JCTC Journal of Chemical Theory and Computation

pubs.acs.org/JCTC Article

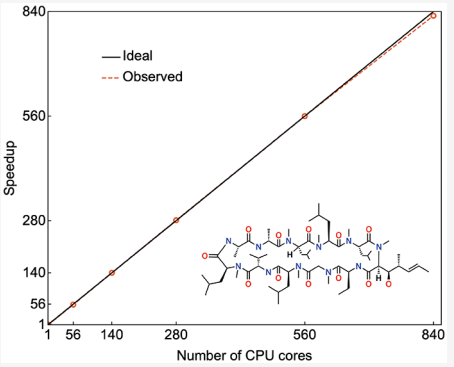
Q-MP2-OS: Møller–Plesset Correlation Energy by Quadrature

Giuseppe M. J. Barca,* Simon C. McKenzie, Nathaniel J. Bloomfield, Andrew T. B. Gilbert, and Peter M. W. Gill*

Cite This: *J. Chem. Theory Comput.* 2020, 16, 1568–1577 Read Online

ACCESS | Metrics & More | Article Recommendations

ABSTRACT: We present a quadrature-based algorithm for computing the opposite-spin component of the MP2 correlation energy which scales quadratically with basis set size and is well-suited to large-scale parallelization. The key ideas, which are rooted in the earlier work of Hirata and co-workers, are to abandon all two-electron integrals, recast the energy as a seven-dimensional integral, approximate that integral by quadrature, and employ a cutoff strategy to minimize the number of intermediate quantities. We discuss our implementation in detail and show that it parallelizes almost perfectly on 840 cores for cyclosporine (a molecule with roughly 200 atoms), exhibits $O(N^2)$ scaling for a sequence of polyglycines, and is principally limited by the accuracy of its quadrature.



1. INTRODUCTION

Møller–Plesset (MP) perturbation theory[1] provides one of the simplest, and most widely used, corrections to the Hartree–Fock (HF) approximation. Despite being introduced to modern quantum chemistry almost 50 years ago,[2] it remains popular today, both in its own right and as a component of double-hybrid density functional theory (DFT) methods.

Its strengths and weaknesses have been thoroughly documented[3−9] and are now well understood. As a result, it is known that, when applied to large-gap systems with medium-size basis sets, the second-order theory (MP2) gives a useful estimate of the correlation energy at a relatively modest computational cost.

Unfortunately, that cost formally scales as $O(N^5)$, where N is the size of the basis set, and in the 1990s, systematic efforts were made to reduce this steep scaling so that MP2 theory could be applied to large systems. Many ideas were explored but the exploitation of the Laplace representation by Almlöf and Häser,[10,11] the application of resolution-of-the-identity (RI) methods by Feyereisen et al.,[12] the examination of pseudospectral techniques by Martínez and Carter,[13] and the systematic use of multipole approximations by Hetzer, Pulay, and Werner[14] are particularly notable.

Over the past decade, driven by the tantalizing potential of petascale computers, such ideas have been supplemented with parallelizable versions of the theory[15−22] and these enabled, for example, the RI-MP2/cc-pVTZ energy of a molecule with 240 atoms to be calculated in less than 5 min on the K supercomputer.[20]

Various groups have explored the advantages of approximating the two-electron integrals in terms of lower-rank quantities. The seminal work of Feyereisen et al. has already been mentioned, but much more recently, Ishimura and Ten-no expressed the two-electron integrals in the molecular orbital (MO) basis as the product of the potential and of the charge distribution of virtual-occupied MO pairs and then evaluated such products by quadrature, as needed.[23] In a related vein, Song and Martínez expanded each integral in the atomic orbital (AO) basis as the tensor product of five second-rank objects and exploited this representation to obtain an algorithm with $O(N^3)$ scaling.[24,25]

Hirata and co-workers have taken an even more extreme path toward the use of low-rank quantities, recasting the MP2 energy as the stochastically estimated integral of a high-dimensional but trivial function.[26−28]

Finally, we note that, by revisiting the screening problem, the Ochsenfeld group has developed a highly efficient scheme for avoiding the evaluation of negligible two-electron AO integrals, eventually obtaining an implementation whose cost scales linearly with N.[29]

Looking to the future, it seems certain that useful MP2 schemes will need to embody many of these ideas: the use of low-rank ingredients to minimize data flow, the application of

Received: November 17, 2019
Published: January 23, 2020

https://dx.doi.org/10.1021/acs.jctc.9b01142
J. Chem. Theory Comput. 2020, 16, 1568–1577

aggressive cutoffs to achieve low cost scaling, and the adoption of strategies with high inherent parallelizability. In the present paper, we discuss a new algorithm for the opposite-spin (OS) component of the MP2 energy which embraces this philosophy. The underlying theory is discussed in section 2, an efficient implementation is discussed in section 3, parallelization, scaling, and accuracy results are discussed in section 4, and conclusions are discussed in section 5. Unless stated otherwise, atomic units are used throughout.

2. THEORY

The total MP2 correlation energy can be separated into opposite- and same-spin components

$$E^{(2)} = E^{(2)}_{\mathrm{OS}} + E^{(2)}_{\mathrm{SS}} \tag{1}$$

and, using Mulliken notation for integrals, the opposite-spin (OS) component can be written

$$E^{(2)}_{\mathrm{OS}} = -\sum_{i}^{N_\mathrm{o}}\sum_{j}^{\bar{N}_\mathrm{o}}\sum_{a}^{N_\mathrm{v}}\sum_{b}^{\bar{N}_\mathrm{v}} \frac{(ia|\bar{j}\,\bar{b})(ia|\bar{j}\,\bar{b})}{\epsilon_a + \bar{\epsilon}_b - \epsilon_i - \bar{\epsilon}_j} \tag{2}$$

where N_o and N_v are the numbers of occupied and virtual α-spin orbitals, ϵ_i and ϵ_a are the occupied and virtual α-spin orbital energies, and the overbar is used to indicate the β-spin counterparts.

The two-electron integrals in the MO basis are related to those in the AO basis by the transformation

$$(ia|\bar{j}\,\bar{b}) = \sum_{\mu\nu\lambda\sigma}^{N} C_{\mu i} C_{\nu a} \bar{C}_{\lambda j} \bar{C}_{\sigma b} (\mu\nu|\lambda\sigma) \tag{3}$$

where C and $\bar{C}$ are the α and β MO coefficient matrices, N is the number of AO basis functions, and the AO integrals are given by

$$(\mu\nu|\lambda\sigma) = \iint \frac{\phi_\mu(\mathbf{r}_1)\,\phi_\nu(\mathbf{r}_1)\,\phi_\lambda(\mathbf{r}_2)\,\phi_\sigma(\mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|}\,\mathrm{d}\mathbf{r}_1\,\mathrm{d}\mathbf{r}_2 \tag{4}$$

Substituting the Laplace–Almlöf representation

$$\frac{1}{\epsilon_a + \bar{\epsilon}_b - \epsilon_i - \bar{\epsilon}_j} = \int_0^\infty \exp[(-\epsilon_a - \bar{\epsilon}_b + \epsilon_i + \bar{\epsilon}_j)t]\,\mathrm{d}t \tag{5}$$

the opposite-spin energy can be written

$$E^{(2)}_{\mathrm{OS}} = -\sum_{\mu\nu\lambda\sigma}^{N}\sum_{\alpha\beta\gamma\delta}^{N} \int_0^\infty X_{\mu\alpha}(t)\,Y_{\nu\beta}(t)\,\bar{X}_{\lambda\gamma}(t)\,\bar{Y}_{\sigma\delta}(t)\,(\mu\nu|\lambda\sigma)(\alpha\beta|\gamma\delta)\,\mathrm{d}t \tag{6}$$

where the density-matrix-like functions are

$$X_{\mu\alpha}(t) = \sum_{i}^{N_\mathrm{o}} C_{\mu i} C_{\alpha i} \exp(+\epsilon_i t) \tag{7}$$

$$Y_{\nu\beta}(t) = \sum_{a}^{N_\mathrm{v}} C_{\nu a} C_{\beta a} \exp(-\epsilon_a t) \tag{8}$$

$$\bar{X}_{\lambda\gamma}(t) = \sum_{j}^{\bar{N}_\mathrm{o}} \bar{C}_{\lambda j} \bar{C}_{\gamma j} \exp(+\bar{\epsilon}_j t) \tag{9}$$

$$\bar{Y}_{\sigma\delta}(t) = \sum_{b}^{\bar{N}_\mathrm{v}} \bar{C}_{\sigma b} \bar{C}_{\delta b} \exp(-\bar{\epsilon}_b t) \tag{10}$$

Integrating over one of the electrons in each of the two-electron integrals in eq 6, we obtain the seven-dimensional integral

$$E^{(2)}_{\mathrm{OS}} = -\sum_{\mu\nu\lambda\sigma}^{N}\sum_{\alpha\beta\gamma\delta}^{N} \int_0^\infty X_{\mu\alpha}(t)\,Y_{\nu\beta}(t)\,\bar{X}_{\lambda\gamma}(t)\,\bar{Y}_{\sigma\delta}(t)\,\mathrm{d}t \times \int D_{\mu\nu}(\mathbf{r}_1)\,U_{\lambda\sigma}(\mathbf{r}_1)\,\mathrm{d}\mathbf{r}_1 \int U_{\alpha\beta}(\mathbf{r}_2)\,D_{\gamma\delta}(\mathbf{r}_2)\,\mathrm{d}\mathbf{r}_2 \tag{11}$$

where

$$D_{\mu\nu}(\mathbf{r}) = \phi_\mu(\mathbf{r})\,\phi_\nu(\mathbf{r}) \tag{12}$$

$$U_{\lambda\sigma}(\mathbf{r}) = \int \frac{\phi_\lambda(\mathbf{r}')\,\phi_\sigma(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}\,\mathrm{d}\mathbf{r}' \tag{13}$$

Equation 11 is analogous to the 13-dimensional integral of Willow et al.[26] except that it is cast in terms of AO rather than MO quantities. Moreover, in the spirit of Friesner's pseudospectral method,[30] the Coulomb operators have been replaced with potentials making the resulting integrand smoother and more amenable to numerical integration.

We now introduce a quadrature for the t integration

$$\frac{1}{\epsilon_a + \bar{\epsilon}_b - \epsilon_i - \bar{\epsilon}_j} \approx \sum_{k}^{K} \omega_k \exp[(-\epsilon_a - \bar{\epsilon}_b + \epsilon_i + \bar{\epsilon}_j)t_k] \tag{14}$$

and define discretized versions of the X and Y matrices

$$X^k_{\mu\alpha} = {\omega_k}^{1/4} X_{\mu\alpha}(t_k) \qquad \bar{X}^k_{\lambda\gamma} = {\omega_k}^{1/4} \bar{X}_{\lambda\gamma}(t_k) \tag{15}$$

$$Y^k_{\nu\beta} = {\omega_k}^{1/4} Y_{\nu\beta}(t_k) \qquad \bar{Y}^k_{\sigma\delta} = {\omega_k}^{1/4} \bar{Y}_{\sigma\delta}(t_k) \tag{16}$$

Likewise, we introduce a quadrature for the spatial integrations over the density-potential products

$$\int D_{\mu\nu}(\mathbf{r})\,U_{\lambda\sigma}(\mathbf{r})\,\mathrm{d}\mathbf{r} \approx \sum_{p}^{G} D^p_{\mu\nu} U^p_{\lambda\sigma} \tag{17}$$

where

$$D^p_{\mu\nu} = \sqrt{w_p}\,D_{\mu\nu}(\mathbf{r}_p) \tag{18}$$

$$U^p_{\lambda\sigma} = \sqrt{w_p}\,U_{\lambda\sigma}(\mathbf{r}_p) \tag{19}$$

and w_p are the quadrature weights.

Combining these quadratures with eq 11 yields the 11-fold sum

$$E^{(2)}_{\mathrm{OS}} \approx -\sum_{k}^{K}\sum_{\mu\nu\lambda\sigma}^{N}\sum_{\alpha\beta\gamma\delta}^{N} X^k_{\mu\alpha} Y^k_{\nu\beta} \bar{X}^k_{\lambda\gamma} \bar{Y}^k_{\sigma\delta} \sum_{p}^{G} D^p_{\mu\nu} U^p_{\lambda\sigma} \sum_{q}^{G} U^q_{\alpha\beta} D^q_{\gamma\delta} \tag{20}$$

which can be reordered to give

$$E^{(2)}_{\mathrm{OS}} \approx -\sum_{k}^{K}\sum_{pq}^{G}\sum_{\alpha\nu}^{N} \left(\sum_{\mu}^{N} X^k_{\mu\alpha} D^p_{\mu\nu}\right)\left(\sum_{\beta}^{N} Y^k_{\nu\beta} U^q_{\alpha\beta}\right) \sum_{\lambda\delta}^{N} \left(\sum_{\gamma}^{N} \bar{X}^k_{\gamma\lambda} D^q_{\gamma\delta}\right)\left(\sum_{\sigma}^{N} \bar{Y}^k_{\delta\sigma} U^p_{\lambda\sigma}\right) \tag{21}$$

https://dx.doi.org/10.1021/acs.jctc.9b01142
J. Chem. Theory Comput. 2020, 16, 1568−1577

$$\approx -\sum_{k}^{K}\sum_{pq}^{G}\left(\sum_{\alpha\nu}^{N} S_{\alpha\nu}^{kp} V_{\nu\alpha}^{kq}\right)\left(\sum_{\lambda\delta}^{N} \bar{S}_{\lambda\delta}^{kq} \bar{V}_{\delta\lambda}^{kp}\right) \tag{22}$$

$$\approx -\sum_{k}^{K}\sum_{pq}^{G} T_{pq}^{k} \bar{T}_{qp}^{k} \tag{23}$$

where we have introduced

$$S_{\alpha\nu}^{kp} = \sum_{\mu}^{N} X_{\mu\alpha}^{k} D_{\mu\nu}^{p} \tag{24}$$

$$V_{\nu\alpha}^{kq} = \sum_{\beta}^{N} Y_{\nu\beta}^{k} U_{\alpha\beta}^{q} \tag{25}$$

$$T_{pq}^{k} = \sum_{\nu\alpha}^{N} S_{\alpha\nu}^{kp} V_{\nu\alpha}^{kq} \tag{26}$$

Equations 21−23 encode a quadrature-based scheme, which we term Q-MP2-OS, for calculating the opposite-spin MP2 energy.

In addition to the connection to pseudospectral methods, Q-MP2-OS shares similarities with the tensor hypercontraction (THC) approach to computing the opposite-spin MP2 energy. Both rely on the Laplace technique to deal with the orbital energy denominator, and both use grids to decompose the two-electron integral tensor into products of lower-rank objects. They differ, however, on what intermediates are formed and what strategies are used to realize an efficient implementation. THC employs density fitting and partitions the molecule into blocks to allow for localized inversion of the fitting-metric matrix. This reduces the formal scaling to $O(N^3)$, although subcubic, $O(N^{2.5})$, scaling has been demonstrated for systems up to 10 000 basis functions.[25] Optimized, molecule-dependent grids have also been developed to reduce the number of grid points and, thus, the cost prefactor.[31] Q-MP2-OS, on the other hand, relies exclusively on cutoff strategies to reduce the scaling to quadratic, as discussed in section 3.

3. IMPLEMENTATION

In this section, we describe an efficient implementation of our Q-MP2-OS method. Section 3.1 considers the grids adopted for the quadratures introduced in eqs 14 and 17. Section 3.2 introduces several screening techniques that significantly reduce the computational cost. Our algorithm is presented in section 3.3.

3.1. Quadratures. *3.1.1. t Quadrature.* For the t quadrature in eq 14 we use the results of Takatsuka et al.,[32] who fit $1/x$ to a sum of exponential functions

$$\frac{1}{x} \approx \sum_{k}^{K} \omega_k \exp[-\alpha_k x] \tag{27}$$

The roots, α_k, and weights, ω_k, are determined by nonlinear minimax optimizations in the intervals $x \in [1, R]$ for several values of R. A simple scaling of these roots and weights by $1/A$ allows the resulting quadratures to be applied to arbitrary intervals $y = Ax \in [A, AR]$.

We adopt the frozen-core and frozen-virtual approximation and denote the lowest and highest energies of the active orbitals by $\epsilon_{\min}$ and $\epsilon_{\max}$. In principle, we are concerned with the range $[E_{\min}, E_{\max}]$, where

$$E_{\min} = 2(\epsilon_{\mathrm{HOMO}} - \epsilon_{\mathrm{LUMO}}) = 2\Delta \tag{28}$$

$$E_{\max} = 2(\epsilon_{\max} - \epsilon_{\min}) \tag{29}$$

In practice, for the systems considered here, we found little benefit in tuning the domain using these values and, for all calculations, we used a $K = 6$ formula scaled to the range $[0.05, 20]$.

3.1.2. r Quadrature. Molecular quadratures arise frequently in density functional theory (DFT),[33−43] and we take inspiration from this community to develop a quadrature grid for use in Q-MP2-OS, which we denote as MG ("minimal grid").

The molecular grid comprises atom-centered grids, each of which is a product of radial and angular grids defined by Euler−Maclaurin[34] and Lebedev[44,45] (EML) quadrature rules, respectively. Becke weights[33] were used to avoid multiple counting. Adopting the standard grid (SG) approach,[35,41,43] we systematically pruned each Lebedev grid in order to minimize the total number of grid points.

A single MG grid was developed and used for H, C, N, and O atoms. Reference data were $E_{\mathrm{OS}}^{(2)}$ ionization potentials, electron affinities, and atomization energies of the H_nX hydrides, and our accuracy criterion was 1 kcal mol^{-1}. Under these constraints, the optimal parent grid was found to be EML(14,38) with an element-independent Bragg−Slater radius of 1.23. Starting with the core, the 14 Lebedev grids were pruned to

$$\{6, 6, 6, 6, 18, 18, 18, 38, 38, 38, 0, 0, 0, 0\}$$

The resulting MG atomic grid has only 192 points, far fewer than (for example) the 3816 points of the carbon SG-1 atomic grid.

3.2. Screening. Formally, the computational costs of eqs 24 and 25 scale as $O(GN^3)$ and eq 26 scales as $O(G^2N^2)$. In order to reduce these quartic costs, we invoke several screening strategies that avoid the computation of negligible quantities and reduce the scaling of computing the Q-MP2-OS energy to only $O(GN)$.

3.2.1. pq-Screening. In a 1959 study on the one-dimensional Schrödinger equation with a periodic and symmetric potential, Kohn proved[46] that the Wannier functions for such a system decay exponentially and, therefore, the related first-order density matrix $\rho(\mathbf{r}_1, \mathbf{r}_2)$ also decays exponentially. This phenomenon, now often known as the "nearsightedness" of the first-order density matrix, extends not only to n-dimensional crystals and more general potentials, but also to molecular insulators. In particular, for large insulating systems, it has been established[47−59] that the first-order density matrix behaves as

$$\rho(\mathbf{r}_1, \mathbf{r}_2) \sim \exp(-\Lambda r_{12}) \tag{30}$$

where $r_{12} = |\mathbf{r}_1 - \mathbf{r}_2|$ and $\Lambda \propto \sqrt{\Delta}$.[56,60,61]

If we consider the generalized first-order matrix

$$\Xi(\mathbf{r}_1, \mathbf{r}_2, t) = \sum_{\mu\nu}^{N} X_{\mu\nu}(t)\, \phi_\mu(\mathbf{r}_1)\, \phi_\nu(\mathbf{r}_2) \tag{31}$$

https://dx.doi.org/10.1021/acs.jctc.9b01142
J. Chem. Theory Comput. 2020, 16, 1568−1577

we note that $\Xi(\mathbf{r}_1,\mathbf{r}_2,0) = \rho(\mathbf{r}_1,\mathbf{r}_2)$ and observe that $\Xi(\mathbf{r}_1,\mathbf{r}_2,t)$ preserves the exponential decay with respect to r_{12} even for $t \neq 0$, i.e.

$$\Xi(\mathbf{r}_1, \mathbf{r}_2, t) \sim \exp(-\Lambda r_{12}) \tag{32}$$

Equation 32 has important implications for our algorithm. In fact, using eq 31, the OS energy in eq 21 can be recast in the form

$$E_{\mathrm{OS}}^{(2)} \approx -\sum_{pq}^{G} E_{pq} = -\sum_{pq}^{G}\sum_{k}^{K} \Xi(\mathbf{r}_p, \mathbf{r}_q, t_k)\, \Upsilon(\mathbf{r}_p, \mathbf{r}_q, t_k) \tag{33}$$

where $\Upsilon(\mathbf{r}_p, \mathbf{r}_q, t_k)$ is a bounded function with respect to $r_{pq} = |\mathbf{r}_p - \mathbf{r}_q|$. Equations 32 and 33 therefore imply

$$E_{pq} \sim \exp(-\Lambda r_{pq}) \tag{34}$$

which shows that the magnitude of the contributions E_{pq} to the $E_{\mathrm{OS}}^{(2)}$ energy arising from a pair of quadrature points $(\mathbf{r}_p, \mathbf{r}_q)$ decays exponentially with r_{pq}.

Based on eq 34, pq pairs that are sufficiently far apart can be screened and the evaluation of their corresponding E_{pq} contributions skipped. However, screening pq pairs is inefficient, and instead, we partition the total molecular grid into batches P and screen PQ pairs. For simplicity, we choose a batch to be the EML grid centered on a nucleus.

Figure 1 plots the PQ-pair energy

$$E_{PQ} = \sum_{p\in P}\sum_{q\in Q} E_{pq} \tag{35}$$

against the internuclear distance, $r_{PQ} = |\mathbf{r}_P - \mathbf{r}_Q|$, for the linear polyglycine Gly_{18}. Exponential decay is clear.

For each PQ pair, we define a critical radius

$$r_{\mathrm{crit}}^{PQ} = r_{PQ} + \frac{\ln[E_{PQ}/\tau]}{\Lambda} \tag{36}$$

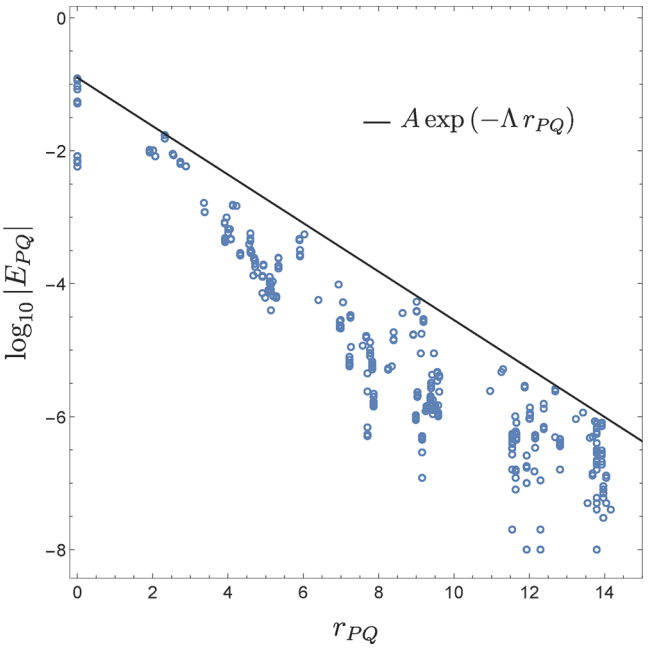


Figure 1. Exponential decay of the energy contributions E_{PQ} arising from PQ pairs in Gly_{18} using the 6-31G* basis set. $\Lambda = 0.840$.

where τ is a user-defined accuracy threshold, while the exponential decay rate, Λ, is an unknown. We determine the parameter Λ by using an adaptive procedure to compute the line of maximum steepness that envelopes all $\log(E_{PQ})$ values as they are evaluated. For given Λ and τ values, there exists a global critical radius, r_{crit}, for which all PQ-batch pairs with $r_{PQ} > r_{\mathrm{crit}}$ can be neglected. We estimate r_{crit} by taking the maximum of the r_{crit}^{PQ} values.

This pq-screening reduces the number of pq pairs that have to be considered from $O(G^2)$ to $O(G)$. Algorithmic details concerning the pq-screening strategy are reported in section 3.3.

3.2.2. k-Screening. The time required to compute the Q-MP2-OS energy is directly proportional to the number K of quadrature roots t_k. For a given PQ batch pair, the contribution to the energy from t_k is

$$E_{PQ}^{k} = \sum_{p\in P}\sum_{q\in Q} T_{pq}^{k}\bar{T}_{qp}^{k} \tag{37}$$

If the roots are in increasing order, one finds that $|E_{PQ}^{k}|$ decays quickly with respect to k. Figure 2 demonstrates this for

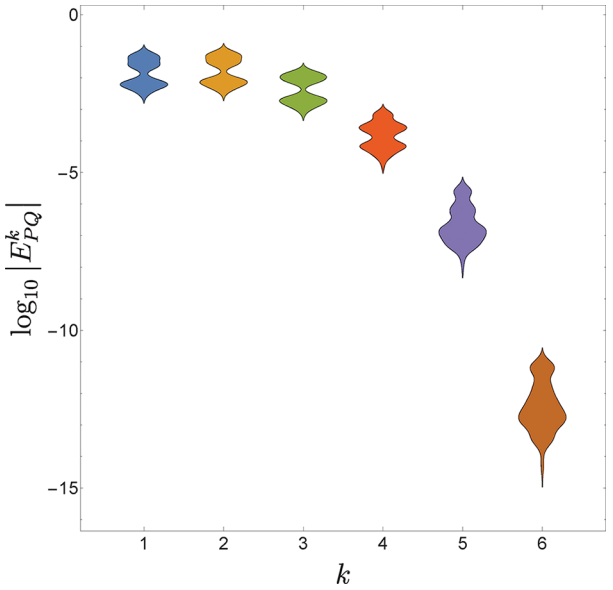


Figure 2. Violin distributions[62] of the energy contributions E_{PQ}^{k} for each PQ pair and t_k in Gly_3 using the 6-31G* basis set.

Gly_3 and shows that the largest t_k can sometimes be skipped during the calculation on a given PQ batch pair. For safety, we always compute the first three t_k's. Subsequently, if the energy contribution E_{PQ}^{k} falls below the threshold τ, we skip all remaining t_k's. In practice, a threshold of 10^{-6} often results in t_6 being skipped.

3.2.3. μν-Screening. Shell-pair screening[63] is employed to reduce the number of shell pairs from $O(N^2)$ to $O(N)$. A shell pair is considered significant if its Gaussian product prefactor

$$G_{\mu\nu} = \exp\left[-\frac{\alpha\beta}{\alpha+\beta}|\mathbf{A}-\mathbf{B}|^2\right] \tag{38}$$

falls above a given cutoff threshold. We use parentheses to denote sums over only significant shell pairs, e.g. $\sum_{(\mu\nu)}$.

3.2.4. (μν)-Screening on P. Further screening is possible because only $O(1)$ shell pairs are significant at a given batch P

https://dx.doi.org/10.1021/acs.jctc.9b01142
J. Chem. Theory Comput. 2020, 16, 1568−1577

of grid points. Thus, for that batch, the $D^{P}_{\mu\nu}$ matrix contains only $O(1)$ significant elements. These elements correspond to basis-function pairs significant on batch P that we denote as $(\mu\nu)_P$ and which satisfy the following equation:

$$\sum_{p}^{P} |D^{p}_{\mu\nu}| \geq \tau \tag{39}$$

3.2.5. αν-Screening. The Kohn decay discussed in section 3.2.1 causes the $S^{kp}_{\alpha\nu}$ to decay exponentially with the distance between α and ν. However, we have found that even more $\alpha\nu$ pairs can be neglected, and indeed, if we keep only the $\alpha\nu$ pairs obtained using traditional shell-pair screening, there is little loss of accuracy in the energy.

To demonstrate this, we consider the error

$$\Delta E_{\alpha\nu} = E^{(2)}_{\mathrm{OS}} - \left[-\sum_{k}^{K} \sum_{pq}^{G} \left(\sum_{\chi\kappa \neq \alpha\nu}^{N} S^{kp}_{\chi\kappa} V^{kq}_{\kappa\chi} \right) \left(\sum_{\lambda\delta \neq \alpha\nu}^{N} \bar{S}^{kq}_{\lambda\delta} \bar{V}^{kp}_{\delta\lambda} \right) \right] \tag{40}$$

introduced by neglecting a single $\alpha\nu$ pair. These errors are plotted against the Gaussian product $G_{\alpha\nu}$ prefactor in Figure 3

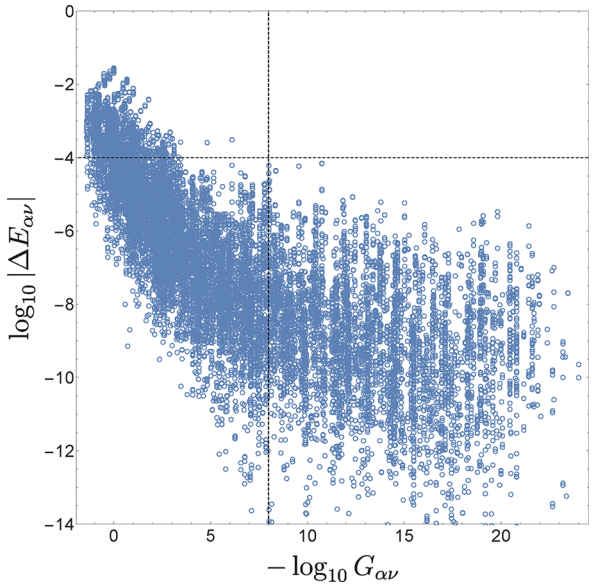


Figure 3. Relationship between energy errors $\Delta E_{\alpha\nu}$ and Gaussian prefactors $G_{\alpha\nu}$ in Gly_3 using an uncontracted 3-21G basis set. Dashed lines show that an $\alpha\nu$ shell-pair cutoff of 10^{-8} yields an energy error of $10^{-4} E_h$.

for triglycine (Gly_3) using the uncontracted 3-21G basis set. The scatter plot initially decays rapidly but then decays much more slowly. As a result, an error of $10^{-4} E_h$ in the energy can be achieved using a shell-pair cutoff threshold of 10^{-8} but, if one desires greater accuracy, all $\alpha\nu$ pairs need to be included.

3.3. Algorithm. Pseudocode for computing the Q-MP2-OS energy is shown in the algorithm in Scheme 1. Our algorithm begins by computing the $X^{k}_{\mu\alpha}$, $\bar{X}^{k}_{\lambda\gamma}$, $Y^{k}_{\nu\beta}$, and $\bar{Y}^{k}_{\sigma\delta}$ matrices via eqs 15 and 16, and stores these in main memory. Each CPU core computes and stores these quantities as the computational time is negligible compared to the remainder of the algorithm. Although our implementation does not do so, these read-only quantities could be computed only once per node. Significant shell-pair data are also distributed to each node.

Scheme 1. Algorithm for Calculating the Q-MP2-OS Energy

1 $E^{(2)}_{\mathrm{OS}} = 0$
2 **forall** k **do**
3 $\quad X^{k}_{\mu\alpha}, \bar{X}^{k}_{\lambda\gamma} \leftarrow$ Eq. (15)
4 $\quad Y^{k}_{\nu\beta}, \bar{Y}^{k}_{\sigma\delta} \leftarrow$ Eq. (16)
5 **end**
6 **forall** P **do**
7 $\quad$ **forall** Q **do**
8 $\quad\quad r_{PQ} = |\mathbf{r}_P - \mathbf{r}_Q|$
9 $\quad$ **end**
10 **end**
11 $PQ \leftarrow$ sort (P, Q) by decreasing r_{PQ}
12 $r_{\mathrm{crit}} = 0$
13 **foreach** CPU core **do**
14 $\quad$ **do**
15 $\quad\quad (P, Q) = \mathrm{pop}(PQ)$
16 $\quad\quad E_{PQ} = 0$
17 $\quad\quad D^{p}_{(\mu\nu)}, U^{p}_{(\mu\nu)} \leftarrow$ Eqs (18,19) $\quad \forall p \in P$
18 $\quad\quad D^{q}_{(\mu\nu)}, U^{q}_{(\mu\nu)} \leftarrow$ Eqs (18,19) $\quad \forall q \in Q$
19 $\quad\quad (\mu\nu)_P \leftarrow$ Form sig $\mu\nu$ on P
20 $\quad\quad (\mu\nu)_Q \leftarrow$ Form sig $\mu\nu$ on Q
21 $\quad\quad$ **forall** k **do**
22 $\quad\quad\quad S^{p}_{(\alpha\nu)_P} = \sum_{\mu} X^{k}_{\mu\alpha} D^{p}_{(\mu\nu)_P}$
23 $\quad\quad\quad V^{q}_{(\nu\alpha)_P} = \sum_{\beta} Y^{k}_{\nu\beta} U^{q}_{(\alpha\beta)_P}$
24 $\quad\quad\quad T_{pq} = \sum_{\nu\alpha} S^{p}_{(\alpha\nu)_P} V^{q}_{(\nu\alpha)_P}$
25 $\quad\quad\quad \bar{S}^{q}_{(\lambda\delta)_Q} = \sum_{\gamma} \bar{X}^{k}_{\gamma\lambda} D^{q}_{(\gamma\delta)_Q}$
26 $\quad\quad\quad \bar{V}^{p}_{(\delta\lambda)_Q} = \sum_{\sigma} \bar{Y}^{k}_{\delta\sigma} U^{p}_{(\lambda\sigma)_Q}$
27 $\quad\quad\quad \bar{T}_{qp} = \sum_{\lambda\delta} \bar{S}^{q}_{(\lambda\delta)_Q} \bar{V}^{p}_{(\delta\lambda)_Q}$
28 $\quad\quad\quad E^{k}_{PQ} = \sum_{pq} T_{pq} \bar{T}_{qp}$
29 $\quad\quad\quad E_{PQ} \mathrel{+}= E^{k}_{PQ}$
30 $\quad\quad\quad$ **if** $|E^{k}_{PQ}| \leq \tau$ **then**
31 $\quad\quad\quad\quad$ **break**
32 $\quad\quad\quad$ **end**
33 $\quad\quad$ **end**
34 $\quad\quad r^{PQ}_{\mathrm{crit}} \leftarrow$ Eq. (36)
35 $\quad\quad r_{\mathrm{crit}} = \max\left[r_{\mathrm{crit}}, r^{PQ}_{\mathrm{crit}} \right]$
36 $\quad\quad E^{(2)}_{\mathrm{OS}} \mathrel{+}= E_{PQ}$
37 $\quad$ **while** $r_{PQ} < r_{\mathrm{crit}}$;
38 **end**

Lines 6−10 of the algorithm in Scheme 1 form a list of all the PQ batch-pair distances, which are simply the internuclear distances. The PQ pairs are then sorted (line 11) by decreasing

https://dx.doi.org/10.1021/acs.jctc.9b01142
J. Chem. Theory Comput. 2020, 16, 1568−1577

value of r_{PQ}. Although trivial, this sort is important because, as discussed in section 3.2.1, the E_{PQ} values decay exponentially with r_{PQ}. We also initialize our estimate of r_{crit} to 0.

Once the PQ list has been formed, each CPU core pops a PQ pair from the bottom of the list, where the PQ pairs with the smallest r_{PQ} values reside.

Each core then forms the $D^{p}_{(\mu\nu)}$, $U^{p}_{(\mu\nu)}$, $D^{q}_{(\mu\nu)}$, and $U^{q}_{(\mu\nu)}$ quantities for its PQ pair, considering only the $O(N)$ significant $(\mu\nu)$ pairs. The storage requirement for each matrix is $O(N) \times O(1)$ (the number of grid points within a batch is constant), and the total FLOP cost of this step for the entire calculation is $O(GN)$, i.e., quadratic with system size.

In lines 22−27 of the algorithm in Scheme 1, each core forms the $S^{p}_{(\alpha\nu)}$, $V^{q}_{(\nu\alpha)}$, T_{pq} and the $\overline{S}^{q}_{(\lambda\delta)}$, $\overline{V}^{q}_{(\delta\lambda)}$, $\overline{T}_{qp}$ matrices for its PQ pair. The t_k's are treated sequentially, allowing reuse of the memory associated with these quantities.

By exploiting the screening described in section 3.2.4, we form S (Scheme 1, line 22) by summing over only the shells associated with significant shell pairs on P, i.e., $(\mu\nu)_P$. Furthermore, by invoking the $\alpha\nu$-screening, we need to form S elements only for $(\alpha\nu)_P$, which we assume to be equivalent to the $(\mu\nu)_P$ list. This assumption becomes less valid with the addition of diffuse functions, and to retain the desired target accuracy, one must tighten the threshold τ in eq 39.

Immediately after forming S, we form V. However, looking ahead to line 24 (Scheme 1), we note that the summation is restricted to $(\alpha\nu)_P$ because $S^{p}_{(\alpha\nu)_P}$ must be significant. For this reason, we only need to form the V elements for $(\alpha\nu)_P$ and we can therefore invoke the $\mu\nu$-screening to sum over only the $(\alpha\beta)_P$ elements of U, on line 23. The storage required for the $S^{p}_{(\alpha\nu)_P}$ and $V^{q}_{(\alpha\nu)_P}$ matrices is only $O(1)$, and their total FLOP cost is only $O(G)$.

On line 24 (Scheme 1), we sum over $\alpha\nu$ pairs to obtain T_{pq}, which is done via a BLAS matrix−matrix multiplication call. Once the T_{pq} and $\overline{T}_{qp}$ matrices are formed, they are summed to give E^{k}_{PQ}, which is the fraction of the E_{PQ} energy arising from the current t_k. In this way, the E^{k}_{PQ} components are accumulated in the batch-pair energy E_{PQ} until either all t_k's have been treated or a specific $|E^{k}_{PQ}|$ becomes smaller than a user-defined threshold τ, triggering the k-screening discussed in section 3.2.2. Each CPU process computes its E_{PQ} contribution to the energy independently of the others.

Finally, we compute the current PQ pair's critical radius, r^{PQ}_{crit}, using eq 36 and estimate the global critical radius, r_{crit}, as the maximum of its previous value and the new r^{PQ}_{crit}. If r_{PQ} is larger than r_{crit}, we gather the energies from all outstanding CPU cores and terminate the calculation; if not, we pop another PQ pair from the bottom of the PQ list and assign it to a now idle CPU core.

4. RESULTS

In this section, we present speedup, scaling, and accuracy results for our Q-MP2-OS implementation in a development version of Q-Chem.[64] All calculations used the following default cutoff thresholds for the screenings described in section 3.2: (i) 10^{-6} for both the pq- and k-screening, (ii) 10^{-10} for both the $\mu\nu$- and $\alpha\nu$-screening, and (iii) 10^{-12} for the $D^{p}_{\mu\nu}$-screening.

Unless otherwise specified, all results were obtained using the 6-31G* basis set and the MG quadrature grids described in section 3.1.2. Furthermore, for all our comparisons with RI-MP2-OS,[65] we employed cc-pVDZ-RI as the density fitting basis.[66]

All calculations were run on the Raijin supercomputer at the Australian National Computational Infrastructure, using Intel Xeon E5-2690v4 @ 2.60 GHz (Broadwell) processors.

4.1. Speedup. Figure 4 shows the speedup (strong scaling) of Q-MP2-OS with respect to the number of cores for

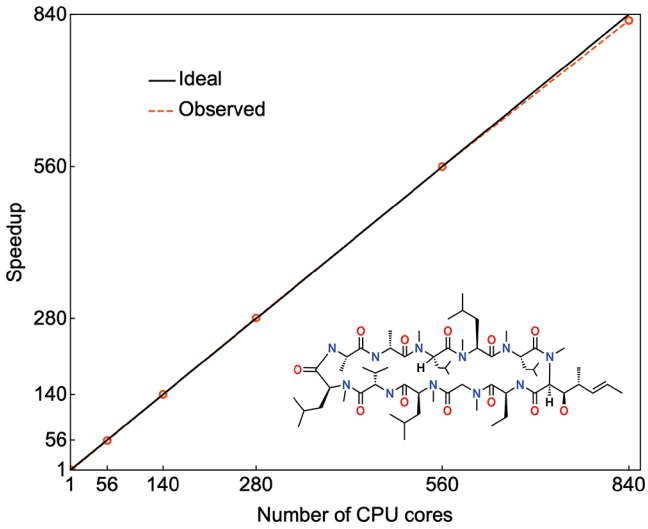


Figure 4. Speedup of the Q-MP2-OS implementation with respect to number of CPU cores for the cyclosporine molecule ($C_{62}H_{111}N_{11}O_{12}$).

cyclosporine ($C_{62}H_{111}N_{11}O_{12}$). These results highlight one of the main strengths of our approach: the Q-MP2-OS algorithm exhibits near-perfect parallelizability.

The calculation performed with 840 cores achieves ∼99% of the ideal speedup, and the 1% deterioration for this calculation arises, not because of increased communication between the CPU processes, but because the number of cores becomes too large compared to the number of significant PQ pairs in this molecule. In such circumstances, each core receives only a few (4−5) PQ pairs and the parallel work load becomes unbalanced.

4.2. Computational Scaling. Figure 5 shows wall times and computational scalings of Q-MP2-OS and its major algorithmic components for a sequence of polyglycines (Gly_n, n = 3, ..., 48). These results show the costs of the D and U matrices are $O(N^2)$, as discussed in section 3.3, and the costs of the S, V, and T matrices are $O(N)$.

For the polyglycines considered, the bottleneck of Q-MP2-OS is the formation of the S and V intermediates and, hence, the total computational time (labeled with black circles in Figure 5) grows subquadratically. However, for sufficiently large molecules, the computation of the U matrices will become the computational bottleneck and the Q-MP2-OS method will scale quadratically with system size.

Figure 6 compares the computational costs of RI-MP2-OS and Q-MP2-OS calculations on Gly_{42}, Gly_{45}, and Gly_{48} using a single core. The RI-MP2-OS[65] calculations were performed using the implementation in Q-Chem and show quartic scaling. The Q-MP2-OS calculations, on the other hand, scale quadratically, but with a large prefactor. These contrasting behaviors lead to a crossover near Gly_{47}, and for larger

https://dx.doi.org/10.1021/acs.jctc.9b01142
J. Chem. Theory Comput. 2020, 16, 1568−1577

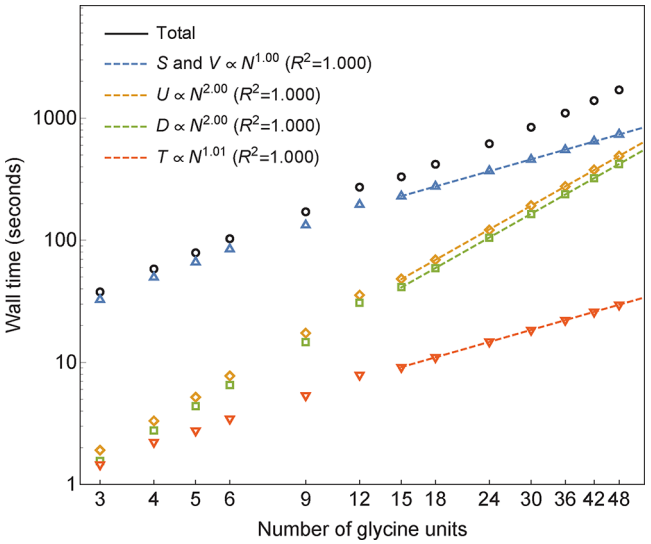


Figure 5. Wall times and computational scaling of our Q-MP2-OS implementation and of its major algorithmic components, for a sequence of polyglycines (Gly_n). Calculations were parallelized over 56 CPU cores.

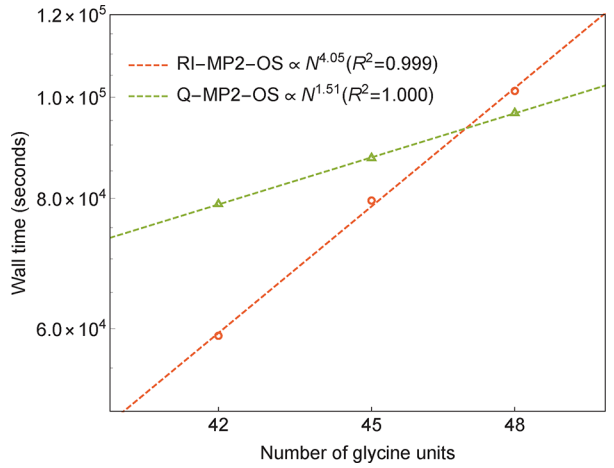


Figure 6. RI-MP2-OS[65] and Q-MP2-OS wall times and scalings for calculations on Gly_{42}, Gly_{45}, and Gly_{48}. Calculations were performed on a single core.

polyglycines, Q-MP2-OS is faster than RI-MP2-OS even before exploiting its embarrassing parallelizability.

In addition to its favorable scaling and parallelizability, Q-MP2-OS also has relatively modest disk and memory requirements. Most implementations of RI-MP2-OS require quadratic memory, and cubic disk to store the auxiliary integrals. These integrals need to be accessed repeatedly, leading to significant disk I/O and degrading the parallelizability of the method. In contrast, our implementation of Q-MP2-OS requires only quadratic memory to store the X and Y matrices and involves no disk access other than reading in the MO coefficients at the start of the calculation.

4.3. Accuracy. The cutoff strategies described in section 3.2 allow us to reduce the computational complexity of the MP2-OS energy from $O(N^4)$ to $O(N^2)$. However, these strategies and the various quadratures introduce errors into the final energies which must be assessed. Our target is chemical accuracy, i.e., <1 kcal mol^{-1}.

We calculated the incremental polymerization energies for the polyglycine polymerization reaction Gly_{n-1} + Gly → Gly_n + H_2O for a range of n values. Results in Table 1 compare

Table 1. Polyglycine Polymerization Energies Using MP2-OS, RI-MP2-OS, and Q-MP2-OS with an MG Grid[a]

n	MP2-OS	ΔRI-MP2-OS	ΔQ-MP2-OS
2	−17.561	+0.002	−0.069
3	−17.754	+0.000	−0.074
4	−18.082	+0.002	−0.055
5	−18.102	+0.002	−0.031
6	−18.172	+0.002	−0.057
7	−18.173	+0.002	−0.040

[a]All energies are quoted in kcal mol^{-1}.

the deviations of the RI-MP2-OS and Q-MP2-OS polymerization energies from those obtained with traditional MP2-OS. The energy errors of Q-MP2-OS are an order of magnitude larger than those of RI-MP2-OS; however, they fall well below our target of chemical accuracy. The Q-MP2-OS absolute error with respect to the total MP2-OS energy for the 3-glycine calculation is 2.1 kcal mol^{-1}, dropping to 0.01 kcal mol^{-1} when the larger SG-1 grid is adopted. These absolute errors can be compared with those of RI-MP2 calculations for the same system using the cc-pVDZ-RI and cc-pVTZ-RI auxiliary basis sets, which are 0.8 and 0.1 kcal mol^{-1}, respectively. Absolute errors in the total energies for both RI-MP2-OS and Q-MP2-OS grow linearly with system size, although this is not shown in Table 1.

To assess the accuracy of Q-MP2-OS over a wider range of reactions, we selected a variety of proton affinities (PA26), problematic DFT reactions (DC13), intramolecular dispersion reactions (IDISP), and isomerization energies of large molecules (ISOL24) from the GMTKN55 database.[67] Results are shown in Table 2. All RI-MP2-OS errors fall well below chemical accuracy. The Q-MP2-OS errors are larger, and although they often achieve our target, there are some exceptions. It is well-known that dispersion and isomerization energies are particularly difficult to evaluate using numerical

Table 2. Variety of Proton Affinities (PA26), Problematic DFT Reactions (DC13), Isomerization Energies of Large Molecules (ISOL24), and Intramolecular Dispersion Reactions (IDISP) Computed Using MP2-OS, RI-MP2-OS, and Q-MP2-OS[a]

			ΔQ-MP2-OS	
	MP2-OS	ΔRI-MP2-OS	MG	SG-1
PA26 #16	+223.615	−0.003	+0.438	+0.029
PA26 #17	+208.198	+0.009	+0.549	+0.042
PA26 #18	+209.830	−0.007	+0.619	−0.055
PA26 #24	+210.297	−0.012	−0.013	−0.035
PA26 #25	+239.910	+0.001	−0.521	−0.246
PA26 #26	+241.587	−0.003	+0.216	+0.051
DC13 #10	+108.696	−0.024	−0.365	+0.040
IDISP #2	−65.848	−0.044	−4.070	+0.132
IDISP #4	+19.331	−0.684	−4.521	−0.076
ISOL24 #1	+67.056	+0.027	+2.505	+0.008
ISOL24 #4	+53.920	−0.154	+10.580	−0.304

[a]The Q-MP2-OS calculations used either the MG or SG-1 quadrature grid. All energies are in kcal mol^{-1}.

https://dx.doi.org/10.1021/acs.jctc.9b01142
J. Chem. Theory Comput. 2020, 16, 1568−1577

quadrature.[35] For the associated IDISP and ISOL24 reactions, the MG grid fails to satisfy our target accuracy of 1 kcal mol^{-1}. However, the larger SG-1 quadrature grid allows us to achieve chemical accuracy across all reactions considered.

5. CONCLUSIONS

We have presented an algorithm for calculating the MP2-OS energy that is well-suited to large-scale parallel machines. Our initial implementation scales quadratically with system size and exhibits almost ideal parallel speedups even with more than 800 cores.

Our algorithm relies on three-dimensional quadrature, and we have discovered that, for our application, relatively modest grids are sufficient compared to those used in DFT calculations. This is consistent with what had been previously observed in the context of pseudospectral methods.[68] We have developed a small grid for the H, C, N, and O atoms that is often capable of yielding chemical accuracy, and we note that further work is required to optimize grids for other accuracy targets and basis sets.

6. FUTURE WORK

Further algorithmic improvements are possible for the Q-MP2-OS implementation. Only the elements of the X and Y matrices associated with significant shell pairs are accessed, and storing only these would reduce the memory footprint to $O(N)$. Furthermore, the D and U matrices can be computed by considering only the $O(1)$ shell pairs which are significant at the current batch of points rather than all $O(N)$ significant shell pairs. This could potentially reduce the Q-MP2-OS cost to $O(N)$.

Furthermore, the present article discussed only the OS part of MP2; this is because the treatment of same-spin (SS) MP2 energies is somewhat more involved, requiring decoupling of the basis function indices of the D matrix elements.

We are currently developing a Q-MP2 algorithm that includes both the above-mentioned computational enhancements and the MP2-SS part, and that can be accelerated using graphics processing units. We are also working on extending our combination of quadrature and screening techniques to self-consistent field (SCF) methods, F12 corrections for MP2, higher-order MP approaches, and coupled-cluster theory.

AUTHOR INFORMATION

Corresponding Authors

Giuseppe M. J. Barca − *Research School of Computer Science, Australian National University, Canberra, Australian Capital Territory 2601, Australia;* orcid.org/0000-0001-5109-4279; Email: giuseppe.barca@anu.edu.au

Peter M. W. Gill − *School of Chemistry, University of Sydney, Sydney, New South Wales 2006, Australia*; Email: p.gill@sydney.edu.au

Authors

Simon C. McKenzie − *School of Chemistry, University of Sydney, Sydney, New South Wales 2006, Australia; Research School of Chemistry, Australian National University, Canberra, Australian Capital Territory 2601, Australia*

Nathaniel J. Bloomfield − *Research School of Chemistry, Australian National University, Canberra, Australian Capital Territory 2601, Australia*

Andrew T. B. Gilbert − *Research School of Chemistry, Australian National University, Canberra, Australian Capital Territory 2601, Australia*

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jctc.9b01142

Funding

P.M.W.G. thanks the Australian Research Council for funding (Grants DP140104071 and DP160100246). This study has been (partially) supported through the grant NEXT No. ANR-10-LABX-0037 in the framework of the "Programme des Investissements d'Avenir".

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

G.M.J.B. and P.M.W.G. thank the National Computational Infrastructure (NCI) for supercomputer time.

REFERENCES

(1) Møller, C.; Plesset, M. S. Note on an approximation treatment for many-electron systems. *Phys. Rev.* **1934**, *46*, 618−622.

(2) Bartlett, R. J.; Silver, D. M. Correlation energy in LiH, BH, and HF with many-body perturbation theory using Slater-type atomic orbitals. *Int. J. Quantum Chem.* **1974**, *8*, 271−276.

(3) Handy, N. C.; Knowles, P. J.; Somasundram, K. On the convergence of the Møller-Plesset perturbation series. *Theor. Chim. Acta* **1985**, *68*, 87.

(4) Laidig, W. D.; Fitzgerald, G.; Bartlett, R. J. Is fifth-order MBPT enough? *Chem. Phys. Lett.* **1985**, *113*, 151−158.

(5) Gill, P. M. W.; Radom, L. Deceptive convergence in Møller-Plesset perturbation energies. *Chem. Phys. Lett.* **1986**, *132*, 16−22.

(6) Gill, P. M. W.; Pople, J. A.; Radom, L.; Nobes, R. H. Why does unrestricted Møller-Plesset perturbation theory converge so slowly for spin-contaminated wave functions? *J. Chem. Phys.* **1988**, *89*, 7307−7314.

(7) Olsen, J.; Christiansen, O.; Koch, H.; Jorgensen, P. Surprising cases of divergent behavior in Møller−Plesset perturbation theory. *J. Chem. Phys.* **1996**, *105*, 5082−5090.

(8) Stillinger, F. H. Møller−Plesset convergence issues in computational quantum chemistry. *J. Chem. Phys.* **2000**, *112*, 9711−9715.

(9) Leininger, M. L.; Allen, W. D.; Schaefer, H. F.; Sherrill, C. D. Is Møller-Plesset perturbation theory a convergent ab initio method? *J. Chem. Phys.* **2000**, *112*, 9213−9222.

(10) Almlöf, J. Elimination of energy denominators in Møller-Plesset perturbation theory by a Laplace transform approach. *Chem. Phys. Lett.* **1991**, *181*, 319−320.

(11) Häser, M. Møller-Plesset (MP2) perturbation theory for large molecules. *Theor. Chim. Acta* **1993**, *87*, 147−173.

(12) Feyereisen, M.; Fitzgerald, G.; Komornicki, A. Use of approximate integrals in ab initio theory. An application in MP2 energy calculations. *Chem. Phys. Lett.* **1993**, *208*, 359−363.

(13) Martínez, T. J.; Carter, E. A. Pseudospectral Møller-Plesset perturbation theory through third order. *J. Chem. Phys.* **1994**, *100*, 3631−3638.

(14) Hetzer, G.; Pulay, P.; Werner, H. J. Multipole approximation of distant pair energies in local MP2 calculations. *Chem. Phys. Lett.* **1998**, *290*, 143−149.

(15) Katouda, M.; Kobayashi, M.; Nakai, H.; Nagase, S. Two-Level hierarchical parallelization of second-order Møller-Plesset perturbation calculations in divide-and-conquer method. *J. Comput. Chem.* **2011**, *32*, 2756−2764.

(16) Del Ben, M.; Hutter, J.; VandeVondele, J. Second-order Møller-Plesset perturbation theory in the condensed phase: An efficient and massively parallel Gaussian and plane waves approach. *J. Chem. Theory Comput.* **2012**, *8*, 4177−4188.

(17) Kristensen, K.; Kjaergaard, T.; Høyvik, I.-M.; Ettenhuber, P.; Jørgensen, P.; Jansik, B.; Reine, S.; Jakowski, J. The divide-expand-consolidate MP2 scheme goes massively parallel. *Mol. Phys.* **2013**, *111*, 1196−1201.
(18) Katouda, M.; Nakajima, T. MPI/OpenMP hybrid parallel algorithm of resolution of identity second-order Møller-Plesset perturbation calculation for massively parallel multicore supercomputers. *J. Chem. Theory Comput.* **2013**, *9*, 5373−5380.
(19) Werner, H.-J.; Knizia, G.; Krause, C.; Schwilk, M.; Dornbach, M. Scalable electron correlation methods I.: PNO-LMP2 with linear scaling in the molecular size and near-inverse-linear scaling in the number of processors. *J. Chem. Theory Comput.* **2015**, *11*, 484−507.
(20) Katouda, M.; Naruse, A.; Hirano, Y.; Nakajima, T. Massively parallel algorithm and implementation of RI-MP2 energy calculation for peta-scale many-core supercomputers. *J. Comput. Chem.* **2016**, *37*, 2623−2633.
(21) Schäfer, T.; Ramberger, B.; Kresse, G. Quartic scaling MP2 for solids: A highly parallelized algorithm in the plane wave basis. *J. Chem. Phys.* **2017**, *146*, 104101.
(22) Martínez-Martínez, L. A.; Amador-Bedolla, C. GPU algorithm for the scaled opposite-spin (SOS) MP2 energy evaluation. *J. Mex. Chem. Soc.* **2017**, *61*, 60−66.
(23) Ishimura, K.; Ten-no, S. MPI/OpenMP hybrid parallel implementation of second-order Møller−Plesset perturbation theory using numerical quadratures. *Theor. Chem. Acc.* **2011**, *130*, 317−321.
(24) Song, C.; Martínez, T. J. Atomic orbital-based SOS-MP2 with tensor hypercontraction. I. GPU-based tensor construction and exploiting sparsity. *J. Chem. Phys.* **2016**, *144*, 174111.
(25) Song, C.; Martínez, T. J. Atomic orbital-based SOS-MP2 with tensor hypercontraction. II. Local tensor hypercontraction. *J. Chem. Phys.* **2017**, *146*, 034104.
(26) Willow, S. Y.; Kim, K. S.; Hirata, S. Stochastic evaluation of second-order many-body perturbation energies. *J. Chem. Phys.* **2012**, *137*, 204122.
(27) Willow, S. Y.; Hermes, M. R.; Kim, K. S.; Hirata, S. Convergence acceleration of parallel Monte Carlo second-order many-body perturbation calculations using redundant walkers. *J. Chem. Theory Comput.* **2013**, *9*, 4396−4402.
(28) Doran, A. E.; Hirata, S. Monte Carlo MP2 on many graphical processing units. *J. Chem. Theory Comput.* **2016**, *12*, 4821−4832.
(29) Maurer, S. A.; Lambrecht, D. S.; Kussmann, J.; Ochsenfeld, C. Efficient distance-including integral screening in linear-scaling Møller-Plesset perturbation theory. *J. Chem. Phys.* **2013**, *138*, 014101.
(30) Friesner, R. A. Solution of self-consistent field electronic structure equations by a pseudospectral method. *Chem. Phys. Lett.* **1985**, *116*, 39−43.
(31) Kokkila Schumacher, S. I. L.; Hohenstein, E. G.; Parrish, R. M.; Wang, L.-P.; Martínez, T. J. Tensor Hypercontraction Second-Order Møller-Plesset Perturbation Theory: Grid Optimization and Reaction Energies. *J. Chem. Theory Comput.* **2015**, *11*, 3042−3052.
(32) Takatsuka, A.; Ten-no, S.; Hackbusch, W. Minimax approximation for the decomposition of energy denominators in Laplace-transformed Møller-Plesset perturbation theories. *J. Chem. Phys.* **2008**, *129*, 044112.
(33) Becke, A. D. A multicenter numerical integration scheme for polyatomic molecules. *J. Chem. Phys.* **1988**, *88*, 2547−2553.
(34) Murray, C. W.; Handy, N. C.; Laming, G. J. Quadrature schemes for integrals of density functional theory. *Mol. Phys.* **1993**, *78*, 997−1014.
(35) Gill, P. M. W.; Johnson, B. G.; Pople, J. A. A standard grid for density functional calculations. *Chem. Phys. Lett.* **1993**, *209*, 506−512.
(36) Treutler, O.; Ahlrichs, R. Efficient molecular numerical integration schemes. *J. Chem. Phys.* **1995**, *102*, 346−354.
(37) Mura, M. E.; Knowles, P. J. Improved radial grids for quadrature in molecular density-functional calculations. *J. Chem. Phys.* **1996**, *104*, 9848−9858.
(38) Gill, P. M. W.; Chien, S. H. Radial quadrature for multiexponential integrands. *J. Comput. Chem.* **2003**, *24*, 732−740.
(39) Köster, A. M.; Flores-Moreno, R.; Reveles, J. U. Efficient and reliable numerical integration of exchange-correlation energies and potentials. *J. Chem. Phys.* **2004**, *121*, 681−690.
(40) El-Sherbiny, A.; Poirier, R. A. An evaluation of the radial part of numerical integration commonly used in DFT. *J. Comput. Chem.* **2004**, *25*, 1378−1384.
(41) Chien, S. H.; Gill, P. M. W. SG-0: A small standard grid for DFT quadrature on large systems. *J. Comput. Chem.* **2006**, *27*, 730−739.
(42) Kakhiani, K.; Tsereteli, K.; Tsereteli, P. A program to generate a basis set adaptive radial quadrature grid for density functional theory. *Comput. Phys. Commun.* **2009**, *180*, 256−268.
(43) Dasgupta, S.; Herbert, J. M. Standard grids for high-precision integration of modern density functionals: SG-2 and SG-3. *J. Comput. Chem.* **2017**, *38*, 869−882.
(44) Lebedev, V. Values of the nodes and weights of ninth to seventeenth order Gauss-Markov quadrature formulae invariant under the octahedron group with inversion. *USSR Comp. Math. Math. Phys.* **1975**, *15*, 44−51.
(45) Lebedev, V. Quadratures on a sphere. *USSR Comp. Math. Math. Phys.* **1976**, *16*, 10−24.
(46) Kohn, W. Analytic properties of Bloch waves and Wannier functions. *Phys. Rev.* **1959**, *115*, 809−821.
(47) Des Cloizeaux, J. Orthogonal orbitals and generalized Wannier functions. *Phys. Rev.* **1963**, *129*, 554−566.
(48) Des Cloizeaux, J. Analytical properties of *n*-dimensional energy bands and Wannier functions. *Phys. Rev.* **1964**, *135*, A698−A707.
(49) Kohn, W. Construction of Wannier functions and applications to energy bands. *Phys. Rev. B* **1973**, *7*, 4388−4398.
(50) Kohn, W.; Onffroy, J. R. Wannier functions in a simple nonperiodic system. *Phys. Rev. B* **1973**, *8*, 2485−2495.
(51) Rehr, J. J.; Kohn, W. Wannier functions in crystals with surfaces. *Phys. Rev. B* **1974**, *10*, 448−455.
(52) Goedecker, S.; Teter, M. Tight-binding electronic-structure calculations and tight-binding molecular dynamics with localized orbitals. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1995**, *51*, 9455−9464.
(53) Stephan, U.; Drabold, D. A. Order-N projection method for first-principles computations of electronic quantities and Wannier functions. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1998**, *57*, 6391−6407.
(54) Maslen, P. E.; Ochsenfeld, C.; White, C. A.; Lee, M. S.; Head-Gordon, M. Locality and sparsity of ab initio one-particle density matrices and localized orbitals. *J. Phys. Chem. A* **1998**, *102*, 2215−2222.
(55) Goedecker, S. Decay properties of the finite-temperature density matrix in metals. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1998**, *58*, 3501−3502.
(56) Ismail-Beigi, S.; Arias, T. A. Locality of the density matrix in metals, semiconductors, and insulators. *Phys. Rev. Lett.* **1999**, *82*, 2127−2130.
(57) He, L.; Vanderbilt, D. Exponential decay properties of Wannier functions and related quantities. *Phys. Rev. Lett.* **2001**, *86*, 5341−5344.
(58) Taraskin, S. N.; Fry, P. A.; Zhang, X.; Drabold, D. A.; Elliott, S. R. Spatial decay of the single-particle density matrix in tight-binding metals: Analytic results in two dimensions. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2002**, *66*, 233101.
(59) Prodan, E.; Kohn, W. Nearsightedness of electronic matter. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 11635−11638.
(60) Kohn, W. Density functional/Wannier function theory for systems of very many atoms. *Chem. Phys. Lett.* **1993**, *208*, 167−172.
(61) Baer, R.; Head-Gordon, M. Sparsity of the density matrix in Kohn-Sham density functional theory and assessment of linear system-size scaling methods. *Phys. Rev. Lett.* **1997**, *79*, 3962−3965.
(62) Hintze, J. L.; Nelson, R. D. Violin Plots: A box plot-density trace synergism. *Am. Stat.* **1998**, *52*, 181−184.
(63) Gill, P. M. W. Molecular integrals over gaussian basis functions. *Adv. Quantum Chem.* **1994**, *25*, 141−205.

https://dx.doi.org/10.1021/acs.jctc.9b01142
J. Chem. Theory Comput. 2020, 16, 1568−1577

(64) Shao, Y.; Gan, Z.; Epifanovsky, E.; Gilbert, A. T. B.; Wormit, M.; Kussmann, J.; Lange, A. W.; Behn, A.; Deng, J.; Feng, X.; Ghosh, D.; Goldey, M.; Horn, P. R.; Jacobson, L. D.; Kaliman, I.; Khaliullin, R. Z.; Kús, T.; Landau, A.; Liu, J.; Proynov, E. I.; Rhee, Y. M.; Richard, R. M.; Rohrdanz, M. A.; Steele, R. P.; Sundstrom, E. J.; Woodcock, H. L., III; Zimmerman, P. M.; Zuev, D.; Albrecht, B.; Alguire, E.; Austin, B.; Beran, G. J. O.; Bernard, Y. A.; Berquist, E.; Brandhorst, K.; Bravaya, K. B.; Brown, S. T.; Casanova, D.; Chang, C.-M.; Chen, Y.; Chien, S. H.; Closser, K. D.; Crittenden, D. L.; Diedenhofen, M.; DiStasio, R. A., Jr.; Do, H.; Dutoi, A. D.; Edgar, R. G.; Fatehi, S.; Fusti-Molnar, L.; Ghysels, A.; Golubeva-Zadorozhnaya, A.; Gomes, J.; Hanson-Heine, M. W. D.; Harbach, P. H. P.; Hauser, A. W.; Hohenstein, E. G.; Holden, Z. C.; Jagau, T.-C.; Ji, H.; Kaduk, B.; Khistyaev, K.; Kim, J.; Kim, J.; King, R. A.; Klunzinger, P.; Kosenkov, D.; Kowalczyk, T.; Krauter, C. M.; Lao, K. U.; Laurent, A.; Lawler, K. V.; Levchenko, S. V.; Lin, C. Y.; Liu, F.; Livshits, E.; Lochan, R. C.; Luenser, A.; Manohar, P.; Manzer, S. F.; Mao, S.-P.; Mardirossian, N.; Marenich, A. V.; Maurer, S. A.; Mayhall, N. J.; Oana, C. M.; Olivares-Amaya, R.; O'Neill, D. P.; Parkhill, J. A.; Perrine, T. M.; Peverati, R.; Pieniazek, P. A.; Prociuk, A.; Rehn, D. R.; Rosta, E.; Russ, N. J.; Sergueev, N.; Sharada, S. M.; Sharma, S.; Small, D. W.; Sodt, A.; Stein, T.; Stück, D.; Su, Y.-C.; Thom, A. J. W.; Tsuchimochi, T.; Vogt, L.; Vydrov, O.; Wang, T.; Watson, M. A.; Wenzel, J.; White, A.; Williams, C. F.; Vanovschi, V.; Yeganeh, S.; Yost, S. R.; You, Z.-Q.; Zhang, I. Y.; Zhang, X.; Zhao, Y.; Brooks, B. R.; Chan, G. K. L.; Chipman, D. M.; Cramer, C. J.; Goddard, W. A., III; Gordon, M. S.; Hehre, W. J.; Klamt, A.; Schaefer, H. F., III; Schmidt, M. W.; Sherrill, C. D.; Truhlar, D. G.; Warshel, A.; Xu, X.; Aspuru-Guzik, A.; Baer, R.; Bell, A. T.; Besley, N. A.; Chai, J.-D.; Dreuw, A.; Dunietz, B. D.; Furlani, T. R.; Gwaltney, S. R.; Hsu, C.-P.; Jung, Y.; Kong, J.; Lambrecht, D. S.; Liang, W.; Ochsenfeld, C.; Rassolov, V. A.; Slipchenko, L. V.; Subotnik, J. E.; Van Voorhis, T.; Herbert, J. M.; Krylov, A. I.; Gill, P. M. W.; Head-Gordon, M.; Neuscamman, E.; Yang, J. Advances in molecular quantum chemistry contained in the Q-Chem 4 program package. *Mol. Phys.* **2015**, *113*, 184−215.

(65) Jung, Y.; Lochan, R. C.; Dutoi, A. D.; Head-Gordon, M. Scaled opposite-spin second order Møller-Plesset correlation energy: An economical electronic structure method. *J. Chem. Phys.* **2004**, *121*, 9793−9802.

(66) Weigend, F.; Köhn, A.; Hättig, C. Efficient use of the correlation consistent basis sets in resolution of the identity MP2 calculations. *J. Chem. Phys.* **2002**, *116*, 3175−3183.

(67) Goerigk, L.; Hansen, A.; Bauer, C.; Ehrlich, S.; Najibi, A.; Grimme, S. A look at the density functional theory zoo with the advanced GMTKN55 database for general main group thermochemistry, kinetics and noncovalent interactions. *Phys. Chem. Chem. Phys.* **2017**, *19*, 32184−32215.

(68) Martínez, T. J.; Carter, E. A. Pseudospectral Møller-Plesset perturbation theory through third order. *J. Chem. Phys.* **1994**, *100*, 3631−3638.

https://dx.doi.org/10.1021/acs.jctc.9b01142
J. Chem. Theory Comput. 2020, 16, 1568−1577