

# Stochastic Resolution of Identity for Correlation Energy Prediction via Doubles Connected Moments Expansion

Chongxiao Zhao<sup>1,2,a)</sup> and Wenjie Dou<sup>1,2,b),\*</sup>

1. Department of Chemistry, School of Science, Westlake University, Hangzhou, Zhejiang 310024, China

2. Institute of Natural Sciences, Westlake Institute for Advanced Study, Hangzhou, Zhejiang 310024, China

<sup>a)</sup>Email: zhaochongxiao@westlake.edu.cn

<sup>b)</sup>Authors to whom correspondence should be addressed: douwenjie@westlake.edu.cn

## Abstract

The recently developed Doubles Connected Moments (DCM) expansion offers a tractable approach for computing correlation energy, exhibiting an noniterative  $O(N^6)$  scaling with system size  $N$ . Benchmark calculations on a set of molecules demonstrate that the DCM can outperform CCSD in terms of accuracy. To further enhance its efficiency, we present a stochastic variant of DCM by introducing a stochastic resolution-of-identity (sRI) technique, which decomposes the essential four-index intermediates. The resulting sRI-DCM scheme only involves one  $O(N^6)$  step, while all other steps do not exceed  $O(N^4)$  at each recursion, and reliably reproduces the results of conventional DCM. Our sRI-DCM achieves an overall experimental scaling of  $O(N^{4.46})$  for series hydrogen dimer chains, demonstrating that it is attractive and practical for large systems containing hundreds of electrons.

## 1. INTRODUCTION

The accurate determination of correlation energy, defined as the difference between the exact energy and the approximate energy from HF wavefunction, constitutes a central problem in quantum chemistry.<sup>[1]</sup> Traditional post-HF methods, such as Møller–Plesset perturbation theory (MPn),<sup>[2–4]</sup> configuration interaction (CI),<sup>[5,6]</sup> and coupled-cluster (CC) theory,<sup>[7,8]</sup> systematically improve the wavefunction mainly by truncating excitations to a certain level, thereby balancing accuracy with computational cost<sup>[9]</sup> These approaches have achieved remarkable success, often reaching chemical accuracy for a wide range of molecular systems. However, these traditional methods face significant difficulties in

treating strongly correlated systems. Moreover, some high-level post-HF methods renders them prohibitively expensive for larger molecules, underscoring the need for alternative computational strategies.

The  $t$ -expansion<sup>[10]</sup> offers an alternative to calculate the correlation energy with moments. This method is inherently non-perturbative and can capture strong correlation effects if the trial wavefunction  $|\Phi\rangle$  has some overlap with the exact ground state. Introduce an auxiliary function

$$\begin{aligned}\langle E(t) \rangle &= \langle \Phi | H | \Phi \rangle = \frac{\langle \Phi | H e^{-tH} | \Phi \rangle}{\langle \Phi | e^{-tH} | \Phi \rangle} \\ &= \sum_{m=0}^{\infty} \frac{(-t)^m}{m!} I_{m+1}\end{aligned}\tag{1}$$

The so-called connected moments  $I_m$  (or cumulants) satisfy the following relationship with moments  $\mu_m$ .

$$\mu_m = \langle \Phi | H^m | \Phi \rangle\tag{2}$$

$$I_{m+1} = \mu_{m+1} - \sum_{n=0}^{m-1} C_m^n I_{n+1} \mu_{m-n}\tag{3}$$

Then the ground state energy  $E_0$  is given as

$$E_0 = \lim_{t \rightarrow \infty} E(t)\tag{4}$$

In practice, the infinite power series in eq 1 needs to be truncated to a certain order. Among all the possible solutions<sup>[11]</sup>, the connected moments expansion (CMX)<sup>[12, 13]</sup> expresses the energy with the lowest  $2n - 1$  connected moments as

$$E_{\text{CMX}(n)} = I_1 - [I_2 \quad I_3 \quad \cdots \quad I_n] \begin{bmatrix} I_3 & I_4 & \cdots & I_{n+1} \\ I_4 & I_5 & \cdots & I_{n+2} \\ \vdots & \vdots & \ddots & \vdots \\ I_{n+1} & I_{n+2} & \cdots & I_{2n-1} \end{bmatrix}^{-1} \begin{bmatrix} I_2 \\ I_3 \\ \vdots \\ I_n \end{bmatrix}\tag{5}$$

Higher-order connected moments are related to lower-order ones via eq 3.

As early as 1987, Cioslowski *et al.* tested CMX(2), which recovers more than 50% of MP3 energy.<sup>[14]</sup> As the CMX method accurately predicts the correlation energy as a noniterative moment-based scheme, however, the computational scaling increases sharply with the order. For example,  $I_4$  contains single, double, triple and quadruple substitutions and gives an  $O(N^7)$  scaling, comparable to MP4. This steep computational scaling ultimately limits this moment-based methods<sup>[15–18]</sup> to relatively low orders, constraining their achievable accuracy.

In the DCM formulation reported by Ganoe and Head-Gordon<sup>[19]</sup>, the method operates solely within the double excitation subspace of the Hilbert space. By constructing four-index intermediates and establishing relationships between

connected moments of different orders, the computational cost can be substantially reduced to  $O(N^6)$ . DCM achieves smooth convergence within approximately  $N=11-14$ , with accuracy surpassing that of CCSD with proper reference. Besides, this moment-based method can handle strong correlated systems, which traditional single-reference methods like CCSD fail to address.

Inspired by the high performance of DCM, we incorporate the stochastic resolution of identity (sRI) technique to further reduce its computational and storage costs. The sRI is a stochastic formulation built upon the deterministic RI framework<sup>[20–22]</sup>. This strategy introduces a set of stochastic orbitals to decouple the high-rank tensors, primarily the four-index electron repulsion integrals (ERIs), into lower-rank integrals, which allows flexible scaling reduction and avoids steep memory need. So far, the sRI technique has been successfully used in many electronic structure methods, such as MP2,<sup>[23–25]</sup> CC2,<sup>[26–29]</sup> DFT and TDDFT,<sup>[30–35]</sup> GF2,<sup>[36–41]</sup> and etc.<sup>[42–44]</sup>

In this work, we integrate the sRI approximation to the DCM formulation, abbreviated as sRI-DCM. This incorporation enables flexible decomposition of intermediate steps and reduces scaling by more than one order of magnitude. Our sRI-DCM is capable of treating systems with hundreds of electrons, and its accuracy is systematically tunable via the number of stochastic orbitals. This method extends the applicational scope of DCM and enables the determination of correlation energy for larger systems.

## 2. THEORY

The standard notation used in this work is summarized in Table 1. The parameters listed in the final column ( $N_{AO}$ ,  $N_{aux}$ ,  $N_{ao}$ ,  $N_{mo}$ ,  $N_{occ}$ , and  $N_{vir}$ ) all scales linearly with the system size  $N$ .

Table 1: Summary of notations in the following equations.

| Items                       | Functions or indices   | Total number |
|-----------------------------|--|--------------|
| AO Gaussian basis functions | $\chi_\alpha(r_1), \chi_\beta(r_1), \chi_\gamma(r_1), \chi_\delta(r_1), \dots$ | $N_{AO}$     |
| Auxiliary basis functions   | $P, Q, R, S, \dots$  | $N_{aux}$    |
| General AOs                 | $\alpha, \beta, \gamma, \delta, \dots$   | $N_{ao}$     |
| General MOs                 | $p, q, r, s, \dots$  | $N_{mo}$     |
| Occupied MOs                | $i, j, k, l, \dots$  | $N_{occ}$    |
| Unoccupied MOs              | $a, b, c, d, \dots$  | $N_{vir}$    |

### 2.1. DCM formulation

In the DCM formulation, a connected moment of the Hamiltonian—for instance, the second one—is calculated using the moments

$$I_2 = \langle H^2 \rangle - \langle H \rangle^2 \quad (6)$$

By approximately inserting the identity  $1 = P + D$ , the equation becomes

$$P = |\Phi_0\rangle \langle \Phi_0| \quad (7)$$

$$D = \frac{1}{4} \sum_{ijab} |\Phi_{ij}^{ab}\rangle \langle \Phi_{ij}^{ab}| \quad (8)$$

$$I_2 = \langle H(P + D)H \rangle - \langle HPH \rangle = \langle HDH \rangle = \langle H^2 \rangle_c \quad (9)$$

Here  $P$  is constructed from the HF reference  $|\Phi_0\rangle$  and  $D$  is the doubles subspace of the orthogonal Hilbert space. Similarly, higher-order connected moments can be expressed in this form. Their structures are governed by the four-index intermediates  $(ij||ab)_n$  as

$$I_{2n-1} = \langle \underbrace{HD \cdots HD}_{nH} \underbrace{H \cdots DH}_{(n-1)H} \rangle = \frac{1}{4} \sum_{ijab} (ij||ab)_n (ij||ab)_{n-1} \quad (10)$$

$$I_{2n} = \langle \underbrace{HD \cdots HD}_{nH} \underbrace{H \cdots DH}_{nH} \rangle = \frac{1}{4} \sum_{ijab} (ij||ab)_n (ij||ab)_n \quad (11)$$

These intermediates satisfy the following recursion relation

$$\begin{aligned} (ij||ab)_{n+1} &= \langle 0|H|2\rangle \langle 2|H|2\rangle_n = \frac{1}{4} \sum_{klcd} (kl||cd)_n \langle \Phi_{kl}^{cd} | H | \Phi_{ij}^{ab} \rangle \\ &= \frac{1}{2} \sum_{cd} \langle cd||ab \rangle (ij||cd)_n + \frac{1}{2} \sum_{kl} \langle ij||kl \rangle (kl||ab)_n \\ &\quad - \Delta_{ij}^{ab} (ij||ab)_n + P(ij)P(ab) \sum_{kc} \langle kb||ic \rangle (kj||ca)_n \end{aligned} \quad (12)$$

Here 0 denotes  $\Phi_0$  while 2 denotes  $\Phi_{\mu_2}$ , and

$$\Delta_{ij}^{ab} = \epsilon_i + \epsilon_j - \epsilon_a - \epsilon_b \quad (13)$$

$$P(pq)f(p, q) = f(p, q) - f(q, p) \quad (14)$$

The primary determining steps for the  $O(N^6)$  scaling of DCM are the contractions of two four-index tensors, the antisymmetrized two-center integrals  $\langle pq||rs \rangle$  and intermediates  $(ij||ab)_n$ , in eq 12. By efficiently decoupling these tensors using sRI, it is possible to reduce this scaling while also lowering storage demands.

## 2.2. RI and sRI formulations

In deterministic RI formulations, the four-index ERI is approximated using a combination of three-index  $(\alpha\beta|P)$  and two-index  $V_{PQ}$  integrals expanded in an auxiliary basis  $\{P\}$ .

$$(\alpha\beta|\gamma\delta) \approx \sum_{PR} (\alpha\beta|P) V_{PR}^{-1} (R|\gamma\delta) \quad (15)$$

Introducing the three-rank tensor  $B_{\alpha\beta}^Q$ , the four-index RI-ERI takes the form

$$B_{\alpha\beta}^Q = \sum_P (\alpha\beta|P) V_{PQ}^{-1/2} \quad (16)$$

$$(\alpha\beta|\gamma\delta) \approx \sum_Q \left[ \sum_P (\alpha\beta|P) V_{PQ}^{-1/2} \right] \left[ \sum_R V_{QR}^{-1/2} (R|\gamma\delta) \right] = \sum_Q B_{\alpha\beta}^Q B_{\gamma\delta}^Q \quad (17)$$

The construction of  $B_{\alpha\beta}^Q$  costs  $O(N_{aux}^2 N_{ao}^2)$  and the final step scales as  $O(N_{aux} N_{ao}^4)$ . Since both  $N_{aux}$  and  $N_{ao}$  grow linearly with system size, the overall computational scaling for RI-ERI is  $O(N^5)$ , with a storage requirement of  $O(N^3)$ .

sRI is a variant of RI that further introduces a set of  $N_s$  stochastic orbitals  $\{\theta^\xi\}$ ,  $\xi = 1, 2, \dots, N_s$ . Each stochastic orbital  $\theta^\xi$  is a column array of length  $N_{aux}$ , with all its element randomly generated to be 1 or -1. Thus, as  $N_s$  increases, these  $\theta^\xi$  satisfy the following equation

$$\langle \theta \otimes \theta \rangle_\xi = \frac{1}{N_s} \sum_{\xi=1}^{N_s} \theta^\xi \otimes (\theta^\xi)^T \approx I \quad (18)$$

This approximate identity can then be inserted into eq 17 to transform one auxiliary basis index  $Q$  into a stochastic orbital index  $\xi$

$$\begin{aligned} (\alpha\beta|\gamma\delta) &= \left\langle \sum_P \left[ (\alpha\beta|P) \sum_Q \left( V_{PQ}^{-1/2} \theta_Q^\xi \right) \right] \sum_R \left[ (R|\gamma\delta) \sum_S \left( V_{SR}^{-1/2} \theta_S^\xi \right) \right] \right\rangle_\xi \quad (19) \\ &= \left\langle R_{\alpha\beta}^\xi R_{\gamma\delta}^\xi \right\rangle_\xi = \frac{1}{N_s} \sum_{\xi=1}^{N_s} R_{\alpha\beta}^\xi R_{\gamma\delta}^\xi \end{aligned}$$

$$R_{\alpha\beta}^\xi = \sum_P (\alpha\beta|P) \sum_Q \left( V_{PQ}^{-1/2} \theta_Q^\xi \right) \quad (20)$$

The construction of  $R_{\alpha\beta}^\xi$  needs two steps, which respectively scale as  $O(N_s N_{aux}^2)$  and  $O(N_s N_{aux} N_{ao}^2)$ . And the evaluation of four-index sRI-ERI costs  $O(N_s N_{ao}^4)$ .

Unlike conventional RI, sRI employs a size-independent parameter  $N_s$ . Consequently, the storage requirement for the intermediate tensor  $R_{\alpha\beta}^\xi$  scales as  $O(N_s N_{ao}^2)$ , which is effectively  $O(N^2)$ . In practice, this reduction by an order of magnitude from the  $O(N^3)$  scaling of  $B_{\alpha\beta}^Q$  enables more flexible and computationally efficient tensor contractions. The primary error of sRI originates from eq 18 and can be systematically controlled by adjusting  $N_s$ , which will be discussed in detail later.

### 2.3. sRI-DCM

To align with Ganoe *et al.*'s equations, we retain the spin orbital form in the following derivation. This allows our code to handle both restricted and

unrestricted spin cases. As the time-determining steps are the contractions of four-index tensors, we mainly focus on the decomposition of these high-order tensors.

First, the antisymmetrized two-center integrals can be constructed with RI or sRI

$$\begin{aligned}
\langle pq||rs\rangle &= \langle pq|rs\rangle - \langle pq|sr\rangle \\
&= \sum_Q (B_{pr}^Q B_{qs}^Q - B_{ps}^Q B_{qr}^Q) \\
&= \langle R_{pr}^\xi R_{qs}^\xi - R_{ps}^\xi R_{qr}^\xi \rangle_\xi
\end{aligned} \tag{21}$$

Apart from the decomposition of four-index ERIs, the four-index intermediates  $(ij||ab)_n$  are also needed to be decoupled. Since these intermediates do not necessarily possess the positive definiteness property like two-electron integrals, additional effort is required to decompose them. To this end, we present two decomposition schemes.

The first scheme is to adopt eigendecomposition. This begins by rearranging  $(ij||ab)_n$  from a four-index to a two-index form, expressed in terms of the composite indices  $(ia)$  and  $(jb)$ .

$$\begin{aligned}
(ij||ab)_n &\triangleq A_{ia,jb} = VEV^T \\
&= \begin{pmatrix} V^{(-)} \\ V^{(+)} \end{pmatrix}^T \left[ \begin{pmatrix} 0 & 0 \\ 0 & E^{(+)} \end{pmatrix} + \begin{pmatrix} E^{(-)} & 0 \\ 0 & 0 \end{pmatrix} \right] \begin{pmatrix} V^{(-)} \\ V^{(+)} \end{pmatrix} \\
&= T_{ia,jb}^{(+)} - T_{ia,jb}^{(-)} = T_{ia,jb}^{(+),\frac{1}{2}} T_{ia,jb}^{(+),\frac{1}{2}} - T_{ia,jb}^{(-),\frac{1}{2}} T_{ia,jb}^{(-),\frac{1}{2}}
\end{aligned} \tag{22}$$

The eigendecomposition itself scales as  $O(N^6)$ . Since  $A_{ia,jb}$  may contain both positive and negative eigenvalues, we partition them into two positive semidefinite blocks and handle each one separately. Below we take the construction of positive eigenvalues contribution as an example.

$$T_{ia,jb}^{(+),\frac{1}{2}} = (VE^{(+),\frac{1}{4}})(VE^{(+),\frac{1}{4}})^T \tag{23}$$

Then transform one index pair of  $o-v$  into sRI space.

$$\begin{aligned}
(ij||ab)_n &\approx \sum_{kc} T_{ia,kc}^{(+),\frac{1}{2}} T_{kc,jb}^{(+),\frac{1}{2}} - \sum_{kc} T_{ia,kc}^{(-),\frac{1}{2}} T_{kc,jb}^{(-),\frac{1}{2}} \\
&= \sum_{kck'c'} T_{ia,kc}^{(+),\frac{1}{2}} \left( \langle \theta \otimes \theta \rangle_\xi \right)_{kc,k'c'} T_{k'c',jb}^{(+),\frac{1}{2}} - \sum_{kck'c'} T_{ia,kc}^{(-),\frac{1}{2}} \left( \langle \theta \otimes \theta \rangle_\xi \right)_{kc,k'c'} T_{k'c',jb}^{(-),\frac{1}{2}} \\
&= \left\langle \left( \sum_{kc} T_{ia,kc}^{(+),\frac{1}{2}} \theta_{kc}^\xi \right) \left( \sum_{k'c'} \theta_{k'c'}^\xi T_{k'c',jb}^{(+),\frac{1}{2}} \right) - \left( \sum_{kc} T_{ia,kc}^{(-),\frac{1}{2}} \theta_{kc}^\xi \right) \left( \sum_{k'c'} \theta_{k'c'}^\xi T_{k'c',jb}^{(-),\frac{1}{2}} \right) \right\rangle_\xi \\
&= \left\langle T_{ia}^{(+),\xi} T_{jb}^{(+),\xi} - T_{ia}^{(-),\xi} T_{jb}^{(-),\xi} \right\rangle_\xi
\end{aligned} \tag{24}$$

Such a subtle decomposition enables us to handle each tensor contraction involving the  $(ij||ab)_n$  by replacing it with the two-rank tensors  $T_{ia}^{(+),\xi}$  and  $T_{ia}^{(-),\xi}$ .

The final expression resembles the sRI formulation of four-index ERIs, with matching computational cost and storage scaling. The difference is that sRI-ERI replaces only one auxiliary basis index with an sRI index, while here we replace a pair of occupied and virtual orbital indices. This shows that sRI is not limited to decoupling four-index ERIs. In fact, sRI has the potential to decompose any positive semidefinite four-index tensor into two-rank tensors, which is an attractive feature.

The second scheme is Cholesky decomposition<sup>[45–47]</sup>. Conventional Cholesky decomposition also scales as  $O(N^6)$ . However, we try this technique in the hope that if the targeted intermediates are sparse, the scaling can be reduced to below  $O(N^4)$ . Since  $A_{ia,jb}$  may not be positive semidefinite, which is a requirement for Cholesky decomposition, we decompose it into two semidefinite components by adding and subtracting a scaled identity matrix  $I$ . Specifically, only the first component then requires Cholesky decomposition and the index  $J$  is proportional to  $N^2$ .

$$\begin{aligned} (ij||ab)_n &\triangleq A_{ia,jb} = (A_{ia,jb} + c * I_{ia,jb}) - c * I_{ia,jb} \\ &= \sum_J L_{ia}^J L_{jb}^J - c * I_{ia,jb} \end{aligned} \quad (25)$$

Here the factor  $c$  is chosen as the the inverse of the lowest negative eigenvalue of  $A_{ia,jb}$ , or  $c = 0$  if all eigenvalues are non-negative. The calculation of  $c$  costs  $O(N^4)$ . We do not set a fixed  $c$  for all calculations, since as we found in our previous sRI implementations<sup>[28,29]</sup>, such treatment sometimes introduce significant stochastic noise when sRI is later applied to the two semi-definite components.

Subsequently, the sRI approximated identity is inserted into the first component, while the second component is constructed directly using stochastic orbitals.

$$\begin{aligned} (ij||ab)_n &\approx \sum_{JJ'} L_{ia}^J \left( \langle \theta \otimes \theta \rangle_\xi \right)_{J,J'} L_{jb}^{J'} - c \left( \langle \theta \otimes \theta \rangle_\xi \right)_{ia,jb} \\ &= \left\langle \left( \sum_J L_{ia}^J \theta_J^\xi \right) \left( \sum_{J'} \theta_{J'}^\xi L_{jb}^{J'} \right) - (\sqrt{c} \theta_{ia}^\xi) (\sqrt{c} \theta_{jb}^\xi) \right\rangle_\xi \\ &= \left\langle M_{ia}^\xi M_{jb}^\xi - N_{ia}^\xi N_{jb}^\xi \right\rangle_\xi \end{aligned} \quad (26)$$

The final expression appears similar to that of the first scheme. Except for the Cholesky decomposition, the scaling of all other steps is also reduced to  $O(N^4)$ . Unfortunately, within our test, the sparse density of  $A_{ia,jb}$  may be quite high, implying that performing the Cholesky decomposition still incurs a cost of  $O(N^6)$ . Finally, the observed scaling of the second scheme is similar to that of the first scheme, while the second scheme shows larger errors. Therefore, in the subsequent sections, we still use the eigendecomposition and we provide some data in the Appendix for the second scheme.

With the sRI approximation, the recurrence relation in eq 12 now becomes

$$\begin{aligned}
(ij||ab)_{n+1} = & \frac{1}{2} \left\langle \sum_{cd} (R_{ca}^{\xi'} R_{db}^{\xi'} - R_{cb}^{\xi'} R_{da}^{\xi'}) (T_{ic}^{(+),\xi} T_{jd}^{(+),\xi} - T_{ic}^{(-),\xi} T_{jd}^{(-),\xi}) \right\rangle_{\xi\xi'} \\
& + \frac{1}{2} \left\langle \sum_{kl} (R_{ki}^{\xi'} R_{lj}^{\xi'} - R_{kj}^{\xi'} R_{li}^{\xi'}) (T_{ka}^{(+),\xi} T_{lb}^{(+),\xi} - T_{ka}^{(-),\xi} T_{lb}^{(-),\xi}) \right\rangle_{\xi\xi'} \\
& - \Delta_{ij}^{ab} (ij||ab)_n \\
& + P(ij)P(ab) \left\langle \sum_{cd} (R_{kj}^{\xi'} R_{bc}^{\xi'} - R_{kc}^{\xi'} R_{bj}^{\xi'}) (T_{ic}^{(+),\xi} T_{ka}^{(+),\xi} - T_{ic}^{(-),\xi} T_{ka}^{(-),\xi}) \right\rangle_{\xi\xi'}
\end{aligned} \tag{27}$$

Here, we use two set of stochastic orbitals, labeled  $\xi$  and  $\xi'$ , because the two four-index tensors are independent. Since all sRI tensors  $R_{pq}^{\xi'}$  and  $T_{ia}^{(\pm),\xi}$  are two-rank tensors and each index appears twice at most in any single multiplication, the contraction of these two-rank tensors scales as  $O(N^3)$  at most. However, the final assembly of  $(ij||ab)_n$  requires  $O(N^4)$  operations. Moreover, the eigendeomposition step scales as  $O(N^6)$ . These two steps prevent our sRI-DCM formulation from reaching a better  $O(N^3)$  scaling. Nevertheless, this scheme remains valuable because it reduces the number of expensive  $O(N^6)$  steps and we can observe scaling reduction in the results. As for storage, the two-rank sRI tensors add extra  $O(N^2)$  memory cost, but this is negligible compared to eliminating the  $O(N^4)$  storage needed for  $\langle pq||rs \rangle$ .

### 3. RESULTS AND DISCUSSION

Using RI and sRI techniques, we develop two code versions, RI-DCM and sRI-DCM, within a developmental version of Q-Chem<sup>[48]</sup> package. We take the RI-DCM results as references and focus on accuracy, time cost, and convergence pattern to evaluate our sRI-DCM method. If not stated otherwise, all calculations in this study use the cc-pVDZ basis set.

Because the sRI technique is stochastic, we average the results of sRI-DCM over ten independent runs with different random seeds. We quantify the stochastic uncertainty using the standard deviation (S.D.). If the reference RI-DCM value falls within the S.D. range, indicating the absolute error is smaller than the S.D., the sRI-DCM result is considered accurate. Other parameters follow Ganoe's work: the maximum DCM order is 20, and the rescale factor  $s$  is 1.1.

All the calculations are carried out in the high performance computing (HPC) center of Westlake University, utilizing an AMD EPYC 7502 (2.5 GHz) node with 64 computational cores.

#### 3.1. Accuracy benchmark

We first benchmark the accuracy of sRI-DCM among a range of small molecules with  $N_s = 5000$  stochastic orbitals. As shown in Table 2, the errors of RI-DCM and sRI-DCM are listed for orders  $n = 5, 10, 15$  and 20. Using

Table 2: Comparison of sRI-DCM and RI-DCM for correlation energy per electron (in  $10^{-3}$  a.u.) among small molecules.

| Molecule                      | DCM(5)    |        | DCM(10)   |        | DCM(15)   |        | DCM(20)   |        |
|-------------------------------|-----------|--------|-----------|--------|-----------|--------|-----------|--------|
|                               | Abs error | S.D.   | Abs error | S.D.   | Abs error | S.D.   | Abs error | S.D.   |
| H <sub>2</sub>                | 0.3525    | 0.7795 | 0.1329    | 0.4485 | 0.1521    | 0.3338 | 0.1020    | 0.3067 |
| H <sub>2</sub> O              | 0.1936    | 0.2159 | 0.1627    | 0.1995 | 0.3266    | 0.3968 | 0.1998    | 0.6000 |
| LiH                           | 0.0152    | 0.1826 | 0.0945    | 0.2206 | 0.0761    | 0.2054 | 0.0851    | 0.2495 |
| LiF                           | 0.0074    | 0.2337 | 0.2489    | 0.2742 | 0.2845    | 0.3938 | 0.4451    | 0.5671 |
| HF                            | 0.1339    | 0.3272 | 0.3086    | 0.7789 | 0.1678    | 0.4896 | 0.2463    | 0.8952 |
| NH <sub>3</sub>               | 0.2706    | 0.2723 | 0.1221    | 0.4214 | 0.3788    | 1.4331 | 0.1133    | 0.9891 |
| C <sub>2</sub> H <sub>2</sub> | 0.1886    | 0.4921 | 0.3353    | 0.6294 | 0.1732    | 0.3166 | 0.1776    | 0.8656 |

water (H<sub>2</sub>O) as an example, the results are shown in Figure 1(a). As the DCM order increases, the RI-DCM curve smoothly approaches the CCSD(T) reference line by around  $n = 10$ . The sRI-DCM curve also converges at a similar order. All RI-DCM points lie within the error bars, confirming the validity of the sRI approximation. In practice, larger  $N_s$  values can be used for greater accuracy, and we will discuss this further in Subsection 3.3. Figure 1(b) includes other molecules, where similar trends are observed.

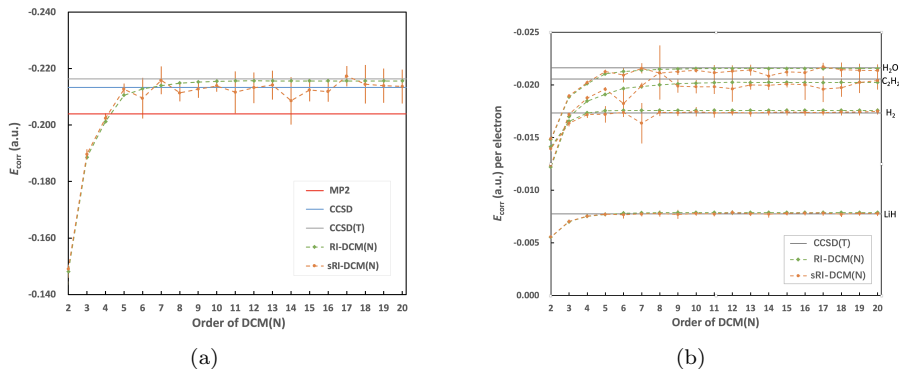


Figure 1: Error comparison of RI-DCM and sRI-DCM among small molecules. Panel (a) is for H<sub>2</sub>O only, and panel (b) adds some other molecules.

### 3.2. Scaling analysis

We then assess the computational time cost of our sRI-DCM within a series of hydrogen dimer chains using sto-3g basis set and  $N_s = 5000$ . In Figure 2, we presents the fitted curves of time consumption as a function of the chain size. As shown in the figure, the experimental scaling for RI-DCM is  $O(N^{5.58})$ , while the scaling of sRI-DCM is  $O(N^{4.46})$ . Overall, our sRI-DCM reduces the

computational cost by approximately one order of magnitude. This improved scaling makes our method more efficient than the baseline for systems smaller than the crossover point at around  $N_e = 100$ .

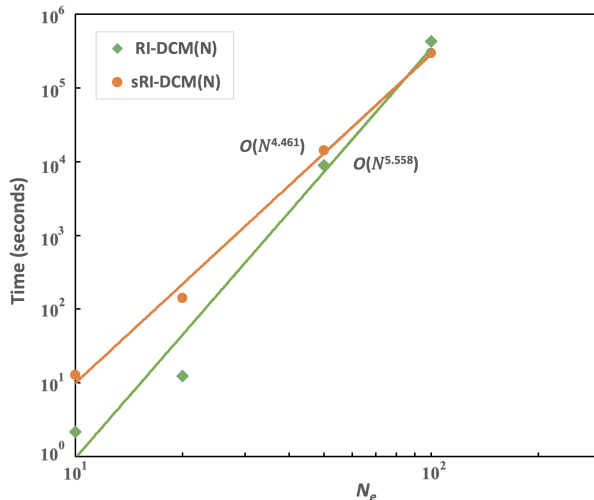


Figure 2: Comparative scaling of computational cost for sRI-DCM and RI-DCM.

### 3.3. Convergence with $N_s$

In Figure 3, we take LiH as an example to evaluate the convergence performance of sRI-CC2 with different  $N_s$  values. For clarity, the x-axis values are slightly shifted for each  $N_s$ . Overall, the sRI-DCM results agree well with RI-DCM across the tested DCM orders. Notably, larger  $N_s$  values lead to smaller error bars (S.D.), which are expected to approach zero. This indicates that using more stochastic orbitals helps reduce stochastic noise and improves accuracy. However, increasing  $N_s$  also raises computational cost. So in practice, we need to choose an appropriate  $N_s$  that balances accuracy and efficiency and in our calculations above, we select  $N_s = 5000$ .

## 4. CONCLUSIONS

In this work, we present a sRI formulation of DCM to estimate correlation energy. Our sRI-DCM achieves good agreement with the original DCM while reducing the computational scaling from  $O(N^6)$  to approximately  $O(N^{4.5})$ . It maintains a similar convergence pattern, and its accuracy can be controlled in practice by adjusting the number of stochastic orbitals  $N_s$ .

The use of sRI in this work extends its application from traditional methods to moment-based calculations. This demonstrates the flexibility and strong scaling reduction capability of sRI in different situations. Our future interest lies

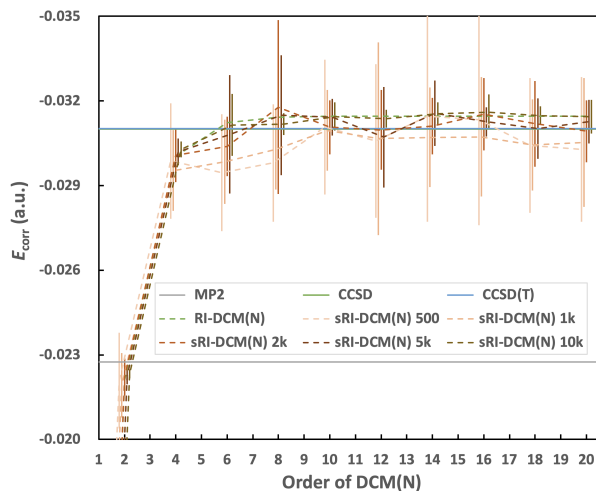


Figure 3: Correlation energy of LiH with RI-DCM and sRI-DCM using different  $N_s$ .

in incorporating other techniques, such as tensor hypercontraction (THC),<sup>[49–51]</sup> to further lower the scaling to  $O(N^3)$  and reduce stochastic noise. Besides, we will use moment expansion to calculate excitation energy.

## AUTHOR DECLARATIONS

### Conflict of Interest

The authors have no conflicts to disclose.

## APPENDIX: RESULTS OBTAINED FROM THE CHOLESKY DECOMPOSITION SCHEME

As mentioned in Section 2.3, using Cholesky decomposition to decouple the crucial four-index intermediate is also a useful scheme. However, since its performance in accuracy is inferior than eigendecomposition, we only provide some data here.

In Figure 4, we plot the correlation energy of some small molecules. Still, the left subfigure corresponds to  $\text{H}_2\text{O}$  and the right includes additional molecules. Moreover, the same axes as in Figure 1 are used to ensure a clear comparison. By comparing Figure 4(a) with Figure 1(a), we can observe that our sRI-DCM with Cholesky decomposition yields larger errors. Nevertheless, its results remain in good agreement with those of RI-DCM, and we therefore regard it as a potentially useful attempt in our theoretical development.

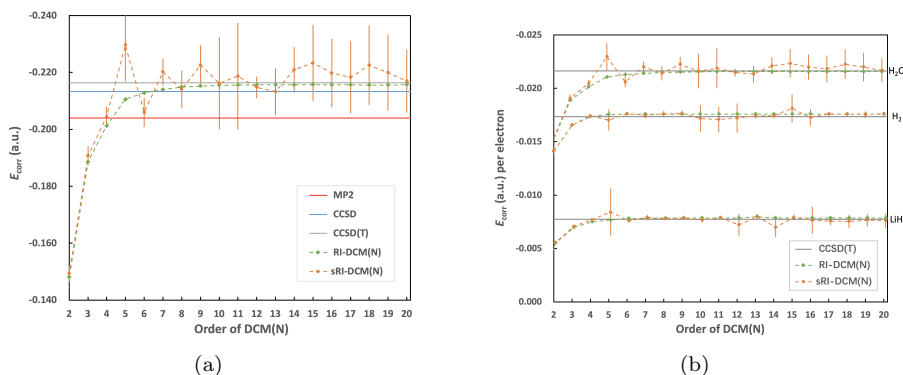


Figure 4: Error comparison of RI-DCM and sRI-DCM among small molecules. Panel (a) is for  $H_2O$  only, and panel (b) adds some other molecules.

## ACKNOWLEDGEMENTS

We acknowledge the high-performance computing (HPC) service from Westlake University. W.D. thanks the funding from National Natural Science Foundation of China (No. 22361142829) and Zhejiang Provincial Natural Science Foundation (No. XHD24B0301). We are grateful for helpful discussions from Joonho Lee.

## DATA AVAILABILITY

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

## References

- [1] A. Szabo and N. S. Ostlund, *Modern quantum chemistry: introduction to advanced electronic structure theory*. Courier Corporation, 2012.
- [2] C. Møller and M. S. Plesset, “Note on an approximation treatment for many-electron systems,” *Physical review*, vol. 46, no. 7, p. 618, 1934.
- [3] R. J. Bartlett and G. D. Purvis, “Many-body perturbation theory, coupled-pair many-electron theory, and the importance of quadruple excitations for the correlation problem,” *International Journal of Quantum Chemistry*, vol. 14, no. 5, pp. 561–581, 1978.
- [4] R. Krishnan and J. A. Pople, “Approximate fourth-order perturbation theory of the electron correlation energy,” *International Journal of Quantum Chemistry*, vol. 14, no. 1, pp. 91–100, 1978.

- [5] J. C. Slater, “The theory of complex spectra,” *Physical review*, vol. 34, no. 10, p. 1293, 1929.
- [6] C. F. Bender and E. R. Davidson, “Studies in configuration interaction: The first-row diatomic hydrides,” *Physical Review*, vol. 183, no. 1, p. 23, 1969.
- [7] F. Coester, “Bound states of a many-particle system,” *Nuclear Physics*, vol. 7, pp. 421–424, 1958.
- [8] F. Coester and H. Kümmel, “Short-range correlations in nuclear wave functions,” *Nuclear Physics*, vol. 17, pp. 477–485, 1960.
- [9] S. M. Blinder and J. E. House, *Mathematical physics in theoretical chemistry*. Elsevier, 2018.
- [10] D. Horn and M. Weinstein, “The t expansion: A nonperturbative analytic tool for hamiltonian systems,” *Physical Review D*, vol. 30, no. 6, p. 1256, 1984.
- [11] C. Stubbins, “Methods of extrapolating the t-expansion series,” *Physical Review D*, vol. 38, no. 6, p. 1942, 1988.
- [12] J. Cioslowski, “Connected moments expansion: a new tool for quantum many-body theory,” *Physical review letters*, vol. 58, no. 2, p. 83, 1987.
- [13] P. J. Knowles, “On the validity and applicability of the connected moments expansion,” *Chemical physics letters*, vol. 134, no. 6, pp. 512–518, 1987.
- [14] J. Cioslowski, M. Kertesz, P. Surjan, and R. Poirier, “Connected moments expansion calculations of the correlation energy in small molecules,” *Chemical physics letters*, vol. 138, no. 6, pp. 516–519, 1987.
- [15] Y. Ai, H. Zhai, J. Tölle, and G. K.-L. Chan, “Quantum many-body linear algebra, hamiltonian moments, and a coupled-cluster inspired framework,” *The Journal of Chemical Physics*, vol. 163, no. 1, p. 011101, 2025.
- [16] A. Zhuravlev, “Cumulant t-expansion for strongly correlated fermions,” *Physics Letters A*, vol. 380, no. 22-23, pp. 1995–1999, 2016.
- [17] —, “Cumulant t-expansion for strongly correlated electrons on a lattice,” *Journal of Experimental and Theoretical Physics*, vol. 131, no. 3, pp. 460–467, 2020.
- [18] —, “Cumulant lanczos tridiagonalization for calculation of the correlation energy in a molecule,” *Computational and Theoretical Chemistry*, vol. 1209, p. 113597, 2022.
- [19] B. Ganoe and M. Head-Gordon, “Doubles connected moments expansion: A tractable approximate horn–weinstein approach for quantum chemistry,” *Journal of Chemical Theory and Computation*, vol. 19, no. 24, pp. 9187–9201, 2023.

- [20] M. Feyereisen, G. Fitzgerald, and A. Komornicki, “Use of approximate integrals in ab initio theory. an application in mp2 energy calculations,” *Chemical physics letters*, vol. 208, no. 5-6, pp. 359–363, 1993.
- [21] K. Eichkorn, O. Treutler, H. Öhm, M. Häser, and R. Ahlrichs, “Auxiliary basis sets to approximate coulomb potentials,” *Chemical physics letters*, vol. 240, no. 4, pp. 283–290, 1995.
- [22] D. E. Bernholdt and R. J. Harrison, “Large-scale correlated electronic structure calculations: the ri-mp2 method on parallel computers,” *Chemical Physics Letters*, vol. 250, no. 5-6, pp. 477–484, 1996.
- [23] D. Neuhauser, E. Rabani, and R. Baer, “Expeditious stochastic approach for mp2 energies in large electronic systems,” *Journal of Chemical theory and Computation*, vol. 9, no. 1, pp. 24–27, 2013.
- [24] Q. Ge, Y. Gao, R. Baer, E. Rabani, and D. Neuhauser, “A guided stochastic energy-domain formulation of the second order møller–plesset perturbation theory,” *The Journal of Physical Chemistry Letters*, vol. 5, no. 1, pp. 185–189, 2014.
- [25] T. Y. Takeshita, W. A. de Jong, D. Neuhauser, R. Baer, and E. Rabani, “Stochastic formulation of the resolution of identity: Application to second order møller–plesset perturbation theory,” *Journal of Chemical Theory and Computation*, vol. 13, no. 10, pp. 4605–4610, 2017.
- [26] C. Zhao, J. Lee, and W. Dou, “Stochastic resolution of identity to cc2 for large systems: Ground state and triplet excitation energy calculations,” *The Journal of Physical Chemistry A*, vol. 128, no. 42, pp. 9302–9310, 2024.
- [27] C. Zhao, Q. Ou, J. Lee, and W. Dou, “Stochastic resolution of identity to cc2 for large systems: excited state properties,” *Journal of Chemical Theory and Computation*, vol. 20, no. 12, pp. 5188–5195, 2024.
- [28] C. Zhao, Q. Ou, C. Li, and W. Dou, “Stochastic resolution of identity to cc2 for large systems: Oscillator strength and ground state gradient calculations,” *The Journal of Chemical Physics*, vol. 163, no. 2, p. 024102, 2025.
- [29] C. Zhao, C. Li, and W. Dou, “Stochastic resolution of identity to cc2 for large systems: Excited-state gradients and derivative couplings,” *The Journal of Chemical Physics*, vol. 163, no. 23, p. 234101, 2025.
- [30] R. Baer, D. Neuhauser, and E. Rabani, “Self-averaging stochastic kohn–sham density-functional theory,” *Physical review letters*, vol. 111, no. 10, p. 106402, 2013.

- [31] D. Neuhauser, R. Baer, and E. Rabani, “Communication: Embedded fragment stochastic density functional theory,” *The Journal of chemical physics*, vol. 141, no. 4, 2014.
- [32] Y. Gao, D. Neuhauser, R. Baer, and E. Rabani, “Sublinear scaling for time-dependent stochastic density functional theory,” *The Journal of chemical physics*, vol. 142, no. 3, 2015.
- [33] D. Neuhauser, E. Rabani, Y. Cytter, and R. Baer, “Stochastic optimally tuned range-separated hybrid density functional theory,” *The Journal of Physical Chemistry A*, vol. 120, no. 19, pp. 3071–3078, 2016.
- [34] N. C. Bradbury, T. Allen, M. Nguyen, and D. Neuhauser, “Deterministic/fragmented-stochastic exchange for large-scale hybrid dft calculations,” *Journal of Chemical Theory and Computation*, vol. 19, no. 24, pp. 9239–9247, 2023.
- [35] M. D. Fabian, E. Rabani, and R. Baer, “Compact gaussian basis sets for stochastic dft calculations,” *Chemical Physics Letters*, vol. 865, p. 141912, 2025.
- [36] D. Neuhauser, R. Baer, and D. Zgid, “Stochastic self-consistent second-order green’s function method for correlation energies of large electronic systems,” *Journal of chemical theory and computation*, vol. 13, no. 11, pp. 5396–5403, 2017.
- [37] T. Y. Takeshita, W. Dou, D. G. Smith, W. A. de Jong, R. Baer, D. Neuhauser, and E. Rabani, “Stochastic resolution of identity second-order matsubara green’s function theory,” *The Journal of chemical physics*, vol. 151, no. 4, 2019.
- [38] W. Dou, T. Y. Takeshita, M. Chen, R. Baer, D. Neuhauser, and E. Rabani, “Stochastic resolution of identity for real-time second-order green’s function: ionization potential and quasi-particle spectrum,” *Journal of chemical theory and computation*, vol. 15, no. 12, pp. 6703–6711, 2019.
- [39] W. Dou, M. Chen, T. Y. Takeshita, R. Baer, D. Neuhauser, and E. Rabani, “Range-separated stochastic resolution of identity: Formulation and application to second-order green’s function theory,” *The Journal of chemical physics*, vol. 153, no. 7, 2020.
- [40] L. Mejía, J. Yin, D. R. Reichman, R. Baer, C. Yang, and E. Rabani, “Stochastic real-time second-order green’s function theory for neutral excitations in molecules and nanostructures,” *Journal of Chemical Theory and Computation*, vol. 19, no. 16, pp. 5563–5571, 2023.
- [41] L. Mejía, S. Sharma, R. Baer, G. K.-L. Chan, and E. Rabani, “Convergence analysis of the stochastic resolution of identity: Comparing hutchinson to hutch++ for the second-order green’s function,” *Journal of Chemical Theory and Computation*, vol. 20, no. 17, pp. 7494–7502, 2024.

- [42] E. Rabani, R. Baer, and D. Neuhauser, “Time-dependent stochastic bethe-salpeter approach,” *Physical Review B*, vol. 91, no. 23, p. 235302, 2015.
- [43] V. Vlček, H. R. Eisenberg, G. Steinle-Neumann, D. Neuhauser, E. Rabani, and R. Baer, “Spontaneous charge carrier localization in extended one-dimensional systems,” *Physical Review Letters*, vol. 116, no. 18, p. 186401, 2016.
- [44] J. Lee and D. R. Reichman, “Stochastic resolution-of-the-identity auxiliary-field quantum monte carlo: Scaling reduction without overhead,” *The Journal of Chemical Physics*, vol. 153, no. 4, 2020.
- [45] N. H. Beebe and J. Linderberg, “Simplifications in the generation and transformation of two-electron integrals in molecular calculations,” *International Journal of Quantum Chemistry*, vol. 12, no. 4, pp. 683–705, 1977.
- [46] E. Epifanovsky, D. Zuev, X. Feng, K. Khistyayev, Y. Shao, and A. I. Krylov, “General implementation of the resolution-of-the-identity and cholesky representations of electron repulsion integrals within coupled-cluster and equation-of-motion methods: Theory and benchmarks,” *The Journal of chemical physics*, vol. 139, no. 13, 2013.
- [47] T. B. Pedersen, S. Lehtola, I. Fdez. Galván, and R. Lindh, “The versatility of the cholesky decomposition in electronic structure theory,” *Wiley interdisciplinary reviews: Computational molecular science*, vol. 14, no. 1, p. e1692, 2024.
- [48] E. Epifanovsky, A. T. Gilbert, X. Feng, J. Lee, Y. Mao, N. Mardirossian, P. Pokhilko, A. F. White, M. P. Coons, A. L. Dempwolff *et al.*, “Software for the frontiers of quantum chemistry: An overview of developments in the q-chem 5 package,” *The Journal of chemical physics*, vol. 155, no. 8, 2021.
- [49] E. G. Hohenstein, R. M. Parrish, and T. J. Martínez, “Tensor hypercontraction density fitting. i. quartic scaling second-and third-order møller-plesset perturbation theory,” *The Journal of chemical physics*, vol. 137, no. 4, 2012.
- [50] E. G. Hohenstein, B. S. Fales, R. M. Parrish, and T. J. Martínez, “Rank-reduced coupled-cluster. iii. tensor hypercontraction of the doubles amplitudes,” *The Journal of Chemical Physics*, vol. 156, no. 5, 2022.
- [51] J. Lee, L. Lin, and M. Head-Gordon, “Systematically improvable tensor hypercontraction: Interpolative separable density-fitting for molecules applied to exact exchange, second-and third-order møller-plesset perturbation theory,” *Journal of chemical theory and computation*, vol. 16, no. 1, pp. 243–263, 2019.