

Interpolative separable density fitting on adaptive real space grids

Hai Zhu,^{1,*} Chia-Nan Yeh,^{2,*} Miguel A. Morales,² Leslie Greengard,^{3,4} Shidong Jiang,^{3,†} and Jason Kaye^{2,3,†}

¹*Institute for Theoretical Sciences, Westlake University, Hangzhou, Zhejiang 310030, China*

²*Center for Computational Quantum Physics, Flatiron Institute, 162 5th Avenue, New York, NY 10010, USA*

³*Center for Computational Mathematics, Flatiron Institute, 162 5th Avenue, New York, NY 10010, USA*

⁴*Courant Institute of Mathematical Sciences, New York University, New York, NY 10012, USA*

We generalize the interpolative separable density fitting (ISDF) method, used for compressing the four-index electron repulsion integral (ERI) tensor, to incorporate adaptive real space grids for potentially highly localized single-particle basis functions. To do so, we employ a fast adaptive algorithm, the recently-introduced dual-space multilevel kernel-splitting method, to solve the Poisson equation for the ISDF auxiliary basis functions. The adaptive grids are generated using a high-order accurate, black-box procedure that satisfies a user-specified error tolerance. Our algorithm relies on the observation, which we prove, that an adaptive grid resolving the pair densities appearing in the ERI tensor can be straightforwardly constructed from one that resolves the single-particle basis functions, with the number of required grid points differing only by a constant factor. We find that the ISDF compression efficiency for the ERI tensor with highly localized basis sets is comparable to that for smoother basis sets compatible with uniform grids. To demonstrate the performance of our procedure, we consider several molecular systems with all-electron basis sets which are intractable using uniform grid-based methods. Our work establishes a pathway for scalable many-body electronic structure simulations with arbitrary smooth basis functions, making simulations of phenomena like core-level excitations feasible on a large scale.

I. INTRODUCTION

The four-index electron repulsion integral (ERI) tensor is a fundamental building block of electronic structure theories, representing the Coulomb interactions between products of N single-particle basis functions $\phi_i(\mathbf{r})$:

$$V_{ijkl} = \int d\mathbf{r} \int d\mathbf{r}' \phi_i(\mathbf{r}) \phi_j(\mathbf{r}) \frac{1}{|\mathbf{r} - \mathbf{r}'|} \phi_k(\mathbf{r}') \phi_l(\mathbf{r}'). \quad (1)$$

The ERI tensor serves as a universal starting point for incorporating quantum many-body effects within an electronic structure simulation based on a single-particle basis. However, building and storing this tensor directly is typically infeasible for large system sizes, and can become a primary bottleneck: it amounts to solving the Poisson equation for N^2 orbital pairs, computing an inner product for all N^4 orbital combinations, and storing N^4 tensor elements. Assuming the number of grid points scales as $\mathcal{O}(N)$, this yields an $\mathcal{O}(N^5)$ total cost.

Even though at first sight the problem of computing and storing V_{ijkl} seems too daunting for large problems, the task is made more manageable when we realize that the tensor is expected to be significantly rank deficient. Here, we show that the rank of V_{ijkl} , interpreted as an $N^2 \times N^2$ matrix, scales linearly with the basis size N . As a result, the direct evaluation of (1) is typically replaced by more efficient formulations that lead to compact representations with reduced computational and storage requirements. Common approaches

include resolution of identity (RI) [1–5], Cholesky decomposition [2, 6, 7], tensor hypercontraction (THC) [8–12], and the canonical polyadic (CP) approximation [13–17]. Other approaches include hierarchical matrix representations of the ERI [18, 19], and tensor network methods to represent the basis functions [20]. Each of these schemes involves a particular compression of the ERI tensor, with different trade-offs in terms of computational cost, accuracy, and applicability.

The most common approach used to compress the ERI tensor in quantum many-body methods, particularly in the quantum chemistry community, is the RI decomposition. It has a good balance of efficiency, accuracy and robustness, which has led to widespread use and quick adoption. It, nonetheless, suffers from serious limitations since it requires $\mathcal{O}(N^3)$ storage and leads to $\mathcal{O}(N^4)$ scaling algorithms when used in correlated many-body calculations. In order to reduce memory requirements and computational costs, more aggressive factorizations are needed. Among competing alternatives, THC stands out for enabling compact ($\mathcal{O}(N^2)$ storage), low-scaling, black-box implementations of a broad range of electronic structure methods, including hybrid density functional theory (DFT) [21–24], Hartree-Fock (HF) [25], coupled-cluster theories [26–30], many-body perturbation theory (e.g., MP2, MP3, *GW*) [8, 9, 31–40], and auxiliary-field quantum Monte Carlo (AFQMC) [41]. One of the leading methods of constructing the THC representation is the interpolative separable density fitting (ISDF) algorithm [10, 11], which scales as $\mathcal{O}(N^3)$, offers controllable accuracy, and relies on standard numerical linear algebra routines. Other methods of constructing the THC representation, such as least-squares (LS) THC [9, 31, 32], have $\mathcal{O}(N^4)$ computational complexity and often require an initial RI/Cholesky decomposition. Notice that

* These authors contributed equally to this work.

† Corresponding authors (sjiang@flatironinstitute.org, jkaye@flatironinstitute.org)

for large systems, particularly in periodic calculations with a large number of k -points, the generation of the RI/Cholesky decomposition is typically unfeasible, limiting the utility of the LS-THC approach.

ISDF uses the interpolative decomposition [42, 43] to construct a low rank representation of the family of pair densities $\rho_{ij}(\mathbf{r}) = \phi_i(\mathbf{r})\phi_j(\mathbf{r})$, in the form of an auxiliary basis $\zeta_\mu(\mathbf{r})$. From these, one can construct a projection of the Coulomb integral operator,

$$V_{\mu\nu} = \int d\mathbf{r} \int d\mathbf{r}' \zeta_\mu(\mathbf{r}) \frac{1}{|\mathbf{r} - \mathbf{r}'|} \zeta_\nu(\mathbf{r}'), \quad (2)$$

which directly leads to a THC decomposition of V_{ijkl} . Sec. II A provides a more detailed description of this procedure. Thus, building the THC decomposition of V_{ijkl} using ISDF requires two main steps: (1) build the auxiliary basis $\zeta_\mu(\mathbf{r})$ from the pair densities $\rho_{ij}(\mathbf{r})$, and (2) compute the Coulomb integrals $V_{\mu\nu}$. The first step is agnostic to the real space grid on which the pair densities are represented. The second step is the solution of the Poisson equation for the auxiliary basis functions $\zeta_\mu(\mathbf{r})$.

In previous works, the THC-ISDF approach has been applied to periodic systems, using a uniform real space grid to represent the pair densities and a standard fast Fourier transform (FFT)-based Poisson solver. This limits the method to single-particle basis functions which are well-resolved on uniform grids with a modest number of points, requiring the use of pseudopotentials. However, many methods, such as the fast multipole method and multi-grid based approaches, have been developed to solve the Poisson equation on adaptive grids which automatically refine into localized features of the density. For the present setting—solving the Poisson equation on a cubic domain—the state of the art algorithms are black-box, linear scaling, and capable of delivering user-specified accuracy with computational throughput (in grid points per second) approaching that of the FFT. Following this approach, we build an adaptive grid resolving the collection of pair densities, use ISDF to construct an auxiliary basis on this grid, and use the recently introduced dual-space multilevel kernel-splitting (DMK) [44] algorithm as an adaptive Poisson solver to compute $V_{\mu\nu}$. This yields a black-box, $\mathcal{O}(N^3)$ algorithm to build a THC decomposition of the ERI tensor which allows for general, potentially highly-localized single-particle basis functions. While our focus in this work is on molecular systems, our framework is equally applicable to periodic systems by imposing periodic boundary conditions in the DMK solver.

The ability to operate on adaptive grids with no assumptions on the analytical form of the basis functions makes ISDF applicable to a wide variety of single-particle basis sets beyond Gaussians, including representations like projector-augmented waves (PAW) and linearized augmented plane waves (LAPW). It enables a cubic-scaling THC construction for all-electron calculations, in which incorporating localized core basis functions is essential for an accurate description of core excitations.

The remainder of this paper is organized as follows. Section II describes THC, ISDF, and our adaptive grid approach. In Section III, we evaluate the performance and accuracy of the method across a range of basis sets, chemical species, and system sizes, and analyze its effectiveness for downstream electronic structure predictions. We conclude with a summary and outlook in Section IV.

II. THC DECOMPOSITION VIA ISDF USING ADAPTIVE GRIDS

The tensor hypercontraction (THC) decomposition of the ERI tensor takes the form

$$V_{ijkl} \approx \sum_{\mu, \nu=1}^R X_{i\mu} X_{j\mu} V_{\mu\nu} X_{k\nu} X_{l\nu}, \quad (3)$$

for $V_{\mu\nu}$ a projection of the original tensor, and $X_{i\mu}$ the collocation matrix. The rank R of the decomposition will be discussed below. Thus the THC decomposition separates the indices i, j, k , and l by introducing two auxiliary indices μ and ν .

The THC form in (3) enables contractions over the ERI tensor to be reorganized into a sequence of tractable matrix-matrix multiplications. This can already be seen from the exchange potential K in the Hartree-Fock approximation:

$$K_{ij} = - \sum_{a,b=1}^N \rho_{ab} V_{iabj} \quad (4a)$$

$$= - \sum_{\mu, \nu=1}^R X_{i\mu} V_{\mu\nu} X_{j\nu} \sum_{a,b=1}^N X_{a\mu} \rho_{ab} X_{b\nu}. \quad (4b)$$

Here ρ is the single-particle density matrix. The complete separation of the orbital indices in (4b) reduces the cost of computing (4a) from $\mathcal{O}(N^4)$ to $\mathcal{O}(RN^2 + R^2N)$.

A similar advantage arises in the GW approximation, in which the dynamic self-energy $\Sigma(\tau)$ involves a similar contraction:

$$\Sigma_{ij}(\tau) = - \sum_{a,b=1}^N G_{ab}(\tau) W_{iabj}(\tau) \quad (5a)$$

$$= - \sum_{\mu, \nu=1}^R X_{i\mu} W_{\mu\nu}(\tau) X_{j\nu} \sum_{a,b=1}^N X_{a\mu} G_{ab}(\tau) X_{b\nu}. \quad (5b)$$

Here $G(\tau)$ is the imaginary time Green's function, and $W(\tau)$ is the screened interaction. We refer to Ref. 45 for further details on the application of THC to both HF and GW .

A. Interpolative separable density fitting

ISDF is an algorithm which constructs the THC decomposition (3). We begin with the pair density

$$\rho_{ij}(\mathbf{r}) = \phi_i(\mathbf{r})\phi_j(\mathbf{r}), \quad (6)$$

for $i, j = 1, \dots, N$. Although there are N^2 pair densities, one might hope that most are numerically linearly dependent, i.e., that the numerical rank of the set of functions $\rho_{ij}(\mathbf{r})$ is $R \ll N^2$. This implies the existence of a decomposition of the form

$$\rho_{ij}(\mathbf{r}) \approx \sum_{\mu=1}^R \phi_i(\mathbf{r}_\mu)\phi_j(\mathbf{r}_\mu)\zeta_\mu(\mathbf{r}). \quad (7)$$

To see this, we consider $\rho_{ij}(\mathbf{r})$ as an $\infty \times N^2$ matrix of rank R (or, if preferred, an $M \times N^2$ matrix with large M , using a dense discretization of $\mathbf{r}$ by a grid of M points), and note that the row rank is equal to the column rank. In other words, $\rho_{ij}(\mathbf{r})$ can be reconstructed from a linear combination of R of its rows, with coefficients $\zeta_\mu(\mathbf{r})$. We

refer to $\mathbf{r}_\mu$ as the interpolating points, and $\zeta_\mu(\mathbf{r})$ as the auxiliary basis.

If one indeed discretizes $\rho_{ij}(\mathbf{r})$ on a grid of M points to obtain an $M \times N^2$ matrix, then (7) constitutes an interpolative decomposition (ID) of that matrix [42, 43]. Several algorithms exist to construct this ID [10, 12, 22, 46]. In this work we adopt the approach described in Refs. 12 and 46, summarized in App. A, which employs a pivoted Cholesky decomposition of the square of the pair density matrix. This yields the decomposition (7), with the rank R chosen to meet a desired approximation accuracy.

As is discussed in Sec. II B and App. B, the size M of the real space grid required to resolve all N^2 pair densities differs by a constant factor (independent of N) from that required to resolve all N single-particle basis functions, so $M = \mathcal{O}(N)$. Since $R \leq \min(M, N^2)$, we have $R = \mathcal{O}(N)$, yielding a compression of $\rho_{ij}(\mathbf{r})$ to $\mathcal{O}(N)$ auxiliary basis functions. The compression factor will be explored numerically in Sec. III.

A THC decomposition of the ERI tensor can be immediately obtained from the compressed representation (7):

$$\begin{aligned} V_{ijkl} &= \int d\mathbf{r} \int d\mathbf{r}' \rho_{ij}(\mathbf{r}) \frac{1}{|\mathbf{r} - \mathbf{r}'|} \rho_{kl}(\mathbf{r}') \\ &\approx \sum_{\mu, \nu=1}^R \phi_i(\mathbf{r}_\mu)\phi_j(\mathbf{r}_\mu) \left[\int d\mathbf{r} \int d\mathbf{r}' \zeta_\mu(\mathbf{r}) \frac{1}{|\mathbf{r} - \mathbf{r}'|} \zeta_\nu(\mathbf{r}') \right] \phi_k(\mathbf{r}_\nu)\phi_l(\mathbf{r}_\nu) \\ &= \sum_{\mu, \nu=1}^R X_{i\mu} X_{j\mu} V_{\mu\nu} X_{k\nu} X_{l\nu}, \end{aligned} \quad (8)$$

where

$$X_{i\mu} = \phi_i(\mathbf{r}_\mu), \quad (9a)$$

$$V_{\mu\nu} = \int d\mathbf{r} \int d\mathbf{r}' \zeta_\mu(\mathbf{r}) \frac{1}{|\mathbf{r} - \mathbf{r}'|} \zeta_\nu(\mathbf{r}'). \quad (9b)$$

The accuracy of this decomposition is controlled by the accuracy of the ID approximation of the pair densities, which can be systematically improved by increasing R . Indeed, if $R = N^2$, then the ID can be made exact, yielding a trivial THC decomposition for which the projected Coulomb matrix contains $R^2 = N^4$ degrees of freedom.

The key steps of the ISDF procedure are therefore (1) the resolution of the pair densities $\rho_{ij}(\mathbf{r})$ on a real space grid of M points, (2) the ID of the resulting $M \times N^2$ matrix to obtain the auxiliary basis $\zeta_\mu(\mathbf{r})$ on the real space grid, and (3) the solution of the Poisson equation for the auxiliary basis functions, and subsequent inner products, to obtain the projected Coulomb matrix $V_{\mu\nu}$ in Eq. 9b. As described in App. B, the ID step is a black-box linear algebra procedure that is agnostic to the real space grid on which the pair densities are represented.

We therefore discuss Steps 1 and 3.

B. Adaptive real space grids and the solution of the Poisson equation

Evaluating the integrals (9b) requires (i) solving the collection of Poisson equations $-\Delta u_\nu(\mathbf{r}) = \zeta_\nu(\mathbf{r})$, and (ii) computing the inner products $\int d\mathbf{r} \zeta_\mu(\mathbf{r}) u_\nu(\mathbf{r})$. The auxiliary basis functions $\zeta_\mu(\mathbf{r})$ are given on a real space grid of M points. For simulations in which core electrons are excluded, e.g., using pseudopotentials or effective core potentials, the single-particle basis functions, and therefore the auxiliary basis functions, tend to be smooth even near nuclei. In this case, a uniform real space grid is sufficient, and the Poisson equation can be solved using the fast Fourier transform (FFT). Then the inner products can be computed using a suitable uniform grid quadrature rule.

Remark 1. For periodic systems, the single-particle basis functions and therefore the auxiliary basis functions

are considered to be periodic, and the Poisson equation can be solved on a single unit cell with periodic boundary conditions under the charge neutrality assumption, i.e., the mean of the electron density is zero. The inner products can then be computed by the trapezoid rule with spectral accuracy [47]. For molecular systems, we assume that all single-particle basis functions have decayed sufficiently by the boundary of the simulation domain, and we solve the Poisson equation in free space. Straightforward use of the FFT to solve the Poisson equation in this setting is inappropriate, since it assumes periodic boundary conditions; in other words, it will introduce slowly-decaying spurious periodic images of the desired solution. Rather, to use uniform grids in a non-periodic setting, the FFT should be used in conjunction with a truncated kernel method, such as that of Ref. [48], which correctly imposes free space boundary conditions. In the present work, we indeed consider the molecular case, but use adaptive rather than uniform grids, so this discussion is not relevant.

In practice, the use of uniform grids is a significant limitation. When core electrons are treated explicitly or hard pseudopotentials are used, the single-particle basis functions have highly localized features, typically requiring denser grids than are affordable (both in terms of memory and computational cost of the various steps). The solution of the Poisson equation on adaptive grids non-uniformly resolving localized structures is a well-studied problem in computational mathematics. Robust, black-box, highly efficient algorithms are available both to construct adaptive grids which automatically resolve a given density to controllable high-order accuracy, and to solve the Poisson equation on the resulting grids. A primary goal of such solvers is linear or quasi-linear scaling in the number of grid points: that is, the cost of solving the Poisson equation should scale as $\mathcal{O}(M)$ or $\mathcal{O}(M \log M)$, where M is the number of points in an adaptive grid discretizing the density.

Many such algorithms are variants of the fast multipole method (FMM) [49–57] or multigrid methods [58–61] (see also [62] for a comparison of the FFT, FMM, and multigrid methods). For free space problems, multigrid-based methods typically require truncating the computational domain and imposing artificial boundary conditions, so we do not consider this approach. Here, we use the recently developed dual-space multilevel kernel-splitting (DMK) framework [44]. DMK shares several characteristics with the FMM, namely its tree-based algorithmic structure and $\mathcal{O}(M)$ computational complexity for evaluating the convolution of a kernel and a function represented on an adaptive grid. However, DMK introduces significant improvements, including a simplified algorithmic framework, applicability to a broader class of kernels, and an acceleration of computationally intensive near-field calculations.

In the present work, we use adaptive octrees with product Chebyshev grids to discretize the basis functions, and the DMK algorithm, which is compatible with such a

discretization, to solve the Poisson equation. Once the adaptive octree discretization has been constructed, the DMK solver can be treated as a black box, so we refer to Ref. 44 for further details on DMK.

To build the adaptive octree discretization, we begin with the single-particle basis functions $\phi_i(\mathbf{r})$, which are typically given in closed form. The tree is built through a recursive subdivision procedure. One begins at the root level of the tree, with a single box covering the full domain. The ϕ_i are evaluated on an $n \times n \times n$ product Chebyshev grid in the box, from which one can obtain Chebyshev polynomial interpolants $p^{(\phi_i, B)}(\mathbf{r}) = \sum_{j,k,l=0}^{n-1} C_{jkl} T_j(\tilde{x}_1) T_k(\tilde{x}_2) T_l(\tilde{x}_3)$. Here T_j is the degree j Chebyshev polynomial of the first kind, $\mathbf{r} = (x_1, x_2, x_3)$, and $\tilde{x}_i = 2(x_i - c_i)/L$ with c_i the center of the box B and L its side length. The resulting approximation is accurate to order $n + 1$, with a fixed, moderate value of n typically chosen. If the maximum interpolation error over all ϕ_i is larger than a user-specified tolerance ε , the box is subdivided into eight equal boxes, and the procedure is repeated for each of the eight new boxes. Here, we use the interpolation error on box B given by

$$E^{(\phi_i, B)} = \frac{\|\phi_i - p^{(\phi_i, B)}\|_{L^2(B)}}{\|\phi_i\|_{L^2}}, \quad (10)$$

where the L^2 norm in the denominator is taken over the full domain. This process continues recursively until $E^{(\phi_i)} = \sqrt{\sum_B (E^{(\phi_i, B)})^2} < \varepsilon$ for $1 \leq i \leq N$, where the sum is taken over all leaf-level boxes. The result of this procedure is a nonuniform collection of the leaf-level boxes, each containing a Chebyshev interpolant. This constitutes a representation of the functions $\phi_i(\mathbf{r})$ accurate to ε (for instance, it can be evaluated at a point by determining the leaf-level box containing that point, and evaluating the corresponding polynomial interpolant). The boxes will be more refined near localized features of the functions $\phi_i(\mathbf{r})$. It can be shown that the total cost of this adaptive grid construction procedure for a single function is $\mathcal{O}(M)$, where M is the total number of points in all leaf-level boxes [63]. An example of such an octree for single-particle orbitals of $(\text{NH}_3)_2$ using the aug-cc-pVTZ basis set is shown in Fig. 1.

We show in App. B that if the single-particle basis functions ϕ_i are well-approximated on each leaf-level box by a Chebyshev polynomial interpolant of degree $n - 1$, the N^2 pair densities $\rho_{ij}(\mathbf{r}) = \phi_i(\mathbf{r})\phi_j(\mathbf{r})$ are well-approximated by Chebyshev polynomial interpolants of degree $2n - 1$. To be more precise, upsampling by increasing the polynomial degree in each box from $n - 1$ to $2n - 1$ guarantees $E^{(\rho_{ij}, B)} \leq C\varepsilon$ for $1 \leq i, j \leq n$ and all leaf boxes B , where the constant C , given explicitly in App. B, depends only on $\max_{1 \leq i \leq N} \|\phi_i(\mathbf{r})\|_\infty$ and on n , which is fixed. We show in App. A that the ISDF auxiliary basis functions $\zeta_\mu(\mathbf{r})$ are linear combinations of the pair densities $\rho_{ij}(\mathbf{r})$, so although this does not prove the same upsampled octree representation is sufficient to resolve them to accuracy ε (this depends on the proper-

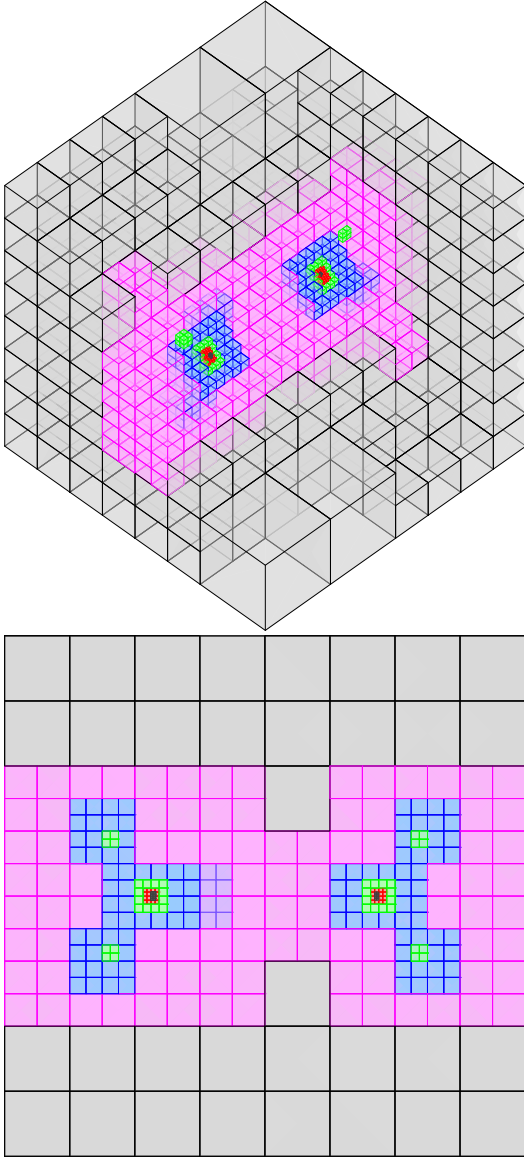


FIG. 1. Adaptive octree resolving single-particle orbitals of $(\text{NH}_3)_2$ using the aug-cc-pVTZ basis set. Top: three-dimensional view of the octree showing the adaptively refined boxes near nuclei. Bottom: slice view of the same octree. Colors indicate different refinement levels.

ties of the matrix A introduced in App. A), we find this to be sufficient in practice.

The DMK Poisson solver is directly compatible with the resulting adaptive grid, taking a density on this grid as input and producing the solution of the Poisson equation as a piecewise polynomial interpolant on the same grid. To compute the inner products appearing in (9b), we compute the integral on each box using Clenshaw-Curtis quadrature rules [64] on the corresponding Chebyshev grids, and then sum the results.

C. Summary of the algorithm and computational complexity

We summarize our full procedure as follows.

1. Build an adaptive real space grid for the single-particle basis functions $\phi_i(\mathbf{r})$, with error tolerance ε , using the algorithm described in Sec. II B. After upsampling, this discretization is used to represent the pair densities $\rho_{ij}(\mathbf{r})$ and the auxiliary basis functions $\zeta_\mu(\mathbf{r})$. The cost of this procedure scales as $\mathcal{O}(MN)$, since N single-particle basis functions are discretized on a grid of M points.
2. Using the ID procedure described in App. A, select the interpolating points $\mathbf{r}_\mu$ and compute the ISDF auxiliary basis functions $\zeta_\mu(\mathbf{r})$ on the M discretization points. The cost of this procedure scales as $\mathcal{O}(R^2M)$.
3. Solve the Poisson equation for each $\zeta_\mu(\mathbf{r})$ on the real space grid using the DMK algorithm. The cost of this step is $\mathcal{O}(M)$ for each auxiliary basis function, or $\mathcal{O}(RM)$ in total. Compute the inner products of the results to obtain $V_{\mu\nu}$, at an $\mathcal{O}(M)$ cost per pair (μ, ν) , or $\mathcal{O}(R^2M)$ in total. Using (9a) and (9b), we can now form the THC decomposition (3) of the ERI tensor V_{ijkl} .

We note that $R > N$ since the single-particle basis functions are linearly independent, so the cost of the adaptive grid construction step is asymptotically subdominant. The total computational complexity is therefore $\mathcal{O}(R^2M)$, with the ID and inner product steps dominating the cost (and notably not the Poisson solves). Since we expect $R = \mathcal{O}(N)$, the complexity in terms of M and N is $\mathcal{O}(MN^2)$, and since furthermore $M = \mathcal{O}(N)$, the complete algorithm is cubic scaling in system size.

III. NUMERICAL RESULTS

In this section, we demonstrate the accuracy and performance of our combination of ISDF with an adaptive Poisson solver. We systematically examine how the error of ISDF-approximated ERIs varies with the adaptive tree tolerance ε (Sec. III A), the size R of the ISDF auxiliary basis (Sec. III B), and the basis set size N and locality (Sec. III C). The improvement over the uniform grid approach is quantified in Sec. III D. We then use the approximated ERIs in electronic structure calculations within the *GW* approximation (Sec. III E).

We use all-electron, atom-centered Gaussian-type orbitals (GTOs) obtained from the PySCF package [65]. The basis sets contain Gaussian exponents ranging from approximately 10^{-2} to 10^5 , with small exponents corresponding to diffuse orbitals and large exponents corresponding to localized core orbitals. This wide range represents an intractable numerical challenge for uniform

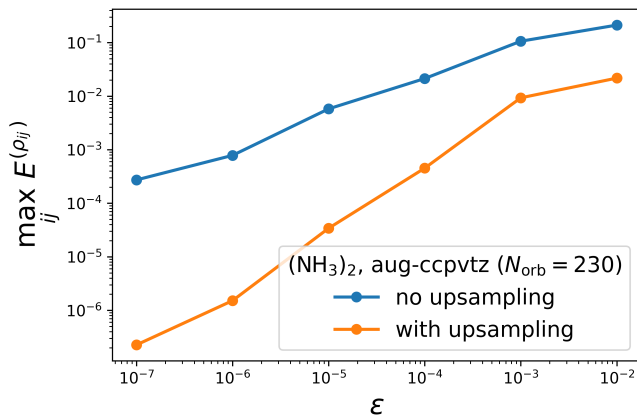


FIG. 2. Maximum L^2 error of adaptive octree grid representation of all pair densities ρ_{ij} , versus adaptive grid tolerance, for the $(\text{NH}_3)_2$ example. We show the error of the interpolant on the original adaptive grid used to resolve the single-particle orbitals (blue, no upsampling), as well as for a grid upsampled by a factor of 1.5 in each dimension (orange, with upsampling). We use the polynomial order $n = \log_{10}(\epsilon^{-1}) + 1$ in the octree grid construction.

grid-based methods. We emphasize that although we use GTOs for the examples here, our proposed framework is compatible with general basis sets, including non-atom-centered, non-Gaussian, and highly localized functions.

The ISDF step and subsequent electronic structure evaluations are performed using CoQuí [66], a software package for electronic structure simulations beyond density functional theory. Within CoQuí, the ISDF procedure incorporates adaptive real space grid generation and DMK Poisson solves using the `dmk` code [67].

A. Accuracy of adaptive grid representation

We first demonstrate our assertion, described in Sec. II B and App. B, that upsampling the adaptive grid used to represent the single-particle orbitals $\phi_i(\mathbf{r})$ effectively resolves the pair densities $\rho_{ij}(\mathbf{r})$. Though it is shown in App. B that an upsampling factor of 2 leads to a rigorous error bound, we find in practice that a factor 1.5 is sufficient, and use this value. We take the polynomial order $n = \log_{10}(\epsilon^{-1}) + 1$ in the adaptive grid construction, so that higher-order polynomials are used for more stringent tolerances.

We consider the ammonia dimer $(\text{NH}_3)_2$ with aug-cc-pVTZ basis set. Fig. 2 shows the maximum interpolation error of any ρ_{ij} using the original adaptive grid resolving the ϕ_i , and the upsampled grid, varying the adaptive grid generation tolerance ϵ . We observe that the tolerance is achieved within one digit for all ρ_{ij} on the upsampled grid.

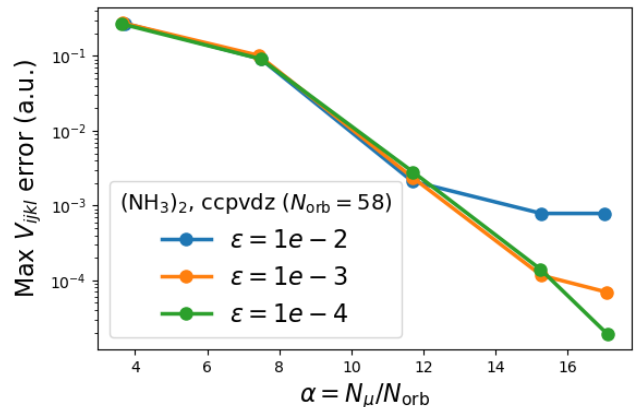


FIG. 3. Maximum error of ERI tensor V_{ijkl} for $(\text{NH}_3)_2$, varying the ISDF truncation rank R of the pair densities and the adaptive grid tolerance ϵ .

B. Accuracy of ISDF truncation

We next examine the convergence of the ISDF-approximated ERI tensor V_{ijkl} with the ISDF rank R . Since we expect the numerical rank of the pair density matrix $\rho_{ij}(\mathbf{r})$ to scale as $R = \mathcal{O}(N)$, we define the proportionality factor $\alpha = R/N$, the ratio of the number of auxiliary basis functions to the number of single-particle basis functions. In practice, the ISDF error is controlled by either varying the ID rank R or the ID error tolerance (using a rank-revealing algorithm), but we report α as a measure of the efficiency of the ISDF approximation in compressing the pair densities.

In Fig. 3, for the same system as above, we plot the maximum error in V_{ijkl} as a function of α for different values of the adaptive grid tolerance ϵ . We observe exponential convergence of the ISDF error with respect to α (equivalently, R , since N is fixed), plateauing at a value determined by ϵ . Reference values of V_{ijkl} are computed analytically with PySCF, which is possible due to the use of GTOs.

C. Increasing basis set size and locality

We next investigate the performance of the algorithm with larger and more localized basis sets, in each case measuring the maximum error of the ERI tensor against α , as in Fig. 3. In Fig. 4, we again examine the $(\text{NH}_3)_2$ molecule, now using the cc-pVXZ series of all-electron basis sets with $X = \text{D, T, Q}$. The adaptive grid tolerance is fixed at $\epsilon = 10^{-4}$. For all three basis sets, the ISDF approximation systematically improves with increasing α , and we observe overall insensitivity of the convergence behavior and compression factor α to the basis set size. This controlled convergence contrasts with resolution-of-identity (RI) or density fitting (DF) methods, in which the accuracy is inherently limited by that of

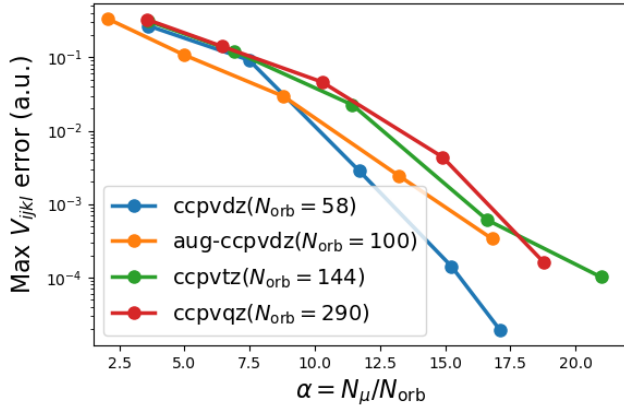


FIG. 4. Maximum error of the ERI tensor V_{ijkl} for $(\text{NH}_3)_2$, varying the ISDF truncation rank of the pair densities, using all-electron cc-PVXZ basis sets with $X = \text{D, T, and Q}$.

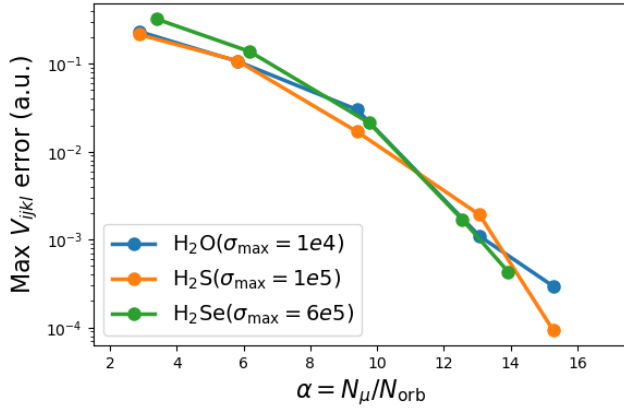


FIG. 5. Maximum error of the ERI tensor V_{ijkl} for chalcogen hydrides with increasing atomic numbers, varying the ISDF truncation rank of the pair densities. The largest GTO exponents σ_{max} for each system are shown in parentheses.

pre-optimized auxiliary basis sets tailored to each single-particle basis, rather than the systematically computed ISDF auxiliary basis [3].

To achieve chemical accuracy across all ERI elements, we find that $\alpha \approx 16$ is sufficient for all three basis sets. We note this value is larger than typical values used in periodic systems with pseudopotentials treated with uniform grids, for which $\alpha \approx 8$ often suffices [12]. For certain systems, we empirically observe a weak dependence of α on the required grid resolution, but we have not systematically studied this scaling. We also note that demanding chemical accuracy in each element of V_{ijkl} is typically an unnecessary stringent criterion, since one is usually interested in observables such as the total energy and orbital energies. Furthermore, in Sec. III E, we describe a hybrid method for electronic structure calculations which allows for smaller α by treating the Hartree potential term without THC.

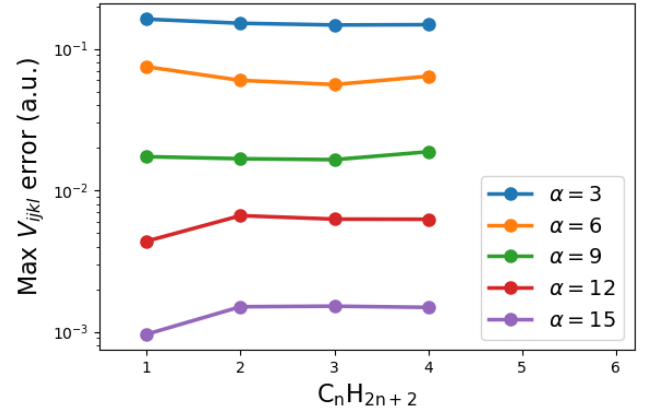


FIG. 6. Maximum error of the ERI tensor V_{ijkl} for alkanes $\text{C}_n\text{H}_{2n+2}$ varying n and the compression factor α .

We next study a sequence of systems for which the single-particle basis functions become increasingly localized. Fig. 5 shows the maximum error of the ERI tensor as a function of α for a series of chalcogen hydrides— H_2O , H_2S , and H_2Se —using the aug-ccPVDZ basis set. As the chalcogen atom becomes heavier, the largest Gaussian exponent in the basis set increases by roughly an order of magnitude, leading to significantly more localized orbitals. We observe consistent convergence behavior across the entire series, and in this case α does not increase with basis set localization, even though the degree of spatial localization in these all-electron GTO basis sets is orders of magnitude stronger than in basis sets used with pseudopotentials or effective core potentials. An FFT-based ISDF implementation is impractical for all three of these systems, due to the large spatial resolution required.

We next investigate performance as the number of atoms in a molecule is increased, using the alkane series $\text{C}_n\text{H}_{2n+2}$, for which increasing n systematically increases the size of the molecule and the number of orbitals in the system. We use the aug-cc-pVDZ basis set. In Fig. 6, we plot the maximum error of the ERI tensor for alkanes ranging from methane ($n = 1$) to pentane ($n = 5$), for different values of α . As n increases, the ISDF error at a fixed α remains roughly constant, demonstrating the scaling $R = \mathcal{O}(N)$.

D. Comparison with uniform grid approach, and timings

In this section, we compare the sizes of the real space grids obtained using our adaptive grid construction with those of a conventional uniform grid. Since all computational bottleneck steps of the ISDF procedure scale linearly (or quasi-linearly, if FFT is used as a Poisson solver) with the grid size M , the ratio of grid sizes provides a reasonable proxy for the speedup achieved by our approach

ε	Setting	Adaptive octree DOF	Uniform grid DOF
10^{-3}	ccECP ($\sigma_{\max} \approx 85$)	588384	Between $256^3 \approx 1.7 \times 10^7$ and $512^3 \approx 1.3 \times 10^8$
	all-electron ($\sigma_{\max} \approx 3 \times 10^6$)	715392	Between $16384^3 \approx 4.4 \times 10^{12}$ and $32768^3 \approx 3.5 \times 10^{13}$
10^{-5}	ccECP ($\sigma_{\max} \approx 85$)	2781864	Between $512^3 \approx 1.3 \times 10^8$ and $1024^3 \approx 1.1 \times 10^9$
	all-electron ($\sigma_{\max} \approx 3 \times 10^6$)	3414636	Between $65536^3 \approx 2.8 \times 10^{14}$ and $131072^3 \approx 2.3 \times 10^{15}$

TABLE I. Number of adaptive real space grid points and estimated number of uniform grid points required to resolve TiO all-electron and effective core potential (ECP) basis sets for different target error tolerances ε .

over an FFT-based ISDF implementation. We also show example wall clock timings for the various steps of our procedure.

We consider two numerical setups: (i) a TiO molecule treated with a cc-PVTZ-ccECP basis set and an effective core potential (ECP) [68, 69], and (ii) the same system treated with an all-electron cc-PVTZ basis set [70]. For the uniform grid, the number of grid points per dimension required to achieve a target accuracy ε is estimated by resolving each basis function along a representative one-dimensional slice, since resolving the localized basis functions with a uniform three-dimensional grid is intractable. In practice, we estimate the number of grid points required per dimension by doubling the grid size until the error falls below the desired tolerance, so we report the corresponding lower and upper bound estimates for each target tolerance.

Table I compares the required grid sizes for target relative errors $\varepsilon = 10^{-3}$ and 10^{-5} . As expected, the ratio between uniform and adaptive grid sizes is several orders of magnitude larger in the all-electron setting compared to the ECP case, since the largest Gaussian exponent increases from approximately 10^2 to 10^6 . This is because the uniform grid size increases in proportion to the largest Gaussian exponent, whereas the adaptive grid size increases only logarithmically. Still, even in the ECP case, the adaptive grid is significantly more compact than the uniform grid.

Operation	Timing (sec)
Octree construction	2.67
ISDF generation of r_μ and $\zeta_\mu(\mathbf{r})$	1313.57
Poisson solves for $\zeta_\mu(\mathbf{r})$	1407.59

TABLE II. Wall clock timings for major steps of our procedure for TiO with the all-electron cc-PVTZ basis set ($N = 98$). The octree tolerance is set to $\varepsilon = 10^{-3}$, and the ISDF auxiliary basis size ($R = 1649$) is chosen so that the maximum error in the V_{ijkl} tensor is below 10^{-3} .

Table II summarizes the single-core timings of the main steps of our workflow for TiO with the all-electron cc-PVTZ basis set and $\varepsilon = 10^{-3}$. The cost of the octree grid construction for the ϕ_i and ρ_{ij} scales only quadratically with the system size and is negligible. For this small example system, the cost of the adaptive Poisson solver step (less than one second per solve) is comparable to that of the ISDF step. We note, however, that the Poisson solver step also scales only quadratically with system

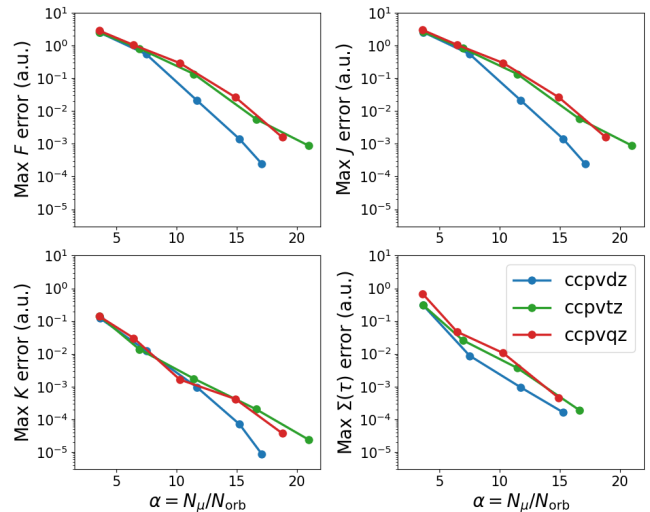


FIG. 7. Convergence of the Fock matrix (top left) with the compression factor α , including Hartree (top right) and exchange (bottom left) contributions, for $(\text{NH}_3)_2$ example, as well as the dynamic GW self-energy (bottom right).

size (and is trivially parallelizable over the auxiliary basis function index), whereas the ISDF step scales cubically. Thus, for larger systems, the cost of the adaptive Poisson solver will become negligible. We therefore conclude that for large systems, the cost *per grid point* of our fully adaptive ISDF framework is similar to that of the standard FFT-based ISDF using uniform grids, as long as the compression factor α does not grow significantly for systems requiring grid adaptivity (as in, for example, the experiment of Fig. 5). We note that the timing of the inner product in (9b) is not shown, since it is small compared to that of the ISDF step and has an identical cost scaling.

E. Correlated electron structure using GW

We next assess the performance of our framework in correlated electronic structure calculations within the GW approximation, using the THC decomposition of the ERI tensor introduced in Eqs. 4 and 5. Implementation details of the THC-GW algorithm are described in Ref. 45.

The GW self-energy consists of two components: the static Hartree-Fock potential (the Fock matrix F), which

provides the mean-field description of electrons, and the dynamic self-energy $\Sigma(\tau)$, a function of imaginary time τ , which captures dynamic electron correlations. As shown at the beginning of Sec. II, both terms depend explicitly on the ERI tensor, and are therefore directly influenced by the accuracy of the ISDF approximation.

We first examine the static mean-field contribution using the $(\text{NH}_3)_2$ dimer as an example. Fig. 7 plots the maximum error in the Fock matrix F as a function of the compression factor α , together with the corresponding errors in its Hartree (J) and exchange (K) components.

The ISDF errors in the Fock matrix F are noticeably larger than those in the ERI tensor V reported in Fig. 4, with chemical accuracy achieved only when $\alpha \approx 18$ or larger. This is consistent with prior observations in THC-based electronic structure methods, for which more auxiliary functions are needed to reach satisfactory HF accuracy, limiting the use of THC for HF calculations. More specifically, our data suggest that the dominant contribution to the Fock matrix error is the Hartree term, whereas the exchange potential shows convergence comparable to or better than that of the ERI tensor itself. This is likely a consequence of the larger overall magnitude of the Hartree potential. This result motivates a practical hybrid scheme for HF calculations: compute the Hartree potential directly using the adaptive grid DMK method and use ISDF only for the exchange term. The total cost of the former scales as $\mathcal{O}(MN^2)$, less than the $\mathcal{O}(MR^2)$ cost of our ISDF scheme. With this hybrid approach, chemical accuracy in the Fock matrix is attainable with $\alpha \approx 10$.

Fig. 7 also shows the maximum error of the dynamic GW self-energy (5) across all orbitals and imaginary times. We observe convergence similar to that for the exchange potential, with $\alpha \approx 12$ sufficient to achieve chemical accuracy. Importantly, computing the GW self-energy involves contractions over the full V_{ijkl} tensor, including the occupied-virtual and virtual-virtual blocks. In contrast, the Hartree-Fock potential only depends on the occupied subspace. The fact that both the dynamic self-energy and the exchange potential exhibit consistent convergence is a consequence of the accurate representation of the full ERI tensor by the ISDF approach, confirming the robustness and transferability of the ISDF approximation beyond mean-field theory.

IV. CONCLUSION

We have developed a cubic-scaling framework for constructing a THC representation of the ERIs for arbitrary smooth single-particle basis functions. The strength of our approach stems from the complementary advantages of ISDF and adaptive grid techniques. ISDF enables the construction of compact, on-the-fly auxiliary bases, circumventing the need for pre-optimized auxiliary basis functions tailored to specific single-particle basis sets. ISDF also scales cubically with system size, in contrast to alternatives like least squares THC, which scale quartically. Using an adaptive Poisson solver allows for fast and accurate evaluation of Coulomb matrix elements in the ISDF auxiliary basis, with a computational cost scaling linearly with the number of real space grid points even for highly localized basis functions. A key component of the method is the adaptive piecewise polynomial representation of the ISDF auxiliary basis functions $\zeta_\mu(\mathbf{r})$, which enables ISDF to handle general basis functions of arbitrary shape. Since the total cost of the adaptive Poisson solves scales only quadratically with the system size and is mild in practice, and the ISDF compression ratio α appears to be at most mildly larger for systems with highly localized orbitals, our results show that ISDF can be performed with a given number of grid points arranged uniformly or adaptively with a comparable cost.

Our approach addresses the primary computational bottleneck in THC-HF and THC-GW calculations involving sharply localized basis functions, yielding a fully cubic-scaling workflow that makes large-scale all-electron calculations practical. The explicit inclusion of core orbitals is necessary for describing core-level excitations and for accurately capturing many-body renormalization effects between core and valence bands. We note that our framework can be straightforwardly generalized to periodic systems, for which the simultaneous separation of k -points and orbital indices makes ISDF even more advantageous. This only requires a modification of the underlying adaptive Poisson solver, and is a topic of our future work.

ACKNOWLEDGMENTS

The Flatiron Institute is a division of the Simons Foundation.

-
- [1] H.-J. Werner, F. R. Manby, and P. J. Knowles, *J. Chem. Phys.* **118**, 8149 (2003).
 - [2] F. Weigend, M. Kattannek, and R. Ahlrichs, *J. Chem. Phys.* **130**, 164106 (2009).
 - [3] X. Ren, P. Rinke, V. Blum, J. Wieferink, A. Tkatchenko, A. Sanfilippo, K. Reuter, and M. Scheffler, *New J. Phys.* **14**, 053020 (2012).
 - [4] Q. Sun, T. C. Berkelbach, J. D. McClain, and G. K.-L. Chan, *J. Chem. Phys.* **147**, 164119 (2017).
 - [5] H.-Z. Ye and T. C. Berkelbach, *J. Chem. Phys.* **154**, 131104 (2021).
 - [6] N. H. F. Beebe and J. Linderberg, *Int. J. Quantum Chem.* **12**, 683 (1977).
 - [7] H. Koch, A. S. de Merás, and T. B. Pedersen, *J. Chem.*

- Phys. **118**, 9481 (2003).
- [8] E. G. Hohenstein, R. M. Parrish, and T. J. Martínez, J. Chem. Phys. **137** (2012), 044103.
 - [9] R. M. Parrish, E. G. Hohenstein, T. J. Martínez, and C. D. Sherrill, J. Chem. Phys. **137**, 224106 (2012).
 - [10] J. Lu and L. Ying, J. Comput. Phys. **302**, 329 (2015).
 - [11] J. Lu and L. Ying, Ann. Math. Sci. Appl. **1**, 321 (2016).
 - [12] C.-N. Yeh and M. A. Morales, J. Chem. Theory Comput. **19**, 6197 (2023), pMID: 37624575.
 - [13] S. R. Chinnamsetty, M. Espig, B. N. Khoromskij, W. Hackbusch, and H.-J. Flad, J. Chem. Phys. **127**, 084110 (2007).
 - [14] U. Benedikt, H. Auer, M. Espig, W. Hackbusch, and A. A. Auer, Mol. Phys. **111**, 2398 (2013).
 - [15] U. Benedikt, K.-H. Böhm, and A. A. Auer, J. Chem. Phys. **139**, 224101 (2013).
 - [16] K.-H. Böhm, A. A. Auer, and M. Espig, J. Chem. Phys. **144**, 244102 (2016).
 - [17] R. Schutski, J. Zhao, T. M. Henderson, and G. E. Scuseria, J. Chem. Phys. **147**, 184113 (2017).
 - [18] X. Xing and E. Chow, SIAM J. Sci. Comput. **42**, A162 (2020).
 - [19] X. Xing, H. Huang, and E. Chow, J. Chem. Phys. **153**, 084119 (2020).
 - [20] N. Jolly, Y. N. n. Fernández, and X. Waintal, Phys. Rev. B **111**, 245115 (2025).
 - [21] W. Hu, L. Lin, and C. Yang, J. Chem. Theory Comput. **13**, 5420 (2017).
 - [22] K. Dong, W. Hu, and L. Lin, J. Chem. Theory Comput. **14**, 1311 (2018).
 - [23] X. Qin, J. Liu, W. Hu, and J. Yang, J. Phys. Chem. A **124**, 5664 (2020).
 - [24] X. Qin, J. Li, W. Hu, and J. Yang, J. Phys. Chem. A **124**, 10066 (2020).
 - [25] S. Sharma, A. F. White, and G. Beylkin, J. Chem. Theory Comput. **18**, 7306 (2022).
 - [26] E. G. Hohenstein, R. M. Parrish, C. D. Sherrill, and T. J. Martínez, J. Chem. Phys. **137** (2012), 221101.
 - [27] E. G. Hohenstein, S. I. L. Kokkila, R. M. Parrish, and T. J. Martínez, J. Chem. Phys. **138** (2013), 124111.
 - [28] E. G. Hohenstein, S. I. L. Kokkila, R. M. Parrish, and T. J. Martínez, J. Phys. Chem. B **117**, 12972 (2013).
 - [29] R. M. Parrish, C. D. Sherrill, E. G. Hohenstein, S. I. L. Kokkila, and T. J. Martínez, J. Chem. Phys. **140** (2014), 181102.
 - [30] E. G. Hohenstein, B. S. Fales, R. M. Parrish, and T. J. Martínez, J. Chem. Phys. **156**, 054102 (2022).
 - [31] R. M. Parrish, E. G. Hohenstein, T. J. Martínez, and C. D. Sherrill, J. Chem. Phys. **138** (2013), 194107.
 - [32] S. I. L. Kokkila Schumacher, E. G. Hohenstein, R. M. Parrish, L.-P. Wang, and T. J. Martínez, J. Chem. Theory Comput. **11**, 3042 (2015).
 - [33] C. Song and T. J. Martínez, J. Chem. Phys. **144** (2016), 174111.
 - [34] C. Song and T. J. Martínez, J. Chem. Phys. **146** (2017), 034104.
 - [35] C. Song and T. J. Martínez, J. Chem. Phys. **147** (2017), 161723.
 - [36] J. Lee, L. Lin, and M. Head-Gordon, J. Chem. Theory Comput. **16**, 243 (2020).
 - [37] D. A. Matthews, J. Chem. Phys. **154** (2021).
 - [38] W. Gao and J. R. Chelikowsky, J. Chem. Theory Comput. **16**, 2216 (2020).
 - [39] H. Ma, L. Wang, L. Wan, J. Li, X. Qin, J. Liu, W. Hu, L. Lin, C. Yang, and J. Yang, J. Phys. Chem. A **125**, 7545 (2021).
 - [40] I. Duchemin and X. Blase, J. Chem. Theory Comput. **17**, 2383 (2021).
 - [41] F. D. Malone, S. Zhang, and M. A. Morales, J. Chem. Theory Comput. **15**, 256 (2019).
 - [42] H. Cheng, Z. Gimbutas, P. G. Martinsson, and V. Rokhlin, SIAM J. Sci. Comput. **26**, 1389 (2005).
 - [43] E. Liberty, F. Woolfe, P.-G. Martinsson, V. Rokhlin, and M. Tygert, Proc. Natl. Acad. Sci. **104**, 20167 (2007).
 - [44] S. Jiang and L. Greengard, Comm. Pure Appl. Math. **78**, 1086 (2025).
 - [45] C.-N. Yeh and M. A. Morales, J. Chem. Theory Comput. **20**, 3184 (2024).
 - [46] D. A. Matthews, J. Chem. Theory Comput. **16**, 1382 (2020).
 - [47] L. N. Trefethen and J. Weideman, SIAM Review **56**, 385 (2014).
 - [48] F. Vico, L. Greengard, and M. Ferrando, J. Comput. Phys. **323**, 191 (2016).
 - [49] L. Greengard, *Rapid evaluation of potential fields in particle systems*, Ph.D. thesis, Yale University, New Haven, CT (USA) (1987).
 - [50] L. Greengard, *The rapid evaluation of potential fields in particle systems* (MIT press, 1988).
 - [51] L. Greengard and V. Rokhlin, J. Comput. Phys. **73**, 325 (1987).
 - [52] H. Cheng, L. Greengard, and V. Rokhlin, J. Comput. Phys. **155**, 468 (1999).
 - [53] Z. Gimbutas and V. Rokhlin, SIAM J. Sci. Comput. **24**, 796 (2003).
 - [54] L. Greengard and V. Rokhlin, in *Acta numerica, 1997*, Acta Numer., Vol. 6 (Cambridge Univ. Press, 1997) pp. 229–269.
 - [55] L. Ying, G. Biros, and D. Zorin, J. Comput. Phys. **196**, 591 (2004).
 - [56] F. Ethridge and L. Greengard, SIAM J. Sci. Comput. **23**, 741 (2001).
 - [57] D. Malhotra and G. Biros, ACM Trans. Math. Softw. **43**, 1 (2016).
 - [58] A. Brandt, Mathematics of computation **31**, 333 (1977).
 - [59] W. Hackbusch, *Multi-grid methods and applications*, Vol. 4 (Springer Science & Business Media, 2013).
 - [60] R. S. Sampath and G. Biros, SIAM J. Sci. Comput. **32**, 1361 (2010).
 - [61] H. Sundar, G. Stadler, and G. Biros, Numer. Linear Algebra Appl. **22**, 664 (2015).
 - [62] A. Gholami, D. Malhotra, H. Sundar, and G. Biros, SIAM J. Sci. Comput. **38**, C280 (2016).
 - [63] L. Greengard, S. Jiang, M. Rachh, and J. Wang, SIAM Review **66** (2024).
 - [64] L. N. Trefethen, SIAM Review **50**, 67 (2008).
 - [65] Q. Sun, X. Zhang, S. Banerjee, P. Bao, M. Barbry, N. S. Blunt, N. A. Bogdanov, G. H. Booth, J. Chen, Z.-H. Cui, J. J. Eriksen, Y. Gao, S. Guo, J. Hermann, M. R. Hermes, K. Koh, P. Koval, S. Lehtola, Z. Li, J. Liu, N. Mardirossian, J. D. McClain, M. Motta, B. Mussard, H. Q. Pham, A. Pulkin, W. Purwanto, P. J. Robinson, E. Ronca, E. R. Sayfutyarova, M. Scheurer, H. F. Schurkus, J. E. T. Smith, C. Sun, S.-N. Sun, S. Upadhyay, L. K. Wagner, X. Wang, A. White, J. D. Whitfield, M. J. Williamson, S. Wouters, J. Yang, J. M. Yu, T. Zhu, T. C. Berkelbach, S. Sharma, A. Y. Sokolov, and G. K.-L. Chan, J. Chem. Phys. **153**, 024109 (2020).

- [66] C.-N. Yeh and M. Morales, CoQuí: Correlated Quantum interface, <https://github.com/AbInitioQHub/coqui/tree/main> (2025).
- [67] R. Blackwell, L. Greengard, S. Jiang, and D. Malhotra, DMK Software Library, <https://github.com/flatironinstitute/dmk> (2025).
- [68] M. C. Bennett, C. A. Melton, A. Annaberdiyev, G. Wang, L. Shulenburger, and L. Mitas, J. Chem. Phys. **147**, 224106 (2017).
- [69] A. Annaberdiyev, G. Wang, C. A. Melton, M. C. Bennett, L. Shulenburger, and L. Mitas, J. Chem. Phys. **149**, 134108 (2018).
- [70] N. B. Balabanov and K. A. Peterson, J. Chem. Phys. **123**, 064107 (2005).
- [71] H. Harbrecht, M. Peters, and R. Schneider, Appl. Numer. Math. **62**, 428 (2012).
- [72] M. J. D. Powell, *Approximation theory and methods* (Cambridge university press, 1981).

Appendix A: Interpolative decomposition of pair densities

This appendix describes the algorithm used to construct the decomposition (7) of the pair density function $\rho_{ij}(\mathbf{r})$. We refer to Ref. 12 for further details. We begin by defining the $N^2 \times M$ pair density matrix

$$\rho_{kl} = \rho_{ij}(\bar{\mathbf{r}}_l) \quad (\text{A1})$$

where the composite index $k = 1, \dots, N^2$ denotes a flattening of the orbital index pair (i, j) , and $\bar{\mathbf{r}}_l$, $l = 1, \dots, M$ is a real space grid sufficient to resolve all pair densities (which can be uniform or adaptive). The desired column-wise interpolative decomposition (ID) [42, 43] of the matrix ρ is given by

$$\rho_{kl} = \sum_{\mu=1}^R \rho_{k,l_\mu} \zeta_{\mu l} \quad (\text{A2})$$

for a subselection l_μ of R grid points, which is equivalent to (7). The R interpolating points $\mathbf{r}_\mu$ correspond to the selected columns, $\mathbf{r}_\mu = \bar{\mathbf{r}}_{l_\mu}$. The ISDF auxiliary basis functions $\zeta_\mu(\mathbf{r})$ can be defined via interpolation of the values $\zeta_\mu(\bar{\mathbf{r}}_l) = \zeta_{\mu l}$ (e.g., via piecewise Chebyshev interpolation in the case of the adaptive octree-based grid described in Sec. II B). The ID (A2) can be constructed in two steps: (i) select the interpolating points $\mathbf{r}_\mu$ from the original grid points $\bar{\mathbf{r}}_l$, yielding ρ_{k,l_μ} , and (ii) solve the collection of M overdetermined $N^2 \times R$ linear systems (A2) to obtain $\zeta_{\mu l}$.

The standard method for column selection in the ID is to perform a rank-revealing column-pivoted QR factorization (QRCP) and to select the columns corresponding to the pivots [42], but this procedure would be quartic-scaling in our case: $\mathcal{O}(RMN^2)$. To avoid this, we consider the $M \times M$ pair density Gram matrix:

$$S_{ll'} = \sum_{k=1}^{N^2} \rho_{kl} \rho_{kl'}. \quad (\text{A3})$$

The interpolating points can then be selected corresponding to the pivots of a pivoted Cholesky factorization of this matrix [71]. This approach has the advantage that the Gram matrix S never needs to be formed explicitly; only its diagonal elements and the R rows corresponding to the pivots are required for an approximation of rank R . This leads to an overall $\mathcal{O}(R^2 M)$ cost for selecting the interpolating points; see Ref. 71 for details. An alternative method is to use randomized QRCP algorithms on the matrix ρ , reducing the cost further to $\mathcal{O}(MN^2 \log N)$ [10] (recall $R > N$), but we have not explored this option in the present work.

Solving the collection of $N^2 \times R$ least squares problems (A2) directly using QR factorization would have an $\mathcal{O}(RMN^2)$ cost. To reduce this cost, we consider the normal equations for the least squares solutions:

$$\sum_{\mu=1}^R A_{\nu\mu} \zeta_{\mu l} = Z_{\nu l} \quad (\text{A4})$$

with

$$\begin{aligned} Z_{\mu l} &= \sum_{k=1}^{N^2} \rho_{k,l_\mu} \rho_{kl} \\ &= \sum_{i,j=1}^N \phi_i(\mathbf{r}_\mu) \phi_j(\mathbf{r}_\mu) \phi_i(\bar{\mathbf{r}}_l) \phi_j(\bar{\mathbf{r}}_l) \\ &= \left(\sum_{i=1}^N \phi_i(\mathbf{r}_\mu) \phi_i(\bar{\mathbf{r}}_l) \right)^2 \end{aligned} \quad (\text{A5})$$

and

$$\begin{aligned} A_{\mu\nu} &= \sum_{k=1}^{N^2} \rho_{k,l_\mu} \rho_{k,l_\nu} \\ &= \sum_{i,j=1}^N \phi_i(\mathbf{r}_\mu) \phi_j(\mathbf{r}_\mu) \phi_i(\mathbf{r}_\nu) \phi_j(\mathbf{r}_\nu) \\ &= \left(\sum_{i=1}^N \phi_i(\mathbf{r}_\mu) \phi_i(\mathbf{r}_\nu) \right)^2. \end{aligned} \quad (\text{A6})$$

Due to the separability of the pair densities, the costs of assembling Z and A are $\mathcal{O}(RMN)$ and $\mathcal{O}(R^2 N)$, respectively. The cost of solving (A4) directly is $\mathcal{O}(R^3 + R^2 M)$. The dominant cost therefore scales as $\mathcal{O}(R^2 M)$.

We note that (A4) gives an expression for the ISDF auxiliary basis functions as linear combinations of the pair densities,

$$\zeta_\mu(\bar{\mathbf{r}}_l) = \zeta_{\mu l} = (A^{-1} Z)_{\mu l} \quad (\text{A7})$$

$$= \sum_{\nu=1}^R \sum_{k=1}^{N^2} (A^{-1})_{\mu\nu} \rho_{k,l_\nu} \rho_{kl} \quad (\text{A8})$$

$$= \sum_{i,j=1}^N \left(\sum_{\nu=1}^R (A^{-1})_{\mu\nu} \rho_{ij}(\mathbf{r}_\nu) \right) \rho_{ij}(\bar{\mathbf{r}}_l), \quad (\text{A9})$$

suggesting that the same piecewise polynomial representation used for the pair densities might also be sufficient for the auxiliary basis functions. Further analysis would be required to establish such a statement rigorously.

Appendix B: Approximation error of pair densities

The following lemma can be found in [72]:

Lemma 1. Let $l_i(x)$, $i = 1, \dots, n$, be the Lagrange polynomials for the interpolation nodes $x_i = \cos((i - 0.5)\pi/n)$, the Chebyshev nodes of the first kind on $[-1, 1]$. Let

$$L_n = \max_{x \in [-1, 1]} \sum_{i=1}^n |l_i(x)|. \quad (\text{B1})$$

Then

$$L_n = \frac{1}{\pi} \sum_{i=1}^n \left| \cot \frac{(i - 1/2)\pi}{2n} \right|. \quad (\text{B2})$$

Furthermore, let $f \in C^n[-1, 1]$ and p^C be its Chebyshev interpolating polynomial of degree less than n . Let p^* be a best polynomial approximation of f in P , the space of polynomials of degree less than n , i.e.,

$$\|f - p^*\|_\infty \leq \|f - p\|_\infty, \quad \text{for any } p \in P. \quad (\text{B3})$$

Then

$$\|f - p^C\|_\infty \leq (1 + L_n) \|f - p^*\|_\infty. \quad (\text{B4})$$

Remark 2. L_n in (B1) is the Lebesgue constant, which is the maximum norm of the Lagrange interpolation operator corresponding to the underlying interpolation nodes. For Chebyshev nodes, numerical calculations show that $L_n < 2.5$ for $n \leq 20$. It can also be shown that

$$L_n \leq \frac{2}{\pi} \log n + 1. \quad (\text{B5})$$

It is straightforward to show that the maximum norm of the Lagrange interpolation operator in d dimensions with tensor product Chebyshev interpolation nodes is simply L_n^d , and the three-dimensional version of (B4) is

$$\|f - p^C\|_\infty \leq (1 + L_n^3) \|f - p^*\|_\infty, \quad (\text{B6})$$

where p^C is the tensor-product Chebyshev interpolating polynomial with degree less than n in each variable.

Theorem 1. Let $\{\phi_i\}_{i=1}^N$ be a set of continuous functions on $B_0 = [-1, 1]^3$. Let $p_i(\mathbf{r}) = \sum_{jkl=0}^{n-1} C_{jkl} T_j(x_1) T_k(x_2) T_l(x_3)$ be the tensor product Chebyshev interpolating polynomial of degree at most

$n - 1$ for $\phi_i(\mathbf{r})$, i.e., $p_i(\mathbf{r}_\nu) = \phi_i(\mathbf{r}_\nu)$ with $\{\mathbf{r}_\nu\}_{\nu=1}^{n^3}$ the tensor product Chebyshev nodes on B_0 . Suppose

$$\|\phi_i - p_i\|_\infty \leq \varepsilon, \quad i = 1, \dots, N. \quad (\text{B7})$$

Let $\{\tilde{\mathbf{r}}_\nu\}_{\nu=1}^{(2n)^3}$ be the $2n \times 2n \times 2n$ tensor product Chebyshev nodes on B_0 . Then the tensor product Chebyshev interpolating polynomial approximation of $\phi_i \phi_j$ at the nodes $\{\tilde{\mathbf{r}}_\nu\}$, denoted by $\tilde{\Phi}_{ij}$, satisfies the error bound

$$\|\phi_i \phi_j - \tilde{\Phi}_{ij}\|_\infty \leq (1 + L_{2n}^3)(1 + 2C_\infty)\varepsilon, \quad (\text{B8})$$

where $C_\infty = \max_i \|\phi_i\|_\infty$.

Proof. Let p_i denote the Chebyshev interpolating polynomial of ϕ_i of degree less than n . By definition, we have

$$\tilde{\Phi}_{ij}(\mathbf{r}) = \sum_{\mu=1}^{n^3} \phi_i(\mathbf{r}_\mu) \phi_j(\mathbf{r}_\mu) l_\mu(\mathbf{r}), \quad i, j = 1, \dots, N, \quad (\text{B9})$$

where $l_\nu(\mathbf{r})$ is the tensor-product Lagrange basis polynomial equal to one at node ν and zero at all other nodes. Thus,

$$\begin{aligned} \|\phi_i \phi_j - p_i p_j\|_\infty &= \|\phi_i \phi_j - \phi_i p_j + \phi_i p_j - p_i p_j\|_\infty \\ &\leq \|\phi_i(\phi_j - p_j) + (\phi_i - p_i)p_j\|_\infty \\ &\leq \|\phi_i\|_\infty \|\phi_j - p_j\|_\infty + \|p_j\|_\infty \|\phi_i - p_i\|_\infty \\ &\leq C_\infty \varepsilon + (C_\infty + \varepsilon) \varepsilon = (\varepsilon + 2C_\infty) \varepsilon \\ &< (1 + 2C_\infty) \varepsilon. \end{aligned} \quad (\text{B10})$$

Let $\tilde{\Phi}_{ij}$ be defined as above, and P^* the best polynomial approximation of $\phi_i \phi_j$ of degree less than $2n$. By (B6), we have

$$\|\phi_i \phi_j - \tilde{\Phi}_{ij}\|_\infty \leq (1 + L_{2n}^3) \|\phi_i \phi_j - P^*\|_\infty. \quad (\text{B11})$$

Since $p_i p_j$ is a particular polynomial approximation of $\phi_i \phi_j$ of degree less than $2n$, we have

$$\|\phi_i \phi_j - P^*\|_\infty \leq \|\phi_i \phi_j - p_i p_j\|_\infty. \quad (\text{B12})$$

Combining (B10) – (B12), we obtain

$$\begin{aligned} \|\phi_i \phi_j - \tilde{\Phi}_{ij}\|_\infty &\leq (1 + L_{2n}^3) \|\phi_i \phi_j - P^*\|_\infty \\ &\leq (1 + L_{2n}^3) \|\phi_i \phi_j - p_i p_j\|_\infty \\ &\leq (1 + L_{2n}^3)(1 + 2C_\infty) \varepsilon. \end{aligned} \quad (\text{B13})$$

□

Remark 3. The statement of Theorem 1 is given in the maximum norm. In practice, we use the relative L^2 norm to measure the approximation error, since the ERI tensor involves integrated quantities. In our numerical experiments, we observe that an upsampling factor of 1.5 per dimension, rather than 2, is sufficient to achieve the desired L^2 accuracy for the pair densities.