

Monte Carlo explicitly correlated second-order many-body perturbation theory

Cole M. Johnson, Alexander E. Doran, Jinmei Zhang, Edward F. Valeev, and So Hirata

Citation: *The Journal of Chemical Physics* **145**, 154115 (2016); doi: 10.1063/1.4964854

View online: <http://dx.doi.org/10.1063/1.4964854>

View Table of Contents: <http://scitation.aip.org/content/aip/journal/jcp/145/15?ver=pdfcov>

Published by the **AIP Publishing**

Articles you may be interested in

[Stochastic many-body perturbation theory for anharmonic molecular vibrations](#)

J. Chem. Phys. **141**, 084105 (2014); 10.1063/1.4892614

[Communication: Stochastic evaluation of explicitly correlated second-order many-body perturbation energy](#)

J. Chem. Phys. **140**, 031101 (2014); 10.1063/1.4862255

[Stochastic evaluation of second-order many-body perturbation energies](#)

J. Chem. Phys. **137**, 204122 (2012); 10.1063/1.4768697

[Second-order many-body perturbation study of ice Ih](#)

J. Chem. Phys. **137**, 204505 (2012); 10.1063/1.4767898

[Analytic first-order properties from explicitly correlated many-body perturbation theory and Gaussian geminal basis](#)

J. Chem. Phys. **108**, 7946 (1998); 10.1063/1.476235



NEW Special Topic Sections

NOW ONLINE
Lithium Niobate Properties and Applications:
Reviews of Emerging Trends

AIP | Applied Physics
Reviews

Monte Carlo explicitly correlated second-order many-body perturbation theory

Cole M. Johnson,¹ Alexander E. Doran,¹ Jinmei Zhang,² Edward F. Valeev,² and So Hirata^{1,a)}

¹Department of Chemistry, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, USA

²Department of Chemistry, Virginia Tech, Blacksburg, Virginia 24061, USA

(Received 8 July 2016; accepted 3 October 2016; published online 21 October 2016)

A stochastic algorithm is proposed and implemented that computes a basis-set-incompleteness (F12) correction to an *ab initio* second-order many-body perturbation energy as a short sum of 6- to 15-dimensional integrals of Gaussian-type orbitals, an explicit function of the electron-electron distance (geminal), and its associated excitation amplitudes held fixed at the values suggested by Ten-no. The integrals are directly evaluated (without a resolution-of-the-identity approximation or an auxiliary basis set) by the Metropolis Monte Carlo method. Applications of this method to 17 molecular correlation energies and 12 gas-phase reaction energies reveal that both the nonvariational and variational formulas for the correction give reliable correlation energies (98% or higher) and reaction energies (within 2 kJ mol⁻¹ with a smaller statistical uncertainty) near the complete-basis-set limits by using just the aug-cc-pVDZ basis set. The nonvariational formula is found to be 2–10 times less expensive to evaluate than the variational one, though the latter yields energies that are bounded from below and is, therefore, slightly but systematically more accurate for energy differences. Being capable of using virtually any geminal form, the method confirms the best overall performance of the Slater-type geminal among 6 forms satisfying the same cusp conditions. Not having to precompute lower-dimensional integrals analytically, to store them on disk, or to transform them in a non-scalable dense-matrix-multiplication algorithm, the method scales favorably with both system size and computer size; the cost increases only as $O(n^4)$ with the number of orbitals (n), and its parallel efficiency reaches 99.9% of the ideal case on going from 16 to 4096 computer processors. *Published by AIP Publishing.* [<http://dx.doi.org/10.1063/1.4964854>]

I. INTRODUCTION

Slow convergence¹ of electron-correlation energies with respect to the size of atomic-orbital (AO) basis set is one of the most serious weaknesses of *ab initio* molecular orbital (MO) theory.^{2–5} Its cause is well recognized: the ability of the products of the AO basis functions to describe the exact wave functions is rather poor at short inter-electronic distances, or even completely lacking in cases where they are cusped at electron-electron coalescence.^{1,6,7} The most direct and effective remedy^{2,3,5,8} is to use a two-electron basis function that depends explicitly on the inter-electronic distance, r_{12} , which is called a correlation factor or geminal.

Slater^{9,10} and Hylleraas^{11,12} were the first to report the results of such calculations for the helium atom, which were much closer to the exactness than those without an explicit function of r_{12} . Kutzelnigg⁸ developed a general framework in which to include the simplest geminal—a linear function of r_{12} —in Gaussian-basis-set second-order many-body perturbation (MP2) theory,^{13,14} pioneering the explicitly correlated MP2 (MP2-R12) method. In this method, a basis-set-incompleteness (R12) correction to reach the complete-basis-set (CBS) limit is written as a short sum of two-electron

integrals over a geminal, an orthogonality projector, and occupied orbitals only. Here, the orthogonality projector is there to prevent a double counting of the same correlation-energy contribution between MP2 and R12.

One may foresee at least three technical challenges in implementing MP2-R12: (1) Integrals of a geminal multiplying Gaussian-type orbitals (GTOs) need to be evaluated preferably analytically, limiting the forms of such factors to a handful of the simplest ones.^{15–21} (2) The excitation amplitudes multiplying these integrals have to be determined through an $O(n^6)$ process (where n is the number of orbitals),²² making MP2-R12 less scalable with system size than $O(n^5)$ of its parent MP2. (3) When the orthogonal projector is expanded, numerous three-electron (9-dimensional) and four-electron (12-dimensional) integrals involving both occupied and virtual orbitals emerge. Computer and human costs of their evaluation would be prohibitive.

Combining with their Gaussian-type geminal method, Persson and Taylor²³ alleviated problems (1) and (3). In their approach, a geminal of any form can be expanded by a small number of Gaussian-type geminals (GTGs), and the resulting many-electron integrals over GTOs and GTGs are directly and analytically evaluated by the McMurchie–Davidson algorithm^{24–26} or by the Rys quadrature.^{27–29} Although a GTG does not satisfy the cusp conditions,^{1,6,7} a geminal expanded by GTGs is still found to capture a majority of

^{a)}Email: sohirata@illinois.edu

the R12 correction, suggesting³⁰ that an accurate description of a correlation hole at intermediate r_{12} is more important than that at $r_{12} = 0$. Furthermore, Szalewicz *et al.*^{31,32} showed that approximating the orthogonal projector by the so-called weak-orthogonality projector makes higher-than-three-electron integrals disappear in the formalism. However, such simplification comes at a price of somewhat poorer performance.^{4,33}

Klopper and Kutzelnigg^{34,35} proposed a practical and now-widely used method that resolved problem (3), in which all three- and four-electron integrals are approximated as linear combinations of two-electron integrals by the resolution-of-the-identity (RI) insertion with an auxiliary basis set (ABS).^{34–37} Also, problem (1) was partially addressed by Ten-no, who introduced^{16,38} a nonlinear function of r_{12} known as the Slater-type geminal, which was found to accelerate the basis-set convergence more than linear r_{12} or a GTG. Ten-no thus established the MP2-F12 method. However, the true merit of the Slater-type geminal may be that its functional form resembles the correlation hole in a wide range of r_{12} so closely that the excitation amplitudes multiplying its geminal integrals can be held fixed at values dictated by the cusp conditions.^{1,6,7} Hence, Ten-no's fixed-amplitude method,³⁹ combined with his Slater-type geminal,^{16,38} brings the cost scaling of MP2-F12 down to the same $O(n^5)$ of MP2, resolving problem (2), because the excitation amplitudes no longer have to be determined through the $O(n^6)$ process.

In this work, we explore a rather different approach—the use of a stochastic algorithm—to resolving all three problems mentioned above as well as the fourth problem: (4) In deterministic implementations, a large number of two-electron integrals are generated from far fewer, higher-dimensional integrals by the RI insertions, which, together with the ordinary two-electron integrals for MP2, need to be pre-calculated, stored (on disks), and used in a series of dense matrix multiplications; such computational steps are non-scalable with system size (in the sense that the cost increases as the fifth power of the number of orbitals) or computer size (because dense matrix multiplications⁴⁰ tend to involve large and frequent inter-processor communications). Here, we present a novel, stochastic MP2-F12 method, which is based on Ten-no's fixed-amplitude method, but otherwise differs from the existing MP2-F12 methods in that it can use virtually any geminal, does not require the RI approximation or an ABS, is not predicated upon numerous molecular integrals precomputed or stored on disks, and thus scales favorably with both system and computer sizes.

It is an extension of the Monte Carlo MP2 (MC-MP2) method,⁴¹ which is a member of the Brueckner–Goldstone quantum Monte Carlo (BGQMC) family¹⁴ of methods for both electrons^{41–46} and vibrations.^{47–50} It wedges diagrammatic many-body perturbation and Green's function theories with the Metropolis Monte Carlo (MC) algorithm,⁵¹ in the same spirit as other recent studies combining *ab initio* MO theory with quantum Monte Carlo (QMC).^{52–62} Unlike more conventional QMC methods,^{63–67} MC-MP2 can compute energy differences (such as correlation energies, correlated ionization potentials

and electron affinities, and quasiparticle energy bands) directly and not as small differences of noisy total energies, does not suffer from any sign problem or fixed-node error, and is systematically convergent at the exact solutions of the Schrödinger equations as the perturbation rank and basis-set size are increased. It is also rigorously (diagrammatically) size-consistent and thus free from a finite-size error. Unlike deterministic MP2, MC-MP2 does not need two-electron integrals precomputed or stored either in the AO or MO basis and is, therefore, more scalable. The operation cost per MC step of MC-MP2 is shown to be linear⁴³ with system size and the cost to achieve a given relative statistical uncertainty is found to be cubic.⁶⁸ It can be easily and efficiently parallelized on many central processing units (CPUs)⁴⁴ or on many graphical processing units (GPUs),⁶⁸ sometimes achieving scalability⁶⁸ unprecedented for *ab initio* electron-correlation theories.

In the explicitly correlated extension called MC-MP2-F12 presented here, we exploit another important advantage of MC-MP2, which is its flexibility with various mathematical forms of basis functions. MC-MP2-F12 can use virtually any geminal forms as it evaluates necessary integrals numerically by the Metropolis MC method. Furthermore, being a sparse integration method, the MC method's relative superiority over quadrature grows with dimension; it was argued that the former is more efficient than the latter when the dimension exceeds eight.⁶⁹ In this sense, three-electron (9-dimensional) and higher-dimensional integrals or even two-electron (6-dimensional) integrals may be more suitably handled by MC integrations than by analytical integrations.^{25,26,28,29} The latter are also more expensive to develop and may have a hard ceiling of applicability when the problem size is too large. The former, in contrast, would execute for a far larger problem, only taking longer to converge and give meaningful results.

Also, the MC method makes it unnecessary to factor high-dimensional integrals into lower-dimensional ones in MC-MP2-F12; it can directly evaluate a short sum of high-dimensional integrals, neither requiring the RI approximation and a large ABS nor involving many dense matrix multiplications, which tend not to be scalable. Consequently, the cost of MC-MP2-F12 to achieve a given relative statistical error is found to scale more favorably [$O(n^4)$] than MP2 or MP2-F12 with Ten-no's fixed amplitudes. Its parallel algorithm of MC-MP2-F12 is also found to exhibit near-perfect scalability up to thousands of processing cores and is essentially free of disk I/O.

In a previous Communication in this journal,⁴¹ Willow and three of the present authors reported a pilot implementation of the MC-MP2-F12 method using the nonvariational V formula (see below) and obtained *total* correlation energies near their CBS limits using Ten-no's fixed-amplitude ansatz³⁹ with the Slater-type geminal.¹⁶ Since this (V) formula is not variational with respect to the form of the geminal or the excitation amplitude multiplying its integrals (which are held fixed³⁹), whether it gives reliable *relative* energies near the CBS limits was not clear. While the cost per MC step for MC-MP2-F12 with the nonvariational (V) formula was shown⁴¹ to be quadratic with the number of orbitals,

the more meaningful cost, i.e., that required to reach a given relative statistical uncertainty, was unknown. Parallel scalability was strongly inferred but never demonstrated, either. Furthermore, the ease of use of any geminal was touted but never exploited. In this paper, we fully develop MC-MP2-F12 using both of the nonvariational (*V*) and variational (*VBX*) formulas and answer all of these unanswered questions.

II. THEORY

A. E_{F12}

The correlation energy of MP2-F12 theory^{8,16,22,39,41} is written as

$$E_{MP2-F12} = E_{MP2} + E_{F12}, \quad (1)$$

where E_{MP2} is the MP2 correlation energy in a finite (AO) basis set¹⁴ and E_{F12} is the correction for the basis-set-incompleteness error. The latter (the R12 or F12 correction) is derived from the Hylleraas functional^{3,5} and consists of three terms,

$$E_{F12} = 2E_{F12}^V + E_{F12}^B + E_{F12}^X, \quad (2)$$

with

$$E_{F12}^V = \sum_{i,j,m,n}^{\text{occ.}} V_{mn}^{ij} (2t_{ij}^{mn} - t_{ji}^{mn}), \quad (3)$$

$$E_{F12}^B = \sum_{i,j,k,l,m,n}^{\text{occ.}} t_{mn}^{kl} B_{kl}^{ij} (2t_{ij}^{mn} - t_{ji}^{mn}), \quad (4)$$

$$E_{F12}^X = - \sum_{i,j,k,l,m,n}^{\text{occ.}} (\epsilon_m + \epsilon_n) t_{mn}^{kl} X_{kl}^{ij} (2t_{ij}^{mn} - t_{ji}^{mn}), \quad (5)$$

where ϵ_p is the p th Hartree–Fock (HF) orbital energy, t_{ij}^{mn} is the so-called geminal amplitude, and all summations run over occupied orbitals spanned by the basis set. The other factors are molecular integrals of two electrons, which are written in the standard physicists' notation as

$$V_{mn}^{ij} = \langle ij | r_{12}^{-1} \hat{Q}_{12} f_{12} | mn \rangle, \quad (6)$$

$$B_{mn}^{ij} = \langle ij | f_{12} \hat{Q}_{12} (\hat{F}_1 + \hat{F}_2) \hat{Q}_{12} f_{12} | mn \rangle, \quad (7)$$

$$X_{mn}^{ij} = \langle ij | f_{12} \hat{Q}_{12} f_{12} | mn \rangle, \quad (8)$$

where f_{12} is the geminal (an explicit function of r_{12}) and \hat{F}_n is the Fock operator of electron n , i.e.,

$$\hat{F}_n = \hat{T}_n + \hat{V}_n + \hat{J}_n - \hat{K}_n. \quad (9)$$

The right-hand side of this equation is the sum of the kinetic-energy operator, nuclear-attraction operator, Coulomb operator, and exchange operator, in this order. Operator \hat{Q}_{12} is the strong-orthogonality projector^{70–73} defined as

$$\hat{Q}_{12} = (1 - \hat{O}_1)(1 - \hat{O}_2) - \hat{V}_1 \hat{V}_2, \quad (10)$$

with

$$\hat{O}_1 = \sum_i^{\text{occ.}} \varphi_i(\mathbf{r}_1) \int d\mathbf{r}_3 \varphi_i^*(\mathbf{r}_3) \hat{P}_{31}, \quad (11)$$

$$\hat{V}_1 = \sum_a^{\text{vir.}} \varphi_a(\mathbf{r}_1) \int d\mathbf{r}_3 \varphi_a^*(\mathbf{r}_3) \hat{P}_{31}, \quad (12)$$

where \hat{P}_{31} replaces the coordinates of electron 1 by 3 in what follows. The summation in Eq. (12) runs over all virtual orbitals spanned by the basis set.

The F12 correction (E_{F12}) is variational with the geminal amplitudes, whose values are, therefore, to be determined by minimizing E_{F12} , through an $O(n^6)$ process,²² where n is the number of orbitals. It can also be shown that at the minimum, $E_{F12}^B + E_{F12}^X = -E_{F12}^V$ and hence,

$$E_{F12} = E_{F12}^V. \quad (13)$$

Ten-no showed³⁹ that an accurate F12 correction can be obtained by holding the geminal amplitudes fixed at the values dictated by the cusp conditions,^{1,6,7} i.e.,

$$t_{ij}^{mn} = \frac{3}{8} \delta_{mi} \delta_{nj} + \frac{1}{8} \delta_{mj} \delta_{ni}, \quad (14)$$

insofar as an appropriate form of the geminal is used, such as Ten-no's Slater-type geminal,

$$f_{12} = \frac{1 - e^{-\gamma r_{12}}}{\gamma}, \quad (15)$$

where γ is an adjustable parameter. With Ten-no's fixed amplitudes, the F12 correction as defined by Eq. (2) (the variational *VBX* formula) is no longer a variational minimum with respect to the "size" (t_{ij}^{mn}) of the geminal, but is still variationally bound from below with respect to its "shape" (γ). The nonvariational (*V*) formula⁴¹ [Eq. (13)] is not variational with either. Since the errors in a variational energy have the same sign and tend to also have similar magnitudes, they may cancel with each other to yield accurate energy differences. Below, we quantify the performance of the nonvariational (*V*) and variational (*VBX*) formulas for energy differences. We will also address another important question: What is special about the Slater-type geminal of Eq. (15) that makes the fixed-amplitude method work? Below, we explore several geminals that have the right asymptote,

$$f_{12} \simeq r_{12} \quad \text{as } r_{12} \rightarrow 0. \quad (16)$$

Substituting Eq. (14) into Eq. (2), we obtain

$$E_{F12}^V = \frac{5}{8} \sum_{i,j}^{\text{occ.}} V_{ij}^{ij} - \frac{1}{8} \sum_{i,j}^{\text{occ.}} V_{ji}^{ij}, \quad (17)$$

$$E_{F12}^B = \frac{7}{32} \sum_{i,j}^{\text{occ.}} B_{ij}^{ij} + \frac{1}{32} \sum_{i,j}^{\text{occ.}} B_{ji}^{ij}, \quad (18)$$

$$E_{F12}^X = -\frac{7}{32} \sum_{i,j}^{\text{occ.}} (\epsilon_i + \epsilon_j) X_{ij}^{ij} - \frac{1}{32} \sum_{i,j}^{\text{occ.}} (\epsilon_j + \epsilon_i) X_{ji}^{ij}. \quad (19)$$

Before converting them into forms suitable for MC integration, we make two additional simplifying approximations to the formalism.

First, we assume the so-called generalized Brillouin condition (GBC) and extended Brillouin condition (EBC),^{17,22} i.e.,

$$\hat{F}_n \varphi_i(\mathbf{r}_n) = \epsilon_i \varphi_i(\mathbf{r}_n), \quad (20)$$

$$\hat{F}_n \varphi_a(\mathbf{r}_n) = \epsilon_a \varphi_a(\mathbf{r}_n), \quad (21)$$

where i and a label an occupied and virtual orbital, respectively. This assumption leads to

$$[\hat{F}_1 + \hat{F}_2, \hat{Q}_{12}] = 0, \quad (22)$$

which, in turn, simplifies the B integrals [Eq. (7)] into

$$B_{ij}^{ij} = \langle ij | f_{12} \hat{Q}_{12} (\hat{F}_1 + \hat{F}_2) Q_{12} f_{12} | ij \rangle \quad (23)$$

$$= \langle ij | f_{12} \hat{Q}_{12} \hat{Q}_{12} (\hat{F}_1 + \hat{F}_2) f_{12} | ij \rangle + \langle ij | f_{12} \hat{Q}_{12} [\hat{F}_1 + \hat{F}_2, \hat{Q}_{12}] f_{12} | ij \rangle \quad (24)$$

$$= \langle ij | f_{12} \hat{Q}_{12} (\hat{F}_1 + \hat{F}_2) f_{12} | ij \rangle \quad (25)$$

and

$$B_{ji}^{ij} = \langle ij | f_{12} \hat{Q}_{12} (\hat{F}_1 + \hat{F}_2) f_{12} | ji \rangle, \quad (26)$$

where the idempotency of \hat{Q}_{12} is used, i.e., $\hat{Q}_{12}^2 = \hat{Q}_{12}$. As will be discussed below, when expanded, each \hat{Q}_{12} introduces a new electron and increases the dimension of an integral by three [see Eqs. (11) and (12)]. Therefore, Eq. (25) is far more computationally tractable than Eq. (23).

Second, using the commutator $[\hat{F}_1 + \hat{F}_2, f_{12}]$, we rewrite Eq. (25) as

$$B_{ij}^{ij} = \langle ij | f_{12} \hat{Q}_{12} (\hat{F}_1 + \hat{F}_2) f_{12} | ij \rangle \quad (27)$$

$$= \langle ij | f_{12} \hat{Q}_{12} [\hat{F}_1 + \hat{F}_2, f_{12}] | ij \rangle + \langle ij | f_{12} \hat{Q}_{12} f_{12} (\hat{F}_1 + \hat{F}_2) | ij \rangle \quad (28)$$

$$= \langle ij | f_{12} \hat{Q}_{12} [\hat{F}_1 + \hat{F}_2, f_{12}] | ij \rangle + (\epsilon_i + \epsilon_j) \langle ij | f_{12} \hat{Q}_{12} f_{12} | ij \rangle \quad (29)$$

$$= \langle ij | f_{12} \hat{Q}_{12} [\hat{F}_1 + \hat{F}_2, f_{12}] | ij \rangle + (\epsilon_i + \epsilon_j) X_{ij}^{ij} \quad (30)$$

and

$$B_{ji}^{ij} = \langle ij | f_{12} \hat{Q}_{12} [\hat{F}_1 + \hat{F}_2, f_{12}] | ji \rangle + (\epsilon_i + \epsilon_j) X_{ji}^{ij}. \quad (31)$$

Note that the GBC is invoked in Eq. (29).

These two approximations reduce E_{F12} to

$$E_{F12} = 2E_{F12}^V + E_{F12}^{BX}, \quad (32)$$

with

$$E_{F12}^{BX} = \frac{7}{32} \sum_{i,j}^{\text{occ.}} (BX)_{ij}^{ij} + \frac{1}{32} \sum_{i,j}^{\text{occ.}} (BX)_{ji}^{ij} \quad (33)$$

and

$$(BX)_{mn}^{ij} = \langle ij | f_{12} \hat{Q}_{12} [\hat{F}_1 + \hat{F}_2, f_{12}] | mn \rangle. \quad (34)$$

The GBC is known to be an accurate approximation, causing errors that are no greater than a milli-Hartree in molecules ranging from H_2 to O_3 .¹⁷ The EBC is less accurate, but still tolerable.^{17,74} For a smaller basis set, in fact, it is recommended^{17,74} that the EBC be invoked when the GBC is used.

B. E_{F12}^V

Expanding \hat{Q}_{12} , i.e., substituting Eq. (10) into Eq. (17), we obtain⁴¹

$$E_{F12}^V = E_{2e}^V + E_{3e}^V + E_{4e}^V, \quad (35)$$

with

$$E_{2e}^V = \frac{5}{8} \sum_{i,j}^{\text{occ.}} \langle ij | \frac{f_{12}}{r_{12}} | ij \rangle - \frac{1}{8} \sum_{i,j}^{\text{occ.}} \langle ij | \frac{f_{12}}{r_{12}} | ji \rangle, \quad (36)$$

$$E_{3e}^V = -\frac{5}{4} \sum_{i,j,k}^{\text{occ.}} \langle ijk | \frac{f_{23}}{r_{12}} | kji \rangle + \frac{1}{4} \sum_{i,j,k}^{\text{occ.}} \langle ijk | \frac{f_{23}}{r_{12}} | kij \rangle, \quad (37)$$

$$E_{4e}^V = \frac{5}{8} \sum_{i,j,k,l}^{\text{occ.}} \langle ijkl | \frac{f_{34}}{r_{12}} | kl ij \rangle - \frac{1}{8} \sum_{i,j,k,l}^{\text{occ.}} \langle ijkl | \frac{f_{34}}{r_{12}} | kl ji \rangle - \frac{5}{8} \sum_{i,j}^{\text{occ.}} \sum_{a,b}^{\text{vir.}} \langle ijab | \frac{f_{34}}{r_{12}} | abij \rangle + \frac{1}{8} \sum_{i,j}^{\text{occ.}} \sum_{a,b}^{\text{vir.}} \langle ijab | \frac{f_{34}}{r_{12}} | abji \rangle, \quad (38)$$

where the subscripts on the left-hand sides indicate the number of electrons involved and thus the dimension of integrals. They are rewritten in the MC-integrable forms as

$$E_{2e}^V = \iint d\mathbf{r}_1 d\mathbf{r}_2 F_{2e}^V(\mathbf{r}_1, \mathbf{r}_2), \quad (39)$$

$$E_{3e}^V = \iiint d\mathbf{r}_1 d\mathbf{r}_2 d\mathbf{r}_3 F_{3e}^V(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3), \quad (40)$$

$$E_{4e}^V = \iiint d\mathbf{r}_1 d\mathbf{r}_2 d\mathbf{r}_3 d\mathbf{r}_4 F_{4e}^V(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \mathbf{r}_4), \quad (41)$$

with

$$F_{2e}^V(\mathbf{r}_1, \mathbf{r}_2) = \frac{5}{8} \frac{f_{12} O_{11} O_{22}}{r_{12}} - \frac{1}{8} \frac{f_{12} O_{12} O_{21}}{r_{12}}, \quad (42)$$

$$F_{3e}^V(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3) = -\frac{5}{4} \frac{f_{23} O_{13} O_{22} O_{31}}{r_{12}} + \frac{1}{4} \frac{f_{23} O_{12} O_{23} O_{31}}{r_{12}}, \quad (43)$$

$$F_{4e}^V(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \mathbf{r}_4) = \frac{5}{8} \frac{f_{34} O_{13} O_{24} O_{31} O_{42}}{r_{12}} - \frac{1}{8} \frac{f_{34} O_{14} O_{23} O_{31} O_{42}}{r_{12}} - \frac{5}{8} \frac{f_{34} O_{13} O_{24} V_{31} V_{42}}{r_{12}} + \frac{1}{8} \frac{f_{34} O_{14} O_{23} V_{31} V_{42}}{r_{12}}, \quad (44)$$

where

$$O_{pq} = \sum_i^{\text{occ.}} \varphi_i^*(\mathbf{r}_p) \varphi_i(\mathbf{r}_q), \quad (45)$$

$$V_{pq} = \sum_a^{\text{vir.}} \varphi_a^*(\mathbf{r}_p) \varphi_a(\mathbf{r}_q). \quad (46)$$

The formalism of this term is unchanged from Ref. 41.

C. E_{F12}^{BX}

The sum of the B and X terms, Eq. (33), is newly considered in this work. Expanding \hat{Q}_{12} in this equation, we find

$$E_{F12}^{BX} = E_{2e}^T + E_{3e}^T + E_{4e}^T + E_{3e}^K + E_{4e}^K + E_{5e}^K, \quad (47)$$

with

$$E_{2e}^T = \frac{7}{32} \sum_{i,j}^{\text{occ.}} \langle ij | f_{12} [\hat{T}_1 + \hat{T}_2, f_{12}] | ij \rangle + \frac{1}{32} \sum_{i,j}^{\text{occ.}} \langle ij | f_{12} [\hat{T}_1 + \hat{T}_2, f_{12}] | ji \rangle, \quad (48)$$

$$E_{3e}^T = -\frac{7}{16} \sum_{i,j,k}^{\text{occ.}} \langle ijk | f_{12} [\hat{T}_2 + \hat{T}_3, f_{23}] | kji \rangle - \frac{1}{16} \sum_{i,j,k}^{\text{occ.}} \langle ijk | f_{12} [\hat{T}_2 + \hat{T}_3, f_{23}] | kij \rangle, \quad (49)$$

$$E_{4e}^T = \frac{7}{32} \sum_{i,j,k,l}^{\text{occ.}} \langle ijkl | f_{12} [\hat{T}_3 + \hat{T}_4, f_{34}] | klji \rangle + \frac{1}{32} \sum_{i,j,k,l}^{\text{occ.}} \langle ijkl | f_{12} [\hat{T}_3 + \hat{T}_4, f_{34}] | klji \rangle - \frac{7}{32} \sum_{i,j}^{\text{occ.}} \sum_{a,b}^{\text{vir.}} \langle ijab | f_{12} [\hat{T}_3 + \hat{T}_4, f_{34}] | abij \rangle - \frac{1}{32} \sum_{i,j}^{\text{occ.}} \sum_{a,b}^{\text{vir.}} \langle ijab | f_{12} [\hat{T}_3 + \hat{T}_4, f_{34}] | abji \rangle \quad (50)$$

and

$$E_{3e}^K = -\frac{7}{32} \sum_{i,j}^{\text{occ.}} \langle ij | f_{12} [\hat{K}_1 + \hat{K}_2, f_{12}] | ij \rangle - \frac{1}{32} \sum_{i,j}^{\text{occ.}} \langle ij | f_{12} [\hat{K}_1 + \hat{K}_2, f_{12}] | ji \rangle, \quad (51)$$

$$E_{4e}^K = +\frac{7}{16} \sum_{i,j,k}^{\text{occ.}} \langle ijk | f_{12} [\hat{K}_2 + \hat{K}_3, f_{23}] | kji \rangle + \frac{1}{16} \sum_{i,j,k}^{\text{occ.}} \langle ijk | f_{12} [\hat{K}_2 + \hat{K}_3, f_{23}] | kij \rangle, \quad (52)$$

$$E_{5e}^K = -\frac{7}{32} \sum_{i,j,k,l}^{\text{occ.}} \langle ijkl | f_{12} [\hat{K}_3 + \hat{K}_4, f_{34}] | klji \rangle - \frac{1}{32} \sum_{i,j,k,l}^{\text{occ.}} \langle ijkl | f_{12} [\hat{K}_3 + \hat{K}_4, f_{34}] | klji \rangle + \frac{7}{32} \sum_{i,j}^{\text{occ.}} \sum_{a,b}^{\text{vir.}} \langle ijab | f_{12} [\hat{K}_3 + \hat{K}_4, f_{34}] | abij \rangle + \frac{1}{32} \sum_{i,j}^{\text{occ.}} \sum_{a,b}^{\text{vir.}} \langle ijab | f_{12} [\hat{K}_3 + \hat{K}_4, f_{34}] | abji \rangle, \quad (53)$$

where we use the fact that only the kinetic-energy (\hat{T}_n) and exchange (\hat{K}_n) operators in the Fock operator [Eq. (9)] do not

commute with f_{12} and, therefore,

$$[\hat{F}_1 + \hat{F}_2, f_{12}] = [\hat{T}_1 + \hat{T}_2, f_{12}] - [\hat{K}_1 + \hat{K}_2, f_{12}]. \quad (54)$$

The subscripts “2e,” “3e,” etc., again refer to the number of electrons that appear in the integrals. It is incremented in the exchange terms [Eqs. (51)–(53)] because the exchange operator introduces one more electron, just like \hat{Q}_{12} (see Appendix A).

After straightforward, but rather tedious algebra, which is partially computerized in this work, we arrive at the following MC-integrable expressions:

$$E_{2e}^T = \iint d\mathbf{r}_1 d\mathbf{r}_2 F_{2e}^{T1}(\mathbf{r}_1, \mathbf{r}_2) + \iint d\mathbf{r}_1 d\mathbf{r}_2 F_{2e}^{T2}(\mathbf{r}_1, \mathbf{r}_2), \quad (55)$$

$$E_{3e}^T = \iiint d\mathbf{r}_1 d\mathbf{r}_2 d\mathbf{r}_3 F_{3e}^{T1}(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3) + \iiint d\mathbf{r}_1 d\mathbf{r}_2 d\mathbf{r}_3 F_{3e}^{T2}(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3), \quad (56)$$

$$E_{4e}^T = \iiint d\mathbf{r}_1 d\mathbf{r}_2 d\mathbf{r}_3 d\mathbf{r}_4 F_{4e}^{T1}(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \mathbf{r}_4) + \iiint d\mathbf{r}_1 d\mathbf{r}_2 d\mathbf{r}_3 d\mathbf{r}_4 F_{4e}^{T2}(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \mathbf{r}_4), \quad (57)$$

$$E_{3e}^K = \iiint d\mathbf{r}_1 d\mathbf{r}_2 d\mathbf{r}_3 F_{3e}^K(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3), \quad (58)$$

$$E_{4e}^K = \iiint d\mathbf{r}_1 d\mathbf{r}_2 d\mathbf{r}_3 d\mathbf{r}_4 F_{4e}^K(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \mathbf{r}_4), \quad (59)$$

$$E_{5e}^K = \int \cdots \int d\mathbf{r}_1 \cdots d\mathbf{r}_5 F_{5e}^K(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \mathbf{r}_4, \mathbf{r}_5). \quad (60)$$

The integrals in the two-electron kinetic-energy contribution [Eq. (55)] are given by

$$F_{2e}^{T1}(\mathbf{r}_1, \mathbf{r}_2) = \frac{7}{32} \frac{f_{12} f_{12}^{(a)} O_{11} O_{22}}{r_{12}} + \frac{7}{32} \frac{f_{12} f_{12}^{(c)} (O_{11} O'_{22} - O'_{11} O_{22})}{r_{12}} + \frac{1}{32} \frac{f_{12} f_{12}^{(a)} O_{12} O_{21}}{r_{12}} + \frac{1}{32} \frac{f_{12} f_{12}^{(c)} (O_{21} O'_{12} - O'_{21} O_{12})}{r_{12}}, \quad (61)$$

$$F_{2e}^{T2}(\mathbf{r}_1, \mathbf{r}_2) = \frac{7}{32} f_{12} f_{12}^{(b)} O_{11} O_{22} + \frac{7}{32} f_{12} f_{12}^{(d)} (O_{11} O'_{22} - O'_{11} O_{22}) + \frac{1}{32} f_{12} f_{12}^{(b)} O_{12} O_{21} + \frac{1}{32} f_{12} f_{12}^{(d)} (O_{21} O'_{12} - O'_{21} O_{12}), \quad (62)$$

with

$$O'_{pq} = \sum_i^{\text{occ.}} \varphi_i^*(\mathbf{r}_p) \mathbf{r}_{12} \cdot \nabla_q \varphi_i(\mathbf{r}_q). \quad (63)$$

See Appendix A for the derivation of Eqs. (61) and (62). The integrand of Eq. (55) is divided into two terms, F_{2e}^{T1} and F_{2e}^{T2} , because they have rather different behavior as $r_{12} \rightarrow 0$,

thus requiring different weight functions in MC integrations (see below). The subscripts of \mathbf{r}_{12} in Eq. (63) are always “12” and independent of p or q . This is related to the fact that, in Eqs. (61) and (62) and throughout our formalism, the singular operator is chosen (by coordinate interchanges) to be always of the form r_{12}^{-1} , so that the same weight function containing r_{12}^{-1} can be consistently applied to these dimensions.

The factors of the geminal, $f^{(a)}$, $f^{(b)}$, $f^{(c)}$, and $f^{(d)}$, are defined as

$$[\hat{T}_1 + \hat{T}_2, f_{12}] = \frac{f_{12}^{(a)}}{r_{12}} + f_{12}^{(b)} - \left\{ \frac{f_{12}^{(c)}}{r_{12}} + f_{12}^{(d)} \right\} \mathbf{r}_{12} \cdot (\nabla_1 - \nabla_2), \quad (64)$$

whose actual forms are compiled in Table I for the 6 geminals⁷⁵ considered in this work. This form is related to the so-called Kutzelnigg’s regularization operator.⁸ The cusp conditions^{6,7} reduce to

$$f_{12}^{(a)} \simeq -2 \quad \text{as } r_{12} \rightarrow 0, \quad (65)$$

which is satisfied by all the geminals in Table I.

The integrands in the three-electron kinetic contribution [Eq. (56)] are given by

$$F_{3e}^{T1}(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3) = -\frac{7}{16} \frac{f_{23} f_{12}^{(a)} O_{31} O_{22} O_{13}}{r_{12}} - \frac{7}{16} \frac{f_{23} f_{12}^{(c)} (O'_{22} O_{31} - O'_{31} O_{22}) O_{13}}{r_{12}} - \frac{1}{16} \frac{f_{23} f_{12}^{(a)} O_{32} O_{21} O_{13}}{r_{12}} - \frac{1}{16} \frac{f_{23} f_{12}^{(c)} (O'_{32} O_{21} - O'_{21} O_{32}) O_{13}}{r_{12}}, \quad (66)$$

$$F_{3e}^{T2}(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3) = -\frac{7}{16} f_{23} f_{12}^{(b)} O_{31} O_{22} O_{13} - \frac{7}{16} f_{23} f_{12}^{(d)} (O'_{22} O_{31} - O'_{31} O_{22}) O_{13} - \frac{1}{16} f_{23} f_{12}^{(b)} O_{32} O_{21} O_{13} - \frac{1}{16} f_{23} f_{12}^{(d)} (O'_{32} O_{21} - O'_{21} O_{32}) O_{13}, \quad (67)$$

whereas those of the four-electron kinetic contribution [Eq. (57)] read

$$F_{4e}^{T1}(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \mathbf{r}_4) = \frac{7}{32} \frac{f_{34} f_{12}^{(a)} O_{31} O_{42} O_{13} O_{24}}{r_{12}} + \frac{7}{32} \frac{f_{34} f_{12}^{(c)} (O'_{42} O_{31} - O'_{31} O_{42}) O_{13} O_{24}}{r_{12}} + \frac{1}{32} \frac{f_{34} f_{12}^{(a)} O_{32} O_{41} O_{13} O_{24}}{r_{12}} + \frac{1}{32} \frac{f_{34} f_{12}^{(c)} (O'_{32} O_{41} - O'_{41} O_{32}) O_{13} O_{24}}{r_{12}} - \frac{7}{32} \frac{f_{34} f_{12}^{(a)} O_{31} O_{42} V_{13} V_{24}}{r_{12}} - \frac{7}{32} \frac{f_{34} f_{12}^{(c)} (O'_{42} O_{31} - O'_{31} O_{42}) V_{13} V_{24}}{r_{12}} - \frac{1}{32} \frac{f_{34} f_{12}^{(a)} O_{32} O_{41} V_{13} V_{24}}{r_{12}} - \frac{1}{32} \frac{f_{34} f_{12}^{(c)} (O'_{32} O_{41} - O'_{41} O_{32}) V_{13} V_{24}}{r_{12}} \quad (68)$$

and

$$F_{4e}^{T2}(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \mathbf{r}_4) = \frac{7}{32} f_{34} f_{12}^{(b)} O_{31} O_{42} O_{13} O_{24} + \frac{7}{32} f_{34} f_{12}^{(d)} (O'_{42} O_{31} - O'_{31} O_{42}) O_{13} O_{24} + \frac{1}{32} f_{34} f_{12}^{(b)} O_{32} O_{41} O_{13} O_{24} + \frac{1}{32} f_{34} f_{12}^{(d)} (O'_{32} O_{41} - O'_{41} O_{32}) O_{13} O_{24} - \frac{7}{32} f_{34} f_{12}^{(b)} O_{31} O_{42} V_{13} V_{24} - \frac{7}{32} f_{34} f_{12}^{(d)} (O'_{42} O_{31} - O'_{31} O_{42}) V_{13} V_{24} - \frac{1}{32} f_{34} f_{12}^{(b)} O_{32} O_{41} V_{13} V_{24} - \frac{1}{32} f_{34} f_{12}^{(d)} (O'_{32} O_{41} - O'_{41} O_{32}) V_{13} V_{24}. \quad (69)$$

The integrands in the exchange contributions [Eqs. (58)–(60)] are found to be

$$F_{3e}^K(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3) = \frac{7}{16} \frac{f_{23}(f_{23} - f_{13})}{r_{12}} O_{12} O_{33} O_{21} + \frac{1}{16} \frac{f_{23}(f_{23} - f_{13})}{r_{12}} O_{13} O_{32} O_{21}, \quad (70)$$

TABLE I. Six geminals considered in this work and the components of their commutator with the kinetic-energy operator in Eq. (64).

Geminal	Name	$f_{12}^{(a)}$	$f_{12}^{(b)}$	$f_{12}^{(c)}$	$f_{12}^{(d)}$
$f_{12}^{(1)} = (1 - e^{-\gamma r_{12}})/\gamma$	Slater	$-2e^{-\gamma r_{12}}$	$\gamma e^{-\gamma r_{12}}$	$e^{-\gamma r_{12}}$	0
$f_{12}^{(2)} = (1 - e^{-\gamma r_{12}^2})/(\gamma r_{12})$	Gauss	$-2e^{-\gamma r_{12}^2}$	$4\gamma r_{12} e^{-\gamma r_{12}^2}$	$(e^{-\gamma r_{12}^2} - 1)/(\gamma r_{12}^2) + 2e^{-\gamma r_{12}^2}$	0
$f_{12}^{(3)} = \gamma r_{12}/(\gamma + r_{12})$	Rational	$-2\gamma^3/(r_{12} + \gamma)^3$	0	$\gamma/(\gamma + r_{12})$	$-\gamma/(\gamma + r_{12})^2$
$f_{12}^{(4)} = \ln(1 + \gamma r_{12})/\gamma$	Logarithm	$-2/(1 + \gamma r_{12})^2$	$-\gamma/(1 + \gamma r_{12})^2$	$1/(1 + \gamma r_{12})$	0
$f_{12}^{(5)} = \arctan(\gamma r_{12})/\gamma$	Arctangent	$-2/(1 + \gamma^2 r_{12}^2)$	0	$1/(1 + \gamma^2 r_{12}^2)$	0
$f_{12}^{(6)} = f_{12}^{(1)}/2 + f_{12}^{(3)}/2$	Hybrid	$-e^{-\gamma r_{12}} - \gamma^3/(r_{12} + \gamma)^3$	$\gamma e^{-\gamma r_{12}}/2$	$e^{-\gamma r_{12}}/2 + \gamma/(2\gamma + 2r_{12})$	$-\gamma/(\gamma + r_{12})^2/2$

$$\begin{aligned}
F_{4e}^K(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \mathbf{r}_4) = & -\frac{7}{16} \frac{f_{24}(f_{23} - f_{13})}{r_{12}} O_{43} O_{21} O_{34} O_{12} \\
& -\frac{1}{16} \frac{f_{24}(f_{23} - f_{13})}{r_{12}} O_{41} O_{23} O_{34} O_{12} \\
& -\frac{7}{16} \frac{f_{34}(f_{14} - f_{24})}{r_{12}} O_{32} O_{44} O_{13} O_{21} \\
& -\frac{1}{16} \frac{f_{34}(f_{14} - f_{24})}{r_{12}} O_{34} O_{42} O_{13} O_{21}, \quad (71)
\end{aligned}$$

and

$$\begin{aligned}
F_{5e}^K(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \mathbf{r}_4, \mathbf{r}_5) \\
= & \frac{7}{16} \frac{f_{35}(f_{14} - f_{24})}{r_{12}} O_{32} O_{54} O_{31} O_{45} O_{21} \\
& + \frac{1}{16} \frac{f_{35}(f_{14} - f_{24})}{r_{12}} O_{34} O_{52} O_{31} O_{45} O_{21} \\
& - \frac{7}{16} \frac{f_{35}(f_{14} - f_{24})}{r_{12}} O_{32} O_{54} V_{31} V_{45} O_{21} \\
& - \frac{1}{16} \frac{f_{35}(f_{14} - f_{24})}{r_{12}} O_{34} O_{52} V_{31} V_{45} O_{21}. \quad (72)
\end{aligned}$$

See Appendix A for the derivation of Eq. (70) as an example.

III. MONTE CARLO ALGORITHM

Each of the contributions to the F12 correction is evaluated by the Metropolis MC method with the redundant-walker convergence-acceleration technique.⁴⁴ Identifying and using an appropriate weight function are essential for the viability (let alone efficiency) of any MC integration. A weight function should be positive everywhere, be analytically integrable, have the same singularity as the integrand, and generally behave like the integrand.

A. E_{F12}^V

The V term [Eq. (35)] is evaluated⁴¹ as

$$E_{2e}^V = \iint d\mathbf{r}_1 d\mathbf{r}_2 F_{2e}^V(\mathbf{r}_1, \mathbf{r}_2) \approx \frac{1}{N} \sum_{n=1}^N \frac{F_{2e}^V(\mathbf{r}_1^{[n]}, \mathbf{r}_2^{[n]})}{w_{2e}(\mathbf{r}_1^{[n]}, \mathbf{r}_2^{[n]})}, \quad (73)$$

$$\begin{aligned}
E_{3e}^V &= \iiint d\mathbf{r}_1 d\mathbf{r}_2 d\mathbf{r}_3 F_{3e}^V(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3) \\
&\approx \frac{1}{N} \sum_{n=1}^N \frac{1}{m} \sum_{k=1}^m \frac{F_{3e}^V(\mathbf{r}_1^{[n]}, \mathbf{r}_2^{[n]}, \mathbf{r}_{3k}^{[n]})}{w_{2e}(\mathbf{r}_1^{[n]}, \mathbf{r}_2^{[n]}) w_{1e}(\mathbf{r}_{3k}^{[n]})}, \quad (74)
\end{aligned}$$

and

$$\begin{aligned}
E_{4e}^V &= \iiint d\mathbf{r}_1 d\mathbf{r}_2 d\mathbf{r}_3 d\mathbf{r}_4 F_{4e}^V(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \mathbf{r}_4) \\
&\approx \frac{1}{N} \sum_{n=1}^N \frac{2!}{m(m-1)} \\
&\quad \times \sum_{k=1}^{m-1} \sum_{l=k+1}^m \frac{F_{4e}^V(\mathbf{r}_1^{[n]}, \mathbf{r}_2^{[n]}, \mathbf{r}_{3k}^{[n]}, \mathbf{r}_{4l}^{[n]})}{w_{2e}(\mathbf{r}_1^{[n]}, \mathbf{r}_2^{[n]}) w_{1e}(\mathbf{r}_{3k}^{[n]}) w_{1e}(\mathbf{r}_{4l}^{[n]})}, \quad (75)
\end{aligned}$$

where N is the total number of MC steps and m is the number of “redundant walkers” (see below for the definition).

Electron pairs with coordinates $\{\mathbf{r}_1^{[n]}, \mathbf{r}_2^{[n]} | 1 \leq n \leq N\}$ are distributed randomly but according to the normalized

two-electron weight function of the form,

$$w_{2e}(\mathbf{r}_1, \mathbf{r}_2) = \frac{1}{N_{2e}} \frac{g(\mathbf{r}_1)g(\mathbf{r}_2)}{r_{12}}, \quad (76)$$

where $g(\mathbf{r})$, in our implementation, is chosen to be a sum of two atom-centered s -type GTOs per atom,

$$g(\mathbf{r}) = \sum_{A=1}^{n_{\text{atom}}} \left(c_A^{(1)} e^{-\zeta_A^{(1)} |\mathbf{r}-\mathbf{r}_A|^2} + c_A^{(2)} e^{-\zeta_A^{(2)} |\mathbf{r}-\mathbf{r}_A|^2} \right). \quad (77)$$

Here, n_{atom} is the number of atoms and \mathbf{r}_A is the position of the A th atom. The normalization coefficient, N_{2e} , can be evaluated analytically.⁷⁶ Although the r_{12}^{-1} singularity in F_{2e}^V [Eq. (42)] is analytically removed by f_{12} in its numerator, we elect to use the above weight function that makes the MC algorithm sample more heavily at short r_{12} distances. This is appropriate or even necessary because F_{2e}^V is expected to vary more rapidly at short r_{12} distances, even after the singularity is removed.⁷⁷⁻⁷⁹ In F_{3e}^V [Eq. (43)] and F_{4e}^V [Eq. (44)], the r_{12}^{-1} singularity remains and the use of the weight function that cancels it is essential. A random distribution according to w_{2e} is achieved by the Metropolis algorithm.^{44,80}

A distribution of one-electron coordinates $\{\mathbf{r}_{3k}^{[n]} | 1 \leq n \leq N\}$ is generated randomly but according to the normalized one-electron weight function,

$$w_{1e}(\mathbf{r}_3) = \frac{g(\mathbf{r}_3)}{N_{1e}}, \quad (78)$$

where N_{1e} is again analytically determined.⁷⁶ Electron 3 or 4 is not strongly coupled with the others in integrand F_{3e}^V or F_{4e}^V , and, therefore, one-electron “walkers” whose distributions resemble the molecule’s electron density are appropriate in these dimensions. In the redundant-walker algorithm,^{41,44} m such independent distributions (redundant walkers) are generated ($1 \leq k \leq m$), so that each of these m one-electron walkers can be used in Eq. (74) and $m(m-1)/2$ distinct pairs of one-electron walkers in Eq. (75). Since generating m distributions increases the cost by no more than a factor of m , but it increases the number of distinct summands (samples) by $O(m^2)$, this algorithm increases the sampling efficiency of Eq. (75) by $O(m)$. However, Eq. (73) is unaffected by the algorithm. Therefore, the overall performance boost by the redundant-walker algorithm in MC-MP2-F12 is hard to predict. The V contribution implemented in this manner was reported in Ref. 41.

B. E_{F12}^{BX}

The sum of the B and X terms [Eq. (47)] consists of two-, three-, and four-electron kinetic-energy contributions as well as three-, four-, and five-electron exchange contributions, which are as high as 15-dimensional. They are evaluated as

$$\begin{aligned}
E_{2e}^T &\approx \frac{1}{N} \sum_{n=1}^N \frac{F_{2e}^{T1}(\mathbf{r}_1^{[n]}, \mathbf{r}_2^{[n]})}{w_{2e}(\mathbf{r}_1^{[n]}, \mathbf{r}_2^{[n]})} + \frac{1}{N} \sum_{n=1}^N \frac{2!}{m(m-1)} \\
&\quad \times \sum_{k=1}^{m-1} \sum_{l=k+1}^m \frac{F_{2e}^{T2}(\mathbf{r}_{1k}^{[n]}, \mathbf{r}_{2l}^{[n]})}{w_{1e}(\mathbf{r}_{1k}^{[n]}) w_{1e}(\mathbf{r}_{2l}^{[n]})}, \quad (79)
\end{aligned}$$

$$E_{3e}^T \approx \frac{1}{N} \sum_{n=1}^N \frac{1}{m} \sum_{k=1}^m \frac{F_{3e}^{T1}(\mathbf{r}_1^{[n]}, \mathbf{r}_2^{[n]}, \mathbf{r}_{3k}^{[n]})}{w_{2e}(\mathbf{r}_1^{[n]}, \mathbf{r}_2^{[n]}) w_{1e}(\mathbf{r}_{3k}^{[n]})} + \frac{1}{N} \sum_{n=1}^N \frac{3!}{m(m-1)(m-2)} \times \sum_{k=1}^{m-2} \sum_{l=k+1}^{m-1} \sum_{h=l+1}^m \frac{F_{3e}^{T2}(\mathbf{r}_{1k}^{[n]}, \mathbf{r}_{2l}^{[n]}, \mathbf{r}_{3h}^{[n]})}{w_{1e}(\mathbf{r}_{1k}^{[n]}) w_{1e}(\mathbf{r}_{2l}^{[n]}) w_{1e}(\mathbf{r}_{3h}^{[n]})}, \quad (80)$$

and

$$E_{4e}^T \approx \frac{1}{N} \sum_{n=1}^N \frac{2!}{m(m-1)} \times \sum_{k=1}^{m-1} \sum_{l=k+1}^m \frac{F_{4e}^{T1}(\mathbf{r}_1^{[n]}, \mathbf{r}_2^{[n]}, \mathbf{r}_{3k}^{[n]}, \mathbf{r}_{4l}^{[n]})}{w_{2e}(\mathbf{r}_1^{[n]}, \mathbf{r}_2^{[n]}) w_{1e}(\mathbf{r}_{3k}^{[n]}) w_{1e}(\mathbf{r}_{4l}^{[n]})} + \frac{1}{N} \sum_{n=1}^N \frac{4!}{m(m-1)(m-2)(m-3)} \sum_{k=1}^{m-3} \sum_{l=k+1}^{m-2} \sum_{h=l+1}^{m-1} \sum_{i=h+1}^m \frac{F_{4e}^{T2}(\mathbf{r}_{1k}^{[n]}, \mathbf{r}_{2l}^{[n]}, \mathbf{r}_{3h}^{[n]}, \mathbf{r}_{4i}^{[n]})}{w_{1e}(\mathbf{r}_{1k}^{[n]}) w_{1e}(\mathbf{r}_{2l}^{[n]}) w_{1e}(\mathbf{r}_{3h}^{[n]}) w_{1e}(\mathbf{r}_{4i}^{[n]})}, \quad (81)$$

where w_{2e} is the weight function for an electron-pair walker for strongly coupled integration variables (\mathbf{r}_1 and \mathbf{r}_2), while w_{1e} is the weight function for m independent one-electron walkers for relatively uncoupled variables. These weight functions are identical to Eqs. (76) and (78), respectively, and hence the same electron-pair and one-electron walkers for the V term can be reused for the BX term (as well as for MC-MP2). The sampling efficiency of the first term of Eq. (79) is unchanged by the redundant-walker algorithm and that of the second term is increased by a factor of $O(m)$. The m -dependence of the number of samplings in individual terms in Eqs. (80) and (81) can be inferred similarly, but that of the overall performance is again hard to predict, but is expected to be small.

The exchange contributions are evaluated as

$$E_{3e}^K \approx \frac{1}{N} \sum_{n=1}^N \frac{1}{m} \sum_{k=1}^m \frac{F_{3e}^K(\mathbf{r}_1^{[n]}, \mathbf{r}_2^{[n]}, \mathbf{r}_{3k}^{[n]})}{w_{2e}(\mathbf{r}_1^{[n]}, \mathbf{r}_2^{[n]}) w_{1e}(\mathbf{r}_{3k}^{[n]})}, \quad (82)$$

$$E_{4e}^K \approx \frac{1}{N} \sum_{n=1}^N \frac{2!}{m(m-1)} \times \sum_{k=1}^{m-1} \sum_{l=k+1}^m \frac{F_{4e}^K(\mathbf{r}_1^{[n]}, \mathbf{r}_2^{[n]}, \mathbf{r}_{3k}^{[n]}, \mathbf{r}_{4l}^{[n]})}{w_{2e}(\mathbf{r}_1^{[n]}, \mathbf{r}_2^{[n]}) w_{1e}(\mathbf{r}_{3k}^{[n]}) w_{1e}(\mathbf{r}_{4l}^{[n]})}, \quad (83)$$

and

$$E_{5e}^K \approx \frac{1}{N} \sum_{n=1}^N \frac{3!}{m(m-1)(m-2)} \sum_{k=1}^{m-2} \sum_{l=k+1}^{m-1} \sum_{h=l+1}^m \frac{F_{5e}^K(\mathbf{r}_1^{[n]}, \mathbf{r}_2^{[n]}, \mathbf{r}_{3k}^{[n]}, \mathbf{r}_{4l}^{[n]}, \mathbf{r}_{5h}^{[n]})}{w_{2e}(\mathbf{r}_1^{[n]}, \mathbf{r}_2^{[n]}) w_{1e}(\mathbf{r}_{3k}^{[n]}) w_{1e}(\mathbf{r}_{4l}^{[n]}) w_{1e}(\mathbf{r}_{5h}^{[n]})}, \quad (84)$$

again reusing the same electron-pair walkers and m independent one-electron walkers. The use of w_{2e} for variables \mathbf{r}_1 and \mathbf{r}_2 and of w_{1e} for the other variables can be rationalized by the structure (occurrence of singularity) in the integrands [Eqs. (70)–(72)].

C. Statistical uncertainty

The MC integrals for both the nonvariational (V) [Eq. (13)] and variational (VBX) [Eq. (32)] formulas can be written in a unified form as

$$E_{F12} = \lim_{N \rightarrow \infty} I_N, \quad (85)$$

$$I_N = \frac{1}{N} \sum_{n=1}^N I^{[n]}, \quad (86)$$

where $I^{[n]}$ collects all summands of the outermost summations in Eqs. (73)–(75), (79)–(81), and/or (82)–(84). The statistical uncertainty σ_N in I_N at the N th MC step can be estimated as

$$\sigma_N^2 = \frac{1}{N^2} \sum_{n=1}^N \{I^{[n]} - I_N\}^2. \quad (87)$$

However, this estimate is well known⁸¹ to be an underestimation of the true statistical uncertainty because the coordinates of a walker are correlated across several MC steps. A more accurate estimate is obtained by the blocking algorithm of Flyvbjerg and Petersen⁸¹ as

$$\sigma_N^2 = \frac{N_b^2}{N^2} \sum_{n=1}^{N/N_b} \left\{ \frac{1}{N_b} \sum_{n'=1}^{N_b} I^{[(n-1)N_b+n']} - I_N \right\}^2, \quad (88)$$

where N_b is the block size, which is to be gradually enlarged until σ_N plateaus. The relative error at the N th MC step is, therefore,

$$\sigma_{\text{rel}} = \frac{\sigma_N}{I_N}. \quad (89)$$

The error in MC-MP2-F12 (relative to deterministic MP2-F12) seems to be the statistical uncertainty only, and there is no bias (i.e., systematic error).⁸² See Sec. IV C for more details on this issue.

IV. RESULTS AND DISCUSSION

A massively parallel MC-MP2-F12 program was implemented using the redundant-walker algorithm. Both nonvariational V formula⁴¹ [Eq. (13)] and variational VBX formula [Eq. (32)] were used, and the notational distinction is made in this article in the parentheses following the method label as in MC-MP2-F12(V) and MC-MP2-F12(VBX).

The number of redundant walkers (m in various equations in Sec. III) was 40 and the block size [N_b in Eq. (88)] was 6. The frozen-core approximation was used in all cases. A F12 correction thus obtained was added to the corresponding MP2 energy obtained by the conventional deterministic algorithm in NWCHEM.⁸³ This is merely to isolate the performance characteristics of MC-MP2-F12 from those of MC-MP2 because the result of the latter becomes immediately available from the former with no extra cost.

The CBS limits of the MP2 energy were extrapolated applying the X^{-3} formula⁸⁴ to the deterministic values obtained with the aug-cc-pVQZ and aug-cc-pV5Z basis sets. Hereafter, we abbreviate the cc-pVXZ basis set as “XZ” and the aug-cc-pVXZ basis set as “AXZ,” where X = D, T, Q, or 5.

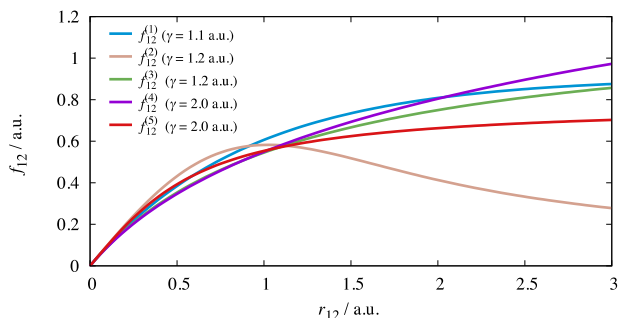


FIG. 1. Five of the 6 geminals considered in this work plotted for the near-optimal value of γ in each case. Geminal 6 (the hybrid geminal) curve is an average of the curves of geminals 1 and 3.

A. Geminals and V versus VBX formulas

The MP2 correlation energies for H_2O and CH_4 at the geometries of Bak *et al.*⁸⁵ were computed with the MC-MP2-F12 method with the DZ and TZ basis sets. Six geminals⁷⁵ listed in Table I were used (whose forms near $r_{12} = 0$ are visualized in Fig. 1). Recall that any differentiable function can be employed as a geminal in MC-MP2-F12(VBX) or any function (even a numerically defined one) in MC-MP2-F12(V). This is one of the unique advantages of MC-MP2-F12, shared by QMC, the latter using the Jastrow factor routinely,^{86–89} some even correlating two electrons and a

nucleus.⁹⁰ See also the work of Nooijen and Bartlett,⁷⁵ who proposed many geminals including those used here, Monkhorst⁹¹ for a rational geminal whose form is dictated by its equation of motion in the first order and asymptotic regions, Grüneis *et al.*,⁹² who used a “Yukawa–Coulomb” geminal with their planewave MP2-F12, and Silkowski *et al.*⁹³ for a range-separated geminal.

Figure 2 shows the γ -dependence of the MC-MP2-F12 energy of H_2O . The nonvariational V data are plotted as green curves, while the variational VBX ones as purple curves. The uncorrected MP2 energies with the DZ and TZ basis sets as well as the CBS limit are indicated as red lines. Figure 3 plots the same for CH_4 .

First, focusing on the bounded VBX curves (purple), we find that all geminals work reasonably well with the exception of geminal 2; the curves from all geminals except 2 are not strongly dependent on γ and close to the CBS limit in a wide range of γ . Geminal 1 is the Slater-type geminal introduced by Ten-no^{16,38} and is widely regarded as one of the best-performing geminals. Geminal 2 is the Gauss-type geminal, which should be distinguished from the Gaussian geminal that does not satisfy the cusp conditions.¹⁶ Its performance is rather poor as its MC-MP2-F12(VBX)/DZ energy falls short of the MP2/TZ energy at any value of γ .

Comparing these performance data with Fig. 1, we notice that only geminal 2 has a rather different long-range behavior than the rest. Unlike the other geminals, which rise

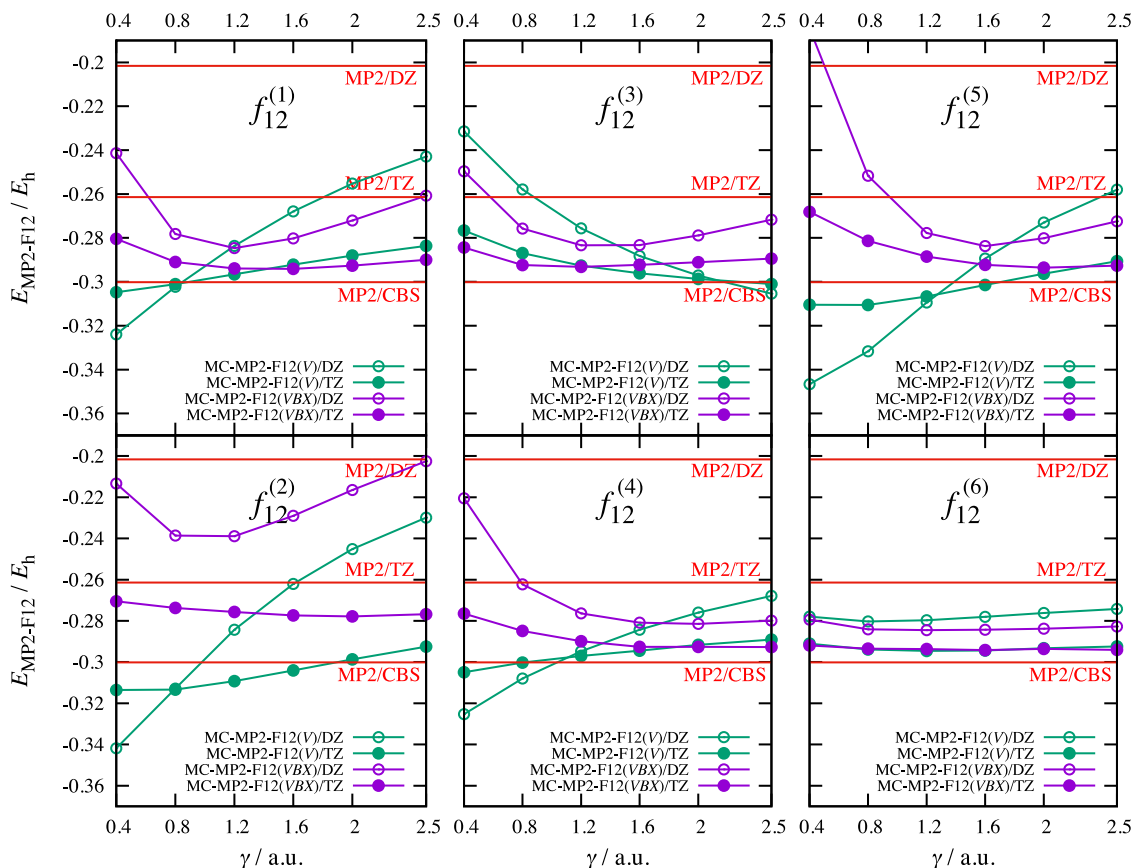


FIG. 2. The MP2 correlation energies of H_2O as a function of γ of the 6 geminals listed in Table I. The geometry was taken from Ref. 85 and the CBS limit was obtained by the X^{-3} extrapolation using the AQZ and ASZ basis sets. The number of MC steps was 8.65×10^7 .

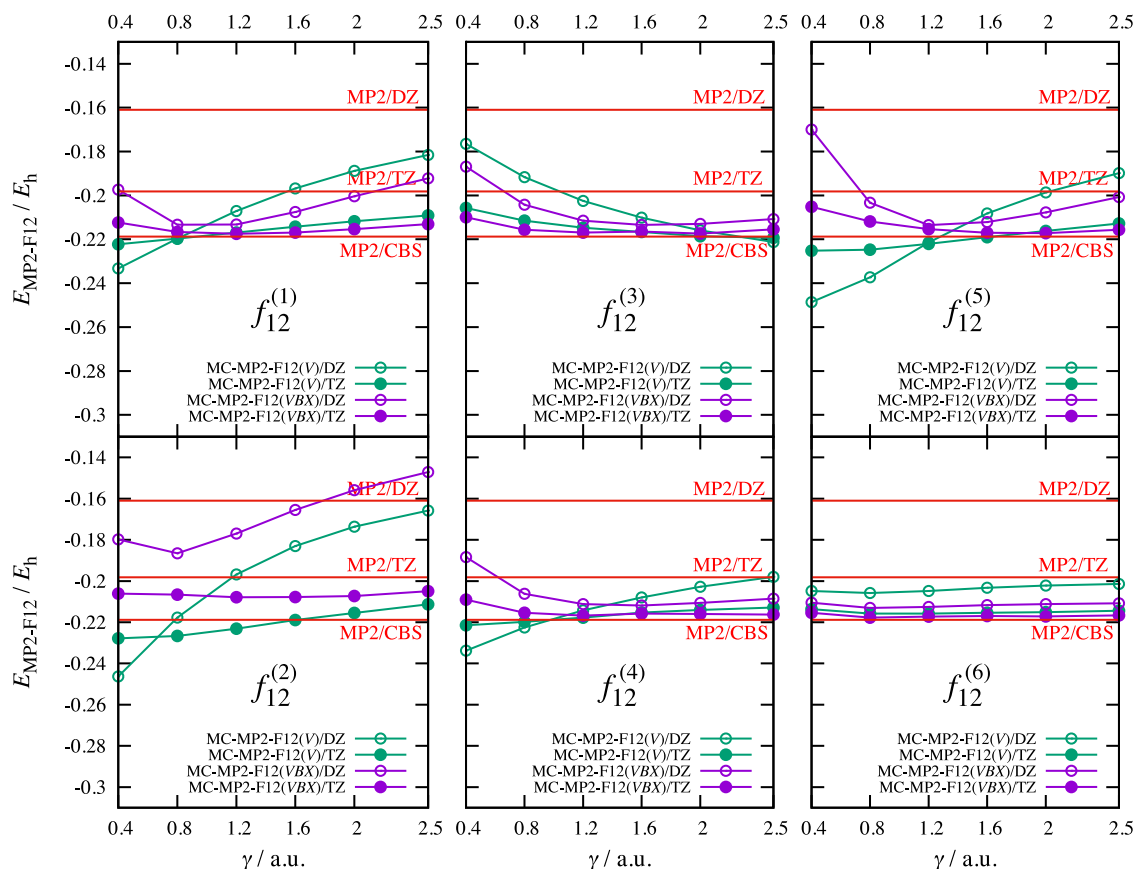


FIG. 3. The MP2 correlation energies of CH₄ as a function of γ of the 6 geminals listed in Table I. The geometry was taken from Ref. 85 and the CBS limit was obtained by the X^{-3} extrapolation using the AQZ and A5Z basis sets. The number of MC steps was 8.65×10^7 .

monotonically, geminal 2 increases and then turns to decrease with r_{12} , meaning that its correlation hole may be $2s$ -orbital-like, when the true correlation hole is $1s$ -orbital-like.⁷⁹ It has been shown^{2,18} that the r_{12}^2 component in the Taylor-series expansion of a geminal plays an important role in describing the correlation hole correctly, but among the 6 geminals studied here, only geminals 2 and 5 lack this component (the VBX curves of geminal 5 are equally unstable with γ as geminal 2, but as will be shown below the relative energies from geminal 5 seem reasonable). These may explain the poor performance of geminal 2.

Next, turning our attention to the nonvariational V curves (green), we observe that they strongly depend on γ and are usually unbounded. For instance, using geminal 5 with $\gamma < 1.2$ a.u., the F12 corrections overshoot the correct values in both H₂O and CH₄, rendering the MC-MP2-F12(V)/DZ energies more negative than the MP2/CBS values, a clear violation of Hylleraas' variational principle. This may or may not become troublesome in chemical applications of MP2-F12, wherein relative energies are often sought; if the errors from the CBS limits have opposite signs let alone greatly different magnitudes across molecules, the errors in relative energies can be amplified, possibly undoing improvements made by F12 in the total correlation energies. In Sec. IV C, we will examine this numerically.

As will be shown in Sec. IV D, the nonvariational V formula has the advantage over the variational VBX formula

of being 2–10 times faster. Can one then use the V formula and reliably determine the CBS limits? We have conceived of two ways to do this. The first is to combine two geminals with opposite γ -dependences, such that the resulting geminal is nearly independent of γ . Geminal 6 is indeed constructed in this way as an average of geminals 1 and 3, whose V curves have opposite γ -dependences. Consequently, the MC-MP2-F12(V) energies of geminal 6 are stable with variation in γ and also much closer to the CBS limits than those from geminal 1 or 3 alone at any γ . We will examine if this construction improves total correlation energies for a wider array of molecules in Sec. IV B and relative energies in Sec. IV C.

The second method attempts to locate the value of γ at which the nonvariational V formula yields the CBS limit. It does so by combining the results of two basis sets. At a given basis set, the VBX energy is closest to the CBS limit at γ_{\min} that minimizes it,

$$E_{\text{MP2/CBS}} \approx E_{\text{MP2/XZ}} + 2E_{\text{F12}}^V(\gamma_{\min})/\text{XZ} + E_{\text{F12}}^{\text{BX}}(\gamma_{\min})/\text{XZ}. \quad (90)$$

Note the approximate equality in the above expression, which is because the left-hand side is a lower bound of the right-hand side. At the minimum, however, the right-hand side should be nearly equal to the V energy as per Eq. (13) insofar as Ten-no's fixed amplitudes³⁹ approximate well the true

variationally optimized geminal amplitudes,

$$2E_{F12}^V(\gamma_{\min})/XZ + E_{F12}^{BX}(\gamma_{\min})/XZ = E_{F12}^V(\gamma_{\min})/XZ. \quad (91)$$

This seems, in fact, borne out in most of the plots of Figs. 2 and 3; the intersection of the VBX and V curves of the same basis set occurs at the minimum of the former (except for the curves of geminal 2). This once again attests to the excellent transferability of Ten-no's fixed amplitudes across all geminals that have a physically reasonable functional form.

If we now assume the equality in Eq. (90) (instead of the approximate equality), we have

$$E_{MP2}/CBS = E_{MP2}/DZ + E_{F12}^V(\gamma_{\min})/DZ \quad (92)$$

$$= E_{MP2}/TZ + E_{F12}^V(\gamma_{\min})/TZ, \quad (93)$$

suggesting that γ_{\min} is at the intersection of the V curves of two different basis sets, at which the energy is the CBS limit. Applying this method to the linearly interpolated V curves of geminal 1, 3, and 4, we recover, respectively, 100.2%, 99.7%, and 99.2% of the CBS limit of H_2O . They are closer to 100% than 94.8% or 98.0% recuperated, respectively, by the VBX method with the DZ and TZ basis sets. Likewise, this intersection method recovers 100.4%, 100.1%, and 100.1% of the CBS limit for CH_4 using geminals 1, 3, and 4, respectively. Notice how closely the two green curves (of geminals 1, 3, and 4) and the MP2/CBS line (red) meet at one and the same point for both H_2O and CH_4 . Although it is at present unclear whether the efficacy of the intersection method occurs universally across many molecules, we have a potentially practical way of capturing nearly 100% of the

CBS limits inexpensively using only the nonvariational V formula.

B. Correlation energies

The F12 corrections and statistical uncertainties were calculated for 17 molecules at the geometries of Bak *et al.*⁸⁵ by MC-MP2-F12 with the ADZ and ATZ basis sets using the V or VBX formula. The MP2 energies were also computed for the same molecules with the AXZ basis set ($X = D, T, Q,$ and 5) using NWCHEM.⁸³

Table II compiles the MP2 correlation energies in the increasing order of their magnitude. The MC-MP2-F12 calculations used the Slater-type geminal with $\gamma = 1.1$ a.u. With the ADZ basis set, the F12 correction brings an average error from the CBS limit from 107.1 mE_h to 6.6 mE_h (with the V formula) or 6.9 mE_h (VBX). The latter two errors are smaller than the average error of 9.9 mE_h in MP2/aug-cc-pV5Z. With the ATZ basis set, MP2 has an average error of 42.5 mE_h , whereas MC-MP2-F12(V) has only 3.0 mE_h and MC-MP2-F12(VBX) 3.5 mE_h . The statistical uncertainties are on the order of a few tenths of 1 mE_h with the largest being 1.2 mE_h after 1.44×10^8 MC steps using 40 redundant walkers. They can furthermore be arbitrarily (albeit slowly) compressed by running a longer MC run.

Figure 4 plots the proportion of the CBS limits recovered by MP2 and MC-MP2-F12(VBX). MP2 recovers on average 74.6%, 90.2%, 95.6%, and 97.7% of the CBS limits with the ADZ, ATZ, AQZ, and A5Z basis sets, respectively. MC-MP2-F12(VBX)/ADZ captures 98.5% of the CBS limits, which

TABLE II. The MP2 correlation energies in E_h . The values in parentheses are statistical uncertainties.

Molecule ^a	MP2					MC-MP2-F12(V) ^b		MC-MP2-F12(VBX) ^b	
	ADZ	ATZ	AQZ	A5Z	CBS ^c	ADZ	ATZ	ADZ	ATZ
H_2	-0.0273	-0.0320	-0.0333	-0.0337	-0.0343	-0.0336(0)	-0.0341(0)	-0.0342(0)	-0.0342(0)
CH_2	-0.1154	-0.1410	-0.1493	-0.1524	-0.1557	-0.1508(1)	-0.1554(1)	-0.1535(1)	-0.1547(2)
CH_4	-0.1677	-0.2008	-0.2111	-0.2149	-0.2188	-0.2134(1)	-0.2189(2)	-0.2165(1)	-0.2186(4)
NH_3	-0.1992	-0.2401	-0.2537	-0.2588	-0.2643	-0.2601(1)	-0.2650(2)	-0.2603(2)	-0.2632(3)
H_2O	-0.2193	-0.2683	-0.2859	-0.2929	-0.3002	-0.3004(1)	-0.3023(1)	-0.2954(2)	-0.2979(3)
C_2H_2	-0.2223	-0.2798	-0.3012	-0.3100	-0.3192	-0.3263(1)	-0.3231(1)	-0.3129(2)	-0.3148(2)
C_2H_4	-0.2628	-0.3137	-0.3311	-0.3376	-0.3445	-0.3337(1)	-0.3444(2)	-0.3389(3)	-0.3438(5)
HF	-0.2847	-0.3399	-0.3583	-0.3651	-0.3722	-0.3622(2)	-0.3714(5)	-0.3673(4)	-0.3700(9)
HNC	-0.2824	-0.3377	-0.3573	-0.3650	-0.3731	-0.3660(3)	-0.3739(3)	-0.3677(7)	-0.3714(5)
HCN	-0.2943	-0.3502	-0.3699	-0.3777	-0.3859	-0.3774(2)	-0.3866(2)	-0.3789(4)	-0.3837(4)
CO	-0.2992	-0.3607	-0.3837	-0.3932	-0.4031	-0.4012(3)	-0.4056(2)	-0.3970(5)	-0.4001(5)
N_2	-0.3173	-0.3796	-0.4018	-0.4109	-0.4203	-0.4147(2)	-0.4218(2)	-0.4127(5)	-0.4170(4)
CH_2O	-0.3327	-0.4021	-0.4271	-0.4370	-0.4474	-0.4439(3)	-0.4492(4)	-0.4405(5)	-0.4436(7)
HNO	-0.3686	-0.4455	-0.4729	-0.4841	-0.4957	-0.4934(3)	-0.4987(3)	-0.4873(5)	-0.4916(7)
H_2O_2	-0.4154	-0.5079	-0.5411	-0.5545	-0.5685	-0.5700(4)	-0.5727(5)	-0.5601(7)	-0.5637(10)
F_2	-0.4280	-0.5361	-0.5758	-0.5927	-0.6105	-0.6253(3)	-0.6193(5)	-0.5978(7)	-0.6023(9)
CO_2	-0.5053	-0.6117	-0.6514	-0.6676	-0.6846	-0.6828(6)	-0.6890(6)	-0.6730(12)	-0.6791(12)
Error ^d	0.1071	0.0425	0.0193	0.0099	0.0000	0.0066(3)	0.0030(3)	0.0069(5)	0.0035(6)

^aGeometries were taken from Ref. 85.

^bOnly the F12 corrections were calculated by the MC-MP2-F12 method using either the nonvariational V formula or the variational VBX formula and the Slater-type geminal (geminal 1) with $\gamma = 1.1$ a.u. The number of MC steps was 1.44×10^8 and the number of the redundant walkers was 40.

^cThe X^{-3} extrapolation using the MP2/AQZ and MP2/A5Z correlation energies.

^dThe standard deviation from the MP2/CBS values (the standard deviation of the statistical uncertainties in parentheses).

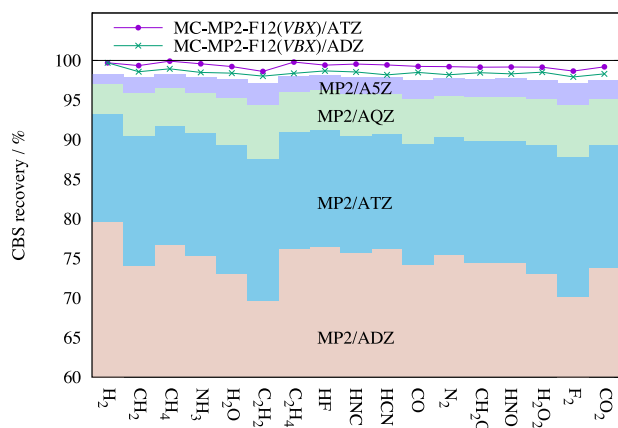


FIG. 4. The percent recovery of the CBS limits of the MP2 correlation energies by various methods. The corresponding numerical data are in Table II.

is already greater than that of MP2/A5Z. The ratio goes up further to 99.3% with MC-MP2-F12(*VBX*)/ATZ.

Therefore, MC-MP2-F12 works exceedingly well for the total correlation energies, with either the *V* or *VBX* formula. Comparing *V* and *VBX*, the former performs slightly better than the latter. This is not surprising because errors from a variational formula are always positive and cannot be accidentally small, while the results from a nonvariational formula can scatter in both higher or lower sides of the correct values and can accidentally agree more accurately with the latter. As shown in Sec. IV C, *VBX* becomes more accurate than *V* for the relative energies because systematic cancellation of positive errors occurs in the former (but not in the latter) when energy differences are taken.

In Table III, we revisit the question of the relative performance of geminals for total correlation energies. With the ADZ basis set, Ten-no's Slater-type geminal (geminal 1) performs the best by a large margin especially in the *V* formula with an average error of 6.6 mE_h . The hybrid geminal (geminal 6), designed to be insensitive to γ in the *V* formula, comes next in performance with an average error of 12.8 mE_h . Geminal 2 is particularly poorly performing. Switching from the nonvariational (*V*) to variational (*VBX*) formula, the average errors of some geminals decrease, while those of the others increase, reflecting the fact that the values of γ are not optimal for either formula. With the ATZ basis set, the hybrid geminal (geminal 6) becomes the most accurate in both the *V* and *VBX* formulas, perhaps supporting its design.

TABLE III. The average error (in mE_h) in the MC-MP2-F12 correlation energies using the geminals listed in Table I from the MP2/CBS values. The 17 molecules used were the same as those in Table II.

Basis set	$f_{12}^{(1)} (\gamma = 1.1 \text{ a.u.})^a$		$f_{12}^{(2)} (\gamma = 1.2 \text{ a.u.})^a$		$f_{12}^{(3)} (\gamma = 1.2 \text{ a.u.})^a$		$f_{12}^{(4)} (\gamma = 2.0 \text{ a.u.})^a$		$f_{12}^{(5)} (\gamma = 2.0 \text{ a.u.})^a$		$f_{12}^{(6)} (\gamma = 1.1 \text{ a.u.})^a$	
	<i>V</i>	<i>VBX</i>	<i>V</i>	<i>VBX</i>	<i>V</i>	<i>VBX</i>	<i>V</i>	<i>VBX</i>	<i>V</i>	<i>VBX</i>	<i>V</i>	<i>VBX</i>
ADZ ^b	6.6 (0.3)	6.9 (0.5)	20.2 (0.1)	63.4 (0.2)	17.6 (0.3)	8.7 (0.7)	17.6 (0.3)	12.3 (0.7)	21.8 (0.2)	12.1 (0.5)	12.8 (0.3)	7.2 (0.6)
ATZ ^b	3.0 (0.3)	3.5 (0.6)	17.0 (0.2)	21.9 (0.3)	4.2 (0.3)	4.0 (0.7)	5.2 (0.4)	4.6 (0.7)	3.0 (0.3)	3.2 (0.5)	2.1 (0.3)	3.2 (0.6)

^aOnly the F12 corrections were calculated by the MC-MP2-F12 method using either the nonvariational *V* formula or the variational *VBX* formula. The number of MC steps was 1.44×10^8 and the number of the redundant walkers was 40.

^bThe standard deviation from the MP2/CBS values (the standard deviation of the statistical uncertainties in parentheses).

TABLE IV. Gas-phase reactions considered in this work.

Reaction	Formula ^a
1	$\text{CO} + \text{H}_2 \rightarrow \text{CH}_2\text{O}$
2	$\text{N}_2 + 3\text{H}_2 \rightarrow 2\text{NH}_3$
3	$\text{C}_2\text{H}_2 + \text{H}_2 \rightarrow \text{C}_2\text{H}_4$
4	$\text{CO}_2 + 4\text{H}_2 \rightarrow \text{CH}_4 + 2\text{H}_2\text{O}$
5	$\text{CH}_2\text{O} + 2\text{H}_2 \rightarrow \text{CH}_4 + \text{H}_2\text{O}$
6	$\text{CO} + 3\text{H}_2 \rightarrow \text{CH}_4 + \text{H}_2\text{O}$
7	$\text{HCN} + 3\text{H}_2 \rightarrow \text{CH}_4 + \text{NH}_3$
8	$\text{HNO} + 2\text{H}_2 \rightarrow \text{NH}_3 + \text{H}_2\text{O}$
9	$\text{C}_2\text{H}_2 + 3\text{H}_2 \rightarrow 2\text{CH}_4$
10	$\text{CH}_2 + \text{H}_2 \rightarrow \text{CH}_4$
11	$\text{F}_2 + \text{H}_2 \rightarrow 2\text{HF}$
12	$2\text{CH}_2 \rightarrow \text{C}_2\text{H}_4$

^aGeometries were taken from Ref. 85.

C. Reaction energies

Energies of 12 gas-phase reactions listed in Table IV were calculated by MP2, MP2-F12, and MC-MP2-F12 using geminal 1 (the Slater-type geminal) with $\gamma = 1.1$ a.u. The results are compiled in Table V. They were computed as the difference in the MP2 or MP2-F12 correlation energy between reactants and products, whose geometries were taken from the work of Bak *et al.*⁸⁵ The MP2 part of the energy was obtained with the conventional deterministic algorithms in NWChem,⁸³ whereas the F12 part was computed either by MC-MP2-F12(*V*), MC-MP2-F12(*VBX*), or deterministic MP2-F12(*VBX*), the latter implemented in MPQC,^{17,74,94,95} which uses neither the GBC nor EBC.

First, focusing on the ADZ results, we confirm that the correlation contributions to the reaction energies of MP2 suffer from excessively large errors from the CBS limits, which are, on average, 12.5 kJ mol^{-1} . They are so large that even the sign is incorrectly predicted (as compared with the CBS limits, if not with the reaction energies) for reactions 2, 5, and 8. Once the F12 corrections from MC-MP2-F12 are added, be they based on the *V* or *VBX* formula, the correlation contributions to the reaction energies are, on average, within 2.6 kJ mol^{-1} of the CBS limits. The statistical uncertainties are no more than 1.7 kJ mol^{-1} after 1.44×10^8 MC steps and are comparable to the intrinsic errors in the F12 method itself, but can be arbitrarily reduced by running longer MC integrations. Therefore, MC-MP2-F12 with either the *V* or *VBX* formula

TABLE V. The MP2 correlation contributions to the reaction energies in kJ mol^{-1} . The values in parentheses are statistical uncertainties.

Reaction ^a	ADZ				ATZ			CBS ^b
	MP2	F12(<i>V BX</i>) ^c	MC-F12(<i>V</i>) ^d	MC-F12(<i>V BX</i>) ^d	MP2	MC-F12(<i>V</i>) ^d	MC-F12(<i>V BX</i>) ^d	MP2
1	-16.1	-24.6	-23.8 (1.0)	-24.2 (1.9)	-24.9	-24.9 (1.2)	-24.5 (2.3)	-26.4
2	1.8	-13.3	-12.1 (0.8)	-13.5 (1.5)	-12.2	-15.3 (1.0)	-18.2 (1.9)	-14.2
3	14.2	16.2	13.6 (0.7)	15.4 (1.3)	15.4	18.7 (1.4)	20.8 (2.6)	17.2
4	21.2	5.2	8.4 (1.7)	7.3 (3.3)	5.8	5.3 (1.9)	3.5 (3.6)	6.1
5	0.5	-8.4	-6.8 (0.8)	-7.5 (1.6)	-7.9	-9.8 (1.2)	-11.9 (2.3)	-8.4
6	-15.7	-33.1	-30.6 (0.7)	-31.7 (1.4)	-32.8	-34.7 (0.9)	-36.5 (1.7)	-34.8
7	24.1	13.7	12.6 (0.6)	12.7 (1.1)	13.9	13.1 (0.9)	11.4 (1.7)	14.5
8	12.1	-0.1	0.5 (0.8)	0.3 (1.5)	2.9	-0.9 (1.1)	-3.0 (2.0)	-0.7
9	24.1	23.9	20.4 (0.6)	22.8 (1.1)	21.3	23.3 (1.2)	23.5 (2.3)	25.2
10	-65.8	-75.4	-76.2 (0.3)	-75.4 (0.5)	-73.0	-77.2 (0.6)	-78.2 (1.1)	-76.0
11	28.3	17.4	16.4 (1.0)	16.1 (2.0)	22.2	18.9 (1.4)	18.3 (2.7)	16.7
12	-141.6	-158.4	-159.2 (0.6)	-158.3 (1.3)	-151.9	-159.1 (1.3)	-159.1 (2.5)	-160.0
Error ^e	11.9	0.0	1.9 (0.9)	1.0 (1.7)	2.7	1.4 (1.2)	2.8 (2.3)	1.1
Error ^f	12.5	1.1	2.6 (0.9)	1.7 (1.7)	3.5	1.3 (1.2)	2.6 (2.3)	0.0

^aSee Table IV.^bSee footnote c of Table II.^cThe deterministic MP2-F12 calculation using the variational *V BX* formula and the Slater-type geminal (geminal 1) with $\gamma = 1.1$ a.u. Unlike MC-MP2-F12, neither the GBC nor EBC was assumed.^dMC-MP2-F12 calculations. See footnote b of Table II.^eThe standard deviation from the deterministic MP2-F12(*V BX*)/ADZ values (the standard deviation of the statistical uncertainties in parentheses).^fThe standard deviation from the MP2/CBS values (the standard deviation of the statistical uncertainties in parentheses).

does work well in practice for relative energies, consistently achieving the chemical accuracy from the CBS limits.

We also find that MC-MP2-F12(*V BX*) gives the results that are systematically closer [than MC-MP2-F12(*V*)] to deterministic MP2-F12(*V BX*) for all reactions barring one. Therefore, we argue that the MC-MP2-F12 method does not have a bias⁸² and is convergent at the correct limit, i.e., the result of the deterministic version of the corresponding method. The statistical uncertainties are also reasonable, but somewhat overestimated as the *V BX* results are within 1σ (statistical uncertainty) from the deterministic values. The *V* results are often outside 3σ from the deterministic results, but this does not mean an underestimation of σ ; it simply means that MC-MP2-F12(*V*) and MP2-F12(*V BX*) are two different methods with slightly different correct results. Note that deterministic MP2-F12(*V BX*) used here does not invoke either the GBC or EBC, unlike MC-MP2-F12(*V BX*), which assumes both. However, we numerically confirmed (not shown) that the differences in the F12 corrections caused by these approximations are much smaller than the typical statistical uncertainties and negligible, except when γ is too small ($\gamma \leq 0.4$ a.u. in the case of the Slater-type geminal).

Furthermore, the errors from the CBS limits are distinctly smaller in MC-MP2-F12(*V BX*) than in MC-MP2-F12(*V*) with the ADZ basis set. This may well be because the intrinsic errors in the F12 method, which are always positive with the variational *V BX* formula, cancel between reactants and products in MC-MP2-F12(*V BX*). However, the statistical uncertainties are nearly twice as large in the *V BX* results as in the *V* results, making it difficult to draw a definitive conclusion. The greater statistical uncertainties in the results of the *V BX* formula are due to the larger number and higher dimension of its integrals and are expected. Nonetheless, the difference

in the overall accuracy between *V* and *V BX* is small, and given its much smaller computational cost (see Sec. IV D) and somewhat smaller statistical uncertainty, the nonvariational *V* formula may be preferred in most applications.

Turning to the results of the ATZ basis set, we find the benefit of the F12 method to diminish rather rapidly with a basis-set extension especially in the MC-MP2-F12 implementation. The average error from the CBS limits is compressed significantly from 12.5 to 3.5 kJ mol^{-1} going from MP2/ADZ to MP2/ATZ. In the meantime, the same quantity of MC-MP2-F12(*V*) decreases from 2.6 to 1.3 kJ mol^{-1} , while the average statistical uncertainty increases from 0.9 to 1.2 kJ mol^{-1} . With MC-MP-F12(*V BX*), the average error from the CBS limits *increases* from 1.7 kJ mol^{-1} (ADZ) to 2.6 kJ mol^{-1} (ATZ), probably owing to the increased statistical uncertainty from 1.7 to 2.3 kJ mol^{-1} . However, we do not consider this to be particularly troubling because it is with a small basis set that the F12 method is most needed. That MC-MP2-F12 is relatively more effective and efficient (in the sense of giving smaller statistical uncertainties) with ADZ than with ATZ which should indeed be considered as a practical advantage.

Table VI summarizes the geminal-dependence of the reaction energies calculated by MC-MP2-F12/ADZ with the *V* or *V BX* formula. Figure 5 plots the same using the *V BX* data and the values from the deterministic counterparts. The statistical uncertainties are comparable across different geminals, but the average errors from the CBS limits vary more greatly and are a better gauge of the geminals' performance.

The Slater-type geminal (geminal 1) is confirmed to be the one with the smallest average errors, either with the *V* or *V BX* formula, and is, therefore, judged to be the most suitable for the fixed-amplitude MP2-F12 method (at least among the

TABLE VI. The correlation contributions to the reaction energies in kJ mol^{-1} calculated by MC-MP2-F12/ADZ using the geminals listed in Table I.

Reaction ^a	$f_{12}^{(1)} (\gamma = 1.1 \text{ a.u.})^b$		$f_{12}^{(2)} (\gamma = 1.2 \text{ a.u.})^b$		$f_{12}^{(3)} (\gamma = 1.2 \text{ a.u.})^b$		$f_{12}^{(4)} (\gamma = 2.0 \text{ a.u.})^b$		$f_{12}^{(5)} (\gamma = 2.0 \text{ a.u.})^b$		$f_{12}^{(6)} (\gamma = 1.1 \text{ a.u.})^b$		CBS ^c
	V	VBX	V	VBX	V	VBX	V	VBX	V	VBX	V	VBX	
1	-23.8	-24.2	-25.0	-22.1	-23.6	-26.7	-22.8	-26.2	-23.3	-25.6	-23.7	-25.7	-26.4
2	-12.1	-13.5	-23.4	-19.8	-9.6	-15.1	-6.5	-12.7	-11.7	-13.6	-9.4	-11.7	-14.2
3	13.6	15.4	4.9	6.8	15.0	16.6	16.5	18.5	12.0	13.2	13.5	14.4	17.2
4	8.4	7.3	2.6	-0.1	13.7	14.5	17.4	15.3	9.7	7.4	11.8	10.7	6.1
5	-6.8	-7.5	-11.1	-10.9	-3.5	-3.7	-4.9	-9.6	-5.2	-5.2	-5.4	-6.0	-8.4
6	-30.6	-31.7	-36.1	-33.0	-27.1	-30.3	-27.7	-35.8	-28.5	-30.8	-29.1	-31.7	-34.8
7	12.6	12.7	3.0	8.1	16.8	16.0	17.8	14.7	13.8	13.9	14.3	13.1	14.5
8	0.5	0.3	-13.8	-13.2	4.6	3.7	5.6	0.9	1.2	2.1	1.6	-0.2	-0.7
9	20.4	22.8	5.4	7.8	22.4	23.6	23.1	20.6	18.6	20.2	21.6	23.3	25.2
10	-76.2	-75.4	-81.3	-75.6	-74.3	-75.7	-73.1	-74.9	-75.2	-76.1	-75.2	-76.0	-76.0
11	16.4	16.1	-3.8	-1.0	19.9	17.9	19.8	12.4	14.1	15.0	18.1	17.4	16.7
12	-159.2	-158.3	-163.1	-152.1	-156.1	-158.4	-152.8	-152.0	-156.9	-159.3	-158.5	-160.9	-160.0
Error ^d	2.6	1.7	10.9	9.5	4.5	3.4	5.7	4.1	3.8	2.6	3.4	2.2	0.0

^aSee Table IV.^bSee footnote b of Table II.^cSee footnote c of Table II.^dThe standard deviation from the MP2/CBS values.

six studied). This may or may not be rationalized by the limiting behavior of the solution of the two-electron equation of motion pointed out by Monkhorst.⁹¹ However, insofar as the short-range r_{12} -dependence is similar to the Slater-type geminal (see Fig. 1), other geminals (barring geminal 2) work almost as effectively. For instance, geminal 3 (the rational geminal) has the average errors that are roughly twice those of geminal 1 (the Slater-type geminal). Geminal 6 (the hybrid geminal) is an average of geminals 1 and 3 and its errors also come in between those of its parent geminals. This geminal may be said to have the advantage that its F12 corrections are nearly independent of γ (see Figs. 2 and 3). If an optimal value of γ is unknown, one may use either the VBX formula or geminal 6.

D. System-size scaling

Here, we determine the asymptotic functional dependence (scaling) of the cost of MC-MP2-F12 on the number of basis functions (n), i.e., the system's spatial size.

Figure 6 plots the cost scaling per MC step. To understand the observation, we briefly review the algorithm first. In

each MC step, an electron-pair walker and m one-electron walkers need to be propagated by the Metropolis algorithm. The cost scaling of this step is $O(mn)$ because a processor must evaluate $O(n)$ weight functions at $O(m)$ coordinates. Next, n AO amplitudes at m accepted walker coordinates are computed at an $O(mn)$ cost. Then, these AO amplitudes are transformed into MO amplitudes at an $O(mn^2)$ cost because at each of $O(m)$ walker coordinates, $O(n^2)$ multiplications of MO coefficients and AO amplitudes occur in this step. Thereupon, O_{pq} [Eq. (45)], V_{pq} [Eq. (46)], and O'_{pq} [Eq. (63)] arrays are constructed at an $O(m^2n)$ cost, for each array is a sum over $O(n)$ orbitals evaluated at two walker coordinates, whose number grows as $O(m^2)$. Once these arrays are constructed, a processor accumulates the F12 correction. The cost of this step is cheap and independent of n (but is heavily dependent on m).

Figure 6 shows that the wall time required for 128 MC steps of the MC-MP2-F12 calculations using the V formula falls accurately on the n^2 line in the large n limit. This is because the overall cost per MC step is dominated by the AO-to-MO transformation of orbital amplitudes, whose cost is an $O(mn^2)$ quantity, when the number of redundant walkers

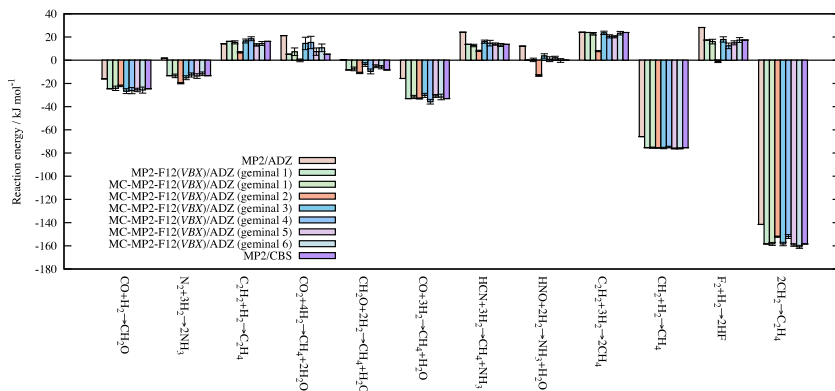


FIG. 5. The correlation contributions to the reaction energies obtained with the MP2, MP2-F12(VBX), and MC-MP2-F12(VBX) methods using the geminals listed in Table I. The basis set is ADZ. The corresponding numerical data are found in Table VI. The error bars are the statistical uncertainties.

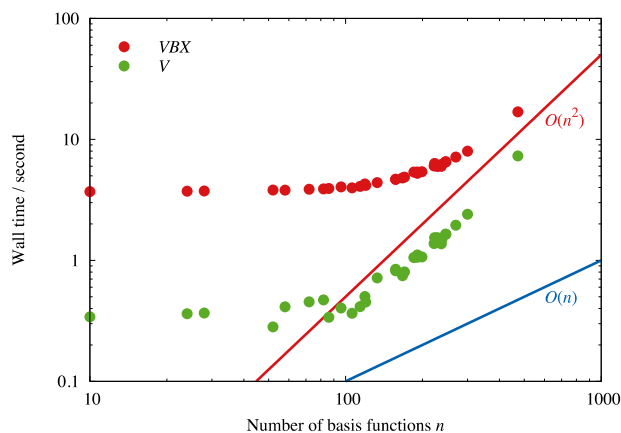


FIG. 6. The wall time (in seconds) spent in 128 MC steps of the MC-MP2-F12(*VBX*) (red dots) and MC-MP2-F12(*V*) (green dots) calculations as a function of the number of basis functions (n). The number of redundant walkers was 40. A set of 31 molecules ranging from water ($n = 10$) to tetrahydrocannabinol ($n = 472$) in the DZ basis set was used. Lines proportional to n and n^2 are superimposed to guide the eyes.

(m) is as small as 40 in this case. For small n , the cost per MC step appeared linear with n ,⁴³ as the $O(mn)$ AO amplitude calculation step is dominant with a large prefactor on the cost function caused by its many exponential evaluations. For larger n , as explored in this work, the $O(mn^2)$ MO amplitude calculation step supersedes the former, despite its small prefactor. When a large value of m is used, the cost per MC step reverts back to linear scaling with n ,⁶⁸ since the $O(m^2n)$ step of constructing O_{pq} , V_{pq} , and O'_{pq} becomes the hotspot. The same logic applies to the cost scaling of MC-MP2. Therefore, the cost scaling of MC-MP2-F12 varies with the balance between n and m , but is observed to be $O(n^2)$ per step in the present implementation and for a wide range of molecular sizes.

The cost function of the *VBX* formula may not have entered an asymptotic region by $n = 472$, but its functional form is almost certain to be $O(n^2)$ because a *VBX* calculation includes all steps of the *V* calculation. For a small molecule, a *VBX* calculation is nearly an order of magnitude more expensive than a *V* calculation. The gap becomes smaller and is only a factor of 2.3 for the largest molecule in the plot with $n = 472$. Hence, in large molecules, the cost advantage of the *V* formula may be considered insignificant, rendering the *VBX* formula preferred for its variational stability; recall that MC-MP2 and MC-MP2-F12 are designed for large molecules as they are not competitive with their deterministic counterparts for smaller ones.

The scaling of cost per MC step is not a useful measure of a stochastic method's efficiency. A more useful measure is the scaling of cost to reach a given accuracy. Figure 7 testifies that the relative statistical uncertainty (σ_{rel}) [as defined by Eq. (89) with I_N understood to include the MP2 correlation energy] after some MC steps (say, N_{rel}) is asymptotically proportional to n ,

$$\sigma_{\text{rel}} = O(n). \quad (94)$$

As is well known, the statistical uncertainty in an MC integral tends to fall off accurately in proportion to $N^{-1/2}$, where N is

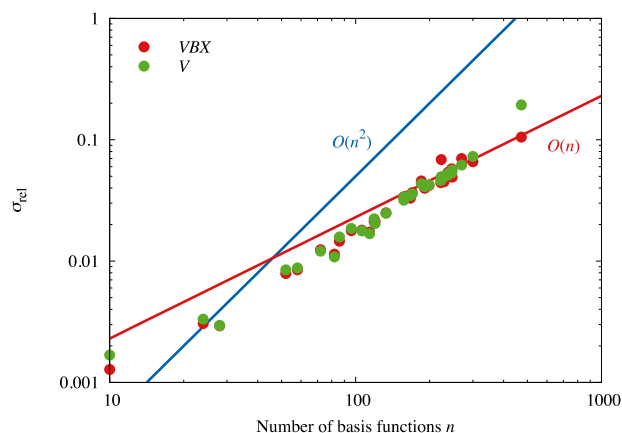


FIG. 7. The relative statistical uncertainty [Eq. (89) with I_N including the MP2 correlation energy] of the MC-MP2-F12(*VBX*) (red dots) and MC-MP2-F12(*V*) (green dots) calculations as a function of the number of basis functions (n). The number of MC steps was 6.55×10^6 , the number of redundant walkers was 40, and the block size was 7. The same set of molecules and basis set to generate Fig. 6 were used. Lines proportional to n and n^2 are superimposed to guide the eyes.

the number of MC steps. That this is also the case with the MC-MP2-F12 results has been confirmed with the data presented in Fig. 8 and elsewhere. Therefore, the relative statistical uncertainty (σ_{fin}) after N_{fin} MC steps is accurately predicted to be

$$\sigma_{\text{fin}} \approx \frac{\sigma_{\text{rel}}}{\sqrt{N_{\text{fin}}/N_{\text{rel}}}} = \frac{O(n)}{\sqrt{N_{\text{fin}}}}. \quad (95)$$

For σ_{fin} to become smaller than a given size-independent tolerance, N_{fin} has to be an $O(n^2)$ quantity. The total cost is, therefore, N_{fin} times the $O(n^2)$ cost per MC step, which is $O(n^4)$. We conclude that the MC-MP2-F12 cost to achieve a given accuracy (as measured by a relative statistical uncertainty) increases as $O(n^4)$. This may be compared with the $O(n^5)$ scaling of deterministic MP2 or MP2-F12 with fixed amplitudes.

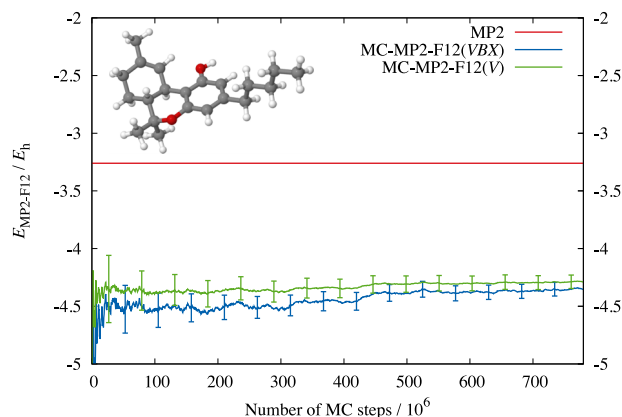


FIG. 8. Convergence of the MP2 correlation energy of tetrahydrocannabinol ($n = 472$) in the DZ basis set with respect to the number of MC steps. The error bars correspond to σ_N of Eq. (88) with the block size of 7. The Slater-type geminal (1) with $\gamma = 1.1$ a.u. was used. The number of redundant walkers was 40. The 4096-way parallel *VBX* calculation with a total of 7.87×10^8 MC steps took 7.17 h of wall time.

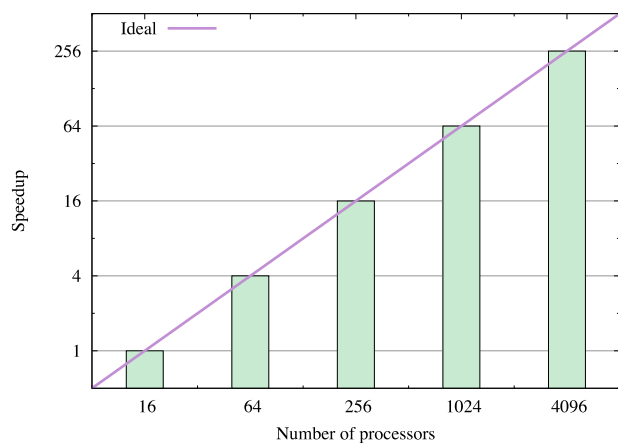


FIG. 9. Speedup (relative to the 16-CPU run) of parallel MC-MP2-F12(*VBX*)/DZ calculations for the water molecule as a function of the number of CPU cores measured in terms of the wall time spent for the same number of samples. A speedup line for the perfect scalability is superimposed.

E. Computer-size scaling

Figure 9 records the parallel scalability of MC-MP2-F12 on many CPUs. The speedup was measured as the rate of compression of the wall time spent in the last 25 000 MC steps of an MC-MP2-F12(*VBX*)/DZ calculation of the water molecule. The number of redundant walkers, m , was 20. The unit of speed was taken as that of an execution on 16 CPU cores or one XE node of Blue Waters at National Center for Supercomputing Applications of University of Illinois.

The MC-MP2-F12 algorithm is naturally parallel by design, in which each processor is tasked with its own Metropolis propagation of walkers and accumulation of the F12 correction with no mandatory interprocessor communications. All processors occasionally (every 128 MC steps) report their snapshot values of the F12 correction and statistical uncertainty to the master process by `MPI_REDUCE`, just to periodically protect the most current calculation results. As a result, it achieves scalability of 99.9% going from 16 CPU cores to 4096 CPU cores or a 255.8-fold speedup for the tiniest problem of the water molecule in the DZ basis set, attesting to its strong scaling. For a larger calculation for tetrahydrocannabinol in the DZ basis set ($n = 472$) shown in Fig. 8, a speedup by a factor of 7.84 (98% scalability) is observed from 512 to 4096 CPU cores (using the *VBX* formula). Furthermore, these calculations are indefinitely restartable in the sense that the F12 corrections and statistical uncertainties from two separate executions on different computers can be concatenated to produce the result for a longer continuous calculation.

V. CONCLUSIONS

Dense matrix multiplications, which have dominated electronic-structure algorithms, may be fundamentally non-scalable with both system and computer sizes. One may, therefore, need to redesign new, scalable algorithms when applications are attempted to large molecules, solids, or even liquids, running on a modern supercomputer, which

nowadays has up to hundreds of thousands or millions of processors. There are at least two such fundamentally scalable algorithms: local-basis (fragment or divide-and-conquer) algorithms^{96–99} and stochastic (Monte Carlo) algorithms.^{51,64,65}

The local-basis algorithms, introduced to *ab initio* electron-correlation theories by Saebø and Pulay,^{96,97} take advantage of a rapid decay of interparticle interactions in a chemical system^{100,101} and the resulting sparsity of the associated interaction matrices in a spatially local basis set. In the stochastic algorithms, the dimensional sparsity of interaction matrices is exploited by sampling increasingly less in higher dimensions. This is achieved by the Metropolis algorithm⁸⁰ with an appropriate weight function that describes the desired dimension-biased distribution of samples.

In this work, we have fully developed a stochastic algorithm of MP2-F12 using both the variational *VBX* formula and nonvariational *V* formula, the latter on the basis of the pilot implementation reported as Ref. 41. On the basis of numerical experimentation, we have established the $O(n^4)$ scaling of the MC-MP2-F12 method with the number of orbitals n , which is one-rank lower than the usual $O(n^5)$ scaling of MP2 or MP2-F12. The former does not need to use the RI approximation with an ABS to lower the dimensions of integrals or to evaluate the resulting long sum-of-products of lower-dimensional integrals, which is not scalable. It does not have to precompute or store any integrals and is free of any significant disk I/O.

Since all integrals (except one-electron integrals available from the HF calculations) are evaluated numerically by the MC method, MC-MP2-F12 can use nearly any mathematical form of geminal. Taking advantage of this, we have quantified the relative performance of the 6 geminals that satisfy the cusp conditions. We have confirmed the overall best performance of the Slater-type geminal. The Gauss-type geminal, which has a noticeably different long-range behavior than the rest, is distinctly poorly performing to the extent that the reaction energies are hardly improved by its F12 correction.

Contrary to our concern that the nonvariational formula may be practically useless for relative energies because its F12 corrections are dependent on the shape of a geminal (γ , in particular) in a manner that is hard to predict, we found it to give reliable results at a computational cost that is 2–10 times smaller than the variational formula. Nevertheless, the relative energies from the variational formula are systematically more accurate because the errors caused by the suboptimal nature of the geminal shape cancel between the two total energies. For larger molecules, the cost differential between the two formulas becomes smaller, and it may thus be recommended to use the variational formula whenever possible. The statistical uncertainties in reaction energies are less than 2 kJ mol⁻¹ after 1.44×10^8 MC steps with the ADZ basis set, which are roughly half of the systematic errors of the F12 method itself (from the CBS limits) with the same basis set. The former can be made arbitrarily smaller by running longer MC calculations.

The parallel scalability is excellent by virtue of the embarrassingly parallel nature, by design, of the algorithm. Each processor carries out its own MC integrations with no

mandatory, frequent, or large interprocessor communications. A speedup by a factor of 255.8 has been achieved upon increasing the number of processors from 16 to 4096, which is 99.9% of the perfect parallel efficiency, which may not be surprising when this method is viewed as a QMC variant. Our algorithm's speed is likely limited only by the hardware size rather than by the software limitations.

MC-MP2-F12 and its sister methods are *not* intended to be a replacement of its deterministic counterparts and are slower and less precise than the latter for smaller problems. Rather, these two classes of implementations have completely different sets of merits and demerits, making them complementary in their applicabilities. MC-MP2-F12 is designed for grand-challenge applications on high-end massively parallel supercomputers, for which existing deterministic implementations may not even start owing to their more rigid resource requirements, higher cost scaling, and lower parallel scaling. Next, we will examine applicability and performance of MC-MP2-F12 for such large problems.

ACKNOWLEDGMENTS

C.M.J. and S.H. have been supported by the U.S. Department of Energy, Office of Science, Basic Energy Sciences through Scientific Discovery through Advanced

Computing (SciDAC) program under Award No. DE-FG02-12ER46875. A.E.D. and S.H. have been supported by the U.S. Department of Energy, Office of Science, Basic Energy Sciences under Award No. DE-FG02-11ER16211. S.H. has been supported by CREST, Japan Science and Technology Agency. J.Z. and E.F.V. acknowledge the support by the U.S. National Science Foundation (Award No. CHE-1362655). It is also part of the Blue Waters sustained-petascale computing project, which is supported by the U.S. National Science Foundation (Award Nos. OCI-0725070 and ACI-1238993) and the state of Illinois. Blue Waters is a joint effort of the University of Illinois at Urbana-Champaign and its National Center for Supercomputing Applications. The authors thank Dr. Soohaeng Y. Willow for providing the computer program used in Ref. 41 and Dr. Seiichiro Ten-no for an illuminating discussion on geminal forms.

APPENDIX A: DERIVATIONS OF EQS. (61), (62), AND (70)

We show that the terms carrying the numerical factor of $7/32$ in Eqs. (61) and (62) arise from the first term of Eq. (48). The other terms in Eqs. (61) and (62) as well as higher-order integrands are derived similarly.

Substituting Eq. (64) into the first term of Eq. (48), we obtain

$$\frac{7}{32} \sum_{i,j}^{\text{occ.}} \langle ij | f_{12} [\hat{T}_1 + \hat{T}_2, f_{12}] | ij \rangle = \frac{7}{32} \sum_{i,j}^{\text{occ.}} \langle ij | f_{12} \left\{ \frac{f_{12}^{(a)}}{r_{12}} + f_{12}^{(b)} \right\} | ij \rangle - \frac{7}{32} \sum_{i,j}^{\text{occ.}} \langle ij | f_{12} \left\{ \frac{f_{12}^{(c)}}{r_{12}} + f_{12}^{(d)} \right\} \mathbf{r}_{12} \cdot (\nabla_1 - \nabla_2) | ij \rangle \quad (\text{A1})$$

$$\begin{aligned} &= \frac{7}{32} \iint d\mathbf{r}_1 d\mathbf{r}_2 \sum_{i,j}^{\text{occ.}} \varphi_i^*(\mathbf{r}_1) \varphi_j^*(\mathbf{r}_2) \left\{ \frac{f_{12}^{(a)}}{r_{12}} + f_{12}^{(b)} \right\} \varphi_i(\mathbf{r}_1) \varphi_j(\mathbf{r}_2) \\ &\quad - \frac{7}{32} \iint d\mathbf{r}_1 d\mathbf{r}_2 \sum_{i,j}^{\text{occ.}} \varphi_i^*(\mathbf{r}_1) \varphi_j^*(\mathbf{r}_2) \left\{ \frac{f_{12}^{(c)}}{r_{12}} + f_{12}^{(d)} \right\} \varphi_j(\mathbf{r}_2) \mathbf{r}_{12} \cdot \nabla_1 \varphi_i(\mathbf{r}_1) \\ &\quad + \frac{7}{32} \iint d\mathbf{r}_1 d\mathbf{r}_2 \sum_{i,j}^{\text{occ.}} \varphi_i^*(\mathbf{r}_1) \varphi_j^*(\mathbf{r}_2) \left\{ \frac{f_{12}^{(c)}}{r_{12}} + f_{12}^{(d)} \right\} \varphi_i(\mathbf{r}_1) \mathbf{r}_{12} \cdot \nabla_2 \varphi_j(\mathbf{r}_2) \end{aligned} \quad (\text{A2})$$

$$\begin{aligned} &= \frac{7}{32} \iint d\mathbf{r}_1 d\mathbf{r}_2 \frac{f_{12} f_{12}^{(a)} O_{11} O_{22}}{r_{12}} + \frac{7}{32} \iint d\mathbf{r}_1 d\mathbf{r}_2 f_{12} f_{12}^{(b)} O_{11} O_{22} \\ &\quad + \frac{7}{32} \iint d\mathbf{r}_1 d\mathbf{r}_2 \frac{f_{12} f_{12}^{(c)} (O_{11} O'_{22} - O'_{11} O_{22})}{r_{12}} \\ &\quad + \frac{7}{32} \iint d\mathbf{r}_1 d\mathbf{r}_2 f_{12} f_{12}^{(d)} (O_{11} O'_{22} - O'_{11} O_{22}), \end{aligned} \quad (\text{A3})$$

where a change in the order of summations and integrations is made in Eq. (A2), an essential step for bringing the expression into a MC-integrable form.

Next, we show that the first term of Eq. (70) with the numerical factor of $7/16$ comes from the first term of Eq. (51).

Substituting the definition of the exchange operator,

$$\hat{K}_1 = \sum_k^{\text{occ.}} \int d\mathbf{r}_3 \varphi_k^*(\mathbf{r}_3) \frac{1}{r_{13}} \hat{P}_{13} \varphi_k(\mathbf{r}_3), \quad (\text{A4})$$

into the \hat{K}_1 contribution of the first term of Eq. (51), we obtain

$$\begin{aligned}
-\frac{7}{32} \sum_{i,j}^{\text{occ.}} \langle ij | f_{12} [\hat{K}_1, f_{12}] | ij \rangle &= -\frac{7}{32} \sum_{i,j}^{\text{occ.}} \iint d\mathbf{r}_1 d\mathbf{r}_2 \varphi_i^*(\mathbf{r}_1) \varphi_j^*(\mathbf{r}_2) f_{12} \sum_k^{\text{occ.}} \int d\mathbf{r}_3 \varphi_k^*(\mathbf{r}_3) \frac{1}{r_{13}} \hat{P}_{13} \varphi_k(\mathbf{r}_3) f_{12} \varphi_i(\mathbf{r}_1) \varphi_j(\mathbf{r}_2) \\
&+ \frac{7}{32} \sum_{i,j}^{\text{occ.}} \iint d\mathbf{r}_1 d\mathbf{r}_2 \varphi_i^*(\mathbf{r}_1) \varphi_j^*(\mathbf{r}_2) f_{12} f_{12} \sum_k^{\text{occ.}} \int d\mathbf{r}_3 \varphi_k^*(\mathbf{r}_3) \frac{1}{r_{13}} \hat{P}_{13} \varphi_k(\mathbf{r}_3) \varphi_i(\mathbf{r}_1) \varphi_j(\mathbf{r}_2) \quad (\text{A5})
\end{aligned}$$

$$\begin{aligned}
&= -\frac{7}{32} \iiint d\mathbf{r}_1 d\mathbf{r}_2 d\mathbf{r}_3 \sum_{i,j,k}^{\text{occ.}} \varphi_i^*(\mathbf{r}_1) \varphi_j^*(\mathbf{r}_2) f_{12} \varphi_k^*(\mathbf{r}_3) \frac{1}{r_{13}} \varphi_k(\mathbf{r}_1) f_{32} \varphi_i(\mathbf{r}_3) \varphi_j(\mathbf{r}_2) \\
&+ \frac{7}{32} \iiint d\mathbf{r}_1 d\mathbf{r}_2 d\mathbf{r}_3 \sum_{i,j,k}^{\text{occ.}} \varphi_i^*(\mathbf{r}_1) \varphi_j^*(\mathbf{r}_2) f_{12} f_{12} \varphi_k^*(\mathbf{r}_3) \frac{1}{r_{13}} \varphi_k(\mathbf{r}_1) \varphi_i(\mathbf{r}_3) \varphi_j(\mathbf{r}_2) \quad (\text{A6})
\end{aligned}$$

$$= -\frac{7}{32} \iiint d\mathbf{r}_1 d\mathbf{r}_2 d\mathbf{r}_3 \frac{O_{13} O_{22} O_{31} f_{12} f_{32}}{r_{13}} + \frac{7}{32} \iiint d\mathbf{r}_1 d\mathbf{r}_2 d\mathbf{r}_3 \frac{O_{13} O_{22} O_{31} f_{12} f_{12}}{r_{13}} \quad (\text{A7})$$

$$= -\frac{7}{32} \iiint d\mathbf{r}_1 d\mathbf{r}_2 d\mathbf{r}_3 \frac{O_{12} O_{33} O_{21} f_{13} f_{23}}{r_{12}} + \frac{7}{32} \iiint d\mathbf{r}_1 d\mathbf{r}_2 d\mathbf{r}_3 \frac{O_{12} O_{33} O_{21} f_{23} f_{23}}{r_{12}}, \quad (\text{A8})$$

in the last step of which a coordinate interchange of $\mathbf{r}_2 \leftrightarrow \mathbf{r}_3$ (additionally $\mathbf{r}_1 \leftrightarrow \mathbf{r}_2$ in the last term) is carried out to ensure that the singularity occurs only at $r_{12} = 0$, which is also essential for bringing the final expression into a form most convenient for MC integration. Furthermore, using a coordinate interchange of $\mathbf{r}_1 \leftrightarrow \mathbf{r}_2$, we can immediately show that the \hat{K}_2 contribution of the first term of Eq. (51) is

$$-\frac{7}{32} \sum_{i,j}^{\text{occ.}} \langle ij | f_{12} [\hat{K}_2, f_{12}] | ij \rangle = -\frac{7}{32} \sum_{i,j}^{\text{occ.}} \langle ij | f_{12} [\hat{K}_1, f_{12}] | ij \rangle, \quad (\text{A9})$$

yielding

$$\begin{aligned}
&-\frac{7}{32} \sum_{i,j}^{\text{occ.}} \langle ij | f_{12} [\hat{K}_1 + \hat{K}_2, f_{12}] | ij \rangle \\
&= \frac{7}{16} \iiint d\mathbf{r}_1 d\mathbf{r}_2 d\mathbf{r}_3 \frac{f_{23}(f_{23} - f_{13})}{r_{12}} O_{12} O_{33} O_{21}. \quad (\text{A10})
\end{aligned}$$

TABLE VII. The weight function [Eq. (77)] for DZ and ADZ.

Atom (A)	$c_A^{(1)}$	$\zeta_A^{(1)}$	$c_A^{(2)}$	$\zeta_A^{(2)}$
H	0.5	0.6	0.05	0.15
C	1.0	1.0	0.10	0.25
N	2.5	1.4	0.25	0.30
O	3.0	1.8	0.30	0.37
F	4.5	1.8	0.45	0.35

TABLE VIII. The weight function [Eq. (77)] for TZ and ATZ.

Atom (A)	$c_A^{(1)}$	$\zeta_A^{(1)}$	$c_A^{(2)}$	$\zeta_A^{(2)}$
H	0.5	0.6	0.05	0.10
C	1.0	0.8	0.10	0.13
N	2.5	1.0	0.25	0.19
O	3.0	1.0	0.30	0.22
F	4.5	1.2	0.45	0.28

The other term of Eq. (51) and higher-rank integrands can be derived similarly.

APPENDIX B: WEIGHT FUNCTIONS

Two types of weight function are used in MC-MP2-F12 calculations. One is the one-electron weight function, $w_{1e}(\mathbf{r})$, in the form of Eq. (78), and the other is the two-electron weight function, $w_{2e}(\mathbf{r}_1, \mathbf{r}_2)$, of Eq. (76). They both are defined in terms of a linear combination, $g(\mathbf{r})$, of two atom-centered s -type GTOs [Eq. (77)]. For each atom, the two exponents and expansion coefficients in $g(\mathbf{r})$ are given in Tables VII and VIII. The only important consideration in defining $g(\mathbf{r})$ is that it should be more diffuse than the most diffuse of the integrands, so that the Metropolis algorithm over-samples rather than under-samples. An incidence of under-sampling can be detected as a vertical jump in the evolution of the statistical uncertainty.

¹W. Kutzelnigg and J. D. Morgan, *J. Chem. Phys.* **96**, 4484 (1992).

²W. Klopper, F. R. Manby, S. Ten-no, and E. F. Valeev, *Int. Rev. Phys. Chem.* **25**, 427 (2006).

³T. Shiozaki, E. F. Valeev, and S. Hirata, *Annu. Rep. Comput. Chem.* **5**, 131 (2009).

⁴C. Hättig, W. Klopper, A. Köhn, and D. P. Tew, *Chem. Rev.* **112**, 4 (2012).

⁵L. Kong, F. A. Bischoff, and E. F. Valeev, *Chem. Rev.* **112**, 75 (2012).

⁶T. Kato, *Commun. Pure Appl. Math.* **10**, 151 (1957).

⁷R. T Pack and W. Byers-Brown, *J. Chem. Phys.* **45**, 556 (1966).

⁸W. Kutzelnigg, *Theor. Chim. Acta* **68**, 445 (1985).

⁹J. C. Slater, *Phys. Rev.* **31**, 333 (1928).

¹⁰J. C. Slater, *Phys. Rev.* **32**, 349 (1928).

¹¹E. A. Hylleraas, *Z. Phys.* **54**, 347 (1929).

¹²E. A. Hylleraas, *Z. Phys.* **65**, 209 (1930).

¹³C. Möller and M. S. Plesset, *Phys. Rev.* **46**, 0618 (1934).

¹⁴S. Hirata, X. He, M. R. Hermes, and S. Y. Willow, *J. Phys. Chem. A* **118**, 655 (2014).

¹⁵C. C. M. Samson, W. Klopper, and T. Helgaker, *Comput. Phys. Commun.* **149**, 1 (2002).

¹⁶S. Ten-no, *Chem. Phys. Lett.* **398**, 56 (2004).

¹⁷A. J. May, E. Valeev, R. Polly, and F. R. Manby, *Phys. Chem. Chem. Phys.* **7**, 2710 (2005).

¹⁸D. P. Tew and W. Klopper, *J. Chem. Phys.* **123**, 074101 (2005).

- ¹⁹F. R. Manby, H. J. Werner, T. B. Adler, and A. J. May, *J. Chem. Phys.* **124**, 094103 (2006).
- ²⁰E. F. Valeev, *J. Chem. Phys.* **125**, 244106 (2006).
- ²¹S. Ten-no and J. Noga, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2**, 114 (2012).
- ²²H.-J. Werner, T. B. Adler, and F. R. Manby, *J. Chem. Phys.* **126**, 164102 (2007).
- ²³B. J. Persson and P. R. Taylor, *J. Chem. Phys.* **105**, 5915 (1996).
- ²⁴L. E. McMurchie and E. R. Davidson, *J. Comput. Phys.* **26**, 218 (1978).
- ²⁵B. J. Persson and P. R. Taylor, *Theor. Chem. Acc.* **97**, 240 (1997).
- ²⁶P. Dahle and P. R. Taylor, *Theor. Chem. Acc.* **105**, 401 (2001).
- ²⁷M. Dupuis, J. Rys, and H. F. King, *J. Chem. Phys.* **65**, 111 (1976).
- ²⁸A. Komornicki and H. F. King, *J. Chem. Phys.* **134**, 244115 (2011).
- ²⁹A. Komornicki and H. F. King, *J. Chem. Phys.* **135**, 129901 (2011).
- ³⁰D. Prendergast, M. Nolan, C. Filippi, S. Fahy, and J. C. Greer, *J. Chem. Phys.* **115**, 1626 (2001).
- ³¹K. Szalewicz, B. Jeziorski, H. J. Monkhorst, and J. G. Zabolitzky, *Chem. Phys. Lett.* **91**, 169 (1982).
- ³²K. Szalewicz, B. Jeziorski, H. J. Monkhorst, and J. G. Zabolitzky, *J. Chem. Phys.* **78**, 1420 (1983).
- ³³P. Dahle, T. Helgaker, D. Jonsson, and P. R. Taylor, *Phys. Chem. Chem. Phys.* **9**, 3112 (2007).
- ³⁴W. Klopper and W. Kutzelnigg, *Chem. Phys. Lett.* **134**, 17 (1987).
- ³⁵W. Kutzelnigg and W. Klopper, *J. Chem. Phys.* **94**, 1985 (1991).
- ³⁶W. Klopper and C. C. M. Samson, *J. Chem. Phys.* **116**, 6397 (2002).
- ³⁷E. F. Valeev, *Chem. Phys. Lett.* **395**, 190 (2004).
- ³⁸S. Ten-no, *J. Chem. Phys.* **126**, 014108 (2007).
- ³⁹S. Ten-no, *J. Chem. Phys.* **121**, 117 (2004).
- ⁴⁰S. Hirata, *J. Phys. Chem. A* **107**, 9887 (2003).
- ⁴¹S. Y. Willow, J. Zhang, E. F. Valeev, and S. Hirata, *J. Chem. Phys.* **140**, 031101 (2014).
- ⁴²S. Y. Willow, K. S. Kim, and S. Hirata, *J. Chem. Phys.* **137**, 204122 (2012).
- ⁴³S. Y. Willow, K. S. Kim, and S. Hirata, *J. Chem. Phys.* **138**, 164111 (2013).
- ⁴⁴S. Y. Willow, M. R. Hermes, K. S. Kim, and S. Hirata, *J. Chem. Theory Comput.* **9**, 4396 (2013).
- ⁴⁵S. Y. Willow and S. Hirata, *J. Chem. Phys.* **140**, 024111 (2014).
- ⁴⁶S. Y. Willow, K. S. Kim, and S. Hirata, *Phys. Rev. B* **90**, 201110 (2014).
- ⁴⁷M. R. Hermes and S. Hirata, *J. Chem. Phys.* **141**, 084105 (2014).
- ⁴⁸M. R. Hermes and S. Hirata, *J. Chem. Phys.* **141**, 244111 (2014).
- ⁴⁹M. R. Hermes and S. Hirata, *J. Chem. Phys.* **143**, 129903 (2015).
- ⁵⁰M. R. Hermes and S. Hirata, *J. Chem. Phys.* **143**, 129904 (2015).
- ⁵¹M. H. Kalos and P. A. Whitlock, *Monte Carlo Methods*, 2nd ed. (Wiley-VCH, Weinheim, 2008).
- ⁵²S. Zhang and H. Krakauer, *Phys. Rev. Lett.* **90**, 136401 (2003).
- ⁵³A. Thom and A. Alavi, *Phys. Rev. Lett.* **99**, 143001 (2007).
- ⁵⁴Y. Ohtsuka and S. Nagase, *Chem. Phys. Lett.* **463**, 431 (2008).
- ⁵⁵G. H. Booth, A. J. W. Thom, and A. Alavi, *J. Chem. Phys.* **131**, 054106 (2009).
- ⁵⁶E. Kozik, K. Van Houcke, E. Gull, L. Pollet, N. Prokof'ev, B. Svistunov, and M. Troyer, *Europhys. Lett.* **90**, 10004 (2010).
- ⁵⁷D. Cleland, G. H. Booth, and A. Alavi, *J. Chem. Phys.* **132**, 041103 (2010).
- ⁵⁸F. R. Petruzielo, A. A. Holmes, H. J. Changlani, M. P. Nightingale, and C. J. Umrigar, *Phys. Rev. Lett.* **109**, 230201 (2012).
- ⁵⁹J. J. Shepherd, G. H. Booth, and A. Alavi, *J. Chem. Phys.* **136**, 244101 (2012).
- ⁶⁰G. H. Booth and G. K.-L. Chan, *J. Chem. Phys.* **137**, 191102 (2012).
- ⁶¹D. Neuhauser, E. Rabani, and R. Baer, *J. Chem. Theory Comput.* **9**, 24 (2013).
- ⁶²S. Ten-no, *J. Chem. Phys.* **138**, 164126 (2013).
- ⁶³D. M. Ceperley and B. J. Alder, *Phys. Rev. Lett.* **45**, 566 (1980).
- ⁶⁴B. L. Hammond, W. A. Lester, and P. J. Reynolds, *Monte Carlo Methods in Ab Initio Quantum Chemistry* (World Scientific, Singapore, 1994).
- ⁶⁵A. Lüchow and J. B. Anderson, *Annu. Rev. Phys. Chem.* **51**, 501 (2000).
- ⁶⁶W. M. C. Foulkes, L. Mitas, R. J. Needs, and G. Rajagopal, *Rev. Mod. Phys.* **73**, 33 (2001).
- ⁶⁷J. Kolorenč and L. Mitas, *Rep. Prog. Phys.* **74**, 026502 (2011).
- ⁶⁸A. E. Doran and S. Hirata, *J. Chem. Theory Comput.* **12**, 4821 (2016).
- ⁶⁹J. Toulouse, R. Assaraf, and C. J. Umrigar, *Adv. Quantum Chem.* **73**, 285 (2016).
- ⁷⁰K.-C. Pan and H. F. King, *J. Chem. Phys.* **53**, 4397 (1970).
- ⁷¹P. Wind, W. Klopper, and T. Helgaker, *Theor. Chem. Acc.* **107**, 173 (2002).
- ⁷²A. J. May and F. R. Manby, *J. Chem. Phys.* **121**, 4479 (2004).
- ⁷³M. Torheyden and E. F. Valeev, *Phys. Chem. Chem. Phys.* **10**, 3410 (2008).
- ⁷⁴E. F. Valeev and C. L. Janssen, *J. Chem. Phys.* **121**, 1214 (2004).
- ⁷⁵M. Nooijen and R. J. Bartlett, *J. Chem. Phys.* **109**, 8232 (1998).
- ⁷⁶S. Obara and A. Saika, *J. Chem. Phys.* **84**, 3963 (1986).
- ⁷⁷F. A. Bischoff, R. J. Harrison, and E. F. Valeev, *J. Chem. Phys.* **137**, 104103 (2012).
- ⁷⁸F. A. Bischoff and E. F. Valeev, *J. Chem. Phys.* **139**, 114106 (2013).
- ⁷⁹S. Hirata, T. Shiozaki, C. M. Johnson, and J. D. Talman, *Mol. Phys.* (published online 2016).
- ⁸⁰N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. N. Teller, and E. Teller, *J. Chem. Phys.* **21**, 1087 (1953).
- ⁸¹H. Flyvbjerg and H. G. Petersen, *J. Chem. Phys.* **91**, 461 (1989).
- ⁸²D. M. Ceperley and B. Bernu, *J. Chem. Phys.* **89**, 6316 (1988).
- ⁸³M. Valiev, E. J. Bylaska, N. Govind, K. Kowalski, T. P. Straatsma, H. J. J. Van Dam, D. Wang, J. Nieplocha, E. Apra, T. L. Windus, and W. A. de Jong, *Comput. Phys. Commun.* **181**, 1477 (2010).
- ⁸⁴T. Helgaker, W. Klopper, H. Koch, and J. Noga, *J. Chem. Phys.* **106**, 9639 (1997).
- ⁸⁵K. L. Bak, P. Jørgensen, J. Olsen, T. Helgaker, and W. Klopper, *J. Chem. Phys.* **112**, 9229 (2000).
- ⁸⁶A. Bijl, *Physica* **7**, 869 (1940).
- ⁸⁷R. Dingle, *Philos. Mag.* **40**, 573 (1949).
- ⁸⁸R. Jastrow, *Phys. Rev.* **98**, 1479 (1955).
- ⁸⁹D. Ceperley, *Phys. Rev. B* **18**, 3126 (1978).
- ⁹⁰C.-J. Huang, C. J. Umrigar, and M. P. Nightingale, *J. Chem. Phys.* **107**, 3007 (1997).
- ⁹¹H. J. Monkhorst, *Mol. Phys.* **103**, 2009 (2005).
- ⁹²A. Grüneis, J. J. Shepherd, A. Alavi, D. P. Tew, and G. H. Booth, *J. Chem. Phys.* **139**, 084112 (2013).
- ⁹³M. Silkowski, M. Lesiuk, and R. Moszynski, *J. Chem. Phys.* **142**, 124102 (2015).
- ⁹⁴C. L. Janssen, I. B. Nielsen, M. L. Leininger, E. F. Valeev, and E. T. Seidl, *The Massively Parallel Quantum Chemistry Program (MPQC), Version 2.2* (Sandia National Laboratories, Livermore, 2003).
- ⁹⁵J. Zhang and E. F. Valeev, *J. Chem. Theory Comput.* **8**, 3175 (2012).
- ⁹⁶S. Saebø and P. Pulay, *Chem. Phys. Lett.* **113**, 13 (1985).
- ⁹⁷S. Saebø and P. Pulay, *Annu. Rev. Phys. Chem.* **44**, 213 (1993).
- ⁹⁸W. Kohn, *Int. J. Quantum Chem.* **56**, 229 (1995).
- ⁹⁹W. Yang, *Phys. Rev. Lett.* **66**, 1438 (1991).
- ¹⁰⁰S. Hirata and Y. Ohnishi, *Phys. Chem. Chem. Phys.* **14**, 7800 (2012).
- ¹⁰¹S. Hirata, M. Keçeli, Y.-Y. Ohnishi, O. Sode, and K. Yagi, *Annu. Rev. Phys. Chem.* **63**, 131 (2012).