Convergence Acceleration of Parallel Monte Carlo Second-Order Many-Body Perturbation Calculations Using Redundant Walkers

Soohaeng Yoo Willow,[†,‡] Matthew R. Hermes,[†] Kwang S. Kim,[‡] and So Hirata*[,†,§]

[†]Department of Chemistry, University of Illinois at Urbana−Champaign, 600 South Mathews Avenue, Urbana, Illinois 61801, United States

[‡]Center for Superfunctional Materials, Department of Chemistry, Pohang University of Science and Technology, San 31, Hyojadong, Namgu, Pohang 790-784, Korea

[§]CREST, Japan Science and Technology Agency, 4-1-8 Honcho, Kawaguchi, Saitama 332-0012, Japan

ABSTRACT: A Monte Carlo (MC) integration of the second-order many-body perturbation (MP2) corrections to energies and self-energies eliminates the usual computational bottleneck of the MP2 algorithm (i.e., the basis transformation of two-electron integrals), thereby achieving near-linear size dependence of its operation cost, a negligible core and disk memory cost, and a naturally parallel computational kernel. In this method, the correlation correction expressions are recast into high-dimensional integrals, which are approximated as the sums of integrands evaluated at coordinates of four electron random walkers guided by a Metropolis algorithm for importance sampling. The gravest drawback of this method, however, is the inevitable statistical uncertainties in its results, which decay slowly as the inverse square-root of the number of MC steps. We propose an algorithm that can increase the number of MC sampling points in each MC step by many orders of magnitude by having $2m$ electron walkers ($m > 2$) and then using $m(m − 1)/2$ permutations of their coordinates in evaluating the integrands. Hence, this algorithm brings an $O(m^2)$-fold increase in the number of MC sampling points at a mere $O(m)$ additional cost of propagating redundant walkers, which is a net $O(m)$-fold enhancement in sampling efficiency. We have demonstrated a large performance increase in the Monte Carlo MP2 calculations for the correlation energies of benzene and benzene dimer as well as for the correlation corrections to the energy, ionization potential, and electron affinity of C_{60}. The calculation for C_{60} has been performed with a parallel implementation of this method running on up to 400 processors of a supercomputer, yielding an accurate prediction of its large electron affinity, which makes its derivative useful as an electron acceptor in bulk heterojunction organic solar cells.
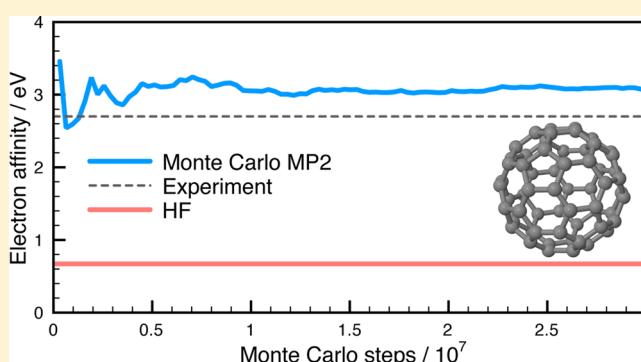
1. INTRODUCTION

Conventional algorithms of *ab initio* electron-correlation theories are almost always based on tensor algebra,[1] which is fundamentally nonscalable with respect to the number of processors. We have sought to redesign these algorithms into ones that are more naturally parallel. As an example of such redesign, we have recently proposed[2,3] Monte Carlo (MC) integrations of the correlation energies and self-energies of second-order many-body perturbation (MP2) theory.[4−7] In this Monte Carlo MP2 (MC-MP2) method,[2,3] the canonical expressions of these perturbation corrections, which are long sums of products of two-electron integrals, are recast into sums of a few high-dimensional integrals. These integrals are then evaluated by an MC method with sampling points generated by the Metropolis algorithm according to a suitably chosen weight function.

MC-MP2 eliminates the hotspot of the conventional MP2 algorithms,[8−10] that is, the atomic-orbital (AO) to molecular-orbital (MO) basis transformation of two-electron integrals, whose operation and memory costs grow respectively as the fifth and fourth power of system size. By virtue of this elimination, the size dependence of the operation cost of MC-MP2 is drastically reduced to quadratic, while the memory cost also becomes negligible. Furthermore, the actual operation costs are found to scale nearly linearly even for such small molecules as H_2O and N_2.[3] This may be the first instance in which the near-linear scaling of MP2 is demonstrated without exploiting the nearsightedness[11] of electron correlation that is meaningful only in a spatially extended molecule. MC-MP2 is, therefore, scalable both with the number of processors and with the number of electrons. The latter does not have to be proportional to the number of atoms or the molecule's spatial extent.

Even over the quantum Monte Carlo (QMC) methods,[12−16] MC-MP2 has several unique advantages. MC-MP2 can compute energy differences (such as correlation energies and self-energies) directly and not as small differences of large noisy

quantities. It does not suffer from the so-called fixed-node errors of the most widely used QMC variant, that is, diffusion Monte Carlo. Its accuracy can be systematically improved in principle by going to higher orders of perturbation theory. In fact, an extension to third-order many-body perturbation (MP3) theory is well underway in our laboratory.

MC-MP2 may have some superficial similarities with other methods that "wed" *ab initio* electron-correlation theories and QMC.[17−26] However, MC-MP2 differs from all these in at least one key aspect: it is not predicated upon the ready availability of MO-based two-electron integrals (or even AO-based integrals), which are expensive to evaluate and store. MC-MP2 achieves this by replacing every cubature by an MC integration. The feasibility of this replacement depends on the effectiveness of our proposed weight function, which is the core intellectual contribution of our previous work on this method.[2]

These unique advantages of MC-MP2, however, come at a price, and it is the statistical errors accompanying its results, which decay at the slow rate of $N^{-1/2}$ with the number of MC steps (N). The errors are somewhat greater than those found in QMC, which can sometimes benefit from the zero-variance weight functions; the weight functions of MC-MP2 do not have this property. Furthermore, the statistical errors in MC-MP2 increase with system size, and in this sense, it should be cautioned that the observed near-linear scaling is for the cost per MC step and not for the cost to reach a certain precision.[22,23] To achieve the precision of a few mE_h for a small molecule by MC-MP2, a large value of N on the order of 10^7 to 10^9 is needed.[2,3] Convergence acceleration is, therefore, imperative.

In this article, we propose and implement an algorithm of MC-MP2 that increases the number of MC sampling points per MC step by many orders of magnitude. It achieves this by having far more walkers than necessary and using all permutations of their coordinates when evaluating the integrands. While the cost of having redundant walkers increases only linearly with their number, the permutations that yield distinct values of the integrands grow in number quadratically, thus resulting in a net performance enhancement. We call this method 'redundant-walker' MC-MP2 (only when such distinction is necessary, as this will be our default algorithm from now on) and demonstrate its effectiveness through applications to benzene, benzene dimer, and C_{60}. We also describe our parallel implementation of MC-MP2, which has been tested on up to 400 processors of a supercomputer for the correlation energies and self-energies of C_{60}.

2. MONTE CARLO MP2

The second-order perturbation correction to the Hartree−Fock (HF) energy of a molecule is written[2] as

$$E^{(2)} = E_A^{(2)} + E_B^{(2)} \tag{1}$$

with

$$E_A^{(2)} = -2 \int \cdots \int d\mathbf{r}_1 \cdots d\mathbf{r}_4 \int_0^\infty d\tau$$
$$\times \frac{o(\mathbf{r}_1, \mathbf{r}_3, \tau)o(\mathbf{r}_2, \mathbf{r}_4, \tau)v(\mathbf{r}_1, \mathbf{r}_3, \tau)v(\mathbf{r}_2, \mathbf{r}_4, \tau)}{r_{12}r_{34}} \tag{2}$$

$$E_B^{(2)} = \int \cdots \int d\mathbf{r}_1 \cdots d\mathbf{r}_4 \int_0^\infty d\tau$$
$$\times \frac{o(\mathbf{r}_1, \mathbf{r}_4, \tau)o(\mathbf{r}_2, \mathbf{r}_3, \tau)v(\mathbf{r}_1, \mathbf{r}_3, \tau)v(\mathbf{r}_2, \mathbf{r}_4, \tau)}{r_{12}r_{34}} \tag{3}$$

and

$$o(\mathbf{r}_1, \mathbf{r}_3, \tau) = \sum_i^{\text{occ.}} \varphi_i^*(\mathbf{r}_1)\varphi_i(\mathbf{r}_3) \exp(\varepsilon_i \tau) \tag{4}$$

$$v(\mathbf{r}_1, \mathbf{r}_3, \tau) = \sum_a^{\text{vir.}} \varphi_a(\mathbf{r}_1)\varphi_a^*(\mathbf{r}_3) \exp(-\varepsilon_a \tau) \tag{5}$$

Equations 2 and 3 are 13-dimensional integrals. In MC-MP2, the 12-dimensional parts (over r_1 through r_4) of these integrals are evaluated by an MC method, while the remaining one-dimensional integrations over τ are carried out with a 21-point Gauss−Kronrod quadrature.[2,27] Hence, each integral is evaluated as

$$I = \int \cdots \int d\mathbf{r}_1 \cdots d\mathbf{r}_4 \int_0^\infty d\tau f(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \mathbf{r}_4, \tau) \tag{6}$$

$$\approx \frac{1}{N} \sum_{n=1}^N \frac{\tilde{f}(\mathbf{r}_1^{[n]}, \mathbf{r}_2^{[n]}, \mathbf{r}_3^{[n]}, \mathbf{r}_4^{[n]})}{w(\mathbf{r}_1^{[n]}, \mathbf{r}_2^{[n]}, \mathbf{r}_3^{[n]}, \mathbf{r}_4^{[n]})} \equiv I_N \tag{7}$$

with

$$\tilde{f}(\mathbf{r}_1^{[n]}, \mathbf{r}_2^{[n]}, \mathbf{r}_3^{[n]}, \mathbf{r}_4^{[n]}) = \sum_{q=1}^{21} w_q f(\mathbf{r}_1^{[n]}, \mathbf{r}_2^{[n]}, \mathbf{r}_3^{[n]}, \mathbf{r}_4^{[n]}, \tau_q) \tag{8}$$

where I is either $E_A^{(2)}$ or $E_B^{(2)}$, I_N is the approximate value of I at the Nth MC step, and f stands for the corresponding integrand of eq 2 or 3. The four electron coordinates (*walkers*) at the nth MC step are denoted by $\{r_1^{[n]}, r_2^{[n]}, r_3^{[n]}, r_4^{[n]}\}$, whereas τ_q and w_q designate Gauss−Kronrod quadrature points and weights, respectively.[2] The statistical uncertainty $\bar{\sigma}$ of the approximate integral I_N is given by

$$\bar{\sigma} = \frac{\sigma}{\sqrt{N}} \tag{9}$$

$$\sigma^2 = \frac{1}{N-1} \sum_{n=1}^N \left\{ \frac{\tilde{f}(\mathbf{r}_1^{[n]}, \cdots, \mathbf{r}_4^{[n]})}{w(\mathbf{r}_1^{[n]}, \cdots, \mathbf{r}_4^{[n]})} - I_N \right\}^2 \tag{10}$$

where I_N is defined in eq 7, σ is the standard deviation [an $O(N^0)$ quantity] of the instantaneous value of the quotient $\tilde{f}/w$, and $\bar{\sigma}$ is the statistical uncertainty associated with I_N, which decays as $O(N^{-1/2})$. In this article, we report the correlation energies and self-energies of MP2 obtained as I_N with their statistical uncertainties as $\bar{\sigma}$.

The viability (let alone the efficiency) of the MC integration depends critically on the weight function, w.[15] It must be analytically integrable,

$$\int\int d\mathbf{r}_1 d\mathbf{r}_2 w(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \mathbf{r}_4) = 1 \tag{11}$$

be positive everywhere,

$$w(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \mathbf{r}_4) > 0 \tag{12}$$

cancel all singularities in the integrand exactly and analytically,

$$\left| \frac{\tilde{f}(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \mathbf{r}_4)}{w(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \mathbf{r}_4)} \right| < \infty \tag{13}$$

and generally behave similar to $\tilde{f}(r_1,r_2,r_3,r_4)$ so that the quotient $\tilde{f}/w$ is a smooth function throughout the domain of integration. Note that the integrands in eqs 2 and 3 have the singularities at $r_{12} = 0$ and $r_{34} = 0$, which can occur anywhere in the 12-dimensional real space.

In our previous work,[2] we have discovered the following weight function, which satisfies all of the above requirements:

$$w(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \mathbf{r}_4) = \tilde{w}(\mathbf{r}_1, \mathbf{r}_2)\tilde{w}(\mathbf{r}_3, \mathbf{r}_4) \tag{14}$$

with

$$\tilde{w}(\mathbf{r}_1, \mathbf{r}_2) = \frac{1}{E_g} \frac{g(\mathbf{r}_1)g(\mathbf{r}_2)}{r_{12}} \tag{15}$$

where $g(r)$ is, typically, but not necessarily, a sum of s-type Gaussian-type orbitals centered at atoms. The normalization factor,

$$E_g = \int d\mathbf{r}_1 \int d\mathbf{r}_2 \frac{g(\mathbf{r}_1)g(\mathbf{r}_2)}{r_{12}} \tag{16}$$

has the form of the so-called Coulomb (J) integral of the HF theory and can be evaluated analytically, when $g(r)$ is a sum of Gaussian-type orbitals.[28]

In each MC step, therefore, two *pairs* of electron walkers, $\{r_1, r_2\}$ and $\{r_3, r_4\}$, must be propagated in such a way that each pair's coordinates are distributed in the 6-dimensional space according to the weight function $\tilde{w}$. This is achieved by the Metropolis−Rosenbluth−Rosenbluth−Teller−Teller algorithm.[29] The nth walker coordinates are the destination of a random-walk step of radius r from the previous coordinates:

$$\mathbf{r}_1^{[n]} = r\mathbf{e}_1 + \mathbf{r}_1^{[n-1]} \tag{17}$$

$$\mathbf{r}_2^{[n]} = r\mathbf{e}_2 + \mathbf{r}_2^{[n-1]} \tag{18}$$

where $\mathbf{e}_1$ and $\mathbf{e}_2$ are two randomly oriented 3-dimensional unit vectors. Let y be the quotient,

$$y = \frac{\tilde{w}(\mathbf{r}_1^{[n]}, \mathbf{r}_2^{[n]})}{\tilde{w}(\mathbf{r}_1^{[n-1]}, \mathbf{r}_2^{[n-1]})} \tag{19}$$

If $y \geq 1$, the new coordinates are accepted. Otherwise, y is compared with a random number x ($0 \leq x \leq 1$); if $y > x$, the new coordinates are accepted and, otherwise, they are rejected. In the latter event, the $(n-1)$th coordinates are retained as the nth coordinates. In the first 10^5 MC steps, the value of r is continuously adjusted so as to make the acceptance ratio as close as 50%. These initial, "equilibration" steps are not included in the summation of eq 7.

Once the two walker pairs' coordinates are determined, the values of MO's at those coordinates and hence $o(r_i, r_j, \tau_q)$ and $v(r_i, r_j, \tau_q)$ at all values of τ_q are evaluated. With these, the values of f and w are obtained and one additional summation step in eq 7 is carried out for both $E_A^{(2)}$ and $E_B^{(2)}$. It should be emphasized that a single MC run informs all diagrammatic contributions at all quadrature grid points.

The operation cost of each MC step is dominated by the evaluation of MO's at newly accepted walker coordinates (rejected moves cost much less because y in eq 19 involves only the weight functions but not MO's and is inexpensive to evaluate). This in turn requires the evaluation of AO's at the same coordinates and transforming them to MO's. The cost of the evaluation of AO's increases linearly with the number of AO's (n), while that of the transformation grows quadratically with n. In theory, therefore, the size dependence of the cost per MC step is $O(n^2)$. In practice, however, the cost of evaluating AO's is much greater than their transformation, and the observed size dependence is nearly linear.[3] The memory cost is negligible because all one has to store are the current and immediate past walker coordinates and the values of MO's at these coordinates, which can be easily stored in core memory. Neither AO- or MO-based two-electron integrals are used anywhere in MC-MP2 and no disk space or access is needed for them.

Essentially the same algorithm can be used to evaluate the second-order perturbation correction to self-energies,[3] which are related to ionization potentials and electron affinities of a molecule or quasiparticle energies of a crystal. Again, a single MC run can yield correlation energies and self-energies simultaneously.

3. REDUNDANT-WALKER ALGORITHM

The gravest drawback of MC-MP2 is the slow $N^{-1/2}$ decay of the statistical uncertainty $\bar{\sigma}$, which estimates the statistical error, $|I − I_N|$, in the approximate integral I_N. Here, we propose convergence acceleration by having a redundant population of walkers. Instead of just two pairs of walkers, $\{r_1, r_2\}$ and $\{r_3, r_4\}$, which are necessary to evaluate $\tilde{f}(r_1, r_2, r_3, r_4)$, we propagate more walker pairs, $\{r_1, r_2\}_k$, with $1 \leq k \leq m$, each distributed according to the weight function $\tilde{w}$.

Using this redundant population, we can now evaluate the integral by substituting all distinct two pairs into eq 7, namely,

$$I \approx \frac{1}{N} \sum_{n=1}^{N} I_n \equiv I_N \tag{20}$$

with

$$I_n = \frac{2}{m(m-1)} \sum_{k=1}^{m-1} \sum_{l=k+1}^{m} \frac{\tilde{f}(\mathbf{r}_{1k}^{[n]}, \mathbf{r}_{2k}^{[n]}, \mathbf{r}_{1l}^{[n]}, \mathbf{r}_{2l}^{[n]})}{w(\mathbf{r}_{1k}^{[n]}, \mathbf{r}_{2k}^{[n]}, \mathbf{r}_{1l}^{[n]}, \mathbf{r}_{2l}^{[n]})} \tag{21}$$

where m is the number of walker pairs and hence the degree of redundancy (with $m = 2$ corresponding to no redundancy), and $r_{1k}^{[n]}$ and $r_{2k}^{[n]}$ are respectively the first and second electron coordinates of the kth walker pair at the nth MC step. The statistical uncertainty $\bar{\sigma}$ in I_N is given by

$$\bar{\sigma} = \frac{\sigma_m}{\sqrt{N}} \tag{22}$$

$$\sigma_m^2 = \frac{1}{N-1} \sum_{n=1}^{N} \{I_n - I_N\}^2 \tag{23}$$

where σ_m is the standard deviation of I_n.

I_N and $\bar{\sigma}$ as evaluated by eqs 20−23 converge much more rapidly than those evaluated by eqs 7−10, especially for large m. This is because of the following. With m walker pairs, we can have $_mC_2 = m(m-1)/2$ distinct two pairs (pairs of pairs) that can be substituted into $\tilde{f}$ and w. This amounts to an increase in the number of MC sampling points per MC step by an $O(m^2)$ factor. To propagate m walker pairs, we only need to spend $O(m)$ operation and memory costs. Hence, we can achieve an $O(m^2)$ increase in the number of distinct pairs of pairs or the number of MC sampling points at an $O(m)$ cost, a net $O(m)$

performance enhancement. For instance, $m = 10$ causes a 5-fold increase in the operation and memory cost (relative to $m = 2$ with no redundancy), but it increases the number of sampling points in each MC step by $_{10}C_2 = 45$. Together, this amounts to a 9-fold increase in overall sampling efficiency. With $m = 400$, this enhancement factor is 200. The value of m is only limited by the core memory size and can be very large and certainly greater than the actual number of electrons in the molecule.

Figure 1 shows the histograms of the instantaneous values of the summand, I_n, defined in eq 21. Despite the enormous
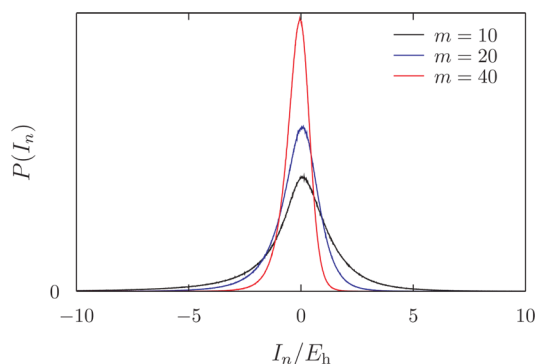

Figure 1. Distribution of I_n (in E_h) in the MC-MP2/cc-pVDZ correlation energy of H_2O as a function of the degree of redundancy (m) with $N = 10^6$ in each case.

widths (a few E_h) of the distributions of I_n, their averages (I_N) are within hundredths or thousandths of E_h of the correct limit, owing to the effectiveness of the weight function.[2,3] It can be noticed that the distributions narrow and σ_m decreases with increasing m. This narrowing is the basis of the convergence acceleration with redundant walkers and is expected because each value of I_n is an average of $m(m-1)/2$ distinct values of the integrand $\bar{\sigma}/w$ and individually approaches I with increasing m.

It is interesting to note that the I_n-distribution seems to deviate systematically more from Gaussian and become more asymmetric with increasing m. The precise cause of this is unclear, but they are sufficiently close to Gaussian that σ_m (and thus $\bar{\sigma}$ also) can serve as a useful, quantitative measure of statistical uncertainty with its interpretation being the standard deviation of a Gaussian distribution. Note that $\bar{\sigma}$ defined by eqs 22 and 23 naturally implements the reblocking method of Flyvbjerg and Petersen,[30] with the blocking length increasing with the degree of redundancy. This is appropriate given a greater degree of serial correlation[31] expected among the integrand $\bar{f}/w$ obtained from various permutations of a larger number of redundant walkers.

The variational Monte Carlo (VMC) method[12,13,32,33] also uses a large number of electron walkers $\{r_1,...,r_m\}$, which are statistically independent sampling points distributed according to the weight function $\rho(r) = |\Psi_T(r)|^2$, where $\Psi_T(r)$ is a trial wave function. The energy in VMC at each MC step can be obtained as the average of the local energy,

$$E_{VMC} = \frac{1}{m} \sum_{k=1}^{m} \frac{\hat{H}\Psi_T(\mathbf{r}_k)}{\Psi_T(\mathbf{r}_k)} \qquad (24)$$

Note that, in VMC, the average is taken over just m distinct contributions at an $O(m)$ cost per MC step, whereas redundant-walker MC-MP2 uses an average over $m(m-1)/$

2 contributions at an $O(m)$ cost per MC step. In this sense, the benefit of redundancy is far greater in MC-MP2 than in VMC.

4. PARALLEL IMPLEMENTATION

Our parallel implementation of redundant-walker MC-MP2 consists in nearly independent instances of essentially a sequential MC-MP2 code running on processors with different random number seeds. Currently, these instances perform MPI_REDUCE every 200 MC steps to record the accumulated correlation energy, self-energies, and their statistical uncertainties. These are the only interprocessor communications, which are, however, not essential and can easily be eliminated to make the whole program free of communication. In the latter case, the number of processors can be dynamically varied, making the whole calculation fault-tolerant and restartable (though these are not pursued in this work). Furthermore, the large number of MC steps (typically, on the order of 10^7) needed even for small molecules makes it unlikely to have an application too small to be scalable because of too large a number of available processors in a modern supercomputer. Figure 2 shows that our initial parallel implementation can achieve a speedup by a factor exceeding 250 on 400 processors, relative to a single processor calculation.
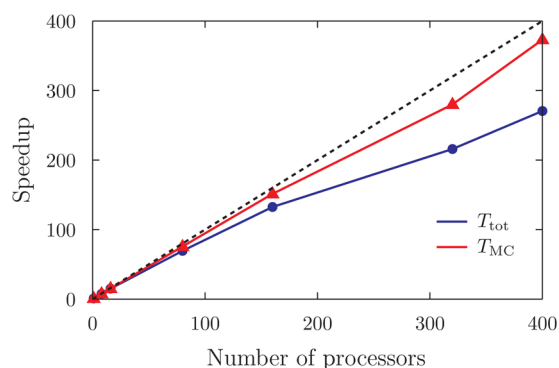

Figure 2. Parallel efficiency of the MC-MP2/cc-pVDZ calculations for C_{60} on the Tachyon II supercomputer of the Korea Institute of Science and Technology Information. The blue curve plots the ratio of the CPU time (T_{tot}) including that spent on the initial equilibration step of the Metropolis algorithm and the final job completion step. The red curve plots the ratio of just the CPU time spent on the MC steps excluding the equilibration step.

The human cost of parallelizing MC-MP2 is extremely small thanks to the intrinsically parallel nature of MC integrations. It took one of us only hours to parallelize MC-MP2, which was to insert a few lines of mpi calls and one line making the random number seed dependent on the processor ID. This is in contrast with a similar effort for conventional MP2, which can take many months or even years to complete with much lower degree of scalability and fault-tolerance.

The HF energies, orbitals, and orbital energies are furnished either by our in-house program for all calculations (except for C_{60}) and by the nwchem program[34] (for C_{60}), with which our parallel MC-MP2 code is interfaced.

5. RESULTS AND DISCUSSION

Table 1 and Figure 3 show the results of MC-MP2/6-31G** on the benzene molecule using the following $g(r)$ in the weight function w:

Table 1. Dependence of the MC-MP2/6-31G** ($N = 10^8$) Correlation Energy $E^{(2)}$ and Uncertainty $\bar{\sigma}$ (both in E_h) of the Benzene Molecule ($r_{CC} = 1.3963$ Å and $r_{CH} = 1.0826$ Å) on the Degree of Redundancy (m)[a]

method	$E^{(2)}$	$\bar{\sigma}$	$_mC_2$	T_1	T_2
MP2	−0.7924				
MC-MP2 ($m = 2$)	−0.6191	0.0492	1	7.4	7.4
MC-MP2 ($m = 5$)	−0.8060	0.0190	10	23	2.3
MC-MP2 ($m = 10$)	−0.7878	0.0077	45	59	1.3

[a]The number of MC sampling points per MC step $_mC_2$, the CPU time T_1 (in hours) per 10^8 MC steps, and the CPU time T_2 (in hours) per 10^8 MC sampling points are also given.
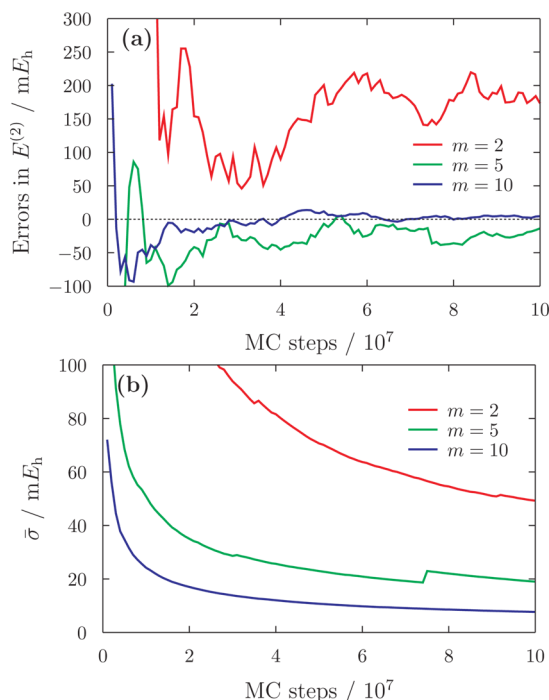

Figure 3. (a) MC-MP2/6-31G** correlation energy $E^{(2)}$ (relative to the MP2/6-31G** value) and (b) the uncertainty $\bar{\sigma}$ (both in mE_h) as a function of the MC steps and the degree of redundancy (m).

$$g_{C_6H_6}(\mathbf{r}) = 4 \sum_{C=1}^{6} g_C(\mathbf{r}) + \sum_{H=1}^{6} g_H(\mathbf{r}) \tag{25}$$

with

$$g_A(\mathbf{r}) = \exp(-\zeta_1 r_A^2) + 0.05 \exp(-\zeta_2 r_A^2) \tag{26}$$

where r_A is the distance from atom A and $(\zeta_1, \zeta_2) = (0.6, 0.15)$ for H and $(0.5, 0.1)$ for C. The prefactor of 4 in eq 25 is the number of valence electrons in the carbon atom. The frozen core approximation is used throughout this work. Note that the plotted $E^{(2)}$ values are the *accumulated* quantity (I_N) up to the Nth MC step and not the instantaneous quantity (I_n) calculated only at the nth MC step. Occasional kinks in the plots of $\bar{\sigma}$ occur owing to large quotient $\bar{f}/w$ where the weight function w is vanishingly small, but some virtual orbitals in $\bar{f}$ are not. It should be reminded that the probability of I_N falling within $4\bar{\sigma}$ of I is 99.8%, assuming a perfect Gaussian distribution of I_n (see Figure 1 for the observed distributions).

Clearly, redundant-walker MC-MP2 with $m = 5$ has much greater efficiency than MC-MP2 with no redundancy ($m = 2$).

Likewise, the results with $m = 10$ display even more rapid convergence than those with $m = 5$. This is ascribed to the increase in the number of MC sampling points by a factor of 10 ($m = 5$) or 45 ($m = 10$) per MC step. Although the redundancy also increases the operation cost per MC step by a factor of 3.1 ($m = 5$) or 8.0 ($m = 10$) (the theoretical values are 2.5 and 5), the benefits far outweigh the additional cost. In fact, the CPU time per 10^8 MC sampling points is reduced from 7.4 h at $m = 2$ to 2.3 h at $m = 5$ or 1.3 h at $m = 10$. Even greater savings in the CPU time to reach a certain precision are expected with larger values of m.

Figure 4 shows the convergence of the correlation energy of a benzene dimer obtained with MC-MP2/6-31G** with two
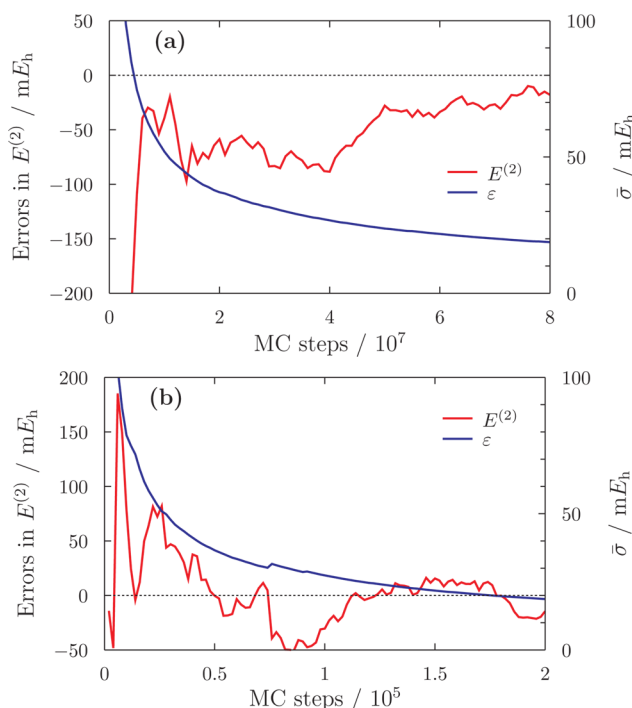

Figure 4. MC-MP2/6-31G** correlation energy $E^{(2)}$ (relative to the MP2/6-31G** value) and the uncertainty $\bar{\sigma}$ (both in mE_h) of the benzene dimer (in an eclipsed configuration with the benzene−benzene distance of 3.8 Å, $r_{CC} = 1.3963$ Å, and $r_{CH} = 1.0826$ Å) as a function of the MC steps with (a) $m = 20$ and (b) $m = 400$. Note the different scales of the horizontal axes of the two plots.

different values of m. The same weight function as eq 25 is placed on each of the two benzene. With $m = 20$, the energy is converged within *ca.* 20 mE_h after 8×10^7 MC steps. Since each step has $_{20}C_2 = 190$ sampling points, this is comparable to nonredundant MC-MP2 ($m = 2$) with 1.5×10^{10} MC steps, which is infeasible. While this is already a significant improvement over $m = 2$, the calculation with $m = 400$ achieves nearly identical convergence with only 2×10^5 MC steps, more than 2 orders of magnitude fewer steps than with $m = 20$. The number of MC sampling points is $_{400}C_2 = 79800$ per MC step or 1.6×10^{10} in total (which explains comparable convergence with $m = 20$ with $N = 8 \times 10^7$). Although a single MC step with $m = 400$ is at least 20 times as expensive as that with $m = 20$, the benefit of the reduction in the number of MC steps by a factor of 400 far exceeds the cost increase.

Figures 5 and 6 plot the correlation corrections to the total energy, highest-occupied MO (HOMO) energy, and lowest-

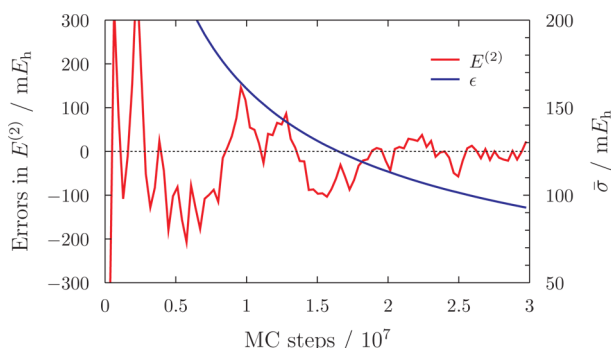unoccupied MO (LUMO) energy of C_{60} as a function of MC steps.

Figure 5. MC-MP2/cc-pVDZ ($m = 120$) correlation energy $E^{(2)}$ (relative to the MP2/cc-pVDZ value) and the uncertainty $\bar{\sigma}$ (both in mE_h) of C_{60} (the diameter of 7.08 Å or $r_{CC} = 1.388$ and 1.452 Å) as a function of the MC steps.
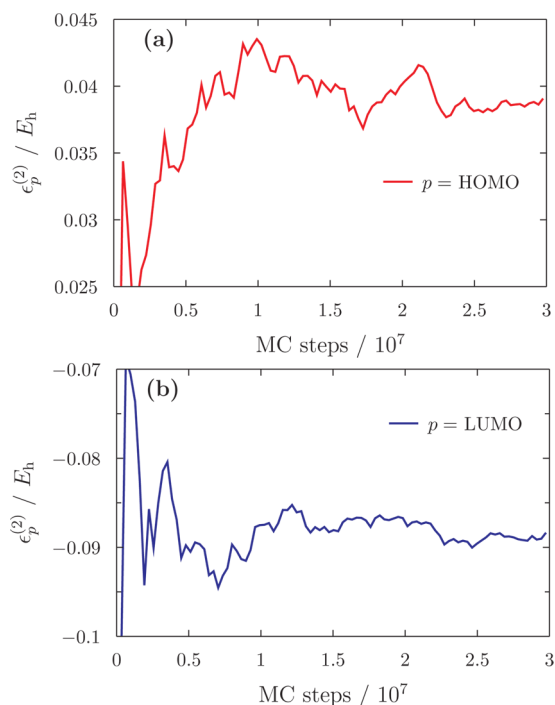

Figure 6. MC-MP2/cc-pVDZ ($m = 120$) correlation corrections (in mE_h) to the HF energies of the (a) HOMO and (b) LUMO of C_{60} as a function of the MC steps.

These are obtained from a single parallel execution of MC-MP2/cc-pVDZ with $m = 120$ using 10 CPU hours of 320 processors in the Blue Waters supercomputer of the National Center for Supercomputing Applications and the following $g(\mathbf{r})$ in the weight function:

$$g_{C_{60}}(\mathbf{r}) = \sum_{C=1}^{60} g_C(\mathbf{r}) \tag{27}$$

where g_C is given by eq 26 with $(\zeta_1, \zeta_2) = (0.5, 0.14)$. Convergence of the MC-MP2 correlation energy toward the correct limit is observed. At 3×10^7 MC steps (which amounts to 2×10^{11} MC sampling points thanks to the 120−fold redundancy), MC-MP2 predicts $E^{(2)} = -7.674\ E_h$ ($\bar{\sigma} = 0.093$

E_h) as compared with the correct limit of $-7.710\ E_h$. The latter value has been obtained with the parallel implementation of the conventional MP2 method in nwchem.[34]

Figure 6 suggests that the MP2 correlation corrections[3] to the energies of HOMO and LUMO are sufficiently converged after 3×10^7 MC steps and are useful for comparison with experimental values. C_{60} and its derivative, phenyl-[6,6]-C_{61}-butyric acid methyl ester (PCBM),[35−38] are widely used as an electron acceptor in bulk heterojunction organic solar cells because of their high electron affinities. Koopmans' approximation in HF theory is well-known to grossly underestimate electron affinities, and C_{60} is no exception. It predicts the value of 0.7 eV as compared with the experimental values of 2.7 eV. MC-MP2 predicts (Table 2) a correlation correction of

Table 2. Ionization Potential and Electron Affinity (in eV) of C_{60}[a]

method	ionization potential	electron affinity
HF/cc-pVDZ[b]	7.8	0.7
MC-MP2/cc-pVDZ[c]	6.8 (0.09)	3.1 (0.09)
experiment	7.61,[d] 7.6[e]	2.7,[f] 2.65[g]

[a]The values in the parentheses are statistical uncertainties $\bar{\sigma}$ (in eV). [b]Koopmans' approximation. [c]A second-order many-body perturbation correction to the Dyson self-energy in the diagonal approximation. [d]Ref 41. [e]Ref 42. [f]Ref 39. [g]Ref 40.

approximately $-0.09\ E_h$ or -2.4 eV to the self-energy of LUMO, bringing the calculated electron affinity to 3.1 eV and in much better agreement with the experimental values.[39,40] The ionization potential of 6.8 eV predicted by MC-MP2 is also consistent with the experimental values,[41,42] although it does not constitute a clear improvement over the HF value (7.8 eV) in this case.

6. CONCLUSION

We have proposed the redundant-walker algorithm that can increase the sampling efficiency of MC-MP2 by many orders of magnitude. We have shown that having a $m/2$-fold redundant population of walkers can lead to an $m(m-1)/2$-fold increase in the number of sampling points at only $m/2$-fold increase in cost, the net performance enhancement by a factor of $(m-1)$. The degree of redundancy (m) is limited only by the core memory space to accommodate the coordinates of m walker pairs. We have considered in this work up to $m = 400$ with a proportionally large performance enhancement, but much greater values of m are allowable.

We have implemented redundant-walker MC-MP2 into a parallel program interfaced with nwchem, the latter furnishing the HF energy, orbital energies, MO's, and AO's. The parallel algorithm consists in nearly completely independent instances of an essentially sequential code running on processors with different random number seeds. Apart from periodic MPI_R-EDUCE calls to accumulate energies and uncertainties across all processors, there are no interprocessor communications. Owing to the inherently naturally parallel nature of this algorithm, a highly scalable code can be written in a short time.

We have demonstrated the efficiency of the proposed method and implemented program for the correlation energies of the benzene molecule and benzene dimer as well as for the correlation corrections to the energy, ionization potential, and electron affinity of C_{60}, the latter obtained from a parallel execution of the aforementioned code on 320 processors of the

Blue Waters supercomputer. The correlation correction brings the calculated electron affinity of 0.7 eV (HF) to 3.1 eV (MC-MP2) and in much better agreement with the experimental value of 2.7 eV. The significance of this calculation lies in the fact that C_{60} is not easily subjected to fragment- or local-basis approach for a calculation speedup. Furthermore, energy differences such as ionization energies are hard to obtain with QMC, whereas they are directly calculable with MC-MP2 applied to self-energies. The demonstrated computational ability to accurately predict electronic properties of C_{60} and even larger organic π-conjugated molecules has considerable contemporary and positive implication for advanced materials research.

■ AUTHOR INFORMATION

Corresponding Author
*E-mail: sohirata@illinois.edu.

Notes
The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

■ REFERENCES

(1) Hirata, S. *J. Phys. Chem. A* **2003**, *107*, 9887−9897.
(2) Willow, S. Y.; Kim, K. S.; Hirata, S. *J. Chem. Phys.* **2012**, *137*, 204122.
(3) Willow, S. Y.; Kim, K. S.; Hirata, S. *J. Chem. Phys.* **2013**, *138*, 164111.
(4) Bartlett, R. J. *Annu. Rev. Phys. Chem.* **1981**, *32*, 359−401.
(5) Szabo, A.; Ostlund, N. S. *Modern Quantum Chemistry*; Dover: New York, 1996.
(6) Helgaker, T.; Jørgensen, P.; Olsen, J. *Molecular Electronic-Structure Theory*; Wiley: Chichester, 2000.
(7) Shavitt, I.; Bartlett, R. J. *Many-Body Methods in Chemistry and Physics*; Cambridge University Press: Cambridge, 2009.
(8) Head-Gordon, M.; Pople, J. A.; Frisch, M. J. *Chem. Phys. Lett.* **1988**, *153*, 503−506.
(9) Frisch, M. J.; Head-Gordon, M.; Pople, J. A. *Chem. Phys. Lett.* **1990**, *166*, 275−280.
(10) Frisch, M. J.; Head-Gordon, M.; Pople, J. A. *Chem. Phys. Lett.* **1990**, *166*, 281−289.
(11) Kohn, W. *Int. J. Quantum Chem.* **1995**, *56*, 229−232.
(12) Ceperley, D. M.; Alder, B. J. *Phys. Rev. Lett.* **1980**, *45*, 566−569.
(13) Hammond, B. L.; Lester Jr., W. A.; Reynolds, P. J. *Monte Carlo Methods in Ab Initio Quantum Chemistry*; World Scientific: Singapore, 1994.
(14) Luchow, A.; Anderson, J. B. *Annu. Rev. Phys. Chem.* **2000**, *51*, 501−526.
(15) Kalos, M. H.; Whitlock, P. A. *Monte Carlo Methods*; Wiley-VCH: Weinheim, 2008.
(16) Kolorenč, J.; Mitas, L. *Rep. Prog. Phys.* **2011**, *74*, 026502.
(17) Zhang, S.; Krakauer, H. *Phys. Rev. Lett.* **2003**, *90*, 136401.
(18) Thom, A.; Alavi, A. *Phys. Rev. Lett.* **2007**, *99*, 143001.
(19) Ohtsuka, Y.; Nagase, S. *Chem. Phys. Lett.* **2008**, *463*, 431−434.
(20) Booth, G. H.; Thom, A. J. W.; Alavi, A. *J. Chem. Phys.* **2009**, *131*, 054106.
(21) Cleland, D.; Booth, G. H.; Alavi, A. *J. Chem. Phys.* **2010**, *132*, 041103.
(22) Petruzielo, F. R.; Holmes, A. A.; Changlani, H. J.; Nightingale, M. P.; Umrigar, C. J. *Phys. Rev. Lett.* **2012**, *109*, 230201.
(23) Shepherd, J. J.; Booth, G. H.; Alavi, A. *J. Chem. Phys.* **2012**, *136*, 244101.
(24) Booth, G. H.; Chan, G. K. L. *J. Chem. Phys.* **2012**, *137*, 191102.
(25) Neuhauser, D.; Rabani, E.; Baer, R. *J. Chem. Theory Comp.* **2013**, *9*, 24−27.
(26) Ten-no, S. *J. Chem. Phys.* **2013**, *138*, 164126.
(27) Kronrod, A. S. *Nodes and Weights of Quadrature Formulas. Sixteen-Place Tables*; Consultants Bureau: New York, 1965.
(28) Obara, S.; Saika, A. *J. Chem. Phys.* **1986**, *84*, 3963−3974.
(29) Metropolis, N.; Rosenbluth, A. W.; Rosenbluth, M. N.; Teller, A. N.; Teller, E. *J. Chem. Phys.* **1953**, *21*, 1087−1092.
(30) Flyvbjerg, H.; Petersen, H. G. *J. Chem. Phys.* **1989**, *91*, 461.
(31) Lee, R. M.; Conduit, G. J.; Nemec, N.; López Ríos, P.; Drummond, N. D. *Phys. Rev. E* **2011**, *83*, 066706.
(32) Ceperley, D.; Alder, B. *Science* **1986**, *231*, 555−560.
(33) Reynolds, P. J.; Ceperley, D. M.; Alder, B. J.; Lester, W. A. *J. Chem. Phys.* **1982**, *77*, 5593−5603.
(34) Valiev, M.; Bylaska, E. J.; Govind, N.; Kowalski, K.; Straatsma, T. P.; Van Dam, H. J. J.; Wang, D.; Nieplocha, J.; Apra, E.; Windus, T. L.; de Jong, W. A. *Comput. Phys. Commun.* **2010**, *181*, 1477−1489.
(35) Gebeyehu, D.; Brabec, C. J.; Padinger, F.; Fromherz, T.; Hummelen, J. C.; Badt, D.; Schindler, H.; Sariciftci, N. S. *Synth. Met.* **2001**, *118*, 1−9.
(36) Shaheen, S. E.; Brabec, C. J.; Sariciftci, N. S.; Padinger, F.; Fromherz, T.; Hummelen, J. C. *Appl. Phys. Lett.* **2001**, *78*, 841−843.
(37) van Duren, J. K. J.; Yang, X. N.; Loos, J.; Bulle-Lieuwma, C. W. T.; Sieval, A. B.; Hummelen, J. C.; Janssen, R. A. J. *Adv. Funct. Mater.* **2004**, *14*, 425−434.
(38) Reyes-Reyes, M.; Kim, K.; Carroll, D. L. *Appl. Phys. Lett.* **2005**, *87*, 083506.
(39) Yang, S. H.; Pettiette, C. L.; Conceicao, J.; Cheshnovsky, O.; Smalley, R. E. *Chem. Phys. Lett.* **1987**, *139*, 233−238.
(40) Wang, L. S.; Conceicao, J.; Jin, C. M.; Smalley, R. E. *Chem. Phys. Lett.* **1991**, *182*, 5−11.
(41) Zimmerman, J. A.; Eyler, J. R.; Bach, S. B. H.; McElvany, S. W. *J. Chem. Phys.* **1991**, *94*, 3556−3562.
(42) Wörgotter, R.; Dünser, B.; Scheier, P.; Märk, T. D. *J. Chem. Phys.* **1994**, *101*, 8674−8679.