

Lawrence Berkeley National Laboratory

Recent Work

Title

Cost-effective description of strong correlation: Efficient implementations of the perfect quadruples and perfect hexuples models.

Permalink

<https://escholarship.org/uc/item/3c33d6zr>

Journal

The Journal of chemical physics, 145(13)

ISSN

0021-9606

Authors

Lehtola, Susi
Parkhill, John
Head-Gordon, Martin

Publication Date

2016-10-01

DOI

10.1063/1.4964317

Peer reviewed

Cost-effective description of strong correlation: Efficient implementations of the perfect quadruples and perfect hexuples models

Susi Lehtola, John Parkhill, and Martin Head-Gordon

Citation: *The Journal of Chemical Physics* **145**, 134110 (2016); doi: 10.1063/1.4964317

View online: <https://doi.org/10.1063/1.4964317>

View Table of Contents: <http://aip.scitation.org/toc/jcp/145/13>

Published by the [American Institute of Physics](#)

Articles you may be interested in

[Incremental full configuration interaction](#)

The Journal of Chemical Physics **146**, 104102 (2017); 10.1063/1.4977727

[Density matrix renormalization group \(DMRG\) method as a common tool for large active-space CASSCF/CASPT2 calculations](#)

The Journal of Chemical Physics **146**, 094102 (2017); 10.1063/1.4976644

[Stochastic multi-reference perturbation theory with application to the linearized coupled cluster method](#)

The Journal of Chemical Physics **146**, 044107 (2017); 10.1063/1.4974177

[Excited states from modified coupled cluster methods: Are they any better than EOM CCSD?](#)

The Journal of Chemical Physics **146**, 144104 (2017); 10.1063/1.4979078

[Strong correlation in incremental full configuration interaction](#)

The Journal of Chemical Physics **146**, 224104 (2017); 10.1063/1.4985566

[Time-dependent N-electron valence perturbation theory with matrix product state reference wavefunctions for large active spaces and basis sets: Applications to the chromium dimer and all-trans polyenes](#)

The Journal of Chemical Physics **146**, 244102 (2017); 10.1063/1.4986975

PHYSICS TODAY

WHITEPAPERS

ADVANCED LIGHT CURE ADHESIVES

Take a closer look at what these environmentally friendly adhesive systems can do

READ NOW

PRESENTED BY
 MASTERBOND
ADHESIVES | SEALANTS | COATINGS

Cost-effective description of strong correlation: Efficient implementations of the perfect quadruples and perfect hexuples models

Susi Lehtola,^{1,a)} John Parkhill,^{2,b)} and Martin Head-Gordon^{1,3,c)}

¹Chemical Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA

²Department of Chemistry, University of Notre Dame, 251 Nieuwland Science Hall, Notre Dame, Indiana 46556, USA

³Department of Chemistry, University of California, Berkeley, California 94720, USA

(Received 30 March 2016; accepted 22 September 2016; published online 7 October 2016)

Novel implementations based on dense tensor storage are presented for the singlet-reference perfect quadruples (PQ) [J. A. Parkhill *et al.*, *J. Chem. Phys.* **130**, 084101 (2009)] and perfect hexuples (PH) [J. A. Parkhill and M. Head-Gordon, *J. Chem. Phys.* **133**, 024103 (2010)] models. The methods are obtained as block decompositions of conventional coupled-cluster theory that are exact for four electrons in four orbitals (PQ) and six electrons in six orbitals (PH), but that can also be applied to much larger systems. PQ and PH have storage requirements that scale as the square, and as the cube of the number of active electrons, respectively, and exhibit quartic scaling of the computational effort for large systems. Applications of the new implementations are presented for full-valence calculations on linear polyenes (C_nH_{n+2}), which highlight the excellent computational scaling of the present implementations that can routinely handle active spaces of hundreds of electrons. The accuracy of the models is studied in the π space of the polyenes, in hydrogen chains (H_{50}), and in the π space of polyacene molecules. In all cases, the results compare favorably to density matrix renormalization group values. With the novel implementation of PQ, active spaces of 140 electrons in 140 orbitals can be solved in a matter of minutes on a single core workstation, and the relatively low polynomial scaling means that very large systems are also accessible using parallel computing. *Published by AIP Publishing.* [<http://dx.doi.org/10.1063/1.4964317>]

I. INTRODUCTION

Efficient handling of strong correlation is one of the still unresolved questions in computational chemistry. Outside the scope of the otherwise useful Kohn–Sham density-functional theory,^{1–3} approaches to strong correlation lean heavily on wave function theories. The conventional way to treat strong correlation is through multiconfigurational self-consistent field (MCSCF) theory, which scales exponentially as the size of the active space grows. As a result, MCSCF cannot be applied to active spaces significantly larger than 16 electrons in 16 orbitals, although the barrier has been recently pushed back through approximative stochastic^{4,5} and adaptive^{6–10} approaches to the full configuration interaction (FCI) problem.

Now, while MCSCF relies on a configuration interaction (CI) type ansatz $|\Psi\rangle = (1 + \hat{T})|\Phi\rangle$ for the wave function, where $|\Phi\rangle$ is the single-particle reference and $|\Psi\rangle$ the true, multiconfigurational wave function, one might ask whether a coupled-cluster¹¹ (CC) type ansatz $|\Psi\rangle = \exp(\hat{T})|\Phi\rangle$ could be used instead to formulate a MCSCF-type model. Often the CC approach is also used to describe dynamic correlation with a CI reference, as in the multireference CC (MR-CC) approach,¹² but MR-CC is much more complicated than single-reference CC, and combining both strong correlation and dynamic correlation within the same machinery would

be clearly advantageous. While in principle CC theory is well able to describe static correlation, the problem that arises upon its application is that the computational effort of the resulting procedure is prohibitively expensive, even if an active space is used.¹³ However, this problem has been recently solved by Parkhill, Lawler, and Head-Gordon, who proposed a truncation of CC for exactness within the active space for a given number of strongly interacting electron pairs,^{14,15} instead of the usual, global truncation based on the excitation level. For example, the perfect quadruples¹⁴ (PQ) model is formed by truncating CC theory with single through quadruple (CCSDTQ) excitations to two pairs by keeping only the amplitudes $t_{i_1 \dots i_n}^{a_1 \dots a_n}$ that satisfy $\{a_k, i_k\} \in \{\text{pair}_1\} \times \{\text{pair}_2\}$. Similarly, the perfect hexuples¹⁵ (PH) model is a three-pair truncation of CC theory with single through hextuple excitations (CCSDTQ56), in which only the amplitudes that satisfy $\{a_k, i_k\} \in \{\text{pair}_1\} \times \{\text{pair}_2\} \times \{\text{pair}_3\}$ are kept. Here, the concept of a *pair* arises from connection to the perfect pairing^{16–20} (PP) theory, in which pair i is described by a quartet of orbitals: the alpha and beta orbitals i and \bar{i} , respectively, that are occupied in the single-particle reference, as well as the corresponding alpha and beta virtual orbitals i^* and \bar{i}^* , respectively. In contrast to MCSCF, PQ and PH exhibit polynomial scaling with respect to the size of the active space. The storage requirements are modest, scaling as $O(N^2)$ for PQ and $O(N^3)$ for PH, N being the number of active orbitals in the calculation. The computational effort scales asymptotically as $O(N^4)$ for both models.

^{a)}Electronic mail: susi.lehtola@alumni.helsinki.fi

^{b)}Electronic mail: john.parkhill@gmail.com

^{c)}Electronic mail: mhg@cchem.berkeley.edu

Another widely used approach for strong correlation is the density matrix renormalization group (DMRG) method,^{21–25} which is exact in principle and which also offers polynomial scaling with respect to system size. However, the method has a significant prefactor: the storage and computational costs scale as^{25,26} $O(M^2N^3)$ and $O(M^3N^3) + O(M^2N^4)$, respectively, where M is the number of states kept in the calculation, the exact solution being recovered at the limit M to ∞ . Furthermore, as a fundamentally one-dimensional approach, DMRG has a history of being hard to use due to issues with orbital ordering, although automatic orbital ordering schemes have recently made DMRG calculations easier to perform. We note that other approaches for a compact description of strong correlation have also been suggested, such as the CC valence bond (CCVB) method,^{27–29} extended CC,^{30–32} and the multifacet graphically contracted function method.^{33,34}

In order to apply MCSCF, DMRG, PQ, or PH to studies of chemical systems, the additional description of dynamic correlation is usually of utmost importance. While MCSCF and DMRG typically rely on a two-pronged approach to dynamic correlation through CI, CC, or perturbative approaches, dynamic correlation can be treated on the same footing as static correlation in PQ and PH,³⁵ enabling the two types of correlation to interact. Note that perturbative approaches can be used as well in combination with PQ³⁶ and PH.

What unites MCSCF, DMRG, PQ, and PH is that the choice of the set of active orbitals is a problem, and often the orbitals need to be optimized in order to fully capture static correlation effects. Orbital optimization is hard for PQ and PH, but earlier experience¹⁵ suggests that in many cases orbitals optimized at the much simpler PP level of theory might be sufficient for a wide variety of applications. Thus, in the present work, only localized Hartree–Fock orbitals and PP optimized orbitals will be considered. Orbitals optimized with CCVB or unrestricted Hartree–Fock natural orbitals^{37–39} might also prove useful. However, because the feasible size of the active space is much larger in PQ and PH than in MCSCF or even DMRG, it is likely that for full-valence calculations, the choice of the active space orbitals should not turn out to be an impediment to the use of the PQ and PH models.

While encouraging results have been achieved with PQ and PH, their existing implementations (while exhibiting the correct $O(N^4)$ scaling with system size) have a too large prefactor to allow their application to studies of chemical problems. Because the previous implementations are based on the use of sparse tensor algebra,⁴⁰ it is worth asking if the prefactor could be reduced by the use of dense tensor storage, instead. Moreover, the existing sparse tensor implementation

does not take advantage of the mathematical structure of the PQ and PH methods, which results in extra work for bookkeeping and prevents efficient compiler optimization from taking place.

In the present work, we describe the dense tensor implementation of the PQ and PH models and benchmark the resulting code against the previous implementation on linear polyenes. Novel applications of the PQ and PH models are demonstrated for the H₅₀ chain, and the results are compared to DMRG. Applications of PQ and PH to the strongly correlated π space of the polyacene series are also presented. The organization of the manuscript is the following. In Sec. II, we illustrate how the truncation of CC theory is performed. Next, in Sec. III, we describe various aspects of how the procedure has been carried out in practice. In Sec. IV, we present the scaling of the new code when applied to linear polyenes, and the energies for the polyenes, the dissociation of the H₅₀ chain, as well as the polyacenes for which natural occupation numbers are also presented. The study concludes with a summary and brief discussion.

II. THEORY

The theory behind the CC approach is well established and shall not be discussed further here.^{41,42} The starting point for the present truncation is the CC equations in spin-integrated form, consisting of tensors that belong to a certain spin block, such as the $\alpha\beta$ block of the double excitation amplitudes, where α and β denote spin-up and spin-down, respectively. As a truncation of CC theory, the models in the present hierarchy inherit all the beneficial properties of CC such as size extensivity, but also its drawbacks such as non-variational behavior in cases where the excitation rank in CC is not high enough to describe the static correlation effects.

In the following, we will omit the spin indices, because the discussion applies to all spin blocks of all tensors that appear in CC theory. We will mostly be focusing on the PQ model, as it already illustrates all of the necessary aspects for a general truncation method without being overly complicated. The division of the orbital space that is used in the present work is illustrated in Figure 1. The system, in a singlet reference state, is composed of a set of inactive core orbitals, a set of N active occupied orbitals, a set of N active virtual orbitals, and a set of inactive virtual orbitals. The tensors such as the excitation amplitudes t , the two-electron integrals v , the de-excitation amplitudes λ , and the one- and two-electron density matrices γ and Γ are truncated using a block decomposition. For instance, for PQ, the double excitation tensor is expanded as

$$t_{ij}^{ab} = \sum_{P=1}^N t_{PP}^{PP} \delta_i P \delta_j P \delta_a P \delta_b P + \sum_{\substack{P,Q=1 \\ P \neq Q}}^N t_{PP}^{PQ} \delta_i P \delta_j P \delta_a P \delta_b Q + \sum_{\substack{P,Q=1 \\ P \neq Q}}^N t_{PP}^{QP} \delta_i P \delta_j P \delta_a Q \delta_b P + \sum_{\substack{P,Q=1 \\ P \neq Q}}^N t_{PQ}^{PP} \delta_i P \delta_j Q \delta_a P \delta_b P + \sum_{\substack{P,Q=1 \\ P \neq Q}}^N t_{PQ}^{PQ} \delta_i P \delta_j Q \delta_a P \delta_b Q + \sum_{\substack{P,Q=1 \\ P \neq Q}}^N t_{QP}^{PQ} \delta_i Q \delta_j P \delta_a P \delta_b Q + \sum_{\substack{P,Q=1 \\ P \neq Q}}^N t_{PP}^{QQ} \delta_i P \delta_j P \delta_a Q \delta_b Q, \quad (1)$$

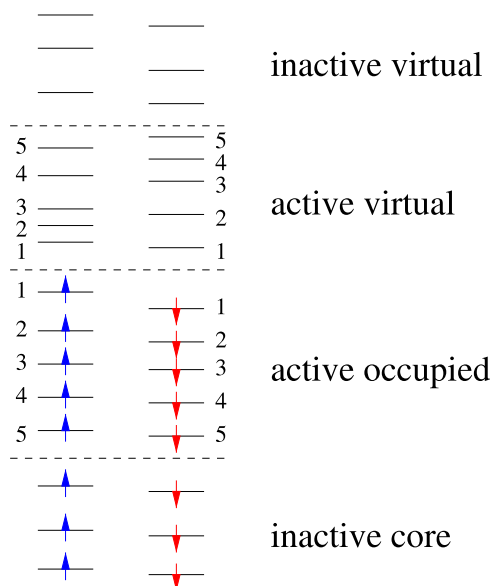


FIG. 1. Division of the orbital space. The active space orbitals are numbered with the value of the pairing label. The hole indices are reversed, because the strongest coupling typically occurs between the highest occupied and lowest unoccupied orbitals.

where i and j denote occupied orbitals, a and b denote virtual orbitals, P and Q are pairing labels that run over the active space, and δ_{PQ} is the Kronecker delta. It is easy to see that Equation (1) for the double excitation amplitudes agrees with the original definition of the PQ model as the set of amplitudes $t_{i_1 \dots i_n}^{a_1 \dots a_n}$ that satisfy $\{a_k, i_k\} \in \{\text{pair}_1\} \times \{\text{pair}_2\}$. As this kind of notation is unusual to most chemists, we will try to clarify the meaning of Equation (1) in the following. In the convention adapted above, all the orbitals in the alpha and beta occupied and virtual spaces are numbered from 1 to N , which is also the range spanned by the pairing labels. Because we have assumed that the spin has been integrated out, all the quantities in Equation (1) are scalar numbers. The indices i, j, a, b are fixed by the left hand side of the equation, which represents the element of the excitation tensor under decomposition. Applying the decomposition to, for example, the element t_{kk}^{cc} with $k \neq c$ will only yield a contribution from the last term of Equation (1) (which equals t_{kk}^{cc}), because the first term would require $k = c$ and all the other terms would contribute only for $P = Q$ which has been excluded in the summations.

The restriction to identically sized active occupied and virtual spaces becomes apparent here, as the same pairing labels cannot be used for spaces of different sizes. Each value of a pairing label corresponds to a quartet of orbitals: the alpha and beta occupied orbital and the alpha and beta virtual orbital. The tensor decomposition in Equation (1) is known in mathematics as a block decomposition.^{43,44} For PH, in addition to the 8 cases shown in Equation (1), one obtains 63 more, corresponding to all the inequivalent ways of distributing three labels to four indices: for instance, in PQ t_{QQ}^{QP} is equivalent to t_{PP}^{PQ} as the dummy summation indices P and Q can be interchanged. Having enumerated the different subtensor representations for all the necessary tensors, one proceeds by inserting them into the CC equations and by

performing the summations over the Kronecker deltas. When this is done, one ends up with a formulation of PQ or PH theory in terms of the dense subtensors. For example, in the case of PH, the original sextuple excitation operator T_{ijklmn}^{abcdef} , which is a rank-12 tensor, is reduced to a set of rank-3 tensors, and any summations that are left over are only with respect to the pairing labels. While, in principle, the doubles contribution to the opposite spin correlation energy

$$E_c \leftarrow \sum_{ijab} t_{ij}^{ab} v_{ij}^{ab} \quad (2)$$

has $8 \times 8 = 64$ distinct subtensor contributions in PQ (and 5041 in PH), it is seen that due to orthogonality of the Kronecker indices, the PQ expression becomes simply

$$\begin{aligned} E_c \leftarrow & \sum_P t_{PP}^{PP} v_{PP}^{PP} + \sum_{\substack{P, Q=1 \\ P \neq Q}}^N t_{PP}^{PQ} v_{PP}^{PQ} \\ & + \sum_{\substack{P, Q=1 \\ P \neq Q}}^N t_{PP}^{QP} v_{PP}^{QP} + \sum_{\substack{P, Q=1 \\ P \neq Q}}^N t_{PQ}^{PP} v_{PQ}^{PP} \\ & + \sum_{\substack{P, Q=1 \\ P \neq Q}}^N t_{QP}^{PP} v_{QP}^{PP} + \sum_{\substack{P, Q=1 \\ P \neq Q}}^N t_{PQ}^{PQ} v_{PQ}^{PQ} \\ & + \sum_{\substack{P, Q=1 \\ P \neq Q}}^N t_{QP}^{PQ} v_{QP}^{PQ} + \sum_{\substack{P, Q=1 \\ P \neq Q}}^N t_{PP}^{QQ} v_{PP}^{QQ}. \end{aligned} \quad (3)$$

The limitation to $P \neq Q$ in Equation (3) and similar equations can be easily lifted by establishing a convention in which all diagonals of the subtensors vanish.

One can identify a set of 8 subtensors in Equations (1) and (3): $t_{PP}^{PP}(P)$, $t_{PP}^{PQ}(P, Q)$, $t_{PP}^{QP}(P, Q)$, $t_{PQ}^{PP}(P, Q)$, $t_{PQ}^{PQ}(P, Q)$, $t_{QP}^{PP}(P, Q)$, $t_{QP}^{PQ}(P, Q)$, and $t_{PP}^{QQ}(P, Q)$. The first term in Equation (1), $t_{PP}^{PP}(P)$, can be identified as the PP amplitude (singles are usually not included in the definition of PP as the orbitals are normally optimized). As the simplest truncation (exactness for two electrons in two orbitals), PP has a special place within the hierarchy of the truncated models. Because PP couples every occupied orbital to a specific virtual orbital, the amplitudes are independent and can be solved for in a closed form. Orbital optimization can then be done in a routine fashion even for large systems.⁴⁵ The terms $t_{PP}^{PP}(P)$, $t_{PQ}^{PQ}(P, Q)$, and $t_{QP}^{PQ}(P, Q)$ in Equation (1) are the basis of the imperfect pairing (IP) model,⁴⁶ which is considerably more complicated than PP but has still been originally derived by hand, and for which orbital optimization is still a routine task.

However, in addition to the terms in Equation (1) that represent double excitations, PQ also includes single, triple, and quadruple excitations, and unlike PP and IP, PQ is not restricted to opposite-spin correlation, so the derivation of the equations for PQ would be a challenging task by hand. Fortunately, it is an ideal task for a computer.

The antisymmetry between same-spin indices can be used to eliminate redundant storage of equivalent contributions (e.g., $t_{PQ}^{QP} = -t_{PQ}^{PQ}$ for the $\alpha\alpha$ and $\beta\beta$ blocks in Equation (1)). The resulting storage costs for the subtensors that appear at the PP, PQ, and PH levels of theory are shown in Table I.

TABLE I. Storage costs for the t and λ amplitudes, one-electron integrals f , two-electron integrals v , and one- and two-particle density matrices γ and Γ , respectively, that appear in unrestricted PP, PQ, and PH. PP includes all rank-1 subtensors, PQ includes all rank-1 and rank-2 subtensors, and PH includes all rank-1 through rank-3 subtensors.

	Rank-1	Rank-2	Rank-3
t/λ	3	17	64
f	10	6	0
v	13	109	96
γ	6	6	0
Γ	13	107	96

For instance, for PQ, the storage cost is $17N^2 + 3N$ for the amplitudes, $113N^2 + 19N$ for the density matrices, and $115N^2 + 23N$ for the integrals. The figures in Table I also include the classes of integrals that are only needed for the calculation of the CC pseudoenergy; for a t amplitude-only program, the amount of integrals would be slightly smaller. Table II details the distribution of the total amplitude subtensors into the individual excitation ranks.

Due to the truncation of the amplitudes and integrals, PP, PQ, and PH are not invariant to rotations in the occupied–occupied or virtual–virtual blocks, unlike their parent CC theories. The same will also apply perfect octuples (POs) theory, which would be a truncation of CC theory with single through octuple excitations (CCSDTQ5678) based on exactness for eight electrons in eight orbitals. This is somewhat counterintuitive, because in contrast to PP, PQ, and PH, the full set of one- and two-electron integrals is used in PO, encoding in principle all there is to know about the system. But, the triple through octuple excitation operators would still not be fully described within PO, as the truncation is based on orbital labels, the used set of orbitals would still be an issue in PO. However, seeing as the issues with the choice of orbitals clearly ease up going from PP to PQ to PH, PO might not be very sensitive to the orbitals after all.

Due to the one-to-one coupling of occupied and virtual orbitals, PP (and IP) is obviously dependent on the relative ordering of the occupied and virtual orbitals. But, the same property also applies to the PQ and PH models. For instance, if one interchanges the virtual orbitals 2 and 3, the PQ doubles amplitude t_{12}^{12} in the original numbering would become t_{12}^{13} in the renumbered system. However, t_{12}^{13} is not included in the PQ model, so the description of the system has changed. By

TABLE II. Number of amplitude subtensors at different levels of truncation, including all spin blocks. PP includes all rank-1 subtensors, PQ includes all rank-1 and rank-2 subtensors, and PH includes all rank-1 through rank-3 subtensors.

	Rank-1	Rank-2	Rank-3
T_1/Λ_1	2	2	
T_2/Λ_2	1	9	8
T_3/Λ_3		5	29
T_4/Λ_4		1	21
T_5/Λ_5			5
T_6/Λ_6			1

a similar analysis, it is easy to see that the dependence on the orbital orderings holds for all tensors that are truncated in the subtensor expansion. Due to this common property, we refer to the hierarchy of the PP, PQ, and PH (as well as higher) models as the *perfect pairing hierarchy* to underline the pairing of the orbital quartets, formed of occupied and virtual alpha and beta orbitals, at each level of theory.

The origin of the dependence on the relative ordering of the occupied and virtual orbitals is easy to understand on physical grounds. In order to reduce the computational scaling, we wished to retain only the strongest interactions between orbitals. The pairing of occupied orbitals to virtual orbitals introduces a sense of locality in the model that results in fast convergence of the truncation, as in the usual local dynamic correlation methods.^{47,48} The models in the PP hierarchy are, however, invariant to swaps between pairs (i.e., quartets) of orbitals.

The models in the PP hierarchy assume a pairing of the occupied and virtual orbitals before the calculation is began. While PP optimized orbitals will produce the best possible pairing in the sense of reproducing the lowest energy at the PP level of theory, it is not clear that the PP orbitals, which usually turn out to be well-localized, will be optimal for the PQ or PH models: for instance, in polyacenes, the optimal orbitals for CCVB (which describes more strong correlation than PP) turn out to be a bit more delocalized than PP orbitals.²⁹ In all our experience so far, the optimal orbitals have turned out to be localized. However, as stated in the Introduction, variational optimization of the orbitals for PQ and PH is hard, which means that one may need to resort to approximate methods of obtaining localized orbitals with matching virtual orbitals. But, this may not be a problem after all: while orbital optimization is necessary at the PP level of theory due to its aggressive truncation (which also makes the optimization problem tractable), it is plausible that the use of optimal orbitals is less important at higher levels of the hierarchy, because at every step the models become closer to FCI which is orbital invariant. Reasonable orbitals can be obtained from a Hartree–Fock or density-functional theory calculation by localizing the occupied orbitals using, e.g., the Foster–Boys⁴⁹ (FB), Edmiston–Ruedenberg⁵⁰ (ER), or Pipek–Mezey^{51,52} (PM) criteria, after which matching virtual orbitals can be obtained, e.g., by using the Sano procedure.⁵³ In the present work, both PP orbitals and generalized Pipek–Mezey orbitals⁵² are used.

III. IMPLEMENTATION

We have written a C++ program that generates C++ source code for the PP, PQ, and PH models, starting from CC equations in spin-orbital form. In the generated code, the subtensors are stored in memory separately as arrays. Subtensor symmetry, like $t_{PQ}^{PQ}(P, Q) = t_{PQ}^{PQ}(Q, P)$ for the same-spin doubles excitation amplitudes, is not used in the storage. As is usual for CC theory, the implementation is based on the use of intermediates in evaluating nested tensor products, such as in the singles contribution to the doubles amplitude

$$t_{ij}^{ab} \leftarrow \sum_{klcd} t_k^{ab} t_l^c t_j^d v_{kl}^{cd}.$$

Pairings are also generated for the intermediates, and the index antisymmetries of the intermediates are used to eliminate labelings that vanish by Pauli exclusion. The truncation of the intermediates and integrals is controlled separately from that of the excitation amplitudes. For PQ the integrals are truncated to 2 pairs while the intermediates can be truncated to either 2 or 3 pairs, and for PH a 3-pair truncation is always used. As a result, the solution of PQ will scale as $O(N^3)$ or $O(N^4)$, depending on which truncation of the intermediates is used, and PH will also scale as $O(N^4)$. In the present work, both PQ and PH use 3-pair intermediates, yielding asymptotic $O(N^4)$ scaling for both models.

The program performs the pairings by *brute force* by looping over all the necessary index permutations of the amplitude diagrams in spin-orbital form, performing the integration over spin into different spin blocks, and performing the pairing for each block separately. The generation of the PQ model was found to take roughly 3 min on a single core of a workstation, while generation of the PH model took 14.5 days on 32 cores on a cluster node. Almost all the time for the PH model was spent in pairing the sextuple t excitation amplitudes.

The equations for the subtensor contributions are saved on disk in human-readable form. OpenMP parallelization is performed over the individual contributions. For example, the PQ model has 4529 different contributions to the t amplitudes, while PH has 213 234 different contributions to the t amplitudes. Altogether, the generated PQ equations take 27 megabytes of disk space in 10 080 files, while the PH equations take four gigabytes of disk space in almost a quarter million files.

Besides merging contributions that require identical sets of intermediates, no attempt has been made to search for an optimal reduced set of contributions. This procedure reduces the amount of separate contributions by roughly one half from the numbers given above. While common subexpression elimination techniques commonly performed in conventional CC theory could clearly be used here, it appears that because of the introduction of a large variety of subtensors, more optimal sets of intermediates could be found with a thorough graph search.

Only a single copy of the integrals is kept in memory. A disk based direct inversion in the iterative subspace⁵⁴ (DIIS) approach is used to accelerate the convergence of the CC amplitudes.⁵⁵

In contrast to conventional CC theory, the pairing of the indices introduces elementwise (Hadamard) products in the pairwise tensor products. Now, while the number of permutations of a rank- n tensor increases as $n!$, the number of possible products of two rank- n tensors would increase as $(n!)^2$, which becomes unmanageable even for relatively small n . For this reason, our implementation is based on hardcoded C++ computation kernels for each permutationally invariant type of contribution, where indices are either fixed (output index appears only on one of the two tensors), elementwise multiplied over (output index appears on both tensors), or contraction indices (index appears on both tensors but is not an output index). The kernels are machine generated in an external library, both as a *for*-loop-only implementation and as

code calling basic linear algebra subprogram (BLAS) routines wherever applicable. When the library is generated, the two kinds of implementations are benchmarked to determine which one to use for a given kernel.

The two-electron integrals are evaluated in the program from B matrices that can be generated by the resolution of the identity⁵⁶ (RI) or Cholesky decomposition⁵⁷ (CD) techniques, as their integral transforms scale more favorably than that of the traditional two-electron integrals, and as the RI/CD representation allows for cherry-picking of the wanted integral elements. The Fock and RI/CD B matrices may currently be obtained from either Q-Chem^{58,59} or ERKALE.^{60,61} Both programs have been utilized in the present work: Q-Chem for calculations using PP orbitals and ERKALE for calculations with localized Hartree–Fock orbitals. In the latter case, the active occupied orbitals are localized using the generalized Pipek–Mezey method⁵² using the intrinsic atomic orbital partial charge estimate.⁶² The gradient descent method used for the localization has been described elsewhere.⁶³ After the occupied space has been localized, corresponding virtual orbitals are obtained for each active occupied orbital using the Sano procedure.⁵³ All calculations performed with Q-Chem used exact integrals for the formation of the Fock matrices, and RI or CD for forming the B matrices for the two-electron integrals for the CC procedure. The calculations performed with ERKALE used only Cholesky integrals for both the Fock and B matrices. To ensure near-exactness, a very small truncation threshold (10^{-8}) was used for the Cholesky procedure both in ERKALE and Q-Chem.

IV. RESULTS

A. Scaling

To demonstrate the scaling of the novel implementations of PQ and PH with respect to the original implementation of the PQ model, we study full-valence calculations on linear all-trans polyenes C_nH_{n+2} with the geometries described in Ref. 26. The STO-3G basis set⁶⁴ and PP optimized orbitals are used, and the two-electron integrals are formed with RI using the auxiliary basis set corresponding to the correlation consistent double- ζ basis set.⁶⁵ The timings shown correspond to the use of a single core on one of the authors' (S.L.) desktop computer with an Intel Core i7-4770 processor and 32 gigabytes of memory. No data are shown for sparse PH, because due to inefficiencies in its implementation, getting sensible scaling data would have required hundreds of gigabytes of memory.

While the use of local PP orbitals in a quasi-1D molecule clearly would favor algorithms employing sparsity, the results shown in Figure 2 emphatically show the tremendous speedup achieved with the dense tensor formalism that does not use screening of small elements at all. The distribution of the computational work with the PQ model is shown in Table III. As is seen, most of the work in the model is performed in the multiplications, with permutations also contributing a nontrivial fraction of the total runtime. The compute kernels and permutations account altogether for 85% of the runtime of the model, with another 11% being spent

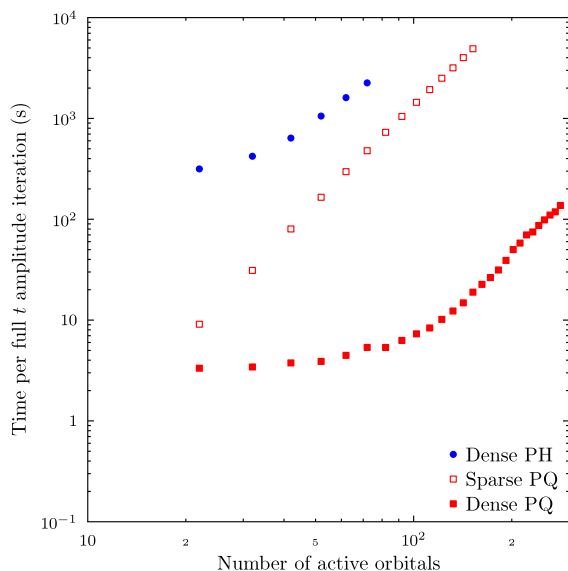


FIG. 2. Scaling of the new dense tensor implementations of PQ and PH compared to the earlier implementations⁴⁰ based on sparse tensors. Note the logarithmic scale.

on dynamic memory management. A small fraction of the total runtime for large systems is lost due to inefficiencies in the present implementation, which are also obvious in Figure 2 for small systems for which the time for solution is quasi-constant. (Due to the larger size of the equations, the inefficiencies are even larger for PH.) The scaling plot also shows some bumps, which are likely artifacts of background processes running simultaneously on the workstation.

The mathematical scaling of the models is determined by the individual multiplication kernels that appear in the model. The individual multiplication kernels for the three-pair intermediate PQ model are shown in Table IV. Here, all the kernels that contain a contraction use BLAS implementations. For instance, the heaviest kernel in Table IV is implemented with a *dgemv* call within a *for*-loop over x , over which there is an elementwise product. Similarly, the second-heaviest kernel is also a *dgemv* call inside a *for*-loop over the y index. The 7th heaviest kernel is just a single *dgemm* call. As can be

TABLE III. Distribution of computational effort in the solution of the PQ model for $C_{52}H_{54}$. The timings include all steps in solving the PQ model: the solution of the t and λ amplitudes as well as the formation of the one- and two-particle density matrices γ and Γ , respectively.

Time spent in multiplication kernels (Table IV)	3617.84
Time spent in addition kernels	42.34
Time spent in permutations	2331.36
Memory allocation	761.73
Integrals	9.27
Diagonal zeroing	39.72
Denominator application	0.17
Sum of above	6802.40
Total elapsed wall time	7047.52

TABLE IV. Multiplication kernels that appear in the three-pair intermediates PQ model with timings for the $C_{52}H_{54}$ molecule.

Kernel	Time (s)	Scaling	BLAS
$o(x) \leftarrow l(x)r(x)$	0.00	N	No
$o(x, y) \leftarrow l(x, y)r(x, y)$	0.01	N^2	No
$o(x) \leftarrow l(x)r(x)$	0.05	N	No
$o(x, y) \leftarrow l(x)r(y)$	0.07	N^2	No
$o(x) \leftarrow \sum_y l(x, y)r(y)$	0.56	N^2	Yes
$o(x, y, z) \leftarrow l(y, z)r(x)$	1.27	N^3	No
$o(x, y) \leftarrow l(x, y)r(y)$	13.51	N^2	No
$o(x) \leftarrow \sum_y l(x, y)r(x, y)$	21.25	N^2	Yes
$o(x, y) \leftarrow l(x, y)r(x)$	23.31	N^2	No
$o(x, y) \leftarrow \sum_z l(x, z)r(y, z)$	24.90	N^3	Yes
$o(x, y, z) \leftarrow l(x, y, z)r(z)$	28.69	N^3	No
$o(x, y) \leftarrow l(x, y)r(x, y)$	47.89	N^2	No
$o(x, y, z) \leftarrow \sum_w l(y, z, w)r(x, w)$	83.04	N^4	Yes
$o(x, y, z) \leftarrow l(x, z)r(y, z)$	110.31	N^3	No
$o(x, y, z) \leftarrow l(x, y, z)r(x)$	125.27	N^3	No
$o(x, y, z) \leftarrow \sum_w l(x, y, w)r(z, w)$	127.99	N^4	Yes
$o(x, y, z) \leftarrow l(x, y, z)r(x, z)$	145.91	N^3	No
$o(x, y, z) \leftarrow l(x, y, z)r(x, y)$	151.28	N^3	No
$o(x, y, z) \leftarrow l(x, y)r(y, z)$	501.45	N^3	No
$o(x, y, z) \leftarrow l(x, y)r(x, z)$	502.06	N^3	No
$o(x, y) \leftarrow \sum_z l(x, y, z)r(y, z)$	521.73	N^3	Yes
$o(x, y) \leftarrow \sum_z l(x, y, z)r(x, z)$	1187.29	N^3	Yes

seen from the kernels, PQ on $C_{52}H_{54}$ with an active space of 202 electrons in 202 orbitals is still in the cubic scaling regime, as the heaviest kernels on the list are cubic scaling. In contrast, PH on $C_{14}H_{16}$ with an active space of 72 electrons in 72 orbitals is on the verge of turning from cubic to quartic scaling, as while the heaviest kernel (3.3 processor hours) is cubic scaling, the second-heaviest kernel (2.7 processor hours) is quartic scaling.

The novel implementation of PQ is fast even for large orbital spaces, and typically much greater effort is spent in the orbital optimization with either the self-consistent field method or PP, as well as the integral transforms: detailed timings for the different steps in the calculation are shown in Table V. As can be seen from Figure 2, the PQ calculation for an active space of 140 electrons in 140 orbitals can be performed in a matter of minutes on a single core with the current version of the code that has not yet been heavily optimized. Even lower scaling could also be obtained by truncating the intermediates in PQ to two pairs, but this would also limit the accuracy of the model.¹⁴

B. Accuracy

To study the accuracy of the generated models in the perfect pairing hierarchy, in the following, we apply the PP, PQ, and PH models on the π space of the polyenes used above for the scaling study, the symmetric and asymmetric dissociation of the H_{50} chain, as well as the strongly correlated π space of polyacene molecules. Hydrogen chains and polyacenes are standard chemical models for testing models of strong correlation.^{66–69} The π space correlation energies for the polyenes are shown in Table VI, from which the accuracy

TABLE V. Comparison of wall timings in seconds for the computational procedure of PQ and PH as a function of active orbitals n_{orb} . The third through fifth column detail the base procedure: SCF solution followed by the PP orbital optimization and the formation of the RI B matrix, while columns 6, 7, 8, and 9 (10, 11, 12, and 13) give the time needed to compute the integrals for PQ (PH), the total time for the t amplitude iterations needed to solve PQ (PH), as well as the fraction of total runtime in percent spent in permutation operations and floating point compute kernels, respectively.

Molecule	n_{orb}	SCF	PP	B matrix	PQ ints	PQ t	PQ perm	PQ flop	PH ints	PH t	PH perm	PH flop
C ₄ H ₆	22	0.23	0.35	0.68	0.00	66.63	1.3	1.8	0.05	6 942.94	4.8	14.2
C ₆ H ₈	32	0.27	0.73	0.86	0.01	75.42	1.8	3.9	0.15	12 228.76	8.8	28.2
C ₈ H ₁₀	42	0.35	1.45	1.13	0.03	90.27	2.7	7.5	0.39	20 436.82	12.8	41.9
C ₁₀ H ₁₂	52	0.44	2.37	1.65	0.06	105.08	4.3	12.7	0.79	39 056.60	16.8	48.9
C ₁₂ H ₁₄	62	0.64	3.65	2.47	0.11	120.68	5.6	17.8	1.49	59 499.11	19.2	54.5
C ₁₄ H ₁₆	72	1.10	5.05	3.89	0.18	144.63	7.1	21.7	2.72	83 323.67	20.4	58.4
C ₁₆ H ₁₈	82	1.71	7.10	5.89	0.26	144.47	9.2	29.8				
C ₁₈ H ₂₀	92	2.48	9.92	8.83	0.37	169.61	11.9	34.9				
C ₂₀ H ₂₂	102	3.54	13.07	13.56	0.48	197.28	13.7	39.6				
C ₂₂ H ₂₄	112	6.98	19.06	18.81	0.68	225.59	15.5	44.0				
C ₂₄ H ₂₆	122	7.83	24.13	26.78	0.89	274.37	18.5	47.5				
C ₂₆ H ₂₈	132	9.11	30.51	37.99	1.11	332.18	20.6	50.0				
C ₂₈ H ₃₀	142	10.53	37.52	55.58	1.40	400.48	22.5	52.8				
C ₃₀ H ₃₂	152	12.07	46.01	70.23	1.73	510.74	26.5	51.5				
C ₃₂ H ₃₄	162	13.64	55.16	104.76	2.11	610.46	27.1	53.4				
C ₃₄ H ₃₆	172	15.62	66.35	130.54	2.51	713.33	27.6	54.9				
C ₃₆ H ₃₈	182	17.26	77.59	161.94	2.99	847.72	28.5	56.1				
C ₃₈ H ₄₀	192	19.32	94.63	242.02	3.50	1055.37	31.7	55.3				
C ₄₀ H ₄₂	202	21.33	107.16	339.07	4.14	1351.42	30.8	51.7				
C ₄₂ H ₄₄	212	25.34	125.60	436.98	4.77	1566.79	32.0	51.7				
C ₄₄ H ₄₆	222	27.52	143.47	586.43	5.48	1888.20	32.5	51.4				
C ₄₆ H ₄₈	232	30.12	166.73	1108.01	6.28	2020.66	32.0	51.7				
C ₄₈ H ₅₀	242	33.16	188.38	1366.80	7.31	2343.37	32.4	51.6				

of the PP hierarchy already becomes apparent: even for the largest system, PH captures 99.3% of the full correlation energy.

Next, while H₅₀ is chemically uninteresting, it is an excellent system for studying models for strong correlation. When the interatomic separation is increased, the system switches from a metallic state to an insulating state with strong multireference character in the intermediate region that is not captured by CC with full single and double (CCSD) or CC with full single and double and perturbative triple excitations

TABLE VI. Polyene π space correlation energy in the STO-3G basis set with generalized Pipek–Mezey localized occupied orbitals paired with Sano virtuals with the PP, PQ, and PH models, compared to DMRG values from Ref. 26. The size of the active space coincides with the number of C atoms in the molecule. Note that PQ and PH are exact for the (4e,4o) system in C₄H₆.

Molecule	PP	PQ	PH	DMRG
C ₄ H ₆	-0.080 460	-0.091 502	-0.091 502	-0.091 502
C ₈ H ₁₀	-0.143 334	-0.172 876	-0.176 908	-0.177 127
C ₁₂ H ₁₄	-0.205 042	-0.253 097	-0.261 466	-0.262 297
C ₁₆ H ₁₈	-0.266 492	-0.333 088	-0.345 899	-0.347 403
C ₂₀ H ₂₂	-0.327 882	-0.413 030	-0.430 310	-0.432 498
C ₂₄ H ₂₆	-0.389 257	-0.492 961	-0.514 717	-0.517 591
C ₂₈ H ₃₀	-0.450 629	-0.572 889	-0.599 123	-0.602 684
C ₃₂ H ₃₄	-0.512 000	-0.652 816	-0.683 528	-0.687 777
C ₃₆ H ₃₈	-0.573 371	-0.732 743	-0.767 934	-0.772 870
C ₄₀ H ₄₂	-0.634 742	-0.812 669	-0.852 338	-0.857 963
C ₄₄ H ₄₆	-0.696 112	-0.892 595	-0.936 743	-0.943 056
C ₄₈ H ₅₀	-0.757 483	-0.972 521	-1.021 147	-1.028 149

(CCSD(T)).²⁶ In the symmetric dissociation, the equispaced chain is stretched apart into noninteracting H atoms. In the asymmetric dissociation, the system is composed of 25 H₂ molecules with bond length $R = 1.4a_0$, where a_0 is the Bohr radius, which are placed on the z axis, and the energy is studied as a function of the intermolecular distance. Here, the STO-6G basis set⁶⁴ and PP optimized orbitals are used. The two-electron integrals for PQ and PH are generated with CD. The DMRG data have been taken from Ref. 26.

The obtained total energies for asymmetric dissociation are shown in Figure 3. For large intermolecular separation, the system is essentially composed of noninteracting H₂ molecules, for which already PP is exact. When the molecules are pushed closer, the differences in the performance of the models become apparent: while PP quickly starts deviating from the exact solution, PQ and PH stay more accurate on a wider range of correlation strength. This is also clearly visible in the fraction of correlation energy captured by the models shown in Figure 4. For the most strongly correlated geometry, in which the atoms are equidistant, PP, PQ, and PH capture 35.2%, 63.9%, and 87.9% of the correlation energy, respectively. PH becomes practically exact (disagreeing from the DMRG correlation energy by at most 1%) at the distance $d = 2.0a_0$, where it captures 99.2% of the correlation energy. For PQ, practical exactness is reached at $d = 2.8a_0$ with 99.4% of the correlation energy retained. PP reaches exactness at $d = 4.2a_0$, where it describes 99.0% of the correlation.

The symmetric dissociation is a much harder problem, but even here the models fare seemingly well, as shown by the total and correlation energy plots in Figures 5 and 6,

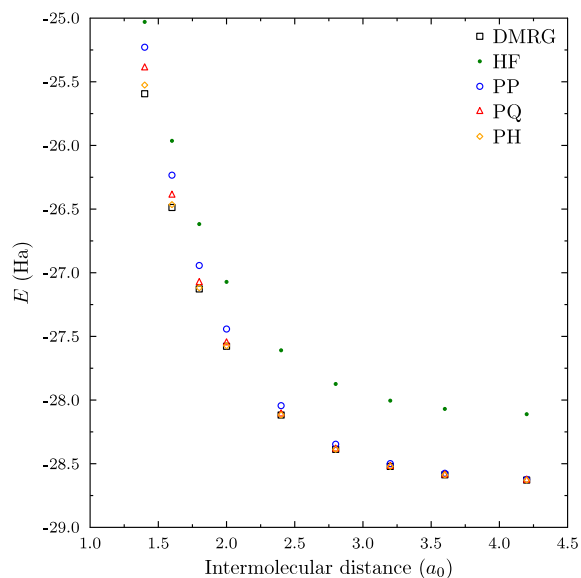


FIG. 3. Asymmetric dissociation of the H_{50} chain into equidistant hydrogen molecules with fixed bond length $R = 1.4a_0$.

respectively. As before with the asymmetric stretch, a clear hierarchy is again seen between the models, with the most strongly correlated geometry showing 22.0%, 47.6%, and 80.9% of the correlation energy being captured by PP, PQ, and PH, respectively. As the chain is pulled apart, the models quickly become more accurate, PH reaching practical exactness at separation $d = 2.8a_0$ where it captures 99.0% of the correlation energy. Unfortunately, for larger interatomic spacings, the CC iterations diverged for PQ and PH (both with the zero and PP guess for the amplitudes); thus, some data points are missing. However, because PP appears to become better and better for large separations, the divergence issue might be solved by the use of a better preconditioner.

The linear polyacenes are known to exhibit multiradical behavior in the π space.⁷⁰ Next, we will try to match the

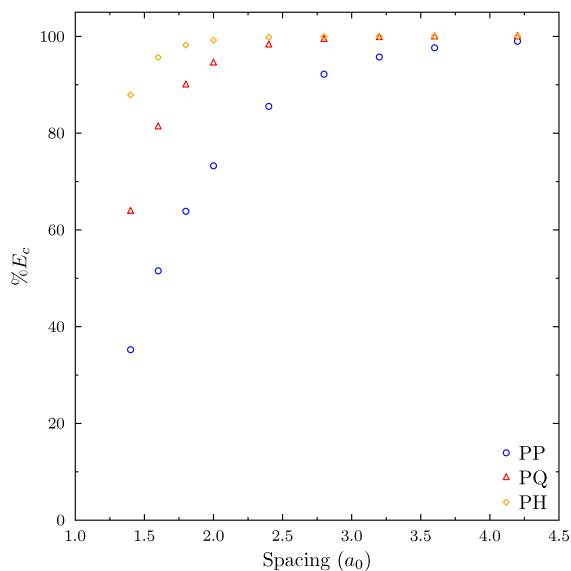


FIG. 4. Fraction of correlation energy captured by the PP, PQ, and PH models for the asymmetric dissociation of the H_{50} chain.

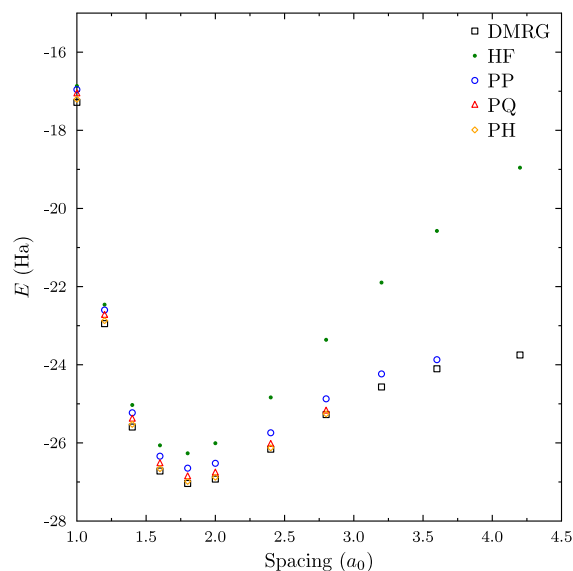


FIG. 5. Symmetric dissociation of the H_{50} chain into equidistant hydrogen atoms.

DMRG energies in the STO-3G basis from Ref. 70 using their geometries. Generalized Pipek–Mezey localized active occupied orbitals paired with Sano virtuals were used. The energies obtained with the PP, PQ, and PH models compared against the DMRG reference are shown in Table VII. Here, the models capture 44%–59%, 69%–88%, and 96%–99% of the DMRG correlation energy for PP, PQ, and PH, respectively. We can also compare the natural occupation numbers produced by PP, PQ, and PH against the ones from DMRG; this is done in Figure 7. While PP and PQ successfully capture a significant amount of the correlation energy, they fail to capture the strong correlation effects in the π space of the polyacenes. For PP, the occupation number plot has an almost step function like character, while PQ reproduces some curvature but is still far from the DMRG reference. In contrast, PH reproduces

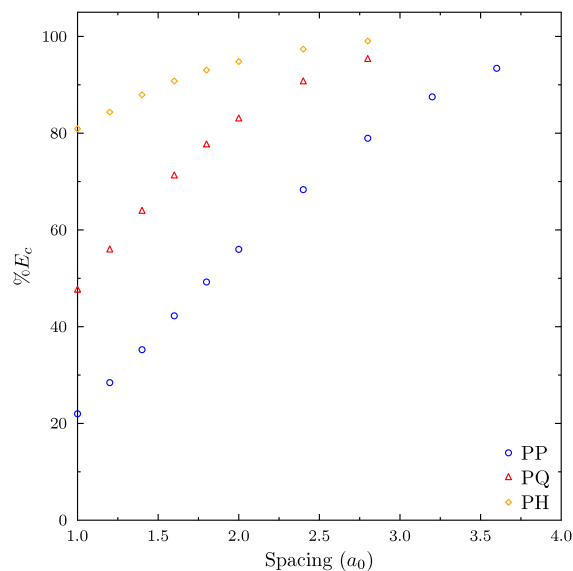


FIG. 6. Fraction of correlation energy captured by the PP, PQ, and PH models for the symmetric dissociation of the H_{50} chain.

TABLE VII. Polyacene π space correlation energy in the STO-3G basis set with generalized Pipek–Mezey localized occupied orbitals paired with Sano virtuals with the PP, PQ, and PH models, compared to DMRG values from Ref. 70. Note that PQ and PH are exact for the (4e,4o) system in C_4H_6 .

Molecule	Active space	PP	PQ	PH	DMRG
2acene	(10e,10o)	-0.104 902	-0.157 684	-0.177 098	-0.178 294
3acene	(14e,14o)	-0.144 544	-0.213 115	-0.249 895	-0.254 307
4acene	(18e,18o)	-0.190 182	-0.279 345	-0.325 491	-0.332 661
5acene	(22e,22o)	-0.228 839	-0.335 088	-0.401 098	-0.412 627
6acene	(26e,26o)	-0.265 084	-0.393 537	-0.479 533	-0.495 683
8acene	(34e,34o)	-0.322 287	-0.495 219	-0.641 147	-0.668 526
10acene	(42e,42o)	-0.382 622	-0.595 840	-0.803 505	-0.838 580
12acene	(50e,50o)	-0.444 834	-0.698 165	-0.968 521	-1.007 378 ^a

^aDMRG value of Ref. 70 was not converged.

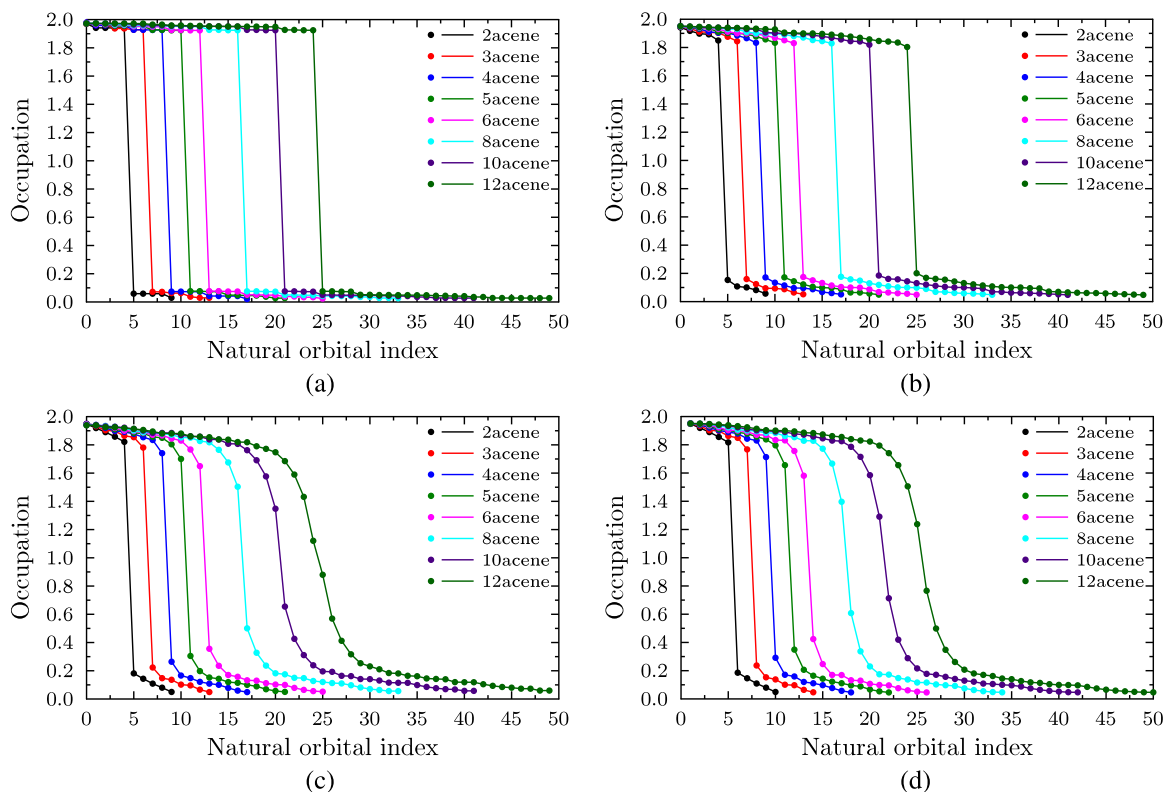


FIG. 7. Polyacene π space natural orbital occupation numbers. (a) PP. (b) PQ. (c) PH. (d) DMRG data from Ref. 70.

a clear signal of strong correlation that is strikingly similar to the DMRG reference values. Quantitative agreement with DMRG is not, however, reached, which we primarily attribute to the use of non-optimal orbitals. Indeed, other choices for the orbital localization method (see the [supplementary material](#)) demonstrate that a different choice for the active space orbitals affects the natural orbital occupation numbers, as well as the fraction of energy captured. Furthermore, as generalized Pipek–Mezey localized Hartree–Fock orbitals were found to yield better energies than PP optimized orbitals at the UPH level of theory, we conclude that as has been seen for CCVB,²⁹ PQ and PH presumably need orbitals that are less localized than those reproduced by PP, and the best choice for the (non-optimized) orbitals is still an open question.

V. SUMMARY AND DISCUSSION

We have described the low-scaling dense tensor implementation of a family of truncated CC methods and generated efficient implementations of the PQ and PH models. High-rank tensors are expressed in terms of a considerable number of subtensors, and summations over Kronecker delta indices are performed before computer code is generated. As straightforward truncations of CC theory, the models inherit all the properties of the parent theory, such as size extensivity, and pre-existing machinery for the calculation of properties could be used.

We have demonstrated impressive speedups (of roughly 200-500 times) compared to the earlier implementation that already had the correct $O(N^4)$ scaling for PQ, and

demonstrated the rapid convergence of the PP, PQ, and PH models in describing strong correlation with applications to linear, all-trans polyenes, the dissociation of the H₅₀ chain, as well as linear polyacenes. Although only minimal basis sets have been used in the present work due to limited availability of high-level reference data, the runtime of the models after integral transformation is basis set independent due to the use of an active space (and the lack of orbital optimization that has not been pursued in the present work). While the models are already computationally fast enough to allow their use in studies of chemical problems, we believe that they can still be made even faster by a variety of approaches. Most importantly, the role of the intermediates in the paired theories is clearly different from the intermediates in full-rank CC theory. A thorough procedure for identifying common factors between different contributions, as well as a graph-type approach for determining the optimal set of intermediates, could plausibly yield a further order of magnitude speedup to the existing implementation. In the case of closed-shell molecules, spin summation of the model would reduce storage costs for the amplitudes by roughly 50% and cut down on computational costs as well. A reorganization of the way the amplitude updates are performed might yield convergence in a smaller number of iterations.⁷¹ Because of the mathematical structure of the models, the storage of the tensors is already distributed, and massively parallel implementations of the models could be pursued.

While we have demonstrated that the truncation hierarchy advocated in the present work allows for applications to large systems, it is further important to note that in systems with little static correlation, it is possible to further speed things up. While in PQ and PH the tensors that correspond to excitations of different ranks have similar storage requirements (within an order of magnitude) as shown in Table II, the runtime tends to be dominated by the highest excitations due to the larger amount of antisymmetrizations necessary in the CC diagrams. Thus, should the higher excitation operators turn out to be insignificant, disabling them can yield further speedups. For example, the PH model with only single and double (and triple and quadruple) excitations enabled is essentially a local CCSD (CCSDTQ) approach, if the treatment is based on localized orbitals.

As we have seen in the application to the acene series, the choice of orbitals may be important for the perfect pairing hierarchy of models, which we will address in future work. In addition, we wish to extend the PQ and PH models to open-shell systems, as well as to include the description of dynamic correlation. The similarities between the PP truncation hierarchy and local CC methods could also be further explored.

SUPPLEMENTARY MATERIAL

See [supplementary material](#) for results on the polyacenes with various choices for the active space orbitals.

ACKNOWLEDGMENTS

S.L. thanks Evgeny Epifanovsky, David Small, and Joonho Lee for discussions, and Johannes Hachmann for

supplying the DMRG natural occupation numbers for the polyacenes. This work was supported by the Director, Office of Basic Energy Sciences, Chemical Sciences, Geosciences, and Biosciences Division of the U.S. Department of Energy, under Contract No. DE-AC02-05CH11231. J.P. thanks the University of Notre Dame's College of Science and Department of Chemistry and Biochemistry for generous start-up funding.

- ¹P. Hohenberg and W. Kohn, *Phys. Rev.* **136**, B864 (1964).
- ²W. Kohn and L. J. Sham, *Phys. Rev.* **140**, A1133 (1965).
- ³Y. Kurzweil, K. V. Lawler, and M. Head-Gordon, *Mol. Phys.* **107**, 2103 (2009).
- ⁴R. E. Thomas, Q. Sun, A. Alavi, and G. H. Booth, *J. Chem. Theory Comput.* **11**, 5316 (2015); e-print [arXiv:1510.03635](#).
- ⁵G. L. Manni, S. D. Smart, and A. Alavi, *J. Chem. Theory Comput.* **12**, 1245 (2016).
- ⁶J. B. Schriber and F. A. Evangelista, *J. Chem. Phys.* **144**, 161106 (2016); e-print [arXiv:1603.08063](#).
- ⁷N. M. Tubman, J. Lee, T. Y. Takeshita, M. Head-Gordon, and K. B. Whaley, *J. Chem. Phys.* **145**, 044112 (2016); e-print [arXiv:1603.02686](#).
- ⁸W. Liu and M. R. Hoffmann, *J. Chem. Theory Comput.* **12**, 1169 (2016).
- ⁹T. Zhang and F. A. Evangelista, *J. Chem. Theory Comput.* **12**, 4326 (2016).
- ¹⁰A. A. Holmes, N. M. Tubman, and C. J. Umrigar, *J. Chem. Theory Comput.* **12**, 3674 (2016).
- ¹¹J. Čížek, *J. Chem. Phys.* **45**, 4256 (1966).
- ¹²D. I. Lyakh, M. Musiał, V. F. Lotrich, and R. J. Bartlett, *Chem. Rev.* **112**, 182 (2012).
- ¹³A. I. Krylov, C. D. Sherrill, E. F. C. Byrd, and M. Head-Gordon, *J. Chem. Phys.* **109**, 10669 (1998).
- ¹⁴J. A. Parkhill, K. Lawler, and M. Head-Gordon, *J. Chem. Phys.* **130**, 084101 (2009).
- ¹⁵J. A. Parkhill and M. Head-Gordon, *J. Chem. Phys.* **133**, 024103 (2010).
- ¹⁶A. C. Hurley, J. Lennard-Jones, and J. A. Pople, *Proc. R. Soc. A* **220**, 446 (1953).
- ¹⁷W. J. Hunt, P. J. Hay, and W. A. Goddard III, *J. Chem. Phys.* **57**, 738 (1972).
- ¹⁸I. I. Ukrainskii, *Theor. Math. Phys.* **32**, 816 (1977).
- ¹⁹W. A. Goddard and L. B. Harding, *Annu. Rev. Phys. Chem.* **29**, 363 (1978).
- ²⁰J. Cullen, *Chem. Phys.* **202**, 217 (1996).
- ²¹S. R. White, *Phys. Rev. Lett.* **69**, 2863 (1992).
- ²²G. K.-L. Chan, J. J. Dorando, D. Ghosh, J. Hachmann, E. Neuscamman, H. Wang, and T. Yanai, *Frontiers in Quantum Systems in Chemistry and Physics* (Springer, Netherlands, Dordrecht, 2008), Vol. 18, pp. 49–65; e-print [arXiv:0711.1398v1](#).
- ²³K. H. Marti and M. Reiher, *Z. Phys. Chem.* **224**, 583 (2010).
- ²⁴S. Wouters and D. Van Neck, *Eur. Phys. J. D* **68**, 272 (2014); e-print [arXiv:1407.2040](#).
- ²⁵R. Olivares-Amaya, W. Hu, N. Nakatani, S. Sharma, J. Yang, and G. K.-L. Chan, *J. Chem. Phys.* **142**, 034102 (2015).
- ²⁶J. Hachmann, W. Cardoan, and G. K.-L. Chan, *J. Chem. Phys.* **125**, 144101 (2006); e-print [arXiv:0606115](#) [cond-mat].
- ²⁷D. W. Small and M. Head-Gordon, *J. Chem. Phys.* **130**, 084103 (2009).
- ²⁸D. W. Small and M. Head-Gordon, *Phys. Chem. Chem. Phys.* **13**, 19285 (2011).
- ²⁹D. W. Small, K. V. Lawler, and M. Head-Gordon, *J. Chem. Theory Comput.* **10**, 2027 (2014).
- ³⁰J. Arponen, *Ann. Phys.* **151**, 311 (1983).
- ³¹J. S. Arponen, R. F. Bishop, and E. Pajanne, *Phys. Rev. A* **36**, 2519 (1987).
- ³²J. S. Arponen, R. F. Bishop, and E. Pajanne, *Phys. Rev. A* **36**, 2539 (1987).
- ³³R. Shepard, G. Gidofalvi, and S. R. Brozell, *J. Chem. Phys.* **141**, 064105 (2014).
- ³⁴R. Shepard, G. Gidofalvi, and S. R. Brozell, *J. Chem. Phys.* **141**, 064106 (2014).
- ³⁵J. A. Parkhill and M. Head-Gordon, *J. Chem. Phys.* **133**, 124102 (2010).
- ³⁶J. A. Parkhill, J. Azar, and M. Head-Gordon, *J. Chem. Phys.* **134**, 154112 (2011).
- ³⁷P. Pulay and T. P. Hamilton, *J. Chem. Phys.* **88**, 4926 (1988).
- ³⁸J. M. Bofill and P. Pulay, *J. Chem. Phys.* **90**, 3637 (1989).
- ³⁹S. Keller, K. Boguslawski, T. Janowski, M. Reiher, and P. Pulay, *J. Chem. Phys.* **142**, 244104 (2015).
- ⁴⁰J. A. Parkhill and M. Head-Gordon, *Mol. Phys.* **108**, 513 (2010).

- ⁴¹T. D. Crawford and H. F. Schaefer, in *Reviews in Computational Chemistry*, edited by K. B. Lipkowitz and D. B. Boyd (Wiley-VCH, John Wiley and Sons, Inc., New York, 2000), Vol. 14, pp. 33–136.
- ⁴²R. Bartlett and M. Musiał, *Rev. Mod. Phys.* **79**, 291 (2007).
- ⁴³L. De Lathauwer, *SIAM J. Matrix Anal. Appl.* **30**, 1033 (2008).
- ⁴⁴L. De Lathauwer, *SIAM J. Matrix Anal. Appl.* **30**, 1022 (2008).
- ⁴⁵G. J. O. Beran, B. Austin, A. Sodt, and M. Head-Gordon, *J. Phys. Chem. A* **109**, 9183 (2005).
- ⁴⁶T. Van Voorhis and M. Head-Gordon, *Chem. Phys. Lett.* **317**, 575 (2000).
- ⁴⁷P. Pulay, *Chem. Phys. Lett.* **100**, 151 (1983).
- ⁴⁸S. Saebo and P. Pulay, *Annu. Rev. Phys. Chem.* **44**, 213 (1993).
- ⁴⁹J. Foster and S. Boys, *Rev. Mod. Phys.* **32**, 300 (1960).
- ⁵⁰C. Edmiston and K. Ruedenberg, *Rev. Mod. Phys.* **35**, 457 (1963).
- ⁵¹J. Pipek and P. G. Mezey, *J. Chem. Phys.* **90**, 4916 (1989).
- ⁵²S. Lehtola and H. Jónsson, *J. Chem. Theory Comput.* **10**, 642 (2014).
- ⁵³T. Sano, *J. Mol. Struct.: THEOCHEM* **528**, 177 (2000).
- ⁵⁴P. Pulay, *Chem. Phys. Lett.* **73**, 393 (1980).
- ⁵⁵G. E. Scuseria, T. J. Lee, and H. F. Schaefer, *Chem. Phys. Lett.* **130**, 236 (1986).
- ⁵⁶O. Vahtras, J. Almlöf, and M. Feyereisen, *Chem. Phys. Lett.* **213**, 514 (1993).
- ⁵⁷N. H. F. Beebe and J. Linderberg, *Int. J. Quantum Chem.* **12**, 683 (1977).
- ⁵⁸E. Epifanovsky, D. Zuev, X. Feng, K. Khistyayev, Y. Shao, and A. I. Krylov, *J. Chem. Phys.* **139**, 134105 (2013).
- ⁵⁹Y. Shao, Z. Gan, E. Epifanovsky, A. T. B. Gilbert, M. Wormit, J. Kussmann, A. W. Lange, A. Behn, J. Deng, X. Feng, D. Ghosh, M. Goldey, P. R. Horn, L. D. Jacobson, I. Kaliman, R. Z. Khaliullin, T. Kus, A. Landau, J. Liu, E. I. Proynov, Y. M. Rhee, R. M. Richard, M. A. Rohrdanz, R. P. Steele, E. J. Sundstrom, H. L. Woodcock III, P. M. Zimmerman, B. A. D. Zuev, E. Alguire, B. Austin, G. J. O. Beran, Y. A. Bernard, E. Berquist, K. Brandhorst, K. B. Bravaya, S. T. Brown, D. Casanova, C.-M. Chang, Y. Chen, S. H. Chien, K. D. Closser, D. L. Crittenden, M. Diedenhofen, R. A. DiStasio, Jr., H. Do, A. D. Dutoi, R. G. Edgar, S. Fatehi, L. Fusti-Molnar, A. Ghysels, A. Golubeva-Zadorozhnaya, J. Gomes, M. W. D. Hanson-Heine, P. H. P. Harbach, A. W. Hauser, E. G. Hohenstein, Z. C. Holden, T.-C. Jagau, B. K. H. Ji, K. Khistyayev, J. Kim, R. A. King, P. Klunzinger, D. Kosenkov, T. Kowalczyk, C. M. Krauter, K. U. Lao, A. Laurent, K. V. Lawler, S. V. Levchenko, C. Y. Lin, F. Liu, E. Livshits, R. C. Lochan, A. Luenser, P. Manohar, S. F. Manzer, S.-P. Mao, N. Mardirossian, A. V. Marenich, S. A. Maurer, N. J. Mayhall, C. M. Oana, R. Olivares-Amaya, D. P. O'Neill, J. A. Parkhill, T. M. Perrine, R. Peverati, A. Prociuk, D. R. Rehn, E. Rosta, N. J. Russ, S. M. Sharada, S. Sharma, D. W. Small, A. Sodt, T. Stein, D. Stück, Y.-C. Su, A. J. W. Thom, T. Tsuchimochi, V. Vanovschi, L. Vogt, O. Vydrov, T. Wang, M. A. Watson, J. Wenzel, A. White, C. F. Williams, S. Yeganeh, S. R. Yost, Z.-Q. You, I. Y. Zhang, X. Zhang, Y. Zhao, B. R. Brooks, G. K. L. Chan, D. M. Chipman, C. J. Cramer, W. A. Goddard III, M. S. Gordon, W. J. Hehre, A. Klamt, H. F. S. III, M. W. Schmidt, C. D. Sherrill, D. G. Truhlar, A. Warshel, X. Xuaf, A. Aspuru-Guzik, R. Baer, A. T. Bell, N. A. Besley, J.-D. Chai, A. Dreuw, B. D. Dunietz, T. R. Furlani, S. R. Gwaltney, C.-P. Hsu, Y. Jung, J. Kong, D. S. Lambrecht, W. Liang, C. Ochsenfeld, V. A. Rassolov, L. V. Slipchenko, J. E. Subotnik, T. V. Voorhis, J. M. Herbert, A. I. Krylov, P. M. W. Gill, and M. Head-Gordon, *Mol. Phys.* **113**, 184 (2015).
- ⁶⁰J. Lehtola, M. Hakala, A. Sakko, and K. Hämäläinen, *J. Comput. Chem.* **33**, 1572 (2012).
- ⁶¹S. Lehtola, ERKALE - HF/DFT from Hel, 2016.
- ⁶²G. Knizia, *J. Chem. Theory Comput.* **9**, 4834 (2013).
- ⁶³S. Lehtola and H. Jónsson, *J. Chem. Theory Comput.* **9**, 5365 (2013).
- ⁶⁴W. J. Hehre, R. F. Stewart, and J. A. Pople, *J. Chem. Phys.* **51**, 2657 (1969).
- ⁶⁵F. Weigend, A. Köhn, and C. Hättig, *J. Chem. Phys.* **116**, 3175 (2002).
- ⁶⁶T. Tsuchimochi and G. E. Scuseria, *J. Chem. Phys.* **131**, 121102 (2009).
- ⁶⁷A. V. Sinititskiy, L. Greenman, and D. A. Mazziotti, *J. Chem. Phys.* **133**, 014104 (2010).
- ⁶⁸J. J. Phillips and D. Zgid, *J. Chem. Phys.* **140**, 241101 (2014).
- ⁶⁹J. Fosso-Tande, T.-S. Nguyen, G. Gidofalvi, and A. E. DePrince, *J. Chem. Theory Comput.* **12**, 2260 (2016).
- ⁷⁰J. Hachmann, J. J. Dorando, M. Avilés, and G. K.-L. Chan, *J. Chem. Phys.* **127**, 134309 (2007).
- ⁷¹D. A. Matthews and J. F. Stanton, *J. Chem. Phys.* **143**, 204103 (2015).