Stochastic evaluation of four-component relativistic second-order many-body perturbation energies: A potentially quadratic-scaling correlation method

J. César Cruz,1 Jorge Garza,1 Takeshi Yanai,2,3 and So Hirata4

1) Departamento de Química, División de Ciencias Básicas e Ingeniería, Universidad Autónoma Metropolitana-Iztapalapa, San Rafael Atlixco 186, Col. Vicentina, Iztapalapa, C. P. 09340, Ciudad de México, México
2) Department of Chemistry, Graduate School of Science, Nagoya University, Furocho, Chikusa Ward, Nagoya, Aichi 464-8601, Japan
3) Institute of Transformative Bio-Molecules (WPI-ITbM), Nagoya University, Furocho, Chikusa Ward, Nagoya, Aichi 464-8601, Japan
4) Department of Chemistry, University of Illinois at Urbana-Champaign, 600 South Mathews Avenue, Urbana, Illinois 61801, USA

(Dated: 28 April 2022)

A second-order many-body perturbation correction to the relativistic Dirac–Hartree–Fock energy is evaluated stochastically by integrating 13-dimensional products of four-component spinors and Coulomb potentials. The integration in the real space of electron coordinates is carried out by the Monte Carlo (MC) method with the Metropolis sampling, whereas the MC integration in the imaginary-time domain is performed by the inverse-CDF (cumulative distribution function) method. The computational cost to reach a given relative statistical error for spatially compact but heavy molecules is observed to be no worse than cubic and possibly quadratic with the number of electrons or basis functions. This is a vast improvement over the quintic scaling of the conventional, deterministic second-order many-body perturbation method. The algorithm is also easily and efficiently parallelized with demonstrated 92% strong scalability going from 64 to 4096 processors.

I. INTRODUCTION

The algorithms of ab initio electron-correlation methods are dominated by matrix multiplications. Implemented naively, such multiplications are often nonscalable with respect to either computer size or system size. The scalability with respect to computer size means the extent to which an algorithm lends itself to a facile and efficient parallel execution on a large number of computing processors, which is, nowadays, on the order of millions. The scalability with system size concerns with the size-dependence of the computational cost, which ranges from \(O(n^5)\) for the second-order many-body perturbation (MP2) method\(^4\) to an exponential function of \(n\) for the full configuration-interaction method\(^2\)\(^,\)\(^3\)\(^,\)\(^5\) where \(n\) is a measure of system size such as the number of electrons or of basis functions.

Consider the four-index integral transformation\(^6\)\(^,\)\(^7\). This is a necessary step for most electron-correlation calculations, and is the rate-determining step of MP2\(^8\)\(^,\)\(^9\). In it, two-electron integrals in the atomic-orbital (AO) basis are transformed to those in the molecular-orbital (MO) basis via four consecutive quarter transformations (which are binary matrix multiplications). This is a hard step to program and even harder to parallelize for the following reasons: Neither the AO-based nor MO-based two-electron integrals fit in core memory, and so they (together with the partially transformed integrals) have to be stored on disks at a large I/O cost. The AOs and MOs have opposing characteristics—AOs are spatially local but non-symmetry-adapted, while MOs are spatially extended and symmetry-adapted—hindering their data-local storage across processors, which is prerequisite for efficient parallelization.

Each of the four quarter transformations comes at a \(O(n^5)\) cost, raising a high barrier for ab initio electron-correlation calculations, as compared to preceding reference mean-field calculations, which can be completed at a \(O(n^2)\) to \(O(n^3)\) cost.

There are at least two general approaches to combating the nonscalability: One is to exploit the sparsity of the matrices, such as that brought to by spatially localizing MOs\(^,\)\(^10\)\(^,\)\(^11\). This will divide a large matrix multiplication into many smaller ones, each associated with a local domain, potentially lowering size-dependence of cost all the way down to \(O(n^l)\). Each of the small matrix multiplications can be taken care of by an individual processor with minimal interprocessor communications, exposing efficient parallelism. The price one pays for this remarkable scalability is its limited applicability only to spatially extended large systems (additionally, to those that have local electronic structures, barring, e.g., metals and charge transfer).

The other approach is to not exhaustively use all matrix elements during the multiplication in favor of stochastically sampling only a small portion of them according to their importance\(^,\)\(^12\)\(^,\)\(^13\)\(^,\)\(^14\). For instance, the Monte Carlo MP2 (MC-MP2) method\(^15\)\(^,\)\(^16\)\(^,\)\(^17\)\(^,\)\(^18\) recasts the MP2 energy formula into the sum of two 13-dimensional integrals, which are evaluated by the MC integration method, eradicating the four-index integral computation, transformation, or storage. Each processor can be tasked with a completely independent MC integration in a parallel execution. Only the final results from surviving processors need to be averaged with near perfect parallel efficiency on up to thousands of CPUs (central processing units) or on hundreds of GPUs (graphical processing units)\(^19\). Size-dependence of cost is not known a priori, as it depends on the suitability of weight functions adopted for importance sampling in the high-dimensional space. Our numerical observation\(^19\) indicated that it is \(O(n^3)\), two ranks lower.
than the deterministic MP2’s $O(n^5)$ cost. Note that some incarnations of quantum Monte Carlo (QMC)\textsuperscript{50-57} solve the Schrödinger equation essentially exactly also at $O(n^3)$ cost. The tradeoff for this greater scalability is the introduction of statistical errors.\textsuperscript{58}

It is of interest to study whether the reduced scaling of stochastic algorithms still holds for spatially compact large systems such as heavy-element compounds. They have a large number of electrons and basis functions, but in a small volume, and therefore the local-correlation approach would not work. Stochastic algorithms such as QMC including MC-MP2 could be among the few that can potentially reduce size-dependence of cost in such crowded systems. For heavy-element compounds, it is imperative to include the effects of special theory of relativity.\textsuperscript{59-63} A relativistic MP2 theory based on the Dirac–Hartree–Fock reference is well established\textsuperscript{64-68} and its conventional algorithmic hotspot is still the four-index integral transformation,\textsuperscript{59,69} which is an $O(n^5)$ procedure. The objective of this article is to demonstrate numerically that the Monte Carlo algorithm of the four-component relativistic MP2 method preserves the reduced scaling for such spatially compact, crowded systems by eradicating the integral transformation step. Numerical evidence obtained in this study suggests that the scaling may be as low as $O(n^2)$.

\section{II. FOUR-COMPONENT RELATIVISTIC MP2 THEORY}

Here, we succinctly review only the salient aspects of the four-component relativistic quantum chemistry\textsuperscript{70,71} as a basis of the relativistic MC-MP2 algorithm described in the next section. The Monte Carlo algorithm is not limited to the four-component relativistic approach and should be equally valid for the two-component (A2C),\textsuperscript{22,72} exact two-component (X2C),\textsuperscript{24,72} and quasi-four-component (Q4C)\textsuperscript{76,77} approaches not to mention the one-component and nonrelativistic ones.\textsuperscript{20,23,24}

The Dirac–Coulomb–(Breit) Hamiltonian\textsuperscript{78} for a molecule with $n_{\text{elec}}$ electrons and $n_{\text{nuc}}$ nucleons is written as

$$\hat{H} = \sum_{i=1}^{n_{\text{elec}}} \hat{h}_D(i) + \sum_{i<j}^{n_{\text{elec}}} \hat{g}_{ij} + \sum_{i<j}^{n_{\text{nuc}}} Z_i Z_j \hat{1}_{ij},$$

where $Z_i$ and $\mathbf{R}_i$ are the charge and position, respectively, of the $i$th nucleus, and $\hat{g}_{ij}$ is the electron-electron interaction operator defined later. Here, $\hat{h}_D(i)$ is the one-electron Dirac Hamiltonian for the $i$th electron given by

$$\hat{h}_D(i) = c \alpha \cdot \mathbf{p}(i) + \beta c^2 + V_{\text{nuc}}(i),$$

where $c$ is the speed of light, $\mathbf{p}(i)$ is the momentum operator that is equal to $-i(\partial/\partial x_i, \partial/\partial y_i, \partial/\partial z_i)$, and $V_{\text{nuc}}(i)$ is the nuclear-attraction potential defined later, while $\alpha = (\alpha_1, \alpha_2, \alpha_3)$ and $\beta$ are $4 \times 4$ matrices of the forms,

$$\alpha_k = \begin{pmatrix} 0_2 & \sigma_k \\ \sigma_k & 0_2 \end{pmatrix}, \quad (k = 1, 2, 3),$$

$$\beta = \begin{pmatrix} \mathbf{1}_2 & 0_2 \\ 0_2 & -\mathbf{1}_2 \end{pmatrix},$$

where $\mathbf{1}_2$ and $\mathbf{0}_2$ are the $2 \times 2$ unit and null matrices, respectively, and $\sigma_k (k = 1, 2, 3)$ are the Pauli matrices.\textsuperscript{28}

The nuclear-attraction potential $V_{\text{nuc}}(i)$ in the point-nucleus approximation is given by

$$V_{\text{nuc}}(i) = -\sum_{i}^{n_{\text{nuc}}} \frac{Z_i}{|\mathbf{R}_i - \mathbf{R}_j|}.$$

The interaction operator $\hat{g}_{ij}$ between the $i$th and $j$th electrons is written as

$$\hat{g}_{ij} = \frac{1}{r_{ij}} \left( \frac{\alpha(i) \cdot \alpha(j)}{2r_{ij}} \right) - \frac{\alpha(i) \cdot \mathbf{r}_{ij} \cdot \alpha(j) \cdot \mathbf{r}_{ij}}{2r_{ij}^3},$$

where $\mathbf{r}_{ij} = \mathbf{r}_i - \mathbf{r}_j$, $r_{ij} = |\mathbf{r}_{ij}|$, and $\alpha(i)$ is the $\alpha$ matrix-vector of Eq. (3) acting on the $i$th electron. The first term is the Coulomb interaction, while the other two terms are the Breit interaction\textsuperscript{61,79}. In this work, only the Coulomb interaction was retained with the Breit interaction neglected. As it will become evident in the next section, however, including the Breit interaction will be straightforward in the MC-MPn methods because the associated integrals are evaluated numerically. Likewise, the use of finite nucleus models\textsuperscript{80} should also be much easier in the MC algorithm than in the deterministic ones. These may be considered in a future study.

An eigenfunction (spinor) of $\hat{h}_D(i)$ is a four-component complex vector,

$$\varphi_\mu = \begin{pmatrix} \varphi_\mu^L \\ \varphi_\mu^S \\ \varphi_\mu^\beta \end{pmatrix},$$

where $L$ and $S$ stand for the large and small components of the spinor, respectively, while $\alpha$ and $\beta$ signify spin. Eigenvalues are divided into two groups: One with positive energies ($E > mc^2$) and the other with negative energies ($E < -mc^2$). The positive energies correspond to electrons with large $L$ amplitudes, while the negative energies to positrons with large $S$ amplitudes.

A wave function of the mean-field (Dirac–Hartree–Fock) theory is the Slater determinant of positive-energy spinors occupied by electrons, which minimizes its expectation value of the Dirac–Coulomb–(Breit) Hamiltonian. The spinors are expanded in a four-component basis set as

$$\varphi_\mu = \sum_\mu \begin{pmatrix} C_{\mu L} \chi_{\mu L}^L \\ C_{\mu L} \chi_{\mu L}^S \\ C_{\mu L} \chi_{\mu L}^\beta \\ C_{\mu L} \chi_{\mu L}^\alpha \end{pmatrix},$$

Each component basis function is a product of a radial Gaussian function and spherical harmonic, although the MC algorithm can use any basis function. The $L$ and $S$ components of a basis function have to satisfy the kinetic balance condition.\textsuperscript{60,81}

$$\begin{pmatrix} \chi_{\mu L}^S \\ \chi_{\mu L}^\beta \end{pmatrix} = i(\mathbf{\alpha} \cdot \mathbf{p}) \begin{pmatrix} \chi_{\mu L}^L \\ \chi_{\mu L}^\alpha \end{pmatrix}.$$
to avoid a variational collapse. This causes the $S$-component basis set to be about twice as large as the $L$-component one, although an innovation was introduced by one of the authors that keeps the sizes of these basis sets equal\textsuperscript{82}. In either case, the number of basis functions is typically two to three times greater in a four-component relativistic calculation than in the corresponding nonrelativistic one, making the integral evaluation and subsequent transformation arduous. Time-reversal symmetry renders each eigenvalue at least doubly degenerate for an even number of electrons so that the corresponding two spinors form a Kramers pair\textsuperscript{38}. This along with spatial symmetry (double group symmetry)\textsuperscript{38} can be exploited to reduce the number of unique integrals and thus its computational cost in a deterministic algorithm. The exploitation of symmetry, however, seems incompatible with the MC algorithm.

The zeroth-order Møller–Plesset perturbation (MP0) energy is the sum of occupied spinor energies, which double counts the electron-electron interaction energy and is thus too negative. The sum of the MP0 and first-order (MP1) energies recovers the Dirac–Hartree–Fock energy. The second-order correction, therefore, accounts for the electron-correlation effect for the first time, and is given by\textsuperscript{64,65,78}

$$E^{(2)} = \frac{1}{2} \sum_{\text{occ.}} \sum_{i,j} \epsilon_i \epsilon_j - \epsilon_a \epsilon_b, \quad (10)$$

where $i$ and $j$ run over all occupied positive-energy spinors, while $a$ and $b$ over all virtual positive-energy spinors, and $\epsilon_p$ is the $p$th spinor energy. That the negative-energy spinors do not enter the summations is called the no-pair approximation\textsuperscript{79,83,84}. For an even number of electrons, this formula can be compressed considerably by using time-reversal symmetry and turned into the so-called Kramers-restricted formula.\textsuperscript{64–66} We do not consider this here because, again, the MC algorithm seems incompatible with symmetry advantages. There are other fast MP2 methods such as the resolution-of-identity\textsuperscript{85} and local approximations\textsuperscript{86} which should be compatible with and complementary to the MC algorithm.

In Eq. (10), $(ia|jb)$ is a two-electron integral in the Dirac–Hartree–Fock spinor basis, defined by\textsuperscript{87}

$$(ia|jb) = \int dr_1 dr_2 \psi_i^+(r_1) \psi_a(r_1) \bar{g}_{12} \psi_j^*(r_2) \psi_b(r_2), \quad (11)$$

which is related to the same in the AO basis functions via MO coefficients \(C_{i,p}^{L} \) and \(C_{i,p}^{S} \) as

$$
(ia|jb) = \sum_{\mu} \sum_{j} \sum_{k} \sum_{l} C_{i,\mu}^{L} C_{\mu,1}^{L} C_{1,\mu}^{S} \langle \mu | L^a | \mu \rangle \langle \mu | L^b | \mu \rangle + \sum_{\mu} \sum_{j} \sum_{k} \sum_{l} C_{i,\mu}^{L} C_{\mu,k}^{S} C_{k,\mu}^{S} \langle \mu | L^a | \mu \rangle \langle \mu | L^b | \mu \rangle + \sum_{\mu} \sum_{j} \sum_{k} \sum_{l} C_{i,\mu}^{L} C_{\mu,k}^{S} C_{k,\mu}^{S} \langle \mu | L^a | \mu \rangle \langle \mu | L^b | \mu \rangle + \ldots \quad (12)$$

with

$$
(\mu | L^a | \mu ) = \int dr_1 dr_2 \chi_\mu^+ (r_1) \chi_\mu (r_2) \bar{g}_{12} \chi_\mu (r_2) \chi_\mu^+ (r_1), \quad (13)
$$

e tc., although the actual implementation of this step is more compact and efficient than these equations may imply.\textsuperscript{35} Nonetheless, the four-index integral transformation [Eq. (12)] is the algorithmic hotspot of the four-component relativistic MP2, having a $O(n^5)$ cost, if we assume $n \propto n_{\text{elec}} \propto n_{\text{bas}}$. In the MC algorithm, this step is altogether eliminated along with the need to store original, partially, or fully transformed integrals. This algorithm is now described.

III. MONTE CARLO RELATIVISTIC MP2 ALGORITHM

Using the following Laplace transform\textsuperscript{88,89}

$$\frac{1}{\epsilon_i + \epsilon_j - \epsilon_a - \epsilon_b} = - \int_0^\infty d\tau \exp \left\{ (\epsilon_i + \epsilon_j - \epsilon_a - \epsilon_b)\tau \right\}, \quad (14)$$

we convert Eq. (10) into the sum of two 13-dimensional integrals\textsuperscript{20} i.e.,

$$
E^{(2)} = -2 \int d\tau_1 dr_1 dr_2 dr_3 dr_4 \int_0^\infty d\tau \ \frac{1}{r_{12} r_{34}}
\times \left\{ \left[ G^*(r_1, r_2, -\tau) \right]^T \cdot \left[ G^*(r_3, r_1, \tau) \right] \right\}
\times \left\{ \left[ G^*(r_2, r_4, -\tau) \right]^T \cdot \left[ G^*(r_4, r_2, \tau) \right] \right\}
+ \int d\tau_1 dr_1 dr_2 dr_3 dr_4 \int_0^\infty d\tau \ \frac{1}{r_{12} r_{34}}
\times \left\{ \left[ G^*(r_1, r_3, -\tau) \right]^T \cdot \left[ G^*(r_4, r_1, \tau) \right] \right\}
\times \left\{ \left[ G^*(r_2, r_4, -\tau) \right]^T \cdot \left[ G^*(r_3, r_2, \tau) \right] \right\}, \quad (15)
$$

where $\tau$ has the physical meaning of imaginary time, and vec[...] is the vectorization operator that reshapes a $m \times n$ matrix into a $mn \times 1$ column vector. Green’s function traces,
\[
G^+(r_d, r_o, \tau) = \sum \varphi_a(r_d) \varphi_a^*(r_o) \exp(-\epsilon_a \tau)
\]
\[
G^-(r_d, r_o, \tau) = \sum \varphi_i(r_d) \varphi_i^*(r_o) \exp(-\epsilon_i \tau)
\]
\[
G^z(r_d, r_o, \tau) = \sum \psi_a(r_d) \psi_a^*(r_o) \exp(-\epsilon_a \tau)
\]

(16)

(17)

For example, one of the factors in Eq. (15) is more explicitly written as

\[
\{\text{vec} [G^+(r_1, r_3, -\tau)]^T \cdot \text{vec} [G^+(r_3, r_1, \tau)]\}
\]
\[
= O^{\lambda a \beta}(r_1, r_3, -\tau)V^{\lambda a \lambda}(r_3, r_1, \tau)
\]
\[
+ O^{\alpha \alpha \beta}(r_1, r_3, -\tau)V^{\alpha \lambda \beta}(r_3, r_1, \tau)
\]
\[
+ O^{\alpha \lambda \alpha}(r_1, r_3, -\tau)V^{\lambda \alpha \beta}(r_3, r_1, \tau)
\]
\[
+ \ldots \ (12 \text{ terms})
\]
\[
+ O^{2\beta \beta \beta}(r_1, r_3, -\tau)V^{2\beta \beta \beta}(r_3, r_1, \tau)
\]

(18)

with

\[
O^{\lambda a \beta}(r_1, r_3, -\tau) = \sum \varphi_i^*(r_1) \varphi_i^\beta(r_3) \exp(\epsilon_i \tau),
\]

(19)

\[
V^{\lambda a \beta}(r_3, r_1, \tau) = \sum a \varphi_a(r_3) \varphi_a^\beta(r_1) \exp(-\epsilon_a \tau),
\]

(20)

e etc. The other factors of Eq. (15) and the remaining 15 \(O\) and 15 \(V\) functions are defined analogously.

Rewriting Eq. (15) as

\[
E^{(2)} = \iiint d \tau d r_1 d r_2 d r_3 d r_4 \int_0^\infty d \tau f(r_1, r_2, r_3, r_4, \tau)
\]

(21)

where \(f\) is the sum of the two integrands, we evaluate the right-hand side by the Monte Carlo (MC) method\(^{20,42}\) as

\[
E^{(2)} \approx \frac{1}{N} \sum_{n=1}^N \frac{f(r_1^{[n]}, r_2^{[n]}, r_3^{[n]}, r_4^{[n]}, \tau^{[n]})}{w(r_1^{[n]}, r_2^{[n]}, r_3^{[n]}, r_4^{[n]}, \tau^{[n]})} \equiv I_N,
\]

(22)

where \(I_N\) is the MC estimate of the integral at the \(N\)th MC step. Electron-pair coordinates at the \(n\)th MC step, \((r_1^{[n]}, r_2^{[n]}, r_3^{[n]}, r_4^{[n]}, \tau^{[n]}),\) are coupled through the Coulomb potential \(1/r_{12}\) and are distributed in the six-dimensional space randomly but according to a judiciously chosen weight function, \(w(r_1^{[n]}, r_2^{[n]}).\) Electron-pair coordinates, \((r_3^{[n]}, r_4^{[n]}),\) are essentially the same, whereas \(w(r_1^{[n]}, \ldots, r_4^{[n]}, \tau^{[n]})\) dictates a random distribution of the imaginary-time coordinate, \(\tau^{[n]},\) at the \(n\)th MC step.\(^{20,23,42}\)

We shall reuse the weight function\(^{20,23,42}\) that have been found effective for the nonrelativistic case. The weight function for electron pairs is the integrand of a two-center Coulomb-repulsion integral for the sum of contracted \(s\)-type Gaussian-type orbitals:

\[
w(r_1^{[n]}, r_2^{[n]}, r_3^{[n]}, r_4^{[n]}) = \frac{g(r_1) g(r_2)}{r_{12}}
\]

(23)

with

\[
g(r) = \sum_{j=1}^n \left[ c_1(j) N_e(j) \exp\left(-\zeta_1(j) |r - R_j|^2\right) \right] + c_2(j) N_e(j) \exp\left(-\zeta_2(j) |r - R_j|^2\right),
\]

(24)

where \(N_e\) is the normalization coefficient of each primitive Gaussian-type orbital. The overall normalization factor,

\[
N_e = \int dr_1 dr_2 \frac{g(r_1) g(r_2)}{r_{12}},
\]

(25)

can be evaluated analytically.\(^{20}\) A random distribution of the electron-pair walkers according to this weight function is achieved by the Metropolis algorithm.\(^{21}\) It may be noticed that the Coulomb potentials, \(1/r_{12}\) and \((1/r_{34}),\) in the integrand \(f,\) which have singularities everywhere, are canceled out by the same in the weight functions.

The weight function in the \(\tau\) dimension is given by\(^{22}\)

\[
w(\tau) = \lambda \exp(-\lambda \tau),
\]

(26)

which is normalized. The exponent \(\lambda\) is parametrized by the energies of the highest-occupied and lowest-unoccupied molecular orbitals (HOMO and LUMO) as

\[
\lambda = 2(\epsilon_{\text{LUMO}} - \epsilon_{\text{HOMO}}).
\]

(27)

A random distribution of the imaginary-time walker according to the weight function, Eq. (26), can be directly generated by the inverse-CDF (cumulative distribution function) method (i.e., without any rejected moves of the Metropolis algorithm). In this case, the inverse CDF can furthermore be obtained analytically.\(^{22,44}\)

The redundant-walker algorithm\(^{28}\) provides an essential speedup for the MC-MP2 algorithm and is invoked in the relativistic MC-MP2 also. It propagates \(m \gg 2\) electron pairs (walkers), \((r_1^{[n]}, r_2^{[n]}), (1 \leq p \leq m),\) and permute their coordinates in all \(mC_2\) distinct ways when evaluating the quotient of Eq. (15), i.e.,

\[
E^{(2)} \approx \frac{1}{N} \sum_{n=1}^N I'_n \equiv I_N
\]

(28)

with

\[
I'_n = \frac{1}{m!} \sum_{p=1}^m \frac{f(r_1^{[n]}, r_2^{[n]}, r_3^{[n]}, r_4^{[n]}, \tau^{[n]})}{w(r_1^{[n]}, r_2^{[n]}, r_3^{[n]}, r_4^{[n]}, \tau^{[n]})},
\]

(29)
It achieves a theoretical speedup by a factor of \((m-1)\) because the sampling is boosted by a factor of \(nC_2\) at an \(m/2\)-fold increase in the cost of walker propagation, although the rate of speed up does not go up indefinitely with \(m\) owing to saturation. See Ref. [32] for the observed performance of this algorithm.

The statistical uncertainty \(\bar{\sigma}_N\) in \(I_N\) is given by

\[
\bar{\sigma}_N = \frac{\sigma_N}{\sqrt{N/N_b}}
\]

where \(\sigma_N\) is the standard deviation concurrently accumulated (excluding the initial burn-in steps) as

\[
\sigma^2_N = \frac{N_b}{N} \sum_{n=1}^{N/N_b} \left[ \frac{1}{N_b} \sum_{n'=1}^{N_b} I_{(n-1)N_b+n'} - I_N \right]^2.
\]

Here, \(N_b\) is the block size in the Flyvbjerg–Petersen algorithm\(^{22}\) preventing underestimation of the standard deviation due to autocorrelations.

The whole algorithm can be easily and efficiently parallelized\(^{23}\) simply, each processor is tasked to perform its own independent MC integration using a distinct random number seed. It periodically records the number of MC steps, \(I_N\), and \(\bar{\sigma}_N\), with no interprocessor communications whatsoever. Only at the end of the calculation, weighted averages of the integrals from all surviving processors need to be taken. In this way, the algorithm is guaranteed to achieve near 100\% parallel efficiency, be fault-tolerant, be asynchronous, and be indefinitely restartable with a variable number of processors.

### IV. NUMERICAL DEMONSTRATIONS

For eight molecules, the relativistic four-component MP2 energy was obtained using the MC method: \(\text{H}_2\) (0.74 Å), \(\text{H}_2\text{O}\) (\(r_{\text{OH}} = 0.96\) Å, \(r_{\text{HH}} = 1.51\) Å), \(\text{CuH}\) (1.54 Å), \(\text{AgH}\) (1.70 Å), \(\text{AuH}\) (1.57 Å), \(\text{Cu}_2\) (2.40 Å), \(\text{Ag}_2\) (2.68 Å), and \(\text{Au}_2\) (2.59 Å). The basis functions used were contracted Gaussian-type orbitals\(^{22,24}\) \(\text{H}\) [8s2p]/[3s2p], \(\text{O}\) [14s9p2d]/[5s4p2d], \(\text{Cu}\) [18s14d]/[8s6p3d], \(\text{Ag}\) [25s21p12d]/[11s9p4d], and \(\text{Au}\) [29s25p15d10f]/[11s8p5d3f]. The weight functions are defined in Table I. The number of walkers \(m\) in Eq. (29) used are given below. The Flyvbjerg–Petersen block size of \(N_b = 100\) was used. A Dirac–Hartree–Fock calculation must precede the MC-MP2 step and was performed deterministically with the UTChem software\(^{25}\), furnishing the MO coefficients and orbital energies. The benchmark deterministic relativistic MP2 calculations were also carried out with UTChem. The MC-MP2 calculations were executed on an Intel core i9-10900K with 5.3 GHz and on the Cori (Haswell) supercomputer of NERSC.

In Table II are presented the relativistic MC-MP2 energies \((E^{(2)})\) for the eight molecules mentioned above, using eight electron-pair walkers \((m = 8)\) at \(10^4\) MC steps. The MC algorithm can reliably reproduce the deterministic results within the estimated statistical uncertainty \(\bar{\sigma}\) even with the relatively short MC runs. Although the MC-MP2 algorithm might most resemble the variational Monte Carlo method among the QMC variety, its energy is not bounded from below by the correct MP2 energy and can be more negative than the latter. The uncertainties systematically increase with the number of electrons. Figure 1 shows a typical convergence behavior of \(E^{(2)}\) as a function of the number of MC steps for \(\text{AgH}\) using four different numbers of the electron-pair walkers \((m = 6, 8, 10,\) and \(12)\). Figure 2 plots the corresponding statistical uncertainty \((\bar{\sigma})\). They attest to the acceleration of convergence with increasing \(m\) by the redundant-walker algorithm.

| Atom  | \(c_1\) | \(c_2\) | \(c_3\) | \(c_4\) |
|-------|--------|--------|--------|--------|
| H     | 0.25   | 0.06   | 0.15   | 0.6    |
| O     | 0.8    | 0.2    | 1.0    | 0.4    |
| Cu    | 0.8    | 0.35   | 2.0    | 0.6    |
| Ag    | 0.1    | 0.1    | 0.8    | 0.6    |
| Au    | 0.05   | 0.6    | 4.0    | 0.8    |

**Table I.** The weight-function parameters [Eq. (23)].

| Molecule | \(E^{(2)}/E_h\) | \(\bar{\sigma}/E_h\) |
|----------|-----------------|---------------------|
| \(\text{H}_2\) | -0.0184 | -0.0183 | 0.0008 |
| \(\text{H}_2\text{O}\) | -0.2252 | -0.2248 | 0.0106 |
| \(\text{CuH}\) | -0.4618 | -0.4634 | 0.0198 |
| \(\text{AgH}\) | -0.3212 | -0.3243 | 0.0152 |
| \(\text{AuH}\) | -0.8424 | -0.8441 | 0.0325 |
| \(\text{Cu}_2\) | -0.8918 | -0.8931 | 0.0371 |
| \(\text{Ag}_2\) | -0.6172 | -0.6146 | 0.0235 |
| \(\text{Au}_2\) | -3.0345 | -3.0316 | 0.1432 |

**Table II.** The relativistic second-order energies, \(E^{(2)}\), and statistical uncertainties, \(\bar{\sigma}\), obtained by the MC-MP2 method after \(10^5\) MC steps with eight electron-pair walkers \((m = 8)\). The benchmark \(E^{(2)}\) obtained by the deterministic MP2 method (as implemented in UTChem) are also shown.

**FIG. 1.** The error in \(E_h\) in the relativistic MC-MP2 energy of \(\text{AgH}\) from the deterministic value as a function of the number of MC steps. Each for four traces uses a different number \((m)\) of electron-pair walkers.
The statistical uncertainty \( \sigma \) (in \( mE_h \)) in the relativistic MC-MP2 energy of AgH as a function of the number of MC steps. Each for four traces uses a different number \((m)\) of electron-pair walkers.

Wall time (in seconds) required for the relativistic MC-MP2 calculation to reach the relative statistical uncertainty \((\sigma/E^{(2)})\) of 0.1 as a function of the number of basis functions, \(n_{bas}\). Two different values of the electron-pair walkers \((m = 8 \text{ and } 12)\) are considered. Lines corresponding to quadratic and cubic cost scaling \(O(n_{bas}^2)\) and \(O(n_{bas}^3)\) are also plotted to guide the eyes.

Number of MC steps taken in 30 minutes of wall clock time as a function of the number of processors executing the MC-MP2 calculation for Au\(_2\) \((m = 8)\) on the Cori (Haswell) supercomputer of NERSC.
ing the number of processors by a factor of 64 (from 64 to 4096 processors), we have observed a 59-fold speedup, i.e., strong scalability of 92% for a fixed job size. This high parallel efficiency is expected because the algorithm is inherently parallel with no interprocessor communications. We can even let the calculation be terminated by the job scheduler with no loss of data or waste of computational efforts. The data from various calculations can furthermore be concatenated to produce more converged energies. Some points in this plot are slightly above the line of perfect scalability relative to 64 processors, displaying small superscaling behavior.

V. SUMMARY

The Monte Carlo four-component relativistic MP2 algorithm has been developed. It is scalable with respect to both system size and computer size. For crowded, but spatially compact molecules with heavy elements, where the local-correlation approach cannot apply, this stochastic algorithm displays tentative $O(n^3_{ele})$ scaling and certainly no worse than $O(n^5_{ele})$, which is a vast improvement over the $O(n^5_{ele})$ scaling of the conventional MP2 algorithms. It can be implemented easily into a nearly perfectly scalable parallel algorithm with no interprocessor communications, which is also fault-tolerant, asynchronous, and indefinitely restartable. We furthermore expect that it lends itself to facile extensions to the Breit interactions, finite nucleus effects, and explicit correlation.

VI. DATA AVAILABILITY STATEMENT

A dataset is retrievable as a compressed file at https://hirata-lab.chemistry.illinois.edu/CruzJCP2022.zip. It contains UTChem basis set and output files as well as spinor expansion coefficients for the eight molecules considered in this study. It also compiles the MC-MP2 output files for the energies and statistical uncertainties as a function of blocked MC steps for the eight molecules. The raw data used for plotting Figures 1 through 4 are also included.

ACKNOWLEDGMENTS

J. C. Cruz thanks CONACYT for the scholarship 620190. T. Yanai thanks JSPS KAKENHI (Grant No. 21H01881 and 21K18931) for financial support. S. Hirata thanks the Center for Scalable, Predictive methods for Excitation and Correlated phenomena (SPEC), which is funded by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, Chemical Sciences, Geosciences, and Biosciences Division, as a part of the Computational Chemical Sciences Program. S. Hirata also thanks Grant No. DE-SC0006028 funded by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences. This research used resources of the National Energy Research Scientific Computing Center (NERSC), a

U.S. Department of Energy Office of Science User Facility located at Lawrence Berkeley National Laboratory, operated under Contract No. DE-AC02-05CH11231 using NERSC award m3196 (2022). The initial phase of this work was performed at Institute for Molecular Science, Okazaki, Japan.

1 C. Møller and M. S. Plessert, “Note on an approximation treatment for many-electron systems,” Phys. Rev. 46, 618–622 (1934).
2 P. Knowles and N. Handy, “A new determinant-based full configuration interaction method,” Chem. Phys. Lett. 111, 315–321 (1984).
3 G. H. Booth, A. J. Thom, and A. Alavi, “Fermion Monte Carlo without interprocessor communications, which is also fault-tolerant, asynchronous, and indefinitely restartable. We furthermore expect that it lends itself to facile extensions to the Breit interactions, finite nucleus effects, and explicit correlation.

VI. DATA AVAILABILITY STATEMENT

A dataset is retrievable as a compressed file at https://hirata-lab.chemistry.illinois.edu/CruzJCP2022.zip. It contains UTChem basis set and output files as well as spinor expansion coefficients for the eight molecules considered in this study. It also compiles the MC-MP2 output files for the energies and statistical uncertainties as a function of blocked MC steps for the eight molecules. The raw data used for plotting Figures 1 through 4 are also included.

ACKNOWLEDGMENTS

J. C. Cruz thanks CONACYT for the scholarship 620190. T. Yanai thanks JSPS KAKENHI (Grant No. 21H01881 and 21K18931) for financial support. S. Hirata thanks the Center for Scalable, Predictive methods for Excitation and Correlated phenomena (SPEC), which is funded by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, Chemical Sciences, Geosciences, and Biosciences Division, as a part of the Computational Chemical Sciences Program. S. Hirata also thanks Grant No. DE-SC0006028 funded by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences. This research used resources of the National Energy Research Scientific Computing Center (NERSC), a

U.S. Department of Energy Office of Science User Facility located at Lawrence Berkeley National Laboratory, operated under Contract No. DE-AC02-05CH11231 using NERSC award m3196 (2022). The initial phase of this work was performed at Institute for Molecular Science, Okazaki, Japan.

1 C. Møller and M. S. Plessert, “Note on an approximation treatment for many-electron systems,” Phys. Rev. 46, 618–622 (1934).
D. Peng, W. Liu, Y. Xiao, and L. Cheng, “Making four- and two-component relativistic density functional methods fully equivalent based on the idea of “from atoms to molecule”, “J. Chem. Phys. 127, 104106 (2007)

K. G. Dyall and F. Knut, Introduction to relativistic quantum chemistry (Oxford University Press, Inc., New York, 2007).

W. Liu, “Advances in relativistic molecular quantum mechanics,” Phys. Rep. 537, 59–89 (2014), advances in Relativistic Molecular Quantum Mechanics.

M. Abe, T. Suzuki, Y. Fujii, and M. Hada, “An ab initio study based on a finite nucleus model for isotope fractionation in the U(III)–U(IV) exchange reaction system,” J. Chem. Phys. 128, 144309 (2008).

Y. Ishikawa, “Basis-set expansion calculations with the Dirac Hamiltonian,” Int. J. Quantum Chem. 26, 375–383 (1984).

T. Yanai, T. Nakajima, Y. Ishikawa, and K. Hirao, “A new computational scheme for the Dirac–Hartree–Fock method employing an efficient integral algorithm,” J. Chem. Phys. 114, 6526–6538 (2001).

J. Sucher, “Foundations of the relativistic theory of many-electron atoms,” Phys. Rev. A 22, 348–362 (1980).

J. Autschbach, “Perspective: Relativistic effects,” J. Chem. Phys. 136, 130902 (2012).

D. E. Bernholdt and R. J. Harrison, “Large-scale correlated electronic structure calculations: The RI-MP2 method on parallel computers,” Chem. Phys. Lett. 250, 477–484 (1996).

M. Schutz, G. Hetzer, and H. J. Werner, “Low-order scaling local electron correlation methods. I. linear scaling local MP2,” J. Chem. Phys. 111, 5691–5705 (1999).

A. Almoukhalalati, S. Knecht, H. J. A. Jensen, K. G. Dyall, and T. Saue, “Electron correlation within the relativistic no-pair approximation,” J. Chem. Phys. 145, 074104 (2016).

J. Almlöf, “Elimination of energy denominators in Möller–Plesset perturbation theory by a Laplace transform approach,” Chem. Phys. Lett. 181, 319–320 (1991).

M. Häser and J. Almlöf, “Laplace transform techniques in Möller–Plesset perturbation theory,” J. Chem. Phys. 96, 489–494 (1992).

S. Obara and A. Saika, “Efficient recursive computation of molecular integrals over Cartesian Gaussian functions,” J. Chem. Phys. 84, 3963 (1986).

N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, “Equation of state calculations by fast computing machines,” J. Chem. Phys. 21, 1087–1092 (1953).

A. E. Doran and S. Hirata, “Monte Carlo second- and third-order many-body Green’s function methods with frequency-dependent, nondiagonal self-energy,” J. Chem. Theory Comput. 15, 6097 (2019).

H. Flyvbjerg and H. G. Petersen, “Error estimates on averages of correlated data,” J. Chem. Phys. 91, 461–466 (1989).

T. Yanai, H. Iikura, T. Nakajima, Y. Ishikawa, and K. Hirao, “A new implementation of four-component relativistic density functional method for heavy-atom polyatomic systems,” J. Chem. Phys. 115, 8267–8273 (2001).

T. Yanai, H. Nakano, T. Nakajima, T. Tsuneda, S. Hirata, Y. Kawashima, Y. Nakao, M. Kamiya, H. Sekino, and K. Hirao, “UTChem: A program for ab initio quantum chemistry,” in Lecture Notes in Computer Science (Springer, Berlin, 2003) pp. 84–95.

S. Ten-no and D. Yamaki, “Communication: Explicitly correlated four-component relativistic second-order Möller–Plesset perturbation theory,” J. Chem. Phys. 137 (2012).