Tensor Hypercontraction Form of the Perturbative Triples Energy in Coupled-Cluster Theory

Andy Jiang, Justin M. Turney,* and Henry F. Schaefer, III*

ABSTRACT: We present the working equations for a reduced-scaling method of evaluating the perturbative triples (T) energy in coupled-cluster theory, through the tensor hypercontraction (THC) of the triples amplitudes ($t_{ijk}^{abc}$). Through our method, we can reduce the scaling of the (T) energy from the traditional $O(N^7)$ to a more modest $O(N^5)$. We also discuss implementation details to aid future research, development, and software realization of this method. Additionally, we show that this method yields submillihartree (mEh) differences from CCSD(T) when evaluating absolute energies and sub-0.1 kcal/mol energy differences when evaluating relative energies. Finally, we demonstrate that this method converges to the true CCSD(T) energy through the systematic increasing of the rank or eigenvalue tolerance of the orthogonal projector, as well as exhibiting sublinear to linear error growth with respect to system size.

1. INTRODUCTION

Coupled-cluster (CC) theory$^{1,2}$ is one of the most important advances of modern quantum chemistry, allowing for a polynomial-time evaluation of the electronic energies and wave function of a molecule, as a size-extensive alternative to truncated configuration interaction (CI) methods.$^{3,4}$ Truncated CC methods also avoid the intractable superexponential scaling of full configuration interaction (FCI), yielding reasonable and chemically accurate relative energies compared to both the FCI limit and to experimental results, especially in the context of CCSD(T), also known as the “gold standard” method in computational quantum chemistry.$^5$ The tractability and accuracy of CC methods make the development of efficient CC methods crucial for the future of quantum chemistry, as evaluation of accurate energies and wave functions is made possible for larger and more complex systems through hardware advances such as massively parallel computing$^{6−22}$ and GPUs.$^{23−29}$

However, there is still a tremendous gap in applicability between coupled-cluster theories (formally scaling at least $O(N^7)$) and lower-scaling methods like Möller–Plesset perturbation theory (MP2)$^{30,31}$ and density functional theory (DFT)$^{32,33}$ (scaling $O(N^3)$ or better). Because of this, DFT and MP2 can be run on system tens or even hundreds of times the size of a system typically evaluated with CC methods.$^{34,35}$ To close the gap between CC and less reliable electron correlation methods, it is useful to devise approximation schemes to CC which reduce the scaling, but also allow a means to systematically control the error compared to the nonapproximated CC method. One such approach involves local correlation,$^{36−47}$ such as used in the DLPNO methods.$^{48,49}$ With large enough molecules, these methods achieve asymptotic linear scaling.

Another approach is the rank reduction of the coupled-cluster amplitudes,$^{50}$ using orthogonal projectors that transform the single and double cluster amplitudes into a smaller basis

$$T^V = U^V_{ia} t^V_{ia}$$

$$T^{WV} = U^W_{ia} t^{WV}_{ia} t^{WV}_{ia}$$

Because of the orthogonal nature of the projectors, getting the full amplitudes from the rank-reduced form is trivial

$$t^V_{ia} = U^V_{ia} T^V$$

$$t^{WV}_{ia} = U^W_{ia} T^{WV} U^W_{ia}$$

As shown by Parrish et al., the size of the V and W indices, also known as the projector rank, can be made directly proportional to the system size, while maintaining a set relative error from the absolute energy of a molecule.$^{50}$ More recently, Hohenstein et al. have shown how to create a tensor hypercontracted (THC) form of the $t^{ab}_{ij}$ amplitudes, through the CANCELCONMP/
PARAFAC (CP) decomposition\textsuperscript{51} of the orthogonal projectors,\textsuperscript{52}
\begin{equation}
U_{ia}^{W} = \sum_{X} y_{ia}^{X} U_{X}^{VW}
\end{equation}
\begin{equation}
t_{ij}^{ab} = \sum_{xy} y_{ia}^{X} U_{X}^{VW} y_{jb}^{Y} T_{XY}^{VW}
\end{equation}
\begin{equation}
T_{XY}^{VW} = \sum_{VW} r_{VW}^{XY} T_{X}^{YY} r_{YY}
\end{equation}
Hohenstein et al. have also shown that, in the context of CCSD, the size of the $X$ index can be made proportional to the system size to maintain a set relative error. Rank-reduction methods have also been applied to coupled-cluster theories involving higher levels of excitation, recently by Lesiuk with the SVD-CCSDT method,\textsuperscript{35} where the concept of orthogonal projectors is used to approximate the triples amplitude in CCSDT theory
\begin{equation}
t_{ijk}^{abc} = U_{ia}^{VW} U_{jb}^{W} U_{kc}^{W} T_{abc}^{VW}
\end{equation}
In the following sections, we combine the concepts of orthogonal projectors and THC to develop working equations for a reduced-scaling variant of the noniterative perturbative triples correction to the CCSD energy.\textsuperscript{5} Recently, Lesiuk derived an $O(N^6)$ approach to the (T) energy with orthogonal projectors which he calls RR-CCSD(T).\textsuperscript{54} In the current paper, we improve upon the work of Lesiuk’s approach utilizing tensor hypercontraction. Similar to how Hohenstein et al. used tensor hypercontraction to improve upon the RR-CCSD method of Parrish et al.,\textsuperscript{55,52} our method uses tensor hypercontraction to improve upon Lesiuk’s RR-CCSD(T) method.\textsuperscript{54} Our new approach, which we name THC-RR-CCSD(T), will commensurately enhance RR-CCSD(T), reducing the scaling of Lesiuk’s from $O(N^6)$ to $O(N^5)$. For consistency, we use many of the same formalisms as Lesiuk\textsuperscript{53,54} and Hohenstein et al.\textsuperscript{52}

2. THEORY

2.1. Notation. We use the following conventions to describe the indices appearing in this work:
- $i, j, k, l$: Occupied molecular orbitals, which range from 1 to $n_{occ}$.
- $a, b, c, d$: Virtual molecular orbitals, which range from 1 to $n_{aux}$.
- $P, Q$: Auxiliary indices of density-fitted/Cholesky-decomposed ERIs, which range from 1 to $n_{aux}$.
- $w, v$: Laplace denominator weight indices, which range from 1 to $n_{aux}$.
- $U, V, W$: Rank-reduced dimensions of the doubles orthogonal projector, which range from 1 to $n_{proj}$.
- $A, B, C$: Rank-reduced dimensions of the triples orthogonal projector, which range from 1 to $n_{proj}$.
- $X, Y, Z$: CP-decomposition ranks of the triples orthogonal projector, which range from 1 to $n_{proj}$.

The relative sizes of the indices are as follows:
\begin{equation}
\begin{array}{c}
n_{occ} < n_{aux} < n_{proj} \approx n_{proj} \\
\end{array}
\end{equation}

Note that $n_{aux}$ does not grow with increasing molecular system size, and therefore, run-time analysis of intermediates with $w, v$ indices will only treat the Laplace index as a prefactor.

In all post-Hartree–Fock computations in this work; i.e., the 1s electrons are not correlated for all first-row atoms. The occupied space $n_{occ}$ always refers to the number of correlated occupied orbitals. The generalized Einstein summation convention is used throughout—all indices appearing on the right-hand side but not on the left-hand side of an expression are summed over.

2.2. Perturbative Triples Correction to CCSD. CCSD is often not sufficient to obtain “chemically reliable” theoretical predictions, and it has been shown that only after triple excitations are considered that relative energies of under 1 kcal/mol can be regularly achieved.\textsuperscript{55–61} However, an explicit treatment of all triples has a very high cost of $O(N^5)$. Therefore, the triples amplitudes are often determined in a perturbative manner, based on the work of Raghavachari et al.\textsuperscript{5} In their formalism, the perturbative triples correction to the CCSD energy is defined as
\begin{equation}
E^{(T)} = E^{[4]} + E^{[5]}_{ST}
\end{equation}
where
\begin{equation}
E^{[4]} = \langle T_4 W, T_4 \rangle
\end{equation}
\begin{equation}
E^{[5]}_{ST} = \langle T_5 W, T_5 \rangle.
\end{equation}
$T_1$, $T_2$, and $T_3$ are known as the “cluster operators” and, in second-quantization formalism, are defined as
\begin{equation}
T_1 = t_1^{ab} E_{ai}
\end{equation}
\begin{equation}
T_2 = t_{ijk}^{ab} E_{ai} E_{bj}
\end{equation}
\begin{equation}
T_3 = t_{ijk}^{ab} E_{ai} E_{bj} E_{ck}
\end{equation}
$E_{ai}$ represents the singlet, spin-adapted excitation operator, and is defined as
\begin{equation}
E_{ai} = a_i^a a_i + \pi_a^i \pi_i
\end{equation}
where the barred creation/annihilation operators refer to the beta spin orbitals and nonbarred refer to the alpha spin orbitals. The accuracy of the (T) method stems from a highly favorable error cancellation between $E_4^{[2]}$ and $E_5^{[3]}$. In restricted, single-reference, closed-shell coupled cluster theory, one can write the equation for the (T) correction as
\begin{equation}
E^{(T)} = \frac{1}{3} \left( 4 W_{ijk}^{abc} W_{ij}^{ab} + W_{ijk}^{abc} W_{ij}^{abcd} (V_{ijk}^{abc} - V_{ijk}^{abcd}) \right)
\end{equation}
where
\begin{equation}
W_{ijk}^{abc} = P_{ijkl} ((ialbd)t_{ij}^{cd} - (ial)l_{ij}^{cd})
\end{equation}
and
\begin{equation}
V_{ijk}^{abc} = W_{ijk}^{abc} + P_{3}(a_{ijklc})
\end{equation}
Following the formalism of Lesiuk,\textsuperscript{53} we define $P_2$ and $P_3$, or the “long” and “short” permutation operations, as
\begin{equation}
P_2 (A_{ijk}^{abc}) = A_{ijk}^{abc} + A_{ijk}^{bac} + A_{ijk}^{abc} + A_{ijkl}^{bca} + A_{ijk}^{cab} + A_{ijkl}^{cda}
\end{equation}
\begin{equation}
P_3 (A_{ijk}^{abc}) = A_{ijk}^{abc} + A_{ijk}^{bac} + A_{ijk}^{cab}
\end{equation}
The perturbative triples amplitude ($t_{ijk}^{abc}$) is defined as
For triples amplitudes, we present two approaches devised by Lesiuk. In his SVD-CCSD(T) algorithm,\textsuperscript{53} he took guess $t_{ijk}^{abc}$ amplitudes, such as from CC3, and applied either a TUCKER-3 decomposition (scaling $O(N^3)$) or an iterative SVD approach (scaling $O(N^5)$), yielding the form of eq 8.

In his RR-CCSD(T) paper, Lesiuk devised an $O(N^5)$ scheme to compute projectors from the form of the perturbative triples amplitudes (eq 22), in a variant of HO-OI (Higher Order-Orthogonal Iteration).\textsuperscript{54} The steps of the algorithm are as follows:

- Start with the a guess of the triples projector $V_{ia}$. This can be done naively by setting $V_{ia} = U_{ia}$ from the doubles amplitudes.
- Evaluate $t_{ia,BC}$ from the current guess of the triples amplitudes, where
  \[ t_{ia,BC} = t_{ijk}^{abc} V_{ia}^{B} V_{ia}^{C} \]  
  By using the explicit expression for $t_{ijk}^{abc}$ and $W_{ijk}^{abc}$, this can be evaluated in $O(N^5)$. The working equations are presented in ref 54.
- Compute the SVD of $t_{ia,BC}$ and take the largest $n_{proj}$ left singular vectors as the next $V_{ia}$. This can be done in $O(N^4)$ time using a modified variant of truncated SVD, given in ref 68. In this algorithm, we save the singular values of this step ($\sigma_{ij}$), when we perform the CP decomposition of the triples projector. The pseudocode for this is presented in Section 4.
- Iterate until convergence. Convergence is defined when the difference between the Frobenius norm of the rank-reduced triples amplitudes $t_{ABC}$, defined as
  \[ t_{ABC} = V_{ia}^{A} t_{ia,BC} \]  
between two successive iterations, falls below $10^{-5}$.

Since the source of the orthogonal projectors is not relevant to the scope of this paper, we only present results from computations utilizing the MP2 projector for the doubles amplitudes, and Lesiuk’s HO-OI approach for the perturbative triples amplitudes.

2.4. Tensor Hypercontraction (THC). Tensor hypercontraction (THC) can be viewed as a “double approximation”, where two auxiliary indices are introduced to fit a high-dimensional tensor instead of just one. The THC form of electron repulsion integrals is defined as\textsuperscript{69}

\[ (pqrs) \approx x_{p}^{l} x_{q}^{m} z_{r}^{j} y_{s}^{k} \]  
This can be derived from the CP decomposition of $B_{ij}^{pq}$ (eq 25)

\[ B_{ij}^{pq} \approx x_{i}^{l} x_{j}^{k} \eta_{ij}^{pq} \]  
\[ z_{ij}^{l} = \eta_{ij}^{kl} \eta_{ij}^{pq} \]  
Similarly, the THC form of coupled-cluster amplitudes can be derived from the tensor hypercontraction of the orthogonal projectors, given by, in the case of the doubles projector\textsuperscript{52}

\[ U_{ia}^{ij} = y_{i}^{a} x_{j}^{a} x_{i}^{a} x_{i}^{a} \]  
For the triples projector, it assumes a very similar form

\[ V_{ia}^{ij} = z_{i}^{a} x_{i}^{a} \theta_{i}^{a} \]
A PARAFAC/CANDENCOMP (CP) decomposition approach on \( V_{ii}^A \) may be used. This approach is not dependent on the source of the projectors, and any of the projector building approaches from Section 2.3 may be used. Here, we use the variant of CP decomposition, first introduced by Hohenstein et al. for the doubles projector,\(^{52}\) where the eigenvalues of the doubles projector are in the CP decomposition, into the alternating least-squares (ALS) iterations.

In our algorithm, for the decomposition of the triples amplitude, instead of using the eigenvalues of the doubles projector, we use the singular values of the \( t_{ia,BC} \) intermediate \((\sigma)\). The functional to minimize is hence

\[
L_{CP} = \sum_{ia} (\sigma(V_{ii}^A - z_i^X z_i^Y \theta_{AX}))^2
\]  

(38)

The update rule for each intermediate is given as

\[
z_i^X = \sum_{aA} \sigma_a^2 V_{ii}^A \sum_{Y} z_Y \theta_{AY} \left[ \sum_{b} z_b^X z_b^Y \sum_{B} \sigma_B^2 \theta_{BX} \theta_{BY} \right]^{-1}
\]  

(39)

\[
z_a^X = \sum_{iA} \sigma_a^2 V_{ii}^A \sum_{Y} z_Y \theta_{AY} \left[ \sum_{j} z_j^X z_j^Y \sum_{b} \sigma_B^2 \theta_{BX} \theta_{BY} \right]^{-1}
\]  

(40)

\[
\theta_{AX} = \sum_{ia} V_{ii}^A \sum_{Y} z_i^X z_Y \left[ \sum_{j} z_j^X z_j^Y \sum_{b} z_b^X z_b^Y \right]^{-1}
\]  

(41)

Note that the update rule for \( \theta \) is the same as in traditional CP decomposition.

Since a CP decomposition does not exactly recreate the original projector, the projectors lose their orthogonal property.\(^{53}\) Therefore, we have to recreate the projectors after the CP decomposition

\[
S_{AB} = V_{ii}^A V_{ii}^B
\]  

(42)

\[
\theta_{AX} = \theta_{BX} S_{AB}^{-1/2}
\]  

(43)

\[
V_{ii}^A = z_i^X z_i^Y \theta_{AX}
\]  

(44)

The \( t_{ia,b} \) amplitudes can now be rewritten as, from eq 8

\[
t_{ijk} = z_i^X z_a^X z_j^Y z_k^Y z_j^Z z_k^Z t_{XYZ}
\]  

(45)

\[
t_{XYZ} = \theta_{by} \theta_{cz} t_{ABC}
\]  

(46)

Recently, Hohenstein et al. have devised an algorithm that takes advantage of the THC form of the \( t_{ia,b} \) amplitudes to develop an \( O(N^3) \) scaling implementation of CCSD.\(^{53}\) In the next section, we show how to extend this to the (T) correction with the THC form of the \( t_{ia,b} \) amplitudes.

3. DERIVATION OF WORKING EQUATIONS

We first define a couple of intermediates. From Lesiuk,\(^{54}\) we define

\[
D_{ij}^{QV} = (B_{ib}^{QV} U_{ij}^{W} - B_{ij}^{QV} T_{ib}^{W}) T^{WW}
\]  

(47)

Next, we define the following chain of intermediates from contracting the polyadic vectors \( (z_i^X \) and \( z_i^Y) \) of the triples projector with the doubles projector, the DF/RI or CD decomposed ERIs, and the D intermediate from eq 47, as well as the \( T_1 \) amplitudes.

\[
\theta_{i}^{X} = \frac{1}{2} \sum_{a,B} \sum_{j} (V_{i}^{A} - z_{i}^{X} z_{j}^{Y} \theta_{AX}) \sum_{b} (V_{j}^{B} - z_{j}^{X} z_{b}^{Y} \theta_{BX})^{-1}
\]  

(51)

\[
D_{i}^{QV} = D_{i}^{QV} z_{i}^{X} z_{i}^{Y}
\]  

(52)

\[
D_{1}^{QVXY} = D_{1}^{QV} z_{1}^{X} z_{1}^{Y}
\]  

(53)

\[
D_{1}^{VXY} = D_{1}^{QV} z_{1}^{X} z_{1}^{Y}
\]  

(54)

\[
D_{1}^{VXY} = D_{1}^{VX} t_{1}^{Y} z_{1}^{X}
\]  

(55)

We then take eqs 23, 19, and 45, and the previously defined intermediates, to arrive at the THC form of the triples energy correction

\[
E^{(T)} + = 8 \cdot D_{VV}^{QV} D_{VQ}^{QV} B^{QZ} t_{XYZ}
\]  

(56)

\[
E^{(T)} + = 4 \cdot D_{XY}^{QV} D_{VQ}^{XX} B^{QZ} t_{XYZ}
\]  

(57)

\[
E^{(T)} - = 4 \cdot D_{XZ}^{QV} D_{VQ}^{XX} B^{QZ} t_{XYZ}
\]  

(58)

\[
E^{(T)} - = 4 \cdot D_{VZ}^{QV} D_{VQ}^{XX} B^{QY} t_{XYZ}
\]  

(59)

\[
E^{(T)} - = 4 \cdot D_{VQ}^{QV} D_{VQ}^{XX} B^{QZ} t_{XYZ}
\]  

(60)

\[
E^{(T)} - = 4 \cdot D_{VQ}^{QV} D_{VQ}^{XX} B^{QY} t_{XYZ}
\]  

(61)

\[
E^{(T)} - = 2 \cdot D_{VQ}^{QV} D_{VQ}^{XX} B^{QZ} t_{XYZ}
\]  

(62)

\[
E^{(T)} + = 2 \cdot D_{VQ}^{QV} D_{VQ}^{XX} B^{QY} t_{XYZ}
\]  

(63)

\[
E^{(T)} + = 2 \cdot D_{VQ}^{QV} D_{VQ}^{XX} B^{QY} t_{XYZ}
\]  

(64)

\[
E^{(T)} + = 2 \cdot D_{VQ}^{QV} D_{VQ}^{XX} B^{QZ} t_{XYZ}
\]  

(65)

Equations 56 and 57 correspond to the first term in eq 23, eqs 58–62 the second term, and eqs 63–65 the third term. All of the contractions can be determined in \( O(N^3) \) time or less.

4. IMPLEMENTATION DETAILS

To aid future research and development, we present pseudocode for some of the algorithms we use for the optimal contraction of intermediate terms to evaluate the THC-RR-CCSD(T) energy. We first present our noniterative SVD algorithm to factorize the \( t_{ia,BC} \) intermediate, inspired by the truncated SVD and diagonalization algorithms given in ref 68. In Algorithm 1, we present a noniterative truncated SVD algorithm to avoid the \( O(N^3) \) scaling of a traditional SVD of the \( t_{ia,BC} \) intermediate. In Algorithms 2–4, we present suggested contraction orders, as well as tensor slicings, for each term of the THC-RR-CCSD(T) energy expression. We try to make the contractions such that highly efficient level 3 BLAS matrix multiplication calls are utilized as much as possible. For each step of each algorithm, the runtime is given, and if a level 3 BLAS matrix multiplication call is possible, then the term (GEMM) is added. Additionally, the \( D_{QVXY} \) intermediate is never fully built to help with memory costs. The runtime of this algorithm is \( O(N^3) \), with \( O(N^4) \) storage costs; the only quartic memory requirements involve the
storage of the $t_{abc}$ and $D_{abc}$ intermediates. It may be possible to reduce the memory cost in future implementations of this method, but that is beyond the scope of this paper.

Algorithm 1 Truncated SVD algorithm for $t_{abc}$.

\begin{verbatim}
$U_{BC} = \text{random}(n_{proj} \sim n_{occ}, n_{proj})$

$Y_{BC} = \text{BEAM}(U_{BC})$

$Q_{abc} = Q(BC)Y_{BC}$

$T_{abc} = \text{Diagonalize}(X_{abc})$

$V_{a} = \text{BEAM}(V_{a}, T_{abc})$

return $V_{a}^{-1}U_{BC}$

\end{verbatim}

Algorithm 2 $\beta^{(1)}$ Contractions (Equations 56 - 57)

\begin{verbatim}
$A_{VZZ} = \beta^{(1)}_{VZZ}$ $\triangleright O(N^4)$, GEMM

$B_{VVZ} = \beta^{(1)}_{VVZ}$ $\triangleright O(N^4)$, GEMM

$C_{VVZ} = \beta^{(1)}_{VVZ}$ $\triangleright O(N^4)$, GEMM

$D_{VZZ} = \beta^{(1)}_{VZZ}$ $\triangleright O(N^4)$

$E^{(1)}_{VZZ} = -4 \cdot \beta^{(1)}_{VZZ}$ $\triangleright O(N^4)$

$F^{(1)}_{VZZ} = -4 \cdot \beta^{(1)}_{VZZ}$ $\triangleright O(N^4)$

$G^{(1)}_{VZZ} = -4 \cdot \beta^{(1)}_{VZZ}$ $\triangleright O(N^4)$

$H^{(1)}_{VZZ} = -4 \cdot \beta^{(1)}_{VZZ}$ $\triangleright O(N^4)$

end for

Algorithm 3 $\beta^{(2)}$ Contractions (Equations 58 - 62)

\begin{verbatim}
for V in [0, n_{proj}) do $\triangleright$ parallelize

$E^{(2)}_{VZZ} = \beta^{(2)}_{VZZ}$ $\triangleright O(N^4)$, GEMM, built on the fly to save storage

$F^{(2)}_{VZZ} = \beta^{(2)}_{VZZ}$ $\triangleright O(N^4)$, GEMM

$G^{(2)}_{VZZ} = \beta^{(2)}_{VZZ}$ $\triangleright O(N^4)$, GEMM

$H^{(2)}_{VZZ} = \beta^{(2)}_{VZZ}$ $\triangleright O(N^4)$, GEMM

$J^{(2)}_{VZZ} = \beta^{(2)}_{VZZ}$ $\triangleright O(N^4)$, GEMM

$K^{(2)}_{VZZ} = \beta^{(2)}_{VZZ}$ $\triangleright O(N^4)$, GEMM

$L^{(2)}_{VZZ} = \beta^{(2)}_{VZZ}$ $\triangleright O(N^4)$, GEMM

end for

for Q in [0, n_{proj}) do $\triangleright$ parallelize

$E^{(2)}_{QZZ} = \beta^{(2)}_{QZZ}$ $\triangleright O(N^4)$, GEMM, built on the fly to save storage

$F^{(2)}_{QZZ} = \beta^{(2)}_{QZZ}$ $\triangleright O(N^4)$, GEMM

$G^{(2)}_{QZZ} = \beta^{(2)}_{QZZ}$ $\triangleright O(N^4)$, GEMM

$H^{(2)}_{QZZ} = \beta^{(2)}_{QZZ}$ $\triangleright O(N^4)$, GEMM

$J^{(2)}_{QZZ} = \beta^{(2)}_{QZZ}$ $\triangleright O(N^4)$, GEMM

$K^{(2)}_{QZZ} = \beta^{(2)}_{QZZ}$ $\triangleright O(N^4)$, GEMM

$L^{(2)}_{QZZ} = \beta^{(2)}_{QZZ}$ $\triangleright O(N^4)$, GEMM

end for

Algorithm 4 $\epsilon^{(2)}$ Contractions (Equations 61 - 65)

\begin{verbatim}
for V in [0, n_{proj}) do $\triangleright$ parallelize

$E^{(2)}_{VZZ} = \epsilon^{(2)}_{VZZ}$ $\triangleright O(N^4)$, GEMM, built on the fly to save storage

$F^{(2)}_{VZZ} = \epsilon^{(2)}_{VZZ}$ $\triangleright O(N^4)$, GEMM

$G^{(2)}_{VZZ} = \epsilon^{(2)}_{VZZ}$ $\triangleright O(N^4)$, GEMM

$H^{(2)}_{VZZ} = \epsilon^{(2)}_{VZZ}$ $\triangleright O(N^4)$, GEMM

$J^{(2)}_{VZZ} = \epsilon^{(2)}_{VZZ}$ $\triangleright O(N^4)$, GEMM

$K^{(2)}_{VZZ} = \epsilon^{(2)}_{VZZ}$ $\triangleright O(N^4)$, GEMM

$L^{(2)}_{VZZ} = \epsilon^{(2)}_{VZZ}$ $\triangleright O(N^4)$, GEMM

end for

end for

end for

end for

end for

The code is implemented in a developmental plugin version of the Psi4 Quantum Chemistry code, following the completion of an exact CCSD computation. Tensor contractions are performed with the help of the EinsumsInCpp software (public on GitHub). The compressed doubles amplitudes $T^{ij}_{abc}$ used to build the triples projector are formed by transforming the exact CCSD amplitudes from the preceding computation by the MP2 projector amplitudes. This method is designed to be fully compatible and used with Hohenstein’s THC-RR-CCSD method. Future studies of using THC-RR-CCSD(T) in conjunction with THC-RR-CCSD are encouraged.

5. RESULTS

5.1. Conformation Energies. We first evaluate our new THC-RR-CCSD(T) method on the CYCONF data set, a set containing 11 different conformations of gaseous cysteine, with 10 corresponding conformation energies, relative to the lowest conformer. We evaluate conformation energies for each of the 10 conformations in CCSD, CCSD(T), and THC-RR-CCSD(T), and for each system, and we use the exact CCSD(T) conformation energy as the reference. We do this using the cc-pVDZ and jun-cc-pVDZ Dunning correlation-consistent basis sets. The basis set jun-cc-pVDZ consists of diffuse functions added to all heavy atoms, except for the basis functions with the highest angular momentum. For the THC-RR-CCSD(T) computations, we set the eigenvalue tolerance of the MP2 projector to be $10^{-4}$. In other words, the ranks ($n_{proj}$) of the doubles and triples projectors are determined from how many eigenvalues of the MP2 $t^{ij}_{abc}$ amplitudes are greater than $10^{-4}$, defined as $r$ from eq 29 in our work. For these computations, $n_{proj}$ is around 400, compared to the max possible rank of 2205 ($n_{occ}n_{virt}$) in the cc-pVDZ basis, yielding a compression ratio of around 18%. Similarly, in the jun-cc-pVDZ basis, the ratio is 440/2793, which is around 16%.

The summary statistics are presented in Table 1, and the results for each individual conformation are presented in Figure 1. In the table, for the THC-RR-CCSD(T) algorithms, the eigenvalue tolerance is given in parentheses. To summarize the findings, THC-RR-CCSD(T) consistently gives lower errors compared to CCSD, for both basis sets, and the errors are on the order of less than 0.1 kcal/mol. The error also does not significantly grow with the addition of diffuse functions, from cc-pVDZ to jun-cc-pVDZ. The (T) correlation energy for these conformers is around 35–36 millihartrees for cc-pVDZ and 37–38 millihartrees for jun-cc-pVDZ. For both basis sets, THC-RR-CCSD(T) recovers around 98.5% of the (T) correlation energy for each conformer. It is further encouraging to note that the absolute energy errors for these sets of computations hover around 0.3–0.4 kcal/mol, such that the evaluation of relative energies benefits from favorable error cancellation. Detailed numbers for how much correlation energy is recovered for each conformer in each basis set are available in the Supporting Information (SI).

5.2. Potential Energy Surface. We perform next, a potential energy surface scan on the benzene–HCN dimer system (compound 19 from the on S22 data set), with the hydrogen atom of HCN pointing toward the $\pi$-bonds in the benzene. We measured the energy of the system at five different interatomic distances, relative to the equilibrium geometry, ranging from 0.9 to 2.0 times the equilibrium geometry length, with the geometries coming from the S22x5 data set. In Figure 2, we plot the shape of the potential energy surface of the THC-RR-CCSD(T) method at an eigenvalue tolerance of $10^{-4}$, as well as using predetermined projector ranks of 400 and 500. For all systems, an eigenvalue tolerance of $10^{-4}$ corresponds to a
Table 1. Errors in Conformation Energy Compared to the Exact CCSD(T) Reference

| Test set                  | Mean error (kcal/mol) | MAE (kcal/mol) | RMSE (kcal/mol) | Std. dev. (kcal/mol) |
|---------------------------|-----------------------|----------------|-----------------|----------------------|
| CCSD/cc-pVDZ             | −0.343                | 0.343          | 0.384           | 0.173                |
| THC-RR-CCSD(T)/cc-pVDZ (10⁻⁴) | −0.072                | 0.072          | 0.075           | 0.023                |
| CCSD/jun-cc-pVDZ         | −0.291                | 0.291          | 0.323           | 0.141                |
| THC-RR-CCSD(T)/jun-cc-pVDZ (10⁻⁴) | −0.076                | 0.076          | 0.082           | 0.031                |

*The number in parenthesis is the eigenvalue tolerance used to determine projector rank.*

Figure 2. Relative energies of benzene–HCN dimer (S22 system 19) evaluated with each method at five different dimer separation distances relative to the equilibrium geometry.

Figure 1. Errors in conformation energies for CCSD and THC-RR-CCSD(T) evaluated on the CYCONF data set, compared to the exact CCSD(T) reference, evaluated in with the cc-pVDZ and jun-cc-pVDZ basis sets, with a 10⁻⁴ eigenvalue tolerance.

5.3. Rank Convergence. Next, to demonstrate the convergence of the THC-RR-CCSD(T) method, compared to the exact CCSD(T) energy, we ran a series of computations of the water dimer from the S22 set, at eigenvalue tolerances from 10⁻³ to 10⁻¹¹. An eigenvalue tolerance of 10⁻¹¹ corresponds to no rank compression for this system. The errors with respect to eigenvalue tolerance and compression ranks are plotted in Figure 3, and it is encouraging to see the errors decrease smoothly to the true CCSD(T) energy, within the DF/RI approximation of the ERIs. The errors are on the order from 0.0 to 0.2 millihartrees, compared to the total (T) correlation energy of around 6.4 millihartrees. We attribute the “kink” in the graph from 10⁻⁴ to 10⁻⁶ as an artifact of the CP decomposition of the triples projector, with the CP error increasing slightly between the projector ranks of 122–156, before going back down. This artifact is well known on studies of the CP decomposition algorithm, where medium CP decomposition ranks suffer larger losses in accuracy compared to small or large ranks. Further studies and work are encouraged to look for ways to mitigate this phenomenon in the context of decomposing CC amplitudes.

5.4. Timings. To establish the lower scaling of the THC-RR-CCSD(T) method compared to CCSD or CCSD(T), we present timings on growing systems of water clusters and linear alkanes, from 1 to 8 heavy atoms, in the cc-pVDZ and jun-cc-pVDZ basis sets. All computations are performed on 48 CPU cores of an Intel Xeon 6136. For each system and basis set combination, we present timings for THC-RR-CCSD(T) at a constant eigenvalue tolerance (10⁻⁴). In Figures 4–7, we present raw timings for the computation of the (T) energy and the noniterative steps in computing the THC-RR-T energy, as well as the raw timings for each iteration of the HO-OI procedure used in forming the triples amplitude projector. Since each computation requires a different number of iterations for convergence in the HO-OI procedure, instead of presenting raw timings for THC-RR-CCSD(T), we instead present an “expected time” for THC-RR-CCSD(T), which is the sum of the time required for the noniterative portion of the THC-RR-T computation, added with the average number of HO-OI iterations multiplied by the per iteration HO-OI time. A different average is computed across every system and basis set combination.

This “expected” time is also what is used when computing the scaling of the THC-RR-CCSD(T) method, as well as the crossover points with CCSD(T) in Tables 3 and 4. To compute
the scaling of each method, we applied a power fit of the run time of each procedure in the form of $t = a \cdot n^b$, where $t$ is the run time, $a$ the prefactor, $n$ the number of basis functions in the molecule, and $b$ the computational scaling. The coefficients $a$ and $b$ are determined through a linear regression of $\log(t)$ as the independent variable and $\log(n)$ as the dependent variable. In our analysis, we only consider timings from systems with three or more heavy atoms, and the $r^2$ coefficient of the linear regression is greater than 0.99 in all cases, showing the values shown in Tables 3 and 4 to be reliable. In all system and basis set combinations, computing the THC-RR-(T) energy scaled significantly better than computing the (T) energy or the preceding CCSD computation. The values are consistent with theoretical considerations, with the computation of the (T) energy scaling empirically around $O(N^6)$, and the computation of the THC-RR-(T) energy scaling around $O(N^4 \cdot p)$. This is consistent with there being few $O(N^5)$ steps in CCSD(T) and few $O(N^5)$ steps in THC-RR-CCSD(T). We note that though our current pilot implementation is not optimized, the crossover points presented for both systems are still reasonable and can be made much lower in future, more optimized implementations.

5.5. Error Growth. Finally, we plot the percentage errors of the (T) correlation energy for each combination of system and basis set used for timings in the previous section and show that the errors do not grow with larger systems, remaining under 3%.

Table 2. Errors in Relative Energies Compared to the Exact CCSD(T) Reference, for a Reference CCSD Computation and THC-RR-CCSD(T) Computations with Varying Parameters

| Test set                  | Mean error (kcal/mol) | MAE (kcal/mol) | RMSE (kcal/mol) | Std. dev. (kcal/mol) |
|---------------------------|-----------------------|----------------|-----------------|----------------------|
| CCSD                      | -0.138                | 0.200          | 0.236           | 0.191                |
| THC-RR-CCSD(T), tol = $10^{-4}$ | -0.100                | 0.103          | 0.132           | 0.086                |
| THC-RR-CCSD(T), $n_{proj} = 400$ | -0.098                | 0.098          | 0.128           | 0.082                |
| THC-RR-CCSD(T), $n_{proj} = 500$ | -0.001                | 0.010          | 0.014           | 0.014                |
in all cases (Figure 8). This shows the size extensivity of the THC-RR-CCSD(T) energy.

6. CONCLUSIONS

In this paper, we present the working equations for the THC-RR-CCSD(T) method, an \( O(N^4) \) scaling approximation to CCSD(T), that allows for systematic control of errors. In our pilot implementation, we show the errors are controllable to the point of maintaining chemical accuracy of less than 0.1 kcal/mol for relative energies, and 1 mEh for absolute energies, while maintaining size extensivity. We also show that the method yields continuous potential energy surfaces that closely match the CCSD(T) surfaces with sufficient projector rank. In addition, we have empirically established that THC-RR-CCSD(T) indeed scales better than CCSD or CCSD(T). In the future, we hope to consider ways to improve the errors of the method at a given eigenvalue tolerance, such as through using other sources for the orthogonal projector. We would also like to look into alternative approaches to the THC factorization of orthogonal projectors. Though a CP decomposition is generally applicable, and relatively easy to implement, it does not assume any underlying form about the amplitudes. One avenue is the extension of the quadrature-based approach of Parrish, Hohenstein, Martinez, and Sherrill with Least-Squares Tensor Hypercontraction (LS-THC) to the triples amplitudes.\(^{79,80}\) Finally, we hope to have a more optimized implementation of this method available in the future, one that yields lower crossover points relative to CCSD(T), and available in the Psi4 package.\(^{70}\)

ASSOCIATED CONTENT

Data Availability Statement
The data that support the findings of this study are available with the article and the Supporting Information.

Supporting Information
The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jctc.2c00996.

Data that support the findings of this study (ZIP)

AUTHOR INFORMATION

Corresponding Authors
Justin M. Turney — Center for Computational Molecular Science and Technology, School of Chemistry and Biochemistry, School of Computational Science and Engineering, Georgia Institute of Technology, Atlanta, Georgia 30332-0400, United States; Email: justin.turney@uga.edu

Henry F. Schaefer, III — Center for Computational Molecular Science and Technology, School of Chemistry and Biochemistry, School of Computational Science and Engineering, Georgia Institute of Technology, Atlanta, Georgia 30332-0400, United States; orcid.org/0000-0003-0252-2083; Email: ccq@uga.edu

Author
Andy Jiang — Center for Computational Quantum Chemistry, Department of Chemistry, University of Georgia, Athens, Georgia 30602, United States; orcid.org/0000-0002-6946-209X

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jctc.2c00996
Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

The authors gratefully acknowledge financial support from the U.S. Department of Energy, Basic Energy Sciences Division, Computational and Theoretical Chemistry (CTC) Grant DE-SC0018164.

REFERENCES

(1) Crawford, T. D.; Schaefer, H. F. An introduction to coupled cluster theory for computational chemists. Reviews in Computational Chemistry 2007, 33, 1–136.

(2) Bartlett, R. J.; Musial, M. Coupled-cluster theory in quantum chemistry. Rev. Mod. Phys. 2007, 79, 291.

(3) Cramer, C. J. Essentials of Computational Chemistry; Wiley, 2002; pp 191–232.

(4) Sherrill, C. D.; Schaefer, H. F. The Configuration Interaction Method: Advances in Highly Correlated Approaches. Adv. Quantum Chem. 1999, 34, 143–269.

(5) Raghavachari, K.; Trucks, G. W.; Pople, J. A.; Head-Gordon, M. A fifth-order perturbation comparison of electron correlation theories. Chem. Phys. Lett. 1989, 157, 479–483.

(6) Hirata, S. Tensor Contraction Engine: Abstraction and Automated Parallel Implementation of Configuration-Interaction, Coupled-Cluster, and Many-Body Perturbation Theories. The Journal of Physical Chemistry A 2003, 107, 9887–9897.

(7) Auer, A. A.; et al. Automatic code generation for many-body electronic structure methods: the tensor contraction engine‡‡. Mol. Phys. 2006, 104, 211–228.

(8) Janowski, T.; Ford, A. R.; Pulay, P. Parallel Calculation of Coupled Clusters Singles and Doubles Wave Functions Using Array Files. Journal of Chemical Theory and Computation 2007, 3, 1368–1377.

(9) Janowski, T.; Pulay, P. Efficient Parallel Implementation of the CCSD External Exchange Operator and the Perturbative Triples (T) Energy Calculation. Journal of Chemical Theory and Computation 2008, 4, 1585–1592.

(10) van Dam, H.; de Jong, W.; Bylaika, E.; Govind, N.; Kowalski, K.; Straatsma, T.; Valiev, M. NWChem: scalable parallel computational chemistry. Wiley Interdisciplinary Reviews: Computational Molecular Science 2011, 1, 888–894.

(11) Deumens, E.; Lotrich, V. F.; Perera, A.; Ponton, M. J.; Sanders, B. A.; Bartlett, R. J. Software design of ACES III with the super instruction architecture. Wiley Interdisciplinary Reviews: Computational Molecular Science 2011, 1, 895–901.

(12) Kobayashi, R.; Bannwarth, C.; Fales, B. S.; Hohenstein, E. G.; Kokkila-Schumacher, S. I. L.; Luehr, N.; Snyder, J. W.; Song, C.; Titov, A. V.; Ulfit, C.; Martinez, T. J. TeraChem: Accelerating electronic structure and ab initio molecular dynamics with graphical processing units. The Journal of Chemical Physics 2012, 137, 120231.

(13) Peng, C.; Calvin, J. A.; Valeev, E. F. Single-precision open-shell CCSD and CCSD(T) calculations on graphics processing units. Phys. Chem. Chem. Phys. 2020, 22, 25103–25111.

(14) Peng, C.; Calvin, J. A.; Valeev, E. F. Coupled-cluster singles, doubles and perturbative triples with density fitting approximation for massively parallel heterogeneous platforms. Int. J. Quantum Chem. 2019, 119, e25894.

(15) Seritan, S.; Bannwarth, C.; Fales, B. S.; Hohenstein, E. G.; Kokkila-Schumacher, S. I. L.; Luehr, N.; Snyder, J. W.; Song, C.; Titov, A. V.; Ulfit, C.; Martinez, T. J. TeraChem: Accelerating electronic structure and ab initio molecular dynamics with graphical processing units. The Journal of Chemical Physics 2012, 137, 120231.

(16) Peng, C.; Calvin, J. A.; Valeev, E. F. Single-precision open-shell CCSD and CCSD(T) calculations on graphics processing units. Phys. Chem. Chem. Phys. 2020, 22, 25103–25111.

(17) Peng, C.; Calvin, J. A.; Valeev, E. F. Massively Parallel Implementation of Explicitly Correlated Coupled-Cluster Singles and Doubles Using TiledArray Framework. The Journal of Physical Chemistry A 2016, 120, 10231–10244.

(18) Peng, C.; Calvin, J. A.; Wang, X.; Clement, M. C.; Pierce, K.; Rishi, V.; Pavoševec, F.; Slattery, S.; Zhang, J.; Teke, N.; Kumar, A.; Masteran, C.; Asadchev, A.; Calvin, J. A.; Valeev, E. F. Massively Parallel Quantum Chemistry: A high-performance research platform for electronic structure. The Journal of Chemical Physics 2020, 153, 044120.

(19) Datta, D.; Gordon, M. S. A Massively Parallel Implementation of the CCSD(T) Method Using the Resolution-of-the-Identity Approximation and a Hybrid Distributed/Shared Memory Parallelization Model. Journal of Chemical Theory and Computation 2021, 17, 4799–4822.

(20) Gyevi-Nagy, L.; Kallay, M.; Nagy, P. R. Accurate Reduced-Cost CCSD(T) Energies: Parallel Implementation, Benchmarks, and Large-Scale Applications. Journal of Chemical Theory and Computation 2021, 17, 860–878.

(21) Kowalski, K.; et al. From NWChem to NWChemEx: Evolving with the Computational Chemistry Landscape. Chemical Reviews 2021, 121, 4962–4998.

(22) Calvin, J. A.; Peng, C.; Rishi, V.; Kumar, A.; Valeev, E. F. Many-Body Quantum Chemistry on Massively Parallel Computers. Chemical Reviews 2021, 121, 1203–1231.

(23) Sentan, S.; Bannwarth, C.; Fales, B. S.; Hohenstein, E. G.; Kokkila-Schumacher, S. I. L.; Luehr, N.; Snyder, J. W.; Song, C.; Titov, A. V.; Ulfit, C.; Martinez, T. J. TeraChem: Accelerating electronic structure and ab initio molecular dynamics with graphical processing units. The Journal of Chemical Physics 2012, 137, 120231.

(24) Wang, Z.; Guo, M.; Wang, F. Single-precision open-shell CCSD and CCSD(T) calculations on graphics processing units. Phys. Chem. Chem. Phys. 2020, 22, 25103–25111.

(25) Peng, C.; Calvin, J. A.; Valeev, E. F. Coupled-cluster singles, doubles and perturbative triples with density fitting approximation for massively parallel heterogeneous platforms. Int. J. Quantum Chem. 2019, 119, e25894.

(26) Dakka, Y.; Shaheen, M.; Nakajima, T.; Ratcliff, L. E.; Genovese, L. Density functional theory calculations of large systems: Interplay between fragments, observables, and computational complexity. WIREs Computational Molecular Science 2022, 12, e1574.

(27) Li, S.; Ma, J.; Jiang, Y. Linear scaling local correlation approach for solving the coupled cluster equations of large systems. J. Comput. Chem. 2002, 23, 237–244.
(37) Li, S.; Shen, J.; Li, W.; Jiang, Y. An efficient implementation of the “cluster-in-molecule” approach for local electron correlation calculations. The Journal of Chemical Physics 2006, 125, 074109.

(38) Li, W.; Piecuch, P.; Gour, J. R.; Li, S. Local correlation calculations using standard and renormalized coupled-cluster approaches. The Journal of Chemical Physics 2009, 131, 114109.

(39) Neese, F.; Wennmohs, F.; Hansen, A. Efficient and accurate local approximations to coupled-electron pair approaches: An attempt to revive the pair natural orbital method. The Journal of Chemical Physics 2009, 130, 114108.

(40) Li, W.; Piecuch, P. Multilevel Extension of the Cluster-in-Molecule Local Correlation Methodology: Merging Coupled-Cluster and Moller–Plesset Perturbation Theories. The Journal of Physical Chemistry A 2010, 114, 6721–6727.

(41) Li, W.; Piecuch, P. Improved Design of Orbital Domains within the Cluster-in-Molecule Local Correlation Framework: Single-Environment Cluster-in-Molecule Ansatz and Its Application to Local Coupled-Cluster Approach with Singles and Doubles. The Journal of Physical Chemistry A 2010, 114, 8644–8657.

(42) Rolik, Z.; Kállay, M. A general-order local coupled-cluster method based on the cluster-in-molecule approach. The Journal of Chemical Physics 2011, 135, 104111.

(43) Rolik, Z.; Szegedy, L.; Ladjánszki, I.; Ladjánszki, B.; Kállay, M. An efficient linear-scaling CCSD(T) method based on local natural orbitals. The Journal of Chemical Physics 2013, 139, 094105.

(44) Riplinger, C.; Neese, F. An efficient and near linear scaling pair natural orbital based local coupled cluster method. The Journal of Chemical Physics 2013, 138, 034106.

(45) Riplinger, C.; Sandhofer, B.; Hansen, A.; Neese, F. Natural triple excitations in local coupled cluster calculations with pair natural orbitals. The Journal of Chemical Physics 2013, 139, 134101.

(46) Liakos, D. G.; Sparta, M.; Kesharwani, M. K.; Martin, J. M. L.; Neese, F. Exploring the Accuracy Limits of Local Pair Natural Orbital Coupled-Cluster Theory. Journal of Chemical Theory and Computation 2015, 11, 1525–1539.

(47) Schwilk, M.; Ma, Q.; Köpl, C.; Werner, H.-J. Scalable Electron Correlation Methods. 3. Efficient and Accurate Parallel Local Coupled Cluster with Pair Natural Orbitals (PNO-LCCSD). Journal of Chemical Theory and Computation 2017, 13, 3650–3675.

(48) Pinski, P.; Riplinger, C.; Valeev, E. F.; Neese, F. Sparse maps—A systematic infrastructure for reduced-scaling electronic structure methods. I. An efficient and simple linear scaling local MP2 method that uses an intermediate basis of pair natural orbitals. The Journal of Chemical Physics 2015, 143, 034108.

(49) Riplinger, C.; Pinski, P.; Becker, U.; Valeev, E. F.; Neese, F. Sparse maps—A systematic infrastructure for reduced-scaling electronic structure methods. II. Linear scaling domain based pair natural orbital coupled cluster theory. The Journal of Chemical Physics 2016, 144, 024109.

(50) Parrish, R. M.; Zhao, Y.; Hohenstein, E. G.; Martínez, T. J. Reduced coupled cluster theory. I. Ground state energies and wavefunctions. The Journal of Chemical Physics 2019, 150, 164118.

(51) Kolda, T. G.; Bader, B. W. Tensor Decompositions and Applications. SIAM Review 2009, 51, 455–500.

(52) Hohenstein, E. G.; Fales, B. S.; Parrish, R. M.; Martínez, T. J. Rank-reduced coupled-cluster. III. Tensor hypercontraction of the doubles amplitudes. The Journal of Chemical Physics 2022, 156, 054102.

(53) Lesiuk, M. Implementation of the Coupled-Cluster Method with Single, Double, and Triple Excitations using Tensor Decompositions. Journal of Chemical Theory and Computation 2020, 16, 453–467.

(54) Lesiuk, M. Quintic-scaling reduced coupled-cluster theory with single and double excitations. The Journal of Chemical Physics 2022, 156, 064103.

(55) Riley, K. E.; Pitonková, M.; Jurečka, P.; Hobza, P. Stabilization and Structure Calculations for Noncovalent Interactions in Extended Molecular Systems Based on Wave Function and Density Functional Theories. Chemical Reviews 2010, 110, 5023–5063.

(56) Karton, A.; Rabinovich, E.; Martin, J. M. L.; Ruscic, B. W4 theory for computational thermochemistry: In pursuit of confident sub-kJ/mol predictions. The Journal of Chemical Physics 2006, 125, 144108.

(57) Tajti, A.; Szalay, P. G.; Császár, A. G.; Kállay, M.; Gauss, J.; Valeev, E. F.; Flowers, B. A.; Vázquez, J.; Stanton, J. F. HEAT: High accuracy extrapolated ab initio thermochemistry. The Journal of Chemical Physics 2004, 121, 11599–11613.

(58) Bak, K. L.; Jørgensen, P.; Olsen, J.; Helgaker, T.; Klopfer, W. Accuracy of atomization energies and reaction enthalpies in standard and extrapolated electronic wave function/basis set calculations. The Journal of Chemical Physics 2000, 112, 9229–9242.

(59) Hopkins, B. W.; Tschumper, G. S. Ab Initio Studies of π···π Interactions: The Effects of Quadrupe Excitations. The Journal of Physical Chemistry A 2004, 108, 2941–2948.

(60) Bartlett, R. J.; Watts, J.; Kucharski, S.; Noga, J. Non-iterative fifth-order triple and quadruple excitation energy corrections in correlated methods. Chem. Phys. Lett. 1990, 165, 513–522.

(61) Dutta, A.; Sherrill, C. D. Full configuration interaction potential energy curves for breaking bonds to hydrogen: An assessment of single-reference correlation methods. The Journal of Chemical Physics 2003, 118, 1610–1619.

(62) Gyevi-Nagy, L.; Kállay, M.; Nagy, P. R. Parallel-Implementation and Parallel Optimization of the CCSD(T) Method: Algorithic Developments and Large-Scale Applications. Journal of Chemical Theory and Computation 2020, 16, 366–384.

(63) Dunlap, B. I.; Connolly, J. W. D.; Sabin, J. R. On some approximations in applications of Xα theory. The Journal of Chemical Physics 1979, 71, 3396–3402.

(64) Weigend, F.; Häser, M.; Patzel, H.; Ahlrichs, R. RI–MP2: optimized auxiliary basis sets and demonstration of efficiency. Chem. Phys. Lett. 1998, 294, 143–152.

(65) Rægen, I.; Johansen, T. Chebyshev decomposition of the two-electron integral matrix in electronic structure calculations. The Journal of Chemical Physics 2008, 128, 194107.

(66) DePrince, A. E.; Sherrill, C. D.; Parrish, R. M.; Martínez, T. J. Tensor hypercontraction density fitting. I. Quartic scaling second- and third-order Moller-Plesset perturbation theory. The Journal of Chemical Physics 2012, 137, 044103.

(67) Häser, M.; Almlof, J. Laplace transform techniques in Möller-Plesset perturbation theory. The Journal of Chemical Physics 1992, 96, 489–494.

(68) Halko, N.; Martinsson, P. G.; Tropp, J. A. Finding Structure with Randomness: Probabilistic Algorithms for Constructing Approximate Matrix Decompositions. SIAM Review 2011, 53, 217–288.

(69) Hohenstein, E. G.; Parrish, R. M.; Martínez, T. J. Tensor hypercontraction density fitting. I. Quartic scaling second- and third-order Möller-Plesset perturbation theory. The Journal of Chemical Physics 2012, 137, 044103.

(70) Smith, D. G. A.; et al. Psi4 1.4: Open-source software for high-throughput quantum chemistry. The Journal of Chemical Physics 2020, 152, 184108.

(71) Wilke, J. J.; Lind, M. C.; Schaefer, H. F.; Császár, A. G.; Allen, W. D. Conformers of Gaseous Cysteine. Journal of Chemical Theory and Computation 2009, 5, 1511–1523.

(72) Goerigk, L.; Grimme, S. A General Database for Main Group Thermochemistry, Kinetics, and Noncovalent Interactions - Assessment of Common and Reparameterized meta-GGA Density Functionals. Journal of Chemical Theory and Computation 2010, 6, 107–126.

(73) Dunning, T. H. Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. The Journal of Chemical Physics 1989, 90, 1007–1023.

(74) Woon, D. E.; Dunning, T. H. Gaussian basis sets for use in correlated molecular calculations. III. The atoms aluminum through argon. The Journal of Chemical Physics 1993, 98, 1358–1371.

(75) Woon, D. E.; Dunning, T. H. Gaussian basis sets for use in correlated molecular calculations. IV. Calculation of static electrical response properties. The Journal of Chemical Physics 1994, 100, 2975–2988.
(76) Papajak, E.; Truhlar, D. G. Convergent Partially Augmented Basis Sets for Post-Hartree–Fock Calculations of Molecular Properties and Reaction Barrier Heights. *Journal of Chemical Theory and Computation* 2011, 7, 10–18.

(77) Jurečka, P.; Šponer, J.; Černý, J.; Hobza, P. Benchmark database of accurate (MP2 and CCSD(T)) complete basis set limit) interaction energies of small model complexes, DNA base pairs, and amino acid pairs. *Phys. Chem. Chem. Phys.* 2006, 8, 1985–1993.

(78) Gráfová, L.; Pitonák, M.; Řezáč, J.; Hobza, P. Comparative Study of Selected Wave Function and Density Functional Methods for Noncovalent Interaction Energy Calculations Using the Extended S22 Data Set. *Journal of Chemical Theory and Computation* 2010, 6, 2365–2376.

(79) Parrish, R. M.; Hohenstein, E. G.; Martínez, T. J.; Sherrill, C. D. Tensor hypercontraction. II. Least-squares renormalization. *The Journal of Chemical Physics* 2012, 137, 224106.

(80) Hohenstein, E. G.; Parrish, R. M.; Sherrill, C. D.; Martínez, T. J. Communication: Tensor hypercontraction. III. Least-squares tensor hypercontraction for the determination of correlated wavefunctions. *The Journal of Chemical Physics* 2012, 137, 221101.