gilbert, A. T. B.; Wormit, M.; Kussmann, J.; Lange, A. W.; Behn, A.; Deng, J.; Feng, X.; Ghosh, D.; Goldey, M.; Horn, P. R.; Jacobson, L. D.; Kaliman, I.; Khaliullin, R. Z.; Kuś, T.; Landau, A.; Liu, J.; Proynov, E. I.; Rhee, Y. M.; Richard, R. M.; Rohrdanz, M. A.; Steele, R. P.; Sundstrom, E. J.; Woodcock III, H. L.; Zimmerman, P. M.; Zuev, D.; Albrecht, B.; Alguire, E.; Austin, B.; Beran, G. J. O.; Bernard, Y. A.; Berquist, E.; Brandhorst, K.; Bravaya, K. B.; Brown, S. T.; Casanova, D.; Chang, C. -M.; Chen, Y.; Chien, S. H.; Closser, K. D.; Crittenden, D. L.; Diedenhofen, M.; DiStasio Jr, R. A.; Do, H.; Dutoi, A. D.; Edgar, R. G.; Fatehi, S.; Fusti-Molnar, L.; Ghysels, A.; Golubeva-Zadorozhnaya, A.; Gomes, J.; Hanson-Heine, M. W. D.; Harbach, P. H. P.; Hauser, A. W.; Hohenstein, E. G.; Holden, Z. C.; Jagau, T. -C.; Ji, H.; Kaduk, B.; Khistyaev, K.; Kim, J.; Kim, J.; King, R. A.; Klunzinger, P.; Kosenkov, D.; Kowalczyk, T.; Krauter, C. M.; Lao, K. U.; Laurent, A. D.; Lawler, K. V.; Levchenko, S. V.; Lin, C. Y.; Liu, F.; Livshits, E.; Lochan, R. C.; Luenser, A.; Manohar, P.; Manzer, S. F.; Mao, S. -P.; Mardirossian, N.; Marenich, A. V.; Maurer, S. A.; Mayhall, N. J.; Neuscamman, E.; Oana, C. M.; Olivares-Amaya, R.; O'Neill, D. P.; Parkhill, J. A.; Perrine, T. M.; Peverati, R.; Prociuk, A.; Rehn, D. R.; Rosta, E.; Russ, N. J.; Sharada, S. M.; Sharada, S.; Small, D. W.; Sodt, A.; Stein, T.; Stück, D.; Su, Y. -C.; Thom, A. J. W.; Tsuchimochi, T.; Vanovschi, V.; Vogt, L.; Vydrov, O.; Wang, T.; Watson, M. A.; Wenzel, J.; White, A.; Williams, C. F.; Yang, J.; Yeganeh, S.; Yost, S. R.; You, Z. -Q.; Zhang, Y.; Zhang, X.; Zhao, Y.; Brooks, B. R.; Chan, G. K. L.; Chipman, D. M.; Cramer, C. J.; Goddard III,

W. A.; Gordon, M. S.; Hehre, W. J.; Klamt, A.; Schaefer III, H. F.; Schmidt, M. W.; Sherrill, C. D.; Truhlar, D. G.; Warshel, A.; Xu, X.; Aspuru- Guzik, A.; Baer, R.; Bell, A. T.; Besley, N. A.; Chai, J. -D.; Dreuw, A.; Dunietz, B. D.; Furlani, T. R.; Gwaltney, S. R.; Hsu, C. -P.; Jung, Y.; Kong, J.; Lambrecht, D. S.; Liang, W.; Ochsenfeld, C.; Rassolov, V. A.; Slipchenko, L. V.; Subotnik, J. E.; Voorhis, T. V.; Herbert, J. M.; Krylov, A. I.; Gill, P. M. W.; Head-Gordo M.Advances in molecular quantum chemistry contained in the Q-Chem 4 program package. *Mol. Phys.* **2000**, *113*, 184-215.

(36) Hättig, C.; Weigend, F. CC2 excitation energy calculations on large molecules using the resolution of the identity approximation. *J. Chem. Phys.* **2000**, *113*, 5154-5161.

(37) Vlček, V.; Eisenberg, H. R.; Steinle-Neumann, G.; Neuhauser, D.; Rabani, E.; Baer, R. Spontaneous charge carrier localization in extended one-dimensional systems. *Phys. Rev. Lett.* **2016**, *116*, 186401.

(38) Vlček, V.; Rabani, E.; Neuhauser, D.; Baer, R. Stochastic GW calculations for molecules. *J. Chem. Theory Comput.* **2017**, *13*, 4997-5003.

(39) Schreiber, M.; Silva-Junior, M. R.; Sauer, S.; Thiel, W. Benchmarks for electronically excited states: CASPT2, CC2, CCSD, and CC3. *J. Chem. Phys.* **2008**, *128*, 134110.

(40) van Setten, M. J.; Caruso, F.; Sharifzadeh, S.; Ren, X.; Scheffler, M.; Liu, F.; Lischner, J.; Lin, L.; Deslippe, J. R.; Louie, S. G.; Yang, C.; Weigend, F.; Neaton, J. B.; Evers, F.; Rinke, P. GW 100: Benchmarking G_0W_0 for molecular systems. *J. Chem. Theory Comput.* **2015**, *11*, 5665-5687.