Title

Sparsity of the electron repulsion integral tensor using different localized virtual orbital representations in local second-order Møller—Plesset theory

Journal

The Journal of Chemical Physics, 158(6)

ISSN

0021-9606

Authors

Wang, Zhenling

Aldossary, Abdulrahman

Head-Gordon, Martin

Publication Date

2023-02-14

DOI

10.1063/5.0134764

Sparsity of the Electron Repulsion Integral Tensor Using Different Localized Virtual Orbital Representations in Local Second Order Møller-Plesset Theory

Zhenling Wang,[†,¶] Abdulrahman Aldossary,[†,¶] and Martin Head-Gordon[*,†,‡]

†*Pitzer Center for Theoretical Chemistry, Department of Chemistry, University of California, Berkeley, California 94720, USA*

‡*Chemical Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA*

¶*Contributed equally to this work*

E-mail: mhg@cchem.berkeley.edu

Abstract

Utilizing localized orbitals, local correlation theory can reduce the unphysically high system-size scaling of post-Hartree-Fock (post-HF) methods to linear scaling in insulating molecules. The sparsity of the 4-index electron repulsion integral (ERI) tensor is central to achieving this reduction. For second-order Møller-Plesset theory (MP2), one of the simplest post-HF methods, only the $(ia|jb)$ ERIs are needed, coupling occupied orbitals i, j, and virtuals a, b. In this paper, we compare the numerical sparsity (called the "ragged list") and another two approaches revealing the low-rank sparsity of the ERI. The ragged list requires only one set

of (localized) virtual orbitals, and we find the orthogonal valence virtual-hard virtual (VV-HV) set of virtuals originally proposed by Subotnik et al. gives the sparsest ERI tensor. To further compress the ERI tensor, the pair natural orbital (PNO) type representation uses different sets of virtual orbitals for different occupied orbital pairs, while the occupied-specific virtual (OSV) approach uses different virtuals for each occupied orbital. Our results indicate that while the low-rank PNO representation achieves significant rank reduction, it also requires more memory than the ragged list. The OSV approach requires similar memory to the ragged list, but involves greater algorithmic complexity. An approximation (called the "fixed sparsity pattern") for solving the local MP2 equations using the numerically sparse ERI tensor is proposed and tested to be sufficiently accurate and to have highly controllable error. A low-scaling local MP2 algorithm based on the ragged list and the fixed sparsity pattern is therefore promising.

Keywords

local correlation theory, orbital localization, pair natural orbitals, linear scaling

1 Introduction

Much of modern quantum chemistry relates to solving the many-body problem accurately to find the ground state wavefunction. The simplest quantum mechanical treatment of the electron-electron interaction is the mean-field Hartree-Fock (HF) method,[1] or the dressed mean-field approach of Kohn-Sham density functional theory (DFT).[2] The latter does a good job of capturing weak (dynamic) electron correlation effects. *Ab initio* quantum chemistry gives some hope of rigorously describing many-Slater determinants and approaching the exact solution. Following this path, the quantity to approximate is the electron correlation energy, defined formally as the difference between the exact and the Hartree-Fock limit energy of the non-relativistic clamped-nuclei Schrodinger equation.[3,4] However, *ab initio* methods are

impeded by high polynomial scaling of computing costs with molecular size, N that prevents their application to large, interesting chemical problems. For example, even second order Møller-Plesset (MP2) theory,[5] one of the simplest approximations to the correlation energy, scales formally as $\mathcal{O}(N^5)$, while MP3 and the simplest useful coupled cluster methods[6] that include double substitutions scale as $\mathcal{O}(N^6)$. This has motivated lower-scaling methods that rely on factorization approximations. For example, using auxiliary basis expansions,[7] the opposite spin MP2 energy can be evaluated with $\mathcal{O}(N^4)$ computational effort,[8] or, using tensor hypercontraction,[9–11] the scaling for the full MP2 energy can be reduced to $\mathcal{O}(N^4)$.

Alternatively, it is valuable to exploit the near-sighted nature of electron-electron interactions,[12] which also applies to dynamic electron correlation. Rather than compressing dense tensors, one can choose appropriate sparsity-revealing orbital representations and then discard numerically insignificant contributions from the tensors of quantum chemistry. Most modern work can be traced back to the pioneering studies of Pulay,[13] and Saebø and Pulay[14] who introduced projected atomic orbitals (PAOs) to represent the virtual space, and employed the Hylleraas functional[15,16] to solve local approximations to MP2 and other more advanced methods. By using PAOs, the electron repulsion integral (ERI) tensor becomes sparse, in fact asymptotically linear with system size. Many have followed to work on local correlation methods.[17–37] At the local MP2 level, perhaps the most widely used linear scaling methods today are the pair natural orbital approaches.[32,34,35]

Every local correlation method starts by choosing a representation for the occupied and virtual (unoccupied) orbitals that makes the ERI tensor sparse. Spatially localized sets of orthogonal occupied orbitals are well-known, such as from Boys localization which rotates the occupied space into a set of minimal spatial variance (*vide infra*).[38] Other common examples are Edmiston-Ruedenberg (ER) localization,[39] and Pipek-Mezek (PM) localization.[40] The virtual space presents a much greater challenge. Perhaps the simplest local virtual representation is that of Pulay[14,41] where the atomic orbitals (AOs) are simply projected onto the virtual space. If the density matrix and the basis functions are localized then so too are the

PAOs.[20] The fact that PAOs are nonorthogonal and rank-deficient means that the overlap matrix must be carried along, which has the potential to cause numerical instabilities, with related challenges in implementing responses of the energy to applied perturbations.[42]

Alternatives to the PAOs to represent the virtual space via localized orbitals do exist. One prominent example was due to Kapuy et al. who developed a localized many-body perturbation theory where virtual orbitals were also localized.[43,44] More recently, Subotnik et al. proposed localizing the valence virtuals via Boys localization separate from the "hard" virtuals which belong to higher-than-valence principal quantum numbers, and can be orthogonalized without losing their AO character.[45] The resulting virtual orbitals, which we refer to as "valence virtual-hard virtual" (VV-HV) orbitals were then employed in local MP2 and CCSD methods that yielded smooth potential energy surfaces.[26,46,47] Aquilante et al. proposed Cholesky decomposing the density matrix or its orthogonal complement to produce a set of orthogonal orbitals,[48] and the Cholesky virtuals have then been used for efficient implementations.[49] These methods and others have been proposed in the literature[50–53] and there is also a major recent review available.[54] The condition number of the atomic orbital (AO) representation as well as the locality of the occupied orbitals influences the extent of localization possible in the virtual space.[55]

In recent years, there has been a renaissance in local correlation due to the very successful revival[27] of the pair natural orbitals (PNO) approach where the pseudo-CI coefficients are diagonalized for each electron pair.[56] Later those were extended to domain-based local PNO (DLPNOs) where the canonical virtuals are replaced by domains of PAOs to accelerate the construction.[32] Chan and co-workers introduced orbital-specific virtuals, finding the most important virtuals that correspond to a given occupied orbital.[30] Werner and co-workers demonstrated linear-scaling for PAOs, OSVs, and PNOs.[21,23,57] These methods offer the advantage of reducing the rank of the ERI tensor but suffer from requiring saving the different sets of orbitals. Yet, the question remains as to which methods of localizing the occupied and virtual orbitals reach linear scaling faster and whether sacrificing the orbital orthogonality

is a price worth paying to hold fewer ERIs in memory.

In this paper, we chose to analyze the sparsity of the ERI tensor using an array of methods including Cholesky, PNO, and PAO. Then we lay out the ragged-list approach where we store the ERIs only above a certain threshold, applying it to evaluate the MP2 energy in a range of alkanes, alkenes, silanes, and other systems of interest. We choose MP2 as the simplest useful correlation method to test the different localization and memory requirements. Previous research suggests this will also serve as a reliable proxy for higher accuracy methods such as coupled cluster theory with singles and doubles (CCSD). Triples may require a separate assessment. The MP2 form of the correlation energy has itself become far more used recently because of its deployment as part of double-hybrid density functional theory approximations.[58–63] Additionally, there has been an exciting improvement in the accuracy of MP2 theory,[64] when the amplitude equations are modified with an energy-dependent regularization.[65] Large MP2 errors for dispersion-dominated intermolecular interactions,[66] as well as transition metal binding energies[64] can be reduced by factors of 3-5 in the regularized κ-MP2 method at negligible additional compute cost.

2 Theory

We use $i, j, k, \cdots$ as the notation for occupied orbitals and $a, b, c, \cdots$ for virtual orbitals. n_{occ} and n_{virt} are the number of occupied and virtual orbitals, respectively. N stands for the size of the system. For simplicity, only spin-restricted systems are considered in this paper.

2.1 Local MP2 Equations

The closed shell RMP2 energy is the dot product of the ERI tensor and the amplitude tensor

$$E_{\text{corr}} = -\sum_{iajb} J_{ab}^{ij} \tau_{ab}^{ij}, \tag{1}$$

where $J^{ij}_{ab} = (ia|jb)$ in a notation where the first two indices refer to electron 1, and the last two indices refer to electron 2. The $\boldsymbol{\tau}$ tensor satisfies a linear system of equations

$$\sum_{i'j'a'b'} \Delta^{iji'j'}_{aba'b'} \tau^{ij}_{ab} = 2J^{ij}_{ab} - K^{ij}_{ab}. \tag{2}$$

where $K^{ij}_{ab} = J^{ij}_{ba}$ and the 8^{th} rank tensor, $\boldsymbol{\Delta}$, is

$$\Delta^{iji'j'}_{aba'b'} = S_{ii'}S_{aa'}S_{jj'}F_{bb'} - S_{ii'}S_{aa'}F_{jj'}S_{bb'} + S_{ii'}F_{aa'}S_{jj'}S_{bb'} - F_{ii'}S_{aa'}S_{jj'}S_{bb'}; \tag{3}$$

here $\mathbf{S}$ is the overlap matrix and $\mathbf{F}$ is the Fock matrix. Notice that if orbitals are canonical and orthogonal, then $\mathbf{S} = \mathbf{I}$ and $\mathbf{F}$ is diagonal, reducing Eq. 2 to a diagonal system and also leaving Eq. 1 in the familiar spin restricted form:

$$E_{\text{corr}} = -\sum_{iajb} \frac{\left(2J^{ij}_{ab} - K^{ij}_{ab}\right) J^{ij}_{ab}}{F_{bb} + F_{aa} - F_{ii} - F_{jj}}. \tag{4}$$

In this case, the $\mathbf{J}$ tensor, having $n^2_{\text{occ}} n^2_{\text{virt}}$ elements in total, is dense as the canonical orbitals are well-known to be delocalized. An $\mathcal{O}(N^5)$ scaling of compute time is needed to build the dense $\mathbf{J}$ tensor, and that becomes the bottleneck for canonical MP2.

Under a localized orbital representation, however, the $\mathbf{J}$ tensor becomes numerically sparse, and therefore lower scaling algorithms are possible. Most local MP2 formalisms can be summarized into the following four steps:

1. Localize occupied and virtual orbitals
2. Build tensors $\mathbf{J}, \mathbf{K}$
3. Solve Eq. 2: $\boldsymbol{\Delta\tau} = 2\mathbf{J} - \mathbf{K}$
4. Calculate the MP2 correlation energy

It should be noted that the bottleneck of local MP2 is not necessarily Step 2 as in the

canonical case. Step 1 and Step 3 can potentially cost more than building the $\mathbf{J}, \mathbf{K}$ tensors, so all steps need to be considered carefully. In this paper, we are going to discuss the performance of different choices of Step 1 as well as a new way to do Step 3. We are actively working on the efficient implementation of Step 2, for which many approaches have been proposed.[21,23,34,67] Since the wall time of our method cannot be properly compared without an efficient implementation of Step 2, we neglect any compute time consideration in this work.

2.2 Representations of Localized Virtual Orbitals

From the local MP2 equation, it is clear that the sparsity of the ERI tensor $\mathbf{J}$ is central to the problem, and the sparsity of $\mathbf{J}$ is dependent on the choice of local orbitals. In this section, we will describe a variety of ways to localize molecular orbitals. Boys (B),[38] Pipek-Mezey (PM),[40] and Edmiston-Ruedenberg (ER)[39] defined functionals that measure the extent of localization for an orthonormal orbital set $\{\phi_i\}$:

$$f_{\mathrm{B}} = \sum_i |\langle \phi_i | \mathbf{r} | \phi_i \rangle|^2 \tag{5}$$

$$f_{\mathrm{PM}} = \sum_i \sum_A |\langle \phi_i | P_A | \phi_i \rangle|^2 \tag{6}$$

$$f_{\mathrm{ER}} = \sum_i (\phi_i \phi_i | \phi_i \phi_i) \tag{7}$$

In f_{PM}, A is an atom and P_A is the projection operator mapping the orbital onto the basis set space of atom A. Boys orbitals minimize the sum of the second central moment of the orbitals; Pipek-Mezey maximizes the locality of Mulliken populations; and Edmiston-Ruedenberg maximizes the self-repulsion of the orbitals. Given the initial orbital space, it is possible to solve for an orthonormal basis that maximizes one of these functional values, and by the construction of these functionals, those orbitals are localized. An iterative method (Jacobi sweep) is usually used to obtain such extrema, yet some other algorithms have also

been developed.[68–71] Edmiston-Ruedenberg normally requires an iterative $\mathcal{O}(N^5)$ scaling because of the integral transformation, which is prohibitive in large molecules. But even though Boys and Pipek-Mezey only cost as iterative $\mathcal{O}(N^3)$, a very large number of iterations are needed for localizing virtual orbitals, meaning that only *occupied* Boys and Pipek-Mezey orbitals are suitable for a local MP2 formalism. Throughout this paper, we will use Boys occupied orbitals (as commonly used already in local MP2[14]) to facilitate the comparison of different virtual representations, which is still a relatively open question in our view.

Pivoted Cholesky[48] and pivoted QR[72] utilize numerical sparsity of the density matrix. A pivoted Cholesky decomposition or pivoted QR decomposition of the density matrix would yield localized orbitals. It can be proved that the resulting orbitals are orthonormal,[48] and the locality results from the elements of the density matrix decaying exponentially as[73,74] $P_o\left(r, r'\right) \sim e^{-\sqrt{G}\left|r-r'\right|}$, where G is the HOMO-LUMO gap. The density-like matrix, $\mathbf{P}_v$ formed by the set of virtual orbital is

$$\mathbf{P}_v = \mathbf{S}^{-1} - \mathbf{P}_o, \tag{8}$$

where $\mathbf{S}$ is the (sparse) overlap matrix of the basis functions, and $\mathbf{P}_o$ is the usual density matrix formed from the occupied orbitals. $\mathbf{S}$ and $\mathbf{P}_o$ are sparse, but $\mathbf{S}^{-1}$ is typically not a sparse matrix (at least in extended basis sets),[75] so $\mathbf{P}_v$ does not have the nice sparsity that $\mathbf{P}_o$ does, and therefore we expect that pivoted Cholesky and pivoted QR could behave poorly in the virtual space without further refinement. This will be tested later.

A natural choice for the virtual space that is most widely used in both legacy and and current local correlation methods are the projected atomic orbitals (PAOs), which retain as much locality of the AOs as possible. PAOs are obtained by directly projecting out the occupied space from each AO (and thus the AO basis set), resulting in a redundant and non-orthogonal set:

$$\mathbf{C}_{\mathrm{PAO}} = \mathbf{I} - \mathbf{P}_o\mathbf{S}. \tag{9}$$

It is possible to truncate and orthogonalize this set without losing too much of its locality,[20] but in this paper, we will stick to the standard definition of PAOs, via Eq. 9.

Subotnik et al. combined the PAO concept with the idea of Boys or Pipek-Mezey localized orbitals and proposed an efficient scheme to localize the virtual space with a set of *orthonormal* localized virtual orbitals.[45] Firstly, it partitions the entire AO basis set space ($\mathcal{A}$) into a minimal basis space ($\mathcal{E}$) and an extra-valence hard virtual space ($\mathcal{H}$):

$$\mathcal{A} = \mathcal{E} \oplus \mathcal{H}. \tag{10}$$

The minimal basis is taken as STO-3G projected into the target basis. Other alternatives include quasi-atomic minimal basis orbitals (QUAMBOs)[76] and intrinsic atomic orbitals (IAOs).[77] $\mathcal{E}$ is then further sub-divided into the occupied space ($\mathcal{O}$) and a valence virtual space ($\mathcal{L}$)

$$\mathcal{A} = \mathcal{O} \oplus \mathcal{L} \oplus \mathcal{H}. \tag{11}$$

Notice that this splitting requires that the large basis set space contains the minimal basis space (i.e. $\mathcal{E} \subset \mathcal{A}$) and that the minimal basis space is sufficient to describe most of the occupied space (i.e. $\mathcal{O} \subset \mathcal{E}$), both of which are normally satisfied. Orbitals in $\mathcal{L}$ and $\mathcal{H}$ are virtual orbitals, and as the $\mathcal{L}$ space is small and chemically meaningful (it typically consists of antibonding orbitals[78]), it can be efficiently localized by either Boys or Pipek-Mezey. Hard virtuals are, however, in principle similar to PAOs and are quite local but not orthogonal to each other. Instead of orthogonalizing the hard virtuals by a somewhat complicated algorithm based on "important classes" as in Ref. 45, we will do a weighted symmetrical orthogonalization[79] using the inverse of their spatial variance. In the rest of this paper, we will refer to these localized virtuals as "valence virtual-hard virtual" (VV-HV) orbitals. Except for Boys, Pipek-Mezey, and Edmiston-Ruedenberg orbitals, all other components described above do not require an initial guess. While the procedure formally scales as $\mathcal{O}(N^3)$, this will not be a computational bottleneck as evaluation of these orthogonal virtuals

requires far less effort than is needed for the Hartree-Fock mean field reference calculation.

It is also possible to use more than one set of virtual orbitals. Pair natural orbitals (PNO) have a set of virtual orbitals specific for each different *pair* of occupied orbitals,[28] and orbital-specific virtuals (OSV) has one set of virtuals created for each *individual* occupied orbital.[29] PNOs can be obtained by performing a singular value decomposition (SVD) on the approximated MP2 pair density matrices (in the virtual space)

$$\mathbf{D}^{ij} = \frac{1}{1+\delta_{ij}} \left(\mathbf{t}^{ij\dagger} \boldsymbol{\tau}^{ij} + \mathbf{t}^{ij} \boldsymbol{\tau}^{ij\dagger} \right). \tag{12}$$

Here we have introduced the opposite spin amplitudes:

$$t_{ab}^{ij} = 2\tau_{ab}^{ij} - \tau_{ba}^{ij} \tag{13}$$

If only singular values above a threshold are retained, the result is a set of orbitals that span the part of the virtual space relevant to the i, j pair. The resulting a, b come from two sets of localized virtual orbitals, normally PAOs, that are close to orbitals i, j, respectively. To facilitate comparison with other localization techniques above that make the $\mathbf{J}$ tensor numerically sparse, we define PNO-like orbitals utilizing the low-rank sparsity of $\mathbf{J}$ as the result of doing SVD on $\mathbf{J}^{ij}$

$$\mathbf{J}^{ij} = \mathbf{U}^{ij} \boldsymbol{\Sigma}^{ij} \mathbf{V}^{ij\dagger}. \tag{14}$$

We then truncate $\boldsymbol{\Sigma}^{ij}$ to $(\boldsymbol{\Sigma}^{ij})'$ such that

$$\max_{a,b} \left| \mathbf{U}^{ij} \left(\boldsymbol{\Sigma}^{ij} \right)' \mathbf{V}^{ij\dagger} - \mathbf{J}^{ij} \right| < \epsilon, \tag{15}$$

where ϵ is a pre-set threshold.

On the other hand, OSVs (for orbital i) can be obtained by doing SVD on $\mathbf{t}^{ii}$ and truncating. A set of OSV-like orbitals (for orbital i) can likewise be defined by diagonalizing

the $\mathbf{J}^{ii}$ block

$$\mathbf{J}^{ii} = \mathbf{W}^{ii}\mathbf{D}^{ii}\mathbf{W}^{ii\dagger}, \tag{16}$$

and keeping the virtuals with eigenvalues larger than a threshold. After that, we compute the elements in off-diagonal blocks of the $\mathbf{J}$ tensor $\mathbf{J}^{ij}_{a_i b_j}$, where a_i, b_j are the OSV-like orbitals for i, j, respectively. We then select significant elements to retain, i.e.,

$$\left|\mathbf{J}^{ij}_{a_i b_j}\right| > \epsilon, \tag{17}$$

and their corresponding exchange integrals $\mathbf{K}^{ij}_{a_i b_j}$. This approach can be thought of as a hybrid between low-rank sparsity and numerical thresholding. The PNO-like scheme and the OSV-like scheme both utilize the low-rank sparsity of the local MP2 problem.

2.3 An approach to solve the local MP2 equation

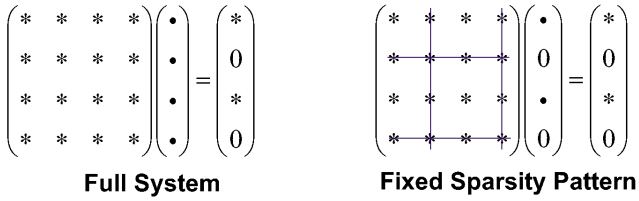

Figure 1: An approximation to simplify the local MP2 equation, which we refer to as the "fixed sparsity pattern" approximation. Stars stand for known elements and dots for unknowns.

To make use of the numerical sparsity of $\mathbf{J}$ in Eq. 2, we will use an approximation that we refer to as the "fixed sparsity pattern". Instead of solving the full set of equations $\boldsymbol{\Delta\tau} = 2\mathbf{J} - \mathbf{K}$, we set all elements of $\mathbf{t}$ that correspond to a below-threshold element in $2\mathbf{J} - \mathbf{K}$ to be zero (Figure 1). With this approximation, we can then neglect the rows and columns corresponding to elements of the right-hand side that are below the threshold. As we will see later, the number of significant elements in $2\mathbf{J} - \mathbf{K}$ scales as $\mathcal{O}(N)$, and that suggests a linear-scaling algorithm is possible if the truncated coefficient matrix still

possesses some sparsity. One must immediately recognize that this fixed sparsity model is not at all mathematically rigorous, in the sense that there is no guarantee of a strict connection between the threshold (ϵ) used for sparsifying $\mathbf{J}$ and the resulting errors in either $\mathbf{t}$ or $E_{\rm corr}$. However, we report tests of the fixed sparsity technique in Section 3.3 on a variety of molecules, and it turns out that the error caused by the fixed sparsity pattern is highly controllable.

When the iterations to solve the linear system converge to a residual error δ, yielding $\boldsymbol{\tau}_\delta$, we have $\boldsymbol{\tau}_\delta = \boldsymbol{\tau} + \mathcal{O}(\delta)$. If we calculate $E_{\rm corr}(\boldsymbol{\tau}_\delta)$ by Eq. 1, the error will be $\mathcal{O}(\delta)$, and we will then have to converge the linear system to a very tight threshold (i.e. very small δ) in order to get an accurate $E_{\rm corr}$, which requires many iterations. In order to improve this, we can employ $\mathbf{t}$ defined in Eq. 13 and calculate $E_{\rm corr}$ by the Hylleraas functional:[15,16]

$$L_{\rm corr} = -2\mathbf{t}\cdot(2\mathbf{J}-\mathbf{K}) + \mathbf{t}\cdot(\Delta\boldsymbol{\tau}). \tag{18}$$

It can be proved that the error, $L_{\rm corr}(\boldsymbol{\tau}_\delta) - E_{\rm corr}(\boldsymbol{\tau})$, is now $\mathcal{O}(\delta^2)$, and thus only a small number of iterations are needed.

However, under the fixed sparsity pattern with threshold ϵ, with exact solution of the linear equations ($\delta = 0$), $L_{\rm corr} \neq E_{\rm corr}$ unless we truncate $\mathbf{J}$ symmetrically. Specifically, we need to keep J^{ij}_{ab} if

$$\left|J^{ij}_{ab}\right| > \epsilon \qquad \text{or} \qquad \left|K^{ij}_{ab}\right| = \left|J^{ij}_{ba}\right| > \epsilon. \tag{19}$$

Since this expands the list of retained elements, the symmetrically truncated $\mathbf{J}$ still satisfies

$$\max_{a,b}\left|(\mathbf{J}^{ij})' - \mathbf{J}^{ij}\right| < \epsilon \qquad \forall\ i,j. \tag{20}$$

As we are keeping a truncated list of elements, we will call the resulting tensor ($\mathbf{J}'$) the "ragged list" representation of $\mathbf{J}$. The term "ragged list" (and its synonym "jagged array") has been widely used in computer science.[80]

3 Results and discussion

3.1 The numerical sparsity of the J tensor

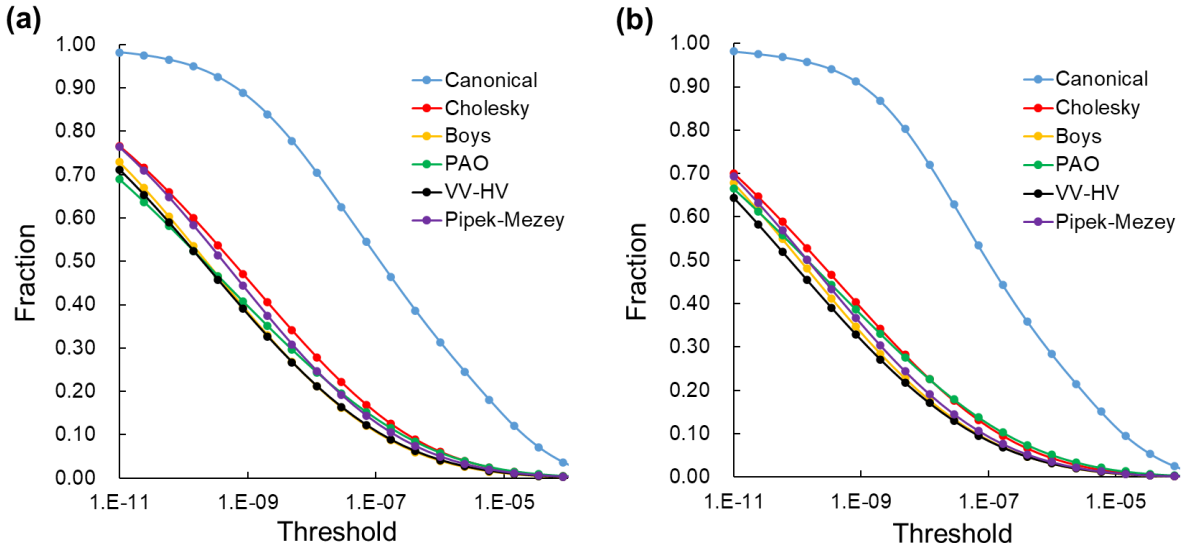

Figure 2: The fraction of significant elements in the J tensor of (a) def2-TZVP and (b) cc-pVTZ linear $C_{20}H_{42}$ with different numerical zero thresholds and different virtual orbitals. Occupied orbitals are Boys localized, and core orbitals are frozen. The geometry chosen is distorted slightly from symmetric, as given in the Supporting Information.

As is known from previous investigations,[20,45,81] different choices of virtual orbitals affect the magnitude of elements in **J**. As an example, we have plotted (Figure 2) the fraction of significant elements for the **J** tensor of def2-TZVP and cc-pVTZ $C_{20}H_{42}$ as a function of truncation threshold (ϵ) for 6 different choices of the virtual orbitals. The ideal extended conformation of $C_{20}H_{42}$ has C_{2h} point group symmetry, which leads to significant sparsity ($\sim$ 25% significant elements) in canonical **J** tensor. Since this is artificial (most large molecules have no symmetry), we have distorted the geometry slightly to break the symmetry. In addition, for the data shown in Figure 2, core orbitals are frozen, and the valence-occupied orbitals are Boys-localized.

From Figure 2, it is clear that any choice of localized virtual orbitals will provide a **J** that is much more numerically sparse than the canonical virtuals. All localized virtuals behave rather similarly, with the VV-HV orbitals winning by a few percentage points. Cholesky

is indeed one of the worst ones because the virtual density matrix $\mathbf{P}_v$ is not as sparse as the occupied one $\mathbf{P}_o$. To further confirm this, we have also plotted the numerical sparsity pattern of $\mathbf{P}_o$ and $\mathbf{P}_v$ of this system in Figure 3. Notice that although three methods (Boys, Pipek-Mezey, and VV-HV) all yield sparse $\mathbf{J}$ tensors for a large enough system, obtaining VV-HVs takes a much shorter time than getting Boys or Pipek-Mezey virtuals. Algorithms for the Boys and Pipek-Mezey localized virtuals may also not converge to a true maximum, or may yield a local maximum.[68,69] Therefore, we will be using the Boys occupied orbitals and VV-HV orbitals for the ragged list representation in the rest of this paper.

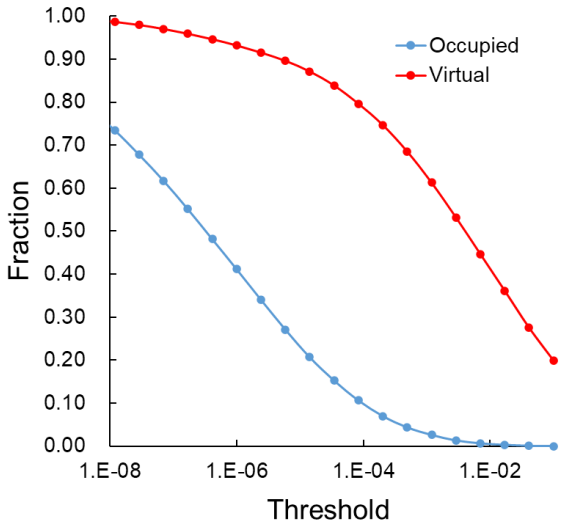

Figure 3: The fraction of significant elements in the occupied and virtual density matrix of cc-pVTZ linear $C_{20}H_{42}$ in the atomic orbital basis as a function of the threshold for defining a numerical zero.

3.2 Comparison of numerical and low-rank representations of J

At the end of Section 2.2, we reviewed how to generate sets of PNO-like and OSV-like virtual orbitals to make use of the low-rank sparsity of $\mathbf{J}$. To compare the effectiveness of using numerical sparsity (the ragged list approach, introduced at the end of Section 2.3) and a low-rank sparse structure (either the PNO-like or the OSV-like approach) to store and manipulate the $\mathbf{J}$ tensor, we have computed the total rank and total memory requirement

for these two data structures for a variety of systems.

The total rank of the PNO-like representation, Eq. 14, is the number of the singular values selected that satisfy Eq. 15, and the memory requirement is the total number of elements of vectors picked in the $\mathbf{U}^{ij}$ and $\mathbf{V}^{ij}$ matrices. The number of rows and columns of the $\mathbf{J}^{ij}$ block is the number of important virtual orbitals for i and j, respectively. A virtual orbital a is considered important to i when the two-electron integral

$$(ia|ia) > \epsilon^2, \tag{21}$$

where ϵ is the same truncation threshold as in Eqs. 19 and 20. Notice that if $(ia|ia) < \epsilon^2$, since normally $(jb|jb) < 1, \quad \forall j, b$, we have

$$(ia|jb) \leq \sqrt{(ia|ia)\,(jb|jb)} < \epsilon \tag{22}$$

by the Schwarz inequality, so we can neglect this i, a pair.

This same argument could also be applied to the OSV-like scheme (Eqs. 16 and 17). In this scheme, we keep all eigenvectors (OSV-like orbitals) whose corresponding eigenvalues are greater than ϵ^2 and also significant elements in $\mathbf{J}$ and $\mathbf{K}$. The total memory requirement for the numerical sparsity structure (the ragged list) is the number of $(ia|jb)$ elements kept according to Eq. 19. The total rank for the OSV-like scheme and the raggist list is the number of unique a and b in block $(\mathbf{J}')^{ij}$ (and also block $(\mathbf{K}')^{ij}$ for the OSV-style) that correspond to a large $(ia|jb)$ integral value, summed over all i, j, and divided by 2.

Figure 4 plots the total rank and total memory requirement for the low-rank (using PNO-like or OSV-like virtuals) and the ragged list (using VV-HVs) representations for linear alkane chains of different lengths. Figure 5 shows the fraction of the memory requirement compared to storing the non-sparse $\mathbf{J}$ of the PNO- and OSV-style low-rank representations and the ragged list representation in def2-TZVP $C_{20}H_{42}$ as a function of truncation threshold (ϵ). The PNO-like virtual orbital representation has a smaller total rank by more than a

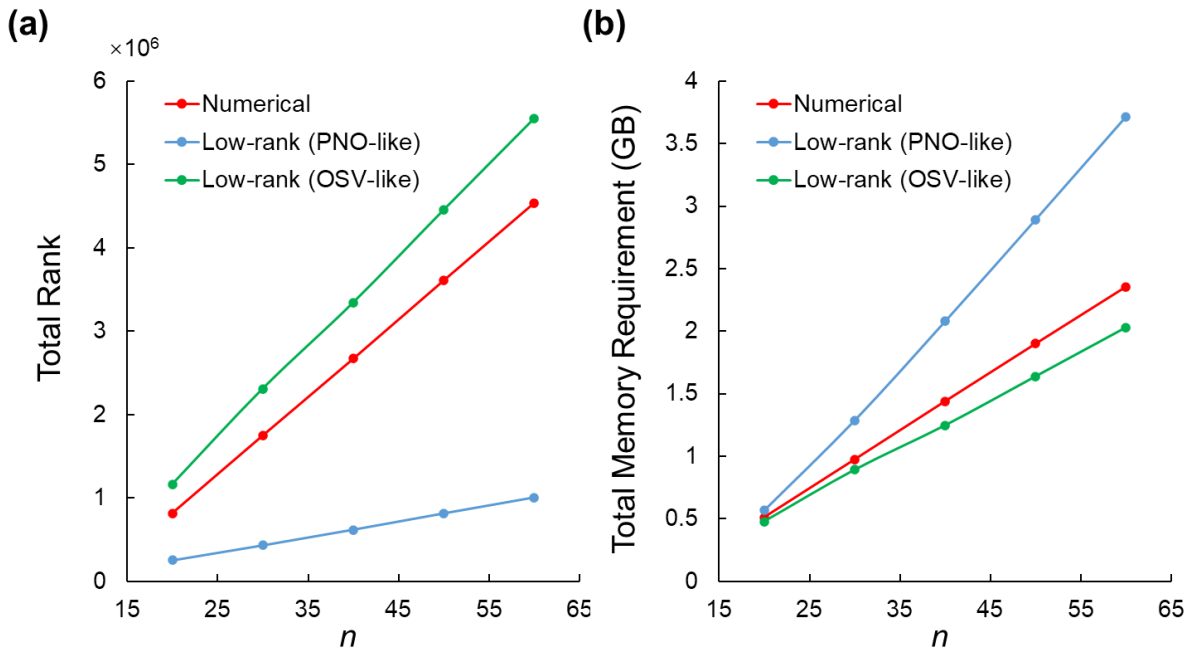

Figure 4: The (a) total rank and (b) total memory requirement in gigabytes for low-rank (using PNO-like and OSV-like virtual orbitals) and ragged list (i.e., numerical, using VV-HV orbitals) representations of the J tensor of def2-SV(P) C_nH_{2n+2} linear alkane chains. Core orbitals are frozen, and $\epsilon = 10^{-7}$. The geometries are provided in the Supporting Information and correspond to extended structures that have been slightly distorted to break point group symmetry.

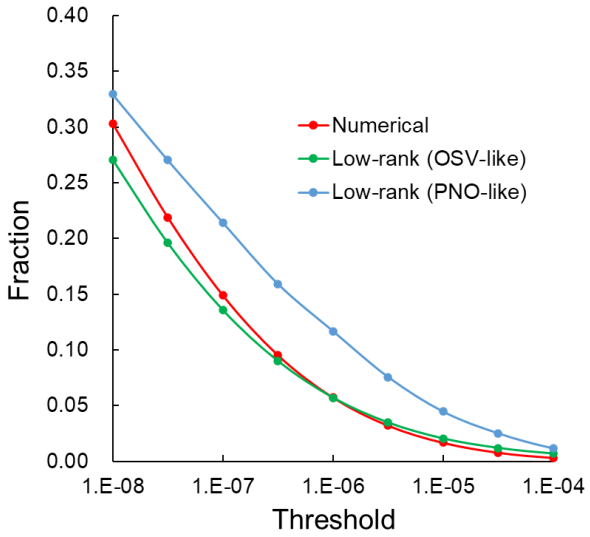

Figure 5: The fraction of the total memory requirement for storing the non-sparse $\mathbf{J}$ using low-rank (using PNO-like or OSV-like virtual orbitals) and ragged list (i.e., numerical, using VV-HV orbitals) representations in def2-TZVP $C_{20}H_{42}$ with different ϵ. Occupied orbitals are Boys localized, and core orbitals are frozen. The geometry chosen is distorted slightly from symmetric, as given in the Supporting Information.

factor of 4, indicating that a roughly 4-fold compression is possible relative to rectangular storage of each $(\mathbf{J}')^{ij}$ sub-block, in the space of significant row and column virtuals. Perhaps surprisingly, the ragged list structure uses a significantly smaller amount of memory than the PNO-like structure. This means that in the sense of computer memory demands, which may set the upper limit of feasible calculations, using PNO might not be superior to simply storing significant elements of the $\mathbf{J}$ tensor. The idea of the OSV-like structure is somewhere between the PNO-like and the ragged list representations as we compress the diagonal blocks of $\mathbf{J}$ but store significant elements of the off-diagonal blocks. As we have to store elements in both $\mathbf{J}, \mathbf{K}$, the OSV-style structure is not quite rank efficient (Figure 4(a)). But from Figure 5, we can see that in the chemically interesting range of truncation thresholds ($\epsilon = 10^{-7} \sim 10^{-5}$), the memory requirement of the OSV-like scheme is close to the ragged list representation. But as using multiple sets of mutually non-orthogonal virtual orbitals significantly increases the complexity of building the $\mathbf{J}, \mathbf{K}$ tensors and solving the local MP2 equations, the advantage of using the OSV-like orbitals is likely to be small.

Data for comparing the PNO style and the VV-HV style of representing **J** tensor on a variety of 1-dimensional and 3-dimensional systems (Figure 6) at $\epsilon = 10^{-5}$ and $\epsilon = 10^{-7}$ are listed in Table 1. We chose $\epsilon = 10^{-5}$ and $\epsilon = 10^{-7}$ as the thresholds because those numbers correspond roughly to a loose accuracy (10^{-3} fractional error) and tight accuracy (10^{-5} fractional error) of local MP2 theory, as will be discussed in the next sub-section. Turning to the results, the pattern is always the same – the PNO approach attains a smaller rank but the ragged list has a smaller memory requirement. Furthermore, the memory saving for the ragged list approach seems to increase for larger systems. This apparently universal pattern suggests that, at a minimum, there is no necessity to use the conceptually and technically more complicated PNO approach.

As an example of the distribution of significant elements in the ragged list, Figure 7 plots the fraction of elements stored in each $\mathbf{J}^{ij}$ block using the ragged list representation of def2-SV(P) vancomycin. Most blocks have either no or only a few nonzero elements when $\epsilon = 10^{-5}$, and even the densest ones have only $\sim 3\%$ significant elements. By contrast, when $\epsilon = 10^{-7}$, the blocks become much denser, with almost no blocks being completely zero. However, numerical sparsity is still significant as the densest blocks only have $\sim 24\%$ of their elements retained as the above threshold. This example illustrates the well-known fact that the onset of linear scaling is earlier for looser thresholds, as well as the fact that significant reductions in memory and/or computational requirements are possible even before the linear scaling regime is reached.

3.3 The accuracy of the fixed sparsity pattern approximation

The fixed sparsity pattern approximation introduced in Section 2.3 allows us to utilize the numerical sparsity of the **J** tensor in solving the local MP2 equations. To assess its accuracy, we have calculated the error caused by this approximation as a function of the truncation threshold ϵ, for a set of example molecules using the def2-SV(P) and def2-TZVP basis sets. As the fixed sparsity pattern solution is slightly different from the true solution and the

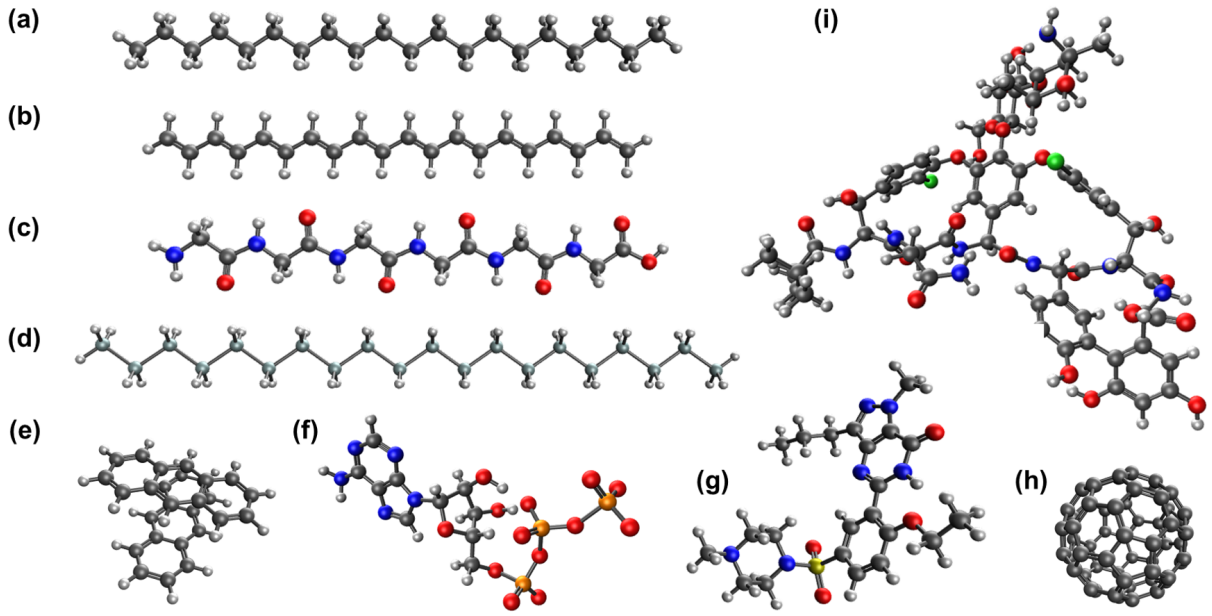

Figure 6: Example 1-D and 3-D systems: (a) $C_{20}H_{42}$, (b) $C_{20}H_{22}$, (c) $(gly)_6$, (d) $Si_{20}H_{42}$, (e) anthracene dimer, (f) ATP^{4-}, (g) sildenafil, (h) fullerene, (i) vancomycin. Color codes for atoms: black, C; white, H; red, O; blue, N; orange, P; yellow, S; green, Cl.

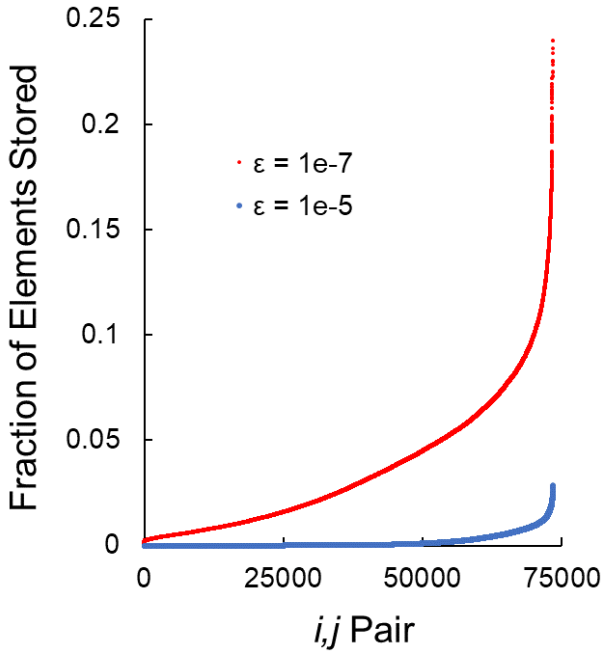

Figure 7: The fraction of elements stored in each block of $\mathbf{J}^{ij}$ in def2-SV(P) vancomycin. Core orbitals are frozen.

Table 1: The fraction of the non-sparse total rank and non-sparse total memory requirement for representing the **J** tensor using low-rank (PNO-style) or numerical sparsity (ragged list approach) for the example systems shown in Figure 6. ϵ is the threshold in Eqs. 15 and 18. Core orbitals are frozen. The def2-TZVP basis set is used except for vancomycin and fullerene, where def2-SV(P) is used.

	$\epsilon = 10^{-5}$				$\epsilon = 10^{-7}$			
	rank %		memory %		rank %		memory %	
	lo-rank	num'l	lo-rank	num'l	lo-rank	num'l	lo-rank	num'l
$C_{20}H_{42}$	5.5	22.0	4.5	1.7	21.5	73.6	21.4	14.9
$C_{20}H_{22}$	32.0	11.1	10.7	2.9	36.8	85.4	36.8	22.5
$(gly)_6$	5.9	21.3	4.6	1.5	22.1	70.5	21.8	13.4
$Si_{20}H_{42}$	4.0	16.8	3.3	1.1	18.3	66.4	18.2	12.1
anthracenes	12.9	61.6	12.9	4.5	55.0	99.5	55.0	52.9
ATP^{4-}	6.6	31.3	6.3	1.9	33.1	89.2	33.1	24.1
sildenafil	6.0	29.3	5.8	1.6	30.0	89.4	30.0	21.0
fullerene	14.6	62.1	14.6	4.4	72.6	100.0	72.6	61.4
vancomycin	1.2	8.1	0.8	0.2	7.6	49.8	7.3	3.6

Hylleraas functional is convex (Eq. 18), the fixed sparsity pattern will always underestimate the correlation energy; i.e., it always overestimates its absolute value.

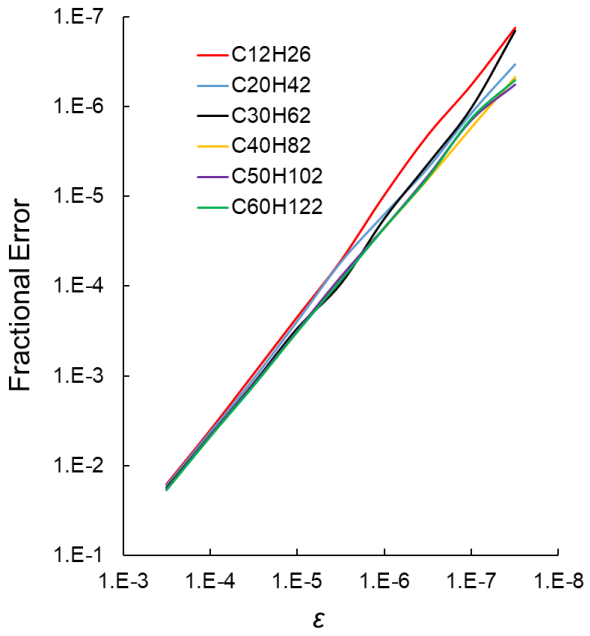

Figure 8: The fractional error in the LMP2/def2-SV(P) correlation energy caused by using the fixed sparsity pattern approximation for C_nH_{2n+2} linear alkane chains as a function of truncation threshold, ϵ. Core orbitals are frozen. The fixed sparsity pattern always underestimates the correlation energy.

In Figure 8, one can see that the fractional error in the LMP2 correlation energy caused by the fixed sparsity pattern is (i) roughly the same regardless of alkane size and (ii) approximately linear in ϵ over a very wide range of ϵ. $C_{12}H_{26}$ looks like an outlier because the system is really small and thus is not in the linear scaling regime when $\epsilon \leq 10^{-6}$. Similar results for different 1-dimensional systems are plotted in Figure 9(a), and the same observation can be made. The reason that the fixed sparsity pattern behaves slightly differently in $Si_{20}H_{42}$ might be attributed to the fact that silicon and carbon atoms have different basis set compositions.

Figure 9(b) plots the error of the fixed sparsity pattern for some 3-dimensional examples. It can be seen that the fixed sparsity pattern behaves quite well except for the anthracene

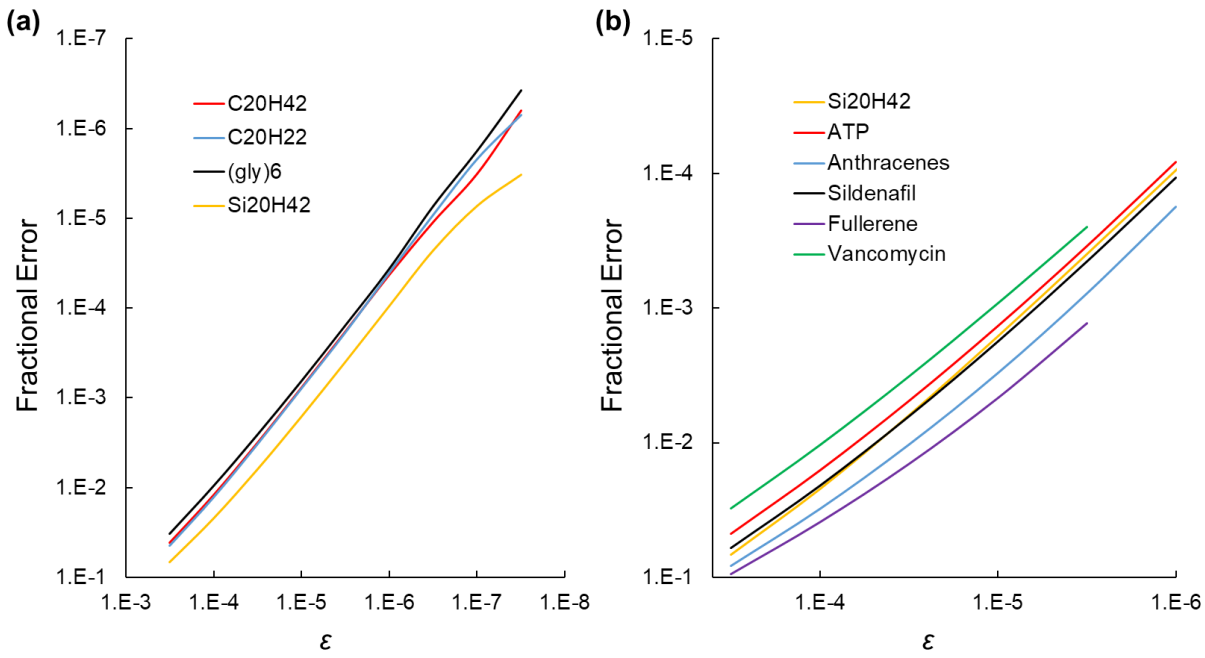

Figure 9: The fractional error in the LMP2 correlation energy caused by using the fixed sparsity pattern approximation for (a) a range of 1-dimensional extended molecules and (b) an additional set of molecules with a range of compactness (see Figure 6). The def2-TZVP basis set is used except for vancomycin and fullerene, where def2-SV(P) is used. All results are shown as a function of truncation threshold ϵ. Core orbitals are frozen. The fixed sparsity pattern always underestimates correlation energy.

dimer and fullerene. Atoms in the anthracene dimer and fullerene are tightly packed (Figure 6). Also, both these systems are extensively conjugated molecules meaning the tails of the localized orbitals are expected to be less well localized. Therefore, numerical sparsity in those molecules is not significant, which is indeed reflected by the relatively large percentage of retained elements in Table 1. When we truncate **J**, compared with other molecules, we will discard a larger number of elements that are only slightly smaller than the threshold, explaining the worse behavior of the fixed sparsity pattern in the two systems. In other cases, the approximation works almost as well for 1-D systems, suggesting that the fixed sparsity pattern approximation will typically be chemically useful.

4 Conclusion

In summary, we have compared two representations of the **J** tensor in the local MP2 problem: the low-rank representation using either PNO-like sets or OSV-like sets of orbitals, and the numerically sparse ("ragged list") representation using orthonormal localized VV-HV orbitals with Boys occupied orbitals. Other sets of virtual orbitals, including PAO, Boys, Pipek-Mezey, and Cholesky virtual orbitals are found to produce numerically denser **J** tensors than the VV-HV one. Tested on a broad variety of molecules, the PNO-style low-rank representation indeed uncovers a much smaller rank than the ragged list representation, but the memory requirement is larger. The OSV-style low-rank representation, however, has a quite similar memory requirement to the ragged list representation. Our results have two interesting related implications. First, since there is no benefit from improved sparsity in using the PAO representation of the virtual space, the complexity of handling the nonorthogonal redundant PAOs can be avoided using the VV-HV orbitals. Second, with orthonormal virtuals, there is also no necessity to utilize PNOs or OSVs for pre-conditioning purposes, and since the storage requirement can be reduced by avoiding their use, this is potentially attractive.

Our other main results consist of assessing an approximation for solving the local MP2 equation utilizing the numerical sparse representation of **J**, termed the "fixed sparsity pattern". This approximation truncates the amplitude tensor identically to the ERI tensor, which is convenient for efficiency purposes, but cannot strictly be justified from a mathematical viewpoint. Nonetheless, the results across a range of molecules indicate that the error caused by this approximation is highly controllable, and therefore a low-scaling local MP2 algorithm based on the ragged list and the fixed sparsity pattern is possible and promising. As the MP3 amplitude equations are very similar to the MP2 ones, we expect that our findings are also applicable to local MP3 theory, and by extension may also apply to local coupled cluster theory.

Conflicts of Interest

M. H.-G. is a part-owner of Q-Chem Inc, whose software was used for all developments and calculations reported here.

Supplementary Material

See Supplementary Material for the geometries of all example molecules in this paper.

Acknowledgements

This work was supported by the U.S. Department of Energy, Office of Science, Office of Advanced Scientific Computing, and Office of Basic Energy Sciences, via the Scientific Discovery through Advanced Computing (SciDAC) program. Additional support was provided by the U.S. Department of Energy, Office of Science, Office of Advanced Scientific Computing Research, under Award DE-SC0022364.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

References

(1) Szabo, A.; Ostlund, N. L. *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*; Dover Publications, 1996.

(2) Mardirossian, N.; Head-Gordon, M. Thirty years of density functional theory in computational chemistry: An overview and extensive assessment of 200 density functionals. *Mol. Phys.* **2017**, *115*, 2315–2372.

(3) Löwdin, P.-O. In *Adv. Chem. Phys.*; Prigogine, I., Ed.; Advances in Chemical Physics; 1958; pp 207–322.

(4) Löwdin, P.-O. O. The historical development of the electron correlation problem. *Int. J. Quantum Chem.* **1995**, *55*, 77–102.

(5) Cremer, D. Møller-Plesset perturbation theory: From small molecule methods to methods for thousands of atoms. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2011**, *1*, 509–530.

(6) Bartlett, R. J.; Musiał, M. Coupled-cluster theory in quantum chemistry. *Rev. Mod. Phys.* **2007**, *79*, 291–352.

(7) Nagy, B.; Jensen, F. Basis sets in quantum chemistry. *Rev. Comput. Chem.* **2017**, *30*, 93–149.

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

(9) Hohenstein, E. G.; Parrish, R. M.; Martínez, T. J. Tensor hypercontraction density fitting. I. Quartic scaling second-and third-order Møller-Plesset perturbation theory. *J. Chem. Phys.* **2012**, *137*, 044103.

(10) Parrish, R. M.; Hohenstein, E. G.; Martínez, T. J.; Sherrill, C. D. Tensor hypercontraction. II. Least-squares renormalization. *J. Chem. Phys.* **2012**, *137*, 224106.

(11) Lee, J.; Lin, L.; Head-Gordon, M. Systematically improvable tensor hypercontraction: Interpolative separable density-fitting for molecules applied to exact exchange, second-and third-order Møller–Plesset perturbation theory. *J. Chem. Theory Comput.* **2019**, *16*, 243–263.

(12) Kohn, W. Density functional and density matrix method scaling linearly with the number of atoms. *Phys. Rev. Lett.* **1996**, *76*, 3168.

(13) Pulay, P. Localizability of dynamic electron correlation. *Chem. Phys. Lett.* **1983**, *100*, 151–154.

(14) Saebo, S.; Pulay, P. Local Treatment of Electron Correlation. *Annu. Rev. Phys. Chem.* **1993**, *44*, 213–236.

(15) Hylleraas, E. A. Über den Grundterm der Zweielektronenprobleme von H-, He, Li+, Be++ usw. *Zeitschrift für Phys.* **1930**, *65*, 209–225.

(16) Hylleraas, E. A. *Adv. Quantum Chem.*; Academic Press, 1964; Vol. 1; pp 1–33.

(17) Hampel, C.; Werner, H. J. Local treatment of electron correlation in coupled cluster theory. *J. Chem. Phys.* **1996**, *104*, 6286–6297.

(18) Maslen, P. E.; Head-Gordon, M. Noniterative local second order Møller–Plesset theory: Convergence with local correlation space. *J. Chem. Phys.* **1998**, *109*, 7093–7099.

(19) Maslen, P.; Head-Gordon, M. Non-iterative local second order Møller–Plesset theory. *Chem. Phys. Lett.* **1998**, *283*, 102–108.

(20) Maslen, P. E.; Ochsenfeld, C.; White, C. A.; Lee, M. S.; Head-Gordon, M. Locality and sparsity of Ab initio one-particle density matrices and localized orbitals. *J. Phys. Chem. A* **1998**, *102*, 2215–2222.

(21) Schütz, M.; Hetzer, G.; Werner, H. J. Low-order scaling local electron correlation methods. I. Linear scaling local MP2. *J. Chem. Phys.* **1999**, *111*, 5691–5705.

(22) Ayala, P. Y.; Scuseria, G. E. Linear scaling second-order Moller–Plesset theory in the atomic orbital basis for large molecular systems. *J. Chem. Phys.* **1999**, *110*, 3660–3671.

(23) Schütz, M.; Werner, H. J. Low-order scaling local electron correlation methods. IV. Linear scaling local coupled-cluster (LCCSD). *J. Chem. Phys.* **2001**, *114*, 661–681.

(24) Ayala, P. Y.; Kudin, K. N.; Scuseria, G. E. Atomic orbital Laplace-transformed second-order Møller–Plesset theory for periodic systems. *J. Chem. Phys.* **2001**, *115*, 9698–9707.

(25) Li, S.; Shen, J.; Li, W.; Jiang, Y. An efficient implementation of the "cluster-in-molecule" approach for local electron correlation calculations. *J. Chem. Phys.* **2006**, *125*, 074109.

(26) Subotnik, J. E.; Sodt, A.; Head-Gordon, M. The limits of local correlation theory: Electronic delocalization and chemically smooth potential energy surfaces. *J. Chem. Phys.* **2008**, *128*, 034103.

(27) Neese, F.; Hansen, A.; Liakos, D. G. Efficient and accurate approximations to the local coupled cluster singles doubles method using a truncated pair natural orbital basis. *J. Chem. Phys.* **2009**, *131*, 064103.

(28) 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. *J. Chem. Phys.* **2009**, *130*, 114108.

(29) Yang, J.; Kurashige, Y.; Manby, F. R.; Chan, G. K.-L. L. Tensor factorizations of local second-order Møller–Plesset theory. *J. Chem. Phys.* **2011**, *134*, 044123.

(30) Kurashige, Y.; Yang, J.; Chan, G. K.-L.; Manby, F. R. Optimization of orbital-specific virtuals in local Mller-Plesset perturbation theory. *J. Chem. Phys.* **2012**, *136*, 124106.

(31) Yang, J.; Chan, G. K.-L.; Manby, F. R.; Schütz, M.; Werner, H. J. The orbital-specific-virtual local coupled cluster singles and doubles method. *J. Chem. Phys.* **2012**, *136*, 144105.

(32) Riplinger, C.; Neese, F. An efficient and near linear scaling pair natural orbital based local coupled cluster method. *J. Chem. Phys.* **2013**, *138*, 34106.

(33) Maurer, S. A.; Lambrecht, D. S.; Kussmann, J.; Ochsenfeld, C. Efficient distance-including integral screening in linear-scaling Møller-Plesset perturbation theory. *J. Chem. Phys.* **2013**, *138*, 01B601.

(34) Werner, H.-J.; Knizia, G.; Krause, C.; Schwilk, M.; Dornbach, M. Scalable Electron Correlation Methods I.: PNO-LMP2 with Linear Scaling in the Molecular Size and Near-Inverse-Linear Scaling in the Number of Processors. *J. Chem. Theory Comput.* **2015**, *11*, 484–507.

(35) 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. *J. Chem. Phys.* **2015**, *143*, 034108.

(36) Wang, Q.; Zou, J.; Xu, E.; Pulay, P.; Li, S. Automatic Construction of the Initial Orbitals for Efficient Generalized Valence Bond Calculations of Large Systems. *J. Chem. Theory Comput.* **2018**, *15*, 141–153.

(37) Bangerter, F. H.; Glasbrenner, M.; Ochsenfeld, C. Low-scaling tensor hypercontraction in the Cholesky molecular orbital basis applied to second-order Møller–Plesset perturbation theory. *J. Chem. Theory Comput.* **2020**, *17*, 211–221.

(38) Boys, S. F. Construction of Molecular orbitals to be minimally variant for changes from one molecule to another. *Rev. Mod. Phys.* **1960**, *32*, 296–299.

(39) Edmiston, C.; Ruedenberg, K. Localized atomic and molecular orbitals. *Rev. Mod. Phys.* **1963**, *35*, 457–464.

(40) Pipek, J.; Mezey, P. G. A fast intrinsic localization procedure applicable for ab initio and semiempirical linear combination of atomic orbital wave functions. *J. Chem. Phys.* **1989**, *90*, 4916–4926.

(41) Pulay, P. Localizability of dynamic electron correlation. *Chem. Phys. Lett.* **1983**, *100*, 151–154.

(42) Stoychev, G. L.; Auer, A. A.; Gauss, J.; Neese, F. DLPNO-MP2 second derivatives for the computation of polarizabilities and NMR shieldings. *J. Chem. Phys.* **2021**, *154*, 164110.

(43) Kapuy, E.; Csépes, Z.; Kozmutza, C. Application of the many-body perturbation theory by using localized orbitals. *Int. J. Quantum Chem.* **1983**, *23*, 981–990.

(44) Kapuy, E.; Bartha, F.; Bogár, F.; Csépes, Z.; Kozmutza, C. Applications of the MBPT in the localized representation. *Int. J. Quantum Chem.* **1990**, *38*, 139–147.

(45) Subotnik, J. E.; Dutoi, A. D.; Head-Gordon, M. Fast localized orthonormal virtual orbitals which depend smoothly on nuclear coordinates. *J. Chem. Phys.* **2005**, *123*, 114108.

(46) Subotnik, J. E.; Head-Gordon, M. A local correlation model that yields intrinsically smooth potential-energy surfaces. *J. Chem. Phys.* **2005**, *123*, 064108.

(47) Subotnik, J. E.; Sodt, A.; Head-Gordon, M. A near linear-scaling smooth local coupled cluster algorithm for electronic structure. *J. Chem. Phys.* **2006**, *125*, 074116.

(48) Aquilante, F.; Bondo Pedersen, T.; Sánchez de Merás, A.; Koch, H. Fast noniterative orbital localization for large molecules. *J. Chem. Phys.* **2006**, *125*, 174101.

(49) Maurer, S. A.; Clin, L.; Ochsenfeld, C. Cholesky-decomposed density MP2 with density fitting: Accurate MP2 and double-hybrid DFT energies for large systems. *The Journal of chemical physics* **2014**, *140*, 224112.

(50) Lehtola, S.; Jónsson, H. Pipek–Mezey orbital localization using various partial charge estimates. *J. Chem. Theory Comput.* **2014**, *10*, 642–649.

(51) Heßelmann, A. Local Molecular Orbitals from a Projection onto Localized Centers. *J. Chem. Theory Comput.* **2016**, *12*, 2720–2741.

(52) Høyvik, I. M.; Olsen, J.; Jørgensen, P. Generalising localisation schemes of orthogonal orbitals to the localisation of non-orthogonal orbitals. *Mol. Phys.* **2017**, *115*, 16–25.

(53) Folkestad, S. D.; Matveeva, R.; Høyvik, I.-M.; Koch, H. Implementation of Occupied and Virtual Edmiston–Ruedenberg Orbitals Using Cholesky Decomposed Integrals. *J. Chem. Theory Comput.* **2022**, *18*, 4733–4744.

(54) Høyvik, I.-M.; Jørgensen, P. Characterization and generation of local occupied and virtual Hartree–Fock orbitals. *Chem. Rev.* **2016**, *116*, 3306–3327.

(55) Høyvik, I.-M. The spectrum of the atomic orbital overlap matrix and the locality of the virtual electronic density matrix. *Mol. Phys.* **2020**, *118*, e1765034.

(56) Meyer, W. Ionization energies of water from PNO-CI calculations. *Int. J. Quantum Chem.* **1971**, *5*, 341–348.

(57) Schütz, M.; Yang, J.; Chan, G. K.-L.; Manby, F. R.; Werner, H.-J. The orbital-specific virtual local triples correction: OSV-L(T). *J. Chem. Phys.* **2013**, *138*, 054109.

(58) Zhao, Y.; Lynch, B. J.; Truhlar, D. G. Doubly Hybrid Meta DFT: New Multi-Coefficient Correlation and Density Functional Methods for Thermochemistry and Thermochemical Kinetics. *J. Phys. Chem. A* **2004**, *108*, 4786–4791.

(59) Zhao, Y.; Lynch, B. J.; Truhlar, D. G. Multi-coefficient extrapolated density functional theory for thermochemistry and thermochemical kinetics. *Phys. Chem. Chem. Phys.* **2005**, *7*, 43.

(60) Grimme, S. Semiempirical hybrid density functional with perturbative second-order correlation. *J. Chem. Phys.* **2006**, *124*, 034108.

(61) Chai, J.-D.; Head-Gordon, M. Long-range corrected double-hybrid density functionals. *J. Chem. Phys.* **2009**, *131*, 174105.

(62) Goerigk, L.; Grimme, S. Double-hybrid density functionals. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2014**, *4*, 576–600.

(63) Mardirossian, N.; Head-Gordon, M. Survival of the most transferable at the top of Jacob's ladder: Defining and testing the ω B97M (2) double hybrid density functional. *The Journal of chemical physics* **2018**, *148*, 241736.

(64) Shee, J.; Loipersberger, M.; Rettig, A.; Lee, J.; Head-Gordon, M. Regularized second-order Møller–Plesset theory: A more accurate alternative to conventional MP2 for noncovalent interactions and transition metal thermochemistry for the same computational cost. *J. Phys. Chem. Lett.* **2021**, *12*, 12084–12097.

(65) Lee, J.; Head-Gordon, M. Regularized orbital-optimized second-order Møller–Plesset perturbation theory: A reliable fifth-order-scaling electron correlation model with orbital energy dependent regularizers. *Journal of chemical theory and computation* **2018**, *14*, 5203–5219.

(66) Nguyen, B. D.; Chen, G. P.; Agee, M. M.; Burow, A. M.; Tang, M. P.; Furche, F. Divergence of many-body perturbation theory for noncovalent interactions of large molecules. *J. Chem. Theory Comput.* **2020**, *16*, 2258–2273.

(67) Werner, H.-J.; Manby, F. R.; Knowles, P. J. Fast linear scaling second-order Møller-Plesset perturbation theory (MP2) using local and density fitting approximations. *J. Chem. Phys.* **2003**, *118*, 8149–8160.

(68) Subotnik, J. E.; Shao, Y.; Liang, W.; Head-Gordon, M. An efficient method for calculating maxima of homogeneous functions of orthogonal matrices: Applications to localized occupied orbitals. *J. Chem. Phys.* **2004**, *121*, 9220–9229.

(69) Subotnik, J. E.; Sodt, A.; Head-Gordon, M. Localized orbital theory and ammonia triborane. *Phys. Chem. Chem. Phys.* **2007**, *9*, 5522.

(70) Lehtola, S.; Jónsson, H. Unitary Optimization of Localized Molecular Orbitals. *J. Chem. Theory Comput.* **2013**, *9*, 5365–5372.

(71) Wang, Z.; Head-Gordon, M. A Linear Surrogate for Optimizing Functions of an Orthogonal Matrix with Applications in Wave Function Theory. *Mol. Phys.* **2022**, In Press.

(72) Damle, A.; Lin, L.; Ying, L. Compressed representation of Kohn–Sham orbitals via selected columns of the density matrix. *J. Chem. Theory Comput.* **2015**, *11*, 1463–1469.

(73) Kohn, W. Analytic properties of Bloch waves and Wannier functions. *Phys. Rev.* **1959**, *115*, 809.

(74) Kohn, W. Construction of Wannier functions and applications to energy bands. *Phys. Rev. B* **1973**, *7*, 4388.

(75) Brandhorst, K.; Head-Gordon, M. Fast sparse Cholesky decomposition and inversion using nested dissection matrix reordering. *J. Chem. Theory Comput.* **2011**, *7*, 351–368.

(76) Lu, W.; Wang, C.; Schmidt, M.; Bytautas, L.; Ho, K.; Ruedenberg, K. Molecule intrinsic minimal basis sets. I. Exact resolution of ab initio optimized molecular orbitals in terms of deformed atomic minimal-basis orbitals. *J. Chem. Phys.* **2004**, *120*, 2629–2637.

(77) Knizia, G. Intrinsic atomic orbitals: An unbiased bridge between quantum theory and chemical concepts. *J. Chem. Theory Comput.* **2013**, *9*, 4834–4843.

(78) Aldossary, A.; Head-Gordon, M. Non-iterative method for constructing valence antibonding molecular orbitals and a molecule-adapted minimum basis. *J. Chem. Phys.* **2022**, *157*, 094102.

(79) Bhatia, R.; Mukherjea, K. K. On weighted Löwdin orthogonalization. *Int. J. Quantum Chem.* **1986**, *29*, 1775–1778.

(80) King, K. *C Programming: A Modern Approach*; W.W. Norton, 2008.

(81) Høyvik, I.-M.; Jansik, B.; Jørgensen, P. Orbital localization using fourth central moment minimization. *J. Chem. Phys.* **2012**, *137*, 224114.