UC Berkeley

Title

Even Faster Exact Exchange for Solids via Tensor Hypercontraction

Authors

Rettig, Adam

Lee, Joonho

Head-Gordon, Martin

Peer reviewed

Even Faster Exact Exchange for Solids via Tensor Hypercontraction

Adam Rettig,[†,‡] Joonho Lee,[∗,¶] and Martin Head-Gordon[∗,†,‡]

†*Department of Chemistry, University of California, Berkeley, California 94720, USA*

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

¶*Department of Chemistry, Columbia University, New York, New York 10027, USA*

E-mail: joonholee@g.harvard.edu; mhg@cchem.berkeley.edu

Abstract

Hybrid density functional theory (DFT) remains intractable for large periodic systems due to the demanding computational cost of exact exchange. We apply the tensor hypercontraction (THC) (or interpolative separable density fitting) approximation to periodic hybrid DFT calculations with Gaussian-type orbitals. This is done to lower the computational scaling with respect to the number of basis functions (N), and $\mathbf{k}$-points (N_k). Additionally, we propose an algorithm to fit only occupied orbital products via THC (i.e. a set of points, N_{ISDF}) to further reduce computation time and memory usage. This algorithm has linear scaling cost with $\mathbf{k}$-points, no explicit dependence of N_{ISDF} on basis set size, and overall cubic scaling with unit cell size. Significant speedups and reduced memory usage may be obtained for moderately sized systems, with additional gains for large systems. Adequate accuracy can be obtained

using THC-oo-K for self-consistent calculations. We perform illustrative hybrid density function theory calculations on the benzene crystal in the basis set and thermodynamic limits to highlight the utility of this algorithm.

1   Introduction

Density functional theory (DFT) has dominated the field of quantum chemistry for the last several decades due to its combination of satisfactory accuracy and relatively low $\mathcal{O}(N^3 - N^5)$ computational cost, where N is the number of basis functions. Within density functional theory, hybrid functionals, which include Hartree-Fock "exact exchange" (shown in Eq. (1)),

$$K_{\mu\nu} = \sum_{\lambda\sigma} (\mu\lambda|\sigma\nu) P_{\lambda\sigma} \tag{1}$$

have emerged as the most accurate and widely used functionals in nearly all applications for molecules,[1–4] with only a mild increase in the computational cost compared to local density functionals. Unsurprisingly, hybrid density functional theory accounts for a large portion of quantum chemical studies performed on molecules today.

Hybrid functionals are known to offer large improvements in accuracy for some systems under periodic boundary conditions (PBC) as well.[5–9] Despite this, local (or pure) density functionals account for the majority of PBC studies due to their computational efficiency. Periodic calculations are often significantly more expensive than typical molecular studies because one needs to reach the thermodynamic limit (TDL). By virtue of translational symmetry, one may reach TDL results by numerically integrating over the first Brillouin zone (i,e, by sampling $\mathbf{k}$-points). This approach leads to lower scaling than equivalent supercell approaches. In local DFT, different $\mathbf{k}$ points can be treated independently, leading to an extra factor in the computational scaling equal to the number of $\mathbf{k}$ points, N_k. Hybrid density functional theory incurs an extra penalty as the different $\mathbf{k}$ points are coupled through the

periodic exact exchange, shown in Eq. (2).

$$K_{\mu\nu}^k = \sum_{\lambda\sigma k'} (\phi_\mu^k \phi_\lambda^{k'} | \phi_\sigma^{k'} \phi_\nu^k) P_{\lambda\sigma}^{k'}. \tag{2}$$

This leads to an extra computational scaling factor of N_k^2, giving an overall computational scaling for periodic exact exchange of $\mathcal{O}(N_k^2 N^3)$ compared to the $\mathcal{O}(N_k N^3)$ of local functionals. **k**-point calculations yield a massive reduction in computational cost for both local and hybrid functionals compared to equivalent supercell calculations that would scale as $\mathcal{O}(N_k^3 N^3)$ in this notation. Despite the computational speedup of **k**-point calculations, hybrid functionals remain considerably more expensive than local functionals for periodic systems with an additional overhead of N_k. Furthermore, hybrid functionals converge slower to the TDL than local functionals, necessitating even larger k-point calculations, further adding to the computational cost.

Efficient exact exchange computation for molecular systems has been the subject of considerable study over the last few decades; many of these ideas have also been transferred to periodic systems. For sufficiently large systems, the number of nonzero elements in the two-electron integral tensor increases only quadratically with the system size, motivating sparsity-aware approaches to lower the scaling of exact exchange to $\mathcal{O}(N)$ for gapped systems.[10–14] In practice, this asymptotic region is hard to reach, limiting the usefulness of such approaches. Approximate factorization of the two-electron integral tensors into lower rank tensors has proved more successful, leading to the pseudospectral method[15] (and the closely related chain of spheres method[16,17]), Cholesky decomposition,[18,19] the resolution of the identity (RI) or density fitting approximation,[20–22] and tensor hypercontraction (THC).[23,24] Of these, RI, has emerged as the most popular in both exact exchange and correlated wavefunction theory due to its large speedup and its relatively robust, low error of $\sim 50\mu E_H$ per atom (which also often cancels in energy differences).

Many periodic codes utilize plane waves as a basis set in contrast to the Gaussian-type orbitals (GTO) used in the molecular case, complicating the direct application of these algorithms to periodic calculations. The adaptively compressed exchange (ACE) algorithm, similar to the occ-RI-K algorithm developed for molecular systems, has become the most common way to speed up exact exchange evaluation in planewave-based periodic quantum chemistry codes.[25] More recently, the interpolative separable density fitting (ISDF)[26–28] approach to THC has been shown to offer considerable speedup for both gamma point and $\mathbf{k}$-point plane wave calculations.[29–32]

We use GTOs for periodic calculations via the Bloch orbital framework, allowing for a more direct application of molecular-based exact exchange algorithms. Our crystalline GTO basis functions are given by:

$$\phi_\mu^{\mathbf{k}}(\mathbf{r}) = \frac{1}{\sqrt{N_k}} \sum_{\mathbf{R}} e^{i\mathbf{k}\cdot\mathbf{R}} \tilde{\phi}_\mu(\mathbf{r} - \mathbf{R}) = \frac{1}{\sqrt{N_k}} e^{i\mathbf{k}\cdot\mathbf{r}} u_\mu^{\mathbf{k}}(\mathbf{r}) \tag{3}$$

where $\mathbf{R}$ are direct lattice vectors, $\mathbf{k}$ is the crystalline momentum, and $\tilde{\phi}_\mu(\mathbf{r})$ is an atomic orbital. $u_\mu^{\mathbf{k}}(\mathbf{r})$ is the cell periodic part of the Bloch orbital, which is periodic along the lattice (i.e., $u_\mu^{\mathbf{k}}(\mathbf{r}) = u_\mu^{\mathbf{k}}(\mathbf{r} + \mathbf{R})$), and is given by:

$$u_\mu^{\mathbf{k}}(\mathbf{r}) = \frac{1}{\sqrt{N_k}} \sum_{\mathbf{R}} e^{i\mathbf{k}\cdot(\mathbf{R}-\mathbf{r})} \tilde{\phi}_\mu(\mathbf{r} - \mathbf{R}) \tag{4}$$

Recently, our group showed the extension of the molecular occ-RI-K algorithm[33] to GTO-based PBC calculations leads to significant speedups for large basis sets, up to two orders of magnitude for the systems studied.[34] ISDF approaches have also been developed for both GTO and numerical atomic orbital (NAO) based Γ-point PBC calculations.[35,36]

In this study, we propose a THC algorithm for periodic exact exchange with $\mathbf{k}$-point sampling in a GTO basis utilizing ISDF. Similar to the analogous plane wave implementa-

tion,[30] the use of ISDF can reduce the computational scaling of exact exchange with respect to the number of $\mathbf{k}$-points. We will additionally propose a new ISDF approach by fitting only products of occupied orbitals, as is done in ACE-ISDF in plane waves to realize further computational savings. We illustrate the scaling improvements of these algorithms and perform an illustrative hybrid DFT study of the benzene crystal cohesive energy.

2 Theory

2.1 Review of Molecular and Γ-point ISDF

In molecular and Γ-point periodic ISDF, the products of atomic orbitals are approximated by a sum of interpolation vectors $\{\xi_P(r)\}$ weighted by the orbitals evaluated at a set of N_{ISDF} interpolation points $\{r_P\}$:

$$\phi_\mu(\mathbf{r})^*\phi_\nu(\mathbf{r}) \approx \sum_P^{N_{\mathrm{ISDF}}} \phi_\mu(\mathbf{r}_P)^*\phi_\nu(\mathbf{r}_P)\xi_P^{[nn]}(\mathbf{r}) \tag{5}$$

In the above, we have labeled the interpolation vectors with the $[nn]$ superscript to denote that these functions are fit to products of two atomic orbitals.

The set of interpolation points $\{\mathbf{r}_P\}$ can be selected in numerous ways; commonly, either a QR decomposition with column pivoting[29] or a centroidal Voronoi tessellation (CVT) k-means algorithm[28] is used. The number of interpolation points is set via a parameter $c_{\mathrm{ISDF}}^{[nn]}$:

$$N_{\mathrm{ISDF}}^{[nn]} = c_{\mathrm{ISDF}}^{[nn]}N. \tag{6}$$

After the interpolation points are chosen, the interpolation vectors $\xi_P^{[nn]}(r)$ are determined

via a least squares fit[24,37] to the orbital products desired:

$$\sum_Q \mathbf{S}_{PQ}^{[nn]} \xi_Q^{[nn]}(\mathbf{r}) = \sum_{\mu\nu} \phi_\mu(\mathbf{r})^* \phi_\nu(\mathbf{r}) \phi_\nu(\mathbf{r}_P)^* \phi_\mu(\mathbf{r}_P) \tag{7}$$

In the above, $\mathbf{S}$ represents the ISDF metric matrix (not to be confused with the AO overlap matrix):

$$S_{PQ}^{[nn]} = \sum_{\mu\nu} \phi_\mu(\mathbf{r}_P)^* \phi_\mu(\mathbf{r}_Q)^* \phi_\nu(\mathbf{r}_P)^* \phi_\nu(\mathbf{r}_Q) \tag{8}$$

The exchange matrix may then be computed using only the basis functions evaluated at a set of interpolation points and the two-electron integrals between interpolation vectors, M_{PQ}, of which there are only N_{ISDF}^2.

$$M_{PQ} = \int \int d\mathbf{r}_1 d\mathbf{r}_2 \frac{\xi_P(\mathbf{r}_1) \xi_Q(\mathbf{r}_2)}{r_{12}} \tag{9}$$

M_{PQ}, given in Eq. (9) may be computed in reciprocal space according to the GPW algorithm. The cost of computing M_{PQ} is $\mathcal{O}(N_g N_{\text{ISDF}}^2)$ due to the contraction of the two interpolation vectors in reciprocal space.

M_{PQ} can be very expensive to compute for large systems, but needs only be computed once at the start of a calculation. It may then be used each iteration to compute $K_{\mu\nu}$ according to Eq. (10).

$$K_{\mu\nu} = \sum_{\lambda\sigma} \phi_\mu(\mathbf{r}_P)^* \phi_\lambda(\mathbf{r}_P) M_{PQ} \phi_\sigma(\mathbf{r}_Q)^* \phi_\nu(\mathbf{r}_Q) P_{\lambda\sigma} \tag{10}$$

This is done most efficiently by first contracting $\phi_\lambda(\mathbf{r}_P)\phi_\sigma(\mathbf{r}_Q)^* P_{\lambda\sigma}$ and taking a Hadamard (element-wise) product with M_{PQ}. This intermediate is then contracted with $\phi_\mu(\mathbf{r}_P)^*$ and then $\phi_\nu(\mathbf{r}_Q)$, which is the bottleneck at $O(N N_{\text{ISDF}}^2)$ cost. As $N << N_g$, the per iteration cost is significantly less than the initial formation of M_{PQ}. Sharma et al. showed that large

speedups were possible with this algorithm for periodic Γ-point calculations.[35]

2.2   Periodic k-point ISDF

We first examine the extension of the Γ-point algorithm presented above to $\mathbf{k}$-point GTO-based calculations. We do this by fitting the ISDF interpolation vectors to only the cell periodic part of our basis functions, $u_\mu^{\mathbf{k}}(r)$:[30]

$$u_\mu^{\mathbf{k}}(\mathbf{r})^* u_\nu^{\mathbf{k}'}(\mathbf{r}) \approx \sum_P u_\mu^{\mathbf{k}}(\mathbf{r}_P)^* u_\nu^{\mathbf{k}'}(\mathbf{r}_P) \xi_P^{[nn]}(\mathbf{r}) \tag{11}$$

The full product of basis functions may then be approximated by adding in the Bloch phase factors:

$$\phi_\mu^{\mathbf{k}}(\mathbf{r})^* \phi_\nu^{\mathbf{k}'}(\mathbf{r}) \approx \sum_P u_\mu^{\mathbf{k}}(\mathbf{r}_P)^* u_\nu^{\mathbf{k}'}(\mathbf{r}_P) e^{i\mathbf{q}\cdot\mathbf{r}} \xi_P^{[nn]}(\mathbf{r}) \tag{12}$$

In the above we have defined $\mathbf{q} = \mathbf{k}' - \mathbf{k}$. Note that our interpolation vectors do not depend on $\mathbf{k}$-points.

The interpolation functions are once again determined via a least squares fit, but we now must include a sum over $\mathbf{k}$-points:

$$\sum_Q \mathbf{S}_{PQ}^{[nn]} \xi_Q^{[nn]}(\mathbf{r}) = \sum_{\mu\nu\mathbf{k}\mathbf{k}'} u_\mu^{\mathbf{k}}(\mathbf{r})^* u_\nu^{\mathbf{k}'}(\mathbf{r}) u_\nu^{\mathbf{k}'}(\mathbf{r}_P)^* u_\mu^{\mathbf{k}}(\mathbf{r}_P) \tag{13}$$

The $\mathbf{S}$ matrix will similarly include a sum over $\mathbf{k}$ points. The $\mathbf{k}$-point $\mathbf{K}$ matrix may then be computed as:

$$K_{\mu\nu}^{\mathbf{k}} = \sum_{\lambda\sigma\mathbf{q}PQ} u_\mu^{\mathbf{k}}(\mathbf{r}_P)^* u_\lambda^{\mathbf{k}+\mathbf{q}}(\mathbf{r}_P) M_{PQ}^{\mathbf{q}} u_\sigma^{\mathbf{k}+\mathbf{q}}(\mathbf{r}_Q)^* u_\nu^{\mathbf{k}}(\mathbf{r}_Q) P_{\lambda\sigma}^{\mathbf{k}+\mathbf{q}} \tag{14}$$

We have folded the Bloch phase factors included in our basis functions into our definition of

7

$\mathbf{M}$, so that it now has a $\mathbf{q}$-dependence:

$$M_{PQ}^{\mathbf{q}} = \int \int d\mathbf{r}d\mathbf{r}' \frac{e^{i\mathbf{q}\cdot\mathbf{r}_1}\xi_P(\mathbf{r}_1)^* e^{-i\mathbf{q}\cdot\mathbf{r}_2}\xi_Q(\mathbf{r}_2)}{r_{12}} \tag{15}$$

However, $\mathbf{M}$ depends only on $\mathbf{q}$, rather than both $\mathbf{k}$ and $\mathbf{k}'$. The number of unique two-electron integrals is, therefore, linear in the number of $\mathbf{k}$-points. $\mathbf{M}$ can be computed via GPW in $\mathrm{O}(N_k N_{\mathrm{ISDF}}^2 N_g)$ time. However, this intermediate may be computed once, and stored for reuse in each iteration.

Naively, evaluating Eq. (14) would lead to $\mathrm{O}(N_k^2 N N_{\mathrm{ISDF}}^2)$ scaling. However, we write $\mathbf{k}'$ as $\mathbf{k} + \mathbf{q}$ to emphasize that the sum over $\mathbf{q}$ done in the contraction of $M_{PQ}^{\mathbf{q}}$ with $u_\lambda^{\mathbf{k}+\mathbf{q}}(\mathbf{r}_P)u_\sigma^{\mathbf{k}+\mathbf{q}}(\mathbf{r}_Q)^* P_{\lambda\sigma}^{\mathbf{k}+\mathbf{q}}$ is a convolution in $\mathbf{k}$ space. Convolutions may be done utilizing FFT in $\mathrm{O}(N_k \log(N_k))$ time. Using this trick, we can lower the computational scaling of this algorithm, which we term THC-AO-K, to $\mathrm{O}(N_k \log(N_k)N N_{\mathrm{ISDF}}^2)$ per iteration. The rate-limiting step of this algorithm is therefore computing $\mathbf{M}$ due to the large value of N_g. In this way, significant savings can be realized for large $\mathbf{k}$-point calculations - the exchange may be computed only cubically in system size and linearly in $N_{\mathbf{k}}$. A summary of the THC-AO-K algorithm is shown in Algorithm 1.

While this algorithm is overall cubic scaling, a significant prefactor is associated with computing the $\mathbf{M}$ matrix. This prefactor was so significant in molecular cases that the THC algorithm does not become competitive with higher scaling approaches unless extremely large systems are studied.[27] In this study, we will analyze the timing of this algorithm to see if the added benefit of reduced $\mathbf{k}$-point scaling will lead to a crossover at useful system sizes. Additionally, one must be concerned with memory usage, as storing the $\mathbf{M}$ matrix requires $\mathrm{O}(N_{\mathrm{ISDF}}^2 N_k)$ space; this can quickly become impractical to store for large systems and large basis sets.

Algorithm 1 THC-AO-K algorithm

$\{\mathbf{r}_P\} \leftarrow$ CVT k-means
Compute $\xi_P^{[nn]}(\mathbf{r})$
$\xi_P^{[nn]}(\mathbf{G}) \leftarrow \text{FFT}(\xi_P^{[nn]}(\mathbf{r}))$
$M_{PQ}^{\mathbf{q}} \leftarrow \sum_{\mathbf{G}} \xi_P^{[nn]}(\mathbf{G}) V^{\mathbf{q}}(\mathbf{G}) \xi_Q^{[nn]}(\mathbf{G})$
while SCF unconverged **do**
    $X_{PQ}^{\mathbf{k}} \leftarrow \sum_{i\mathbf{q}} u_i^{\mathbf{k+q}}(\mathbf{r}_P) M_{PQ}^{\mathbf{q}} u_i^{\mathbf{k+q}}(\mathbf{r}_Q)^*$ (FFT Convolve)
    $K_{\mu\nu}^{\mathbf{k}} \leftarrow \sum_{PQ} u_\mu^{\mathbf{k}}(\mathbf{r}_P)^* X_{PQ}^{\mathbf{k}} u_\nu^{\mathbf{k}}(\mathbf{r}_Q)$
end while

2.3 THC-oo-K

Further reduction in compute time and storage requirements can be realized by fitting only products of occupied orbitals via ISDF as seen in Eq. (16):

$$u_i^{\mathbf{k}}(\mathbf{r})^* u_j^{\mathbf{k}'}(\mathbf{r}) \approx \sum_P u_i^{\mathbf{k}}(\mathbf{r}_P)^* u_j^{\mathbf{k}'}(\mathbf{r}_P) \xi_P^{[oo]}(\mathbf{r}) \tag{16}$$

Accordingly, we label the interpolation vectors with the [oo] superscript to indicate that these functions are fit to products of two occupied orbitals. This approach is similar to what is done with plane-wave ISDF codes using the ACE procedure in an algorithm we term THC-oo-K. THC-oo-K is motivated by the fact that the exchange *energy* depends only on occupied orbitals. Of course, the orbital gradient contributions due to exact exchange depend on the virtual orbitals, but they can be computed by directly differentiating the THC-oo-K energy function. In THC-oo-K, the interpolation vectors fit far fewer quantities - which scale only with the number of occupied orbitals (i.e., independent of basis size!). We therefore define the number of ISDF points similarly to Eq. 6 via:

$$N_{\text{ISDF}}^{[oo]} = c_{\text{ISDF}}^{[oo]} N_{\text{occ}}, \tag{17}$$

9

with the expectation that $N_{\text{ISDF}}^{[oo]} << N_{\text{ISDF}}^{[nn]}$ (where nn refer to atomic orbitals). The ISDF exchange energy is then computed as before via:

$$E_X = -\sum_{ij\mathbf{k}\mathbf{q}PQ} u_i^{\mathbf{k}}(\mathbf{r}_P)^* u_j^{\mathbf{k}+\mathbf{q}}(\mathbf{r}_P) M_{PQ}^{\mathbf{q}} u_j^{\mathbf{k}+\mathbf{q}}(\mathbf{r}_Q)^* u_i^{\mathbf{k}}(\mathbf{r}_Q) \tag{18}$$

While the potential memory and computational savings arising from the decrease in N_{ISDF} seem very promising at first glance, there are several complications that arise from this approach. First, the occupied orbitals change each iteration, necessitating the recalculation of $\xi_P(r)$ and therefore $M_{PQ}^{\mathbf{q}}$ on each iteration. Second, the orbital gradient of the exchange energy is no longer as simple as the THC-AO-K case seen in Eq. (10) due to the dependence of M on the occupied orbitals. The first point shows that there will be a tradeoff in compute time - the computation of M is much cheaper due to the decrease in N_{ISDF}. However, one must compute M each iteration compared to only once for THC-AO-K. We will subsequently investigate whether this tradeoff is favorable for THC-oo-K in normal use cases.

The second point can be addressed by deriving the analytical expression for the vo block of the exchange matrix, given by:

$$
\begin{aligned}
K_{ai}^{\mathbf{k}} &= \sum_{j\mathbf{k}_2 PQ} u_a^{\mathbf{k}}(\mathbf{r}_P)^* u_j^{\mathbf{k}_2}(\mathbf{r}_P) M_{PQ}^{\mathbf{q}} u_j^{\mathbf{k}_2}(\mathbf{r}_Q)^* u_i^{\mathbf{k}}(\mathbf{r}_Q) \\
&+ \sum_{j\mathbf{k}_2\mathbf{q}PQR\mathbf{r}} W_{PQ}^{\mathbf{q}} V_P^{\mathbf{q}}(\mathbf{r}) S_{QR}^{-1} \left[u_a^{\mathbf{k}}(\mathbf{r}_R)^* u_i^{\mathbf{k}}(\mathbf{r}) u_j^{\mathbf{k}_2}(\mathbf{r})^* u_j^{\mathbf{k}_2}(\mathbf{r}_R) + u_a^{\mathbf{k}}(\mathbf{r})^* u_i^{\mathbf{k}}(\mathbf{r}_R) u_j^{\mathbf{k}_2}(\mathbf{r}_R)^* u_j^{\mathbf{k}_2}(\mathbf{r}) \right] \\
&- \sum_{j\mathbf{k}_2\mathbf{q}PQST} W_{PQ}^{\mathbf{q}} S_{QS}^{-1} M_{PT}^{\mathbf{q}} \left[u_a^{\mathbf{k}}(\mathbf{r}_S)^* u_i^{\mathbf{k}}(\mathbf{r}_T) u_j^{\mathbf{k}_2}(\mathbf{r}_S)^* u_j^{\mathbf{k}_2}(\mathbf{r}_T) + u_a^{\mathbf{k}}(\mathbf{r}_T)^* u_i^{\mathbf{k}}(\mathbf{r}_S) u_j^{\mathbf{k}_2}(\mathbf{r}_T)^* u_j^{\mathbf{k}_2}(\mathbf{r}_S) \right]
\end{aligned}
\tag{19}
$$

In the above equation, $W_{PQ}^{\mathbf{q}} = \sum_{ij\mathbf{k}} u_i^{\mathbf{k}}(\mathbf{r}_P)^* u_j^{\mathbf{k}+\mathbf{q}}(\mathbf{r}_P) u_j^{\mathbf{k}+\mathbf{q}}(\mathbf{r}_Q)^* u_i^{\mathbf{k}}(\mathbf{r}_Q)$ where the convolution over $\mathbf{k}$-points may be done using FFT once again. The dependence of $\mathbf{M}$ on the occupied orbitals leads to additional two terms in the $\mathbf{K}$ matrix, but they are relatively easy to compute as most pieces are needed during the computation of the first term anyway. The

effect of these extra terms is an additional prefactor of ~ 3 in the computation time in our experience. These terms may be contracted in a similar manner to THC-AO-K, leading to an overall scaling of $O(N_k N_{\mathrm{ISDF}}^2 N_g)$ once again. A summary of the THC-oo-K algorithm is given in Algorithm 2.

In this study, we will investigate the accuracy of both THC-AO-K and THC-oo-K as a function of N_{ISDF} and analyze their relative computation times across a variety of system sizes.

Algorithm 2 THC-oo-K algorithm

$\{\mathbf{r}_P\} \leftarrow$ CVT k-means
while SCF unconverged **do**
    Compute $\xi_P^{[oo]}(\mathbf{r})$
    $\xi_P^{[oo]}(\mathbf{G}) \leftarrow \mathrm{FFT}(\xi_P^{[oo]}(\mathbf{r}))$
    $M_{PQ}^{\mathbf{q}} \leftarrow \sum_{\mathbf{G}} \xi_P^{[oo]}(\mathbf{G}) V^{\mathbf{q}}(\mathbf{G}) \xi_{\mathbf{Q}}^{[oo]}(\mathbf{G})$
    $X_{PQ}^{\mathbf{k}} \leftarrow \sum_{iq} u_i^{\mathbf{k}+\mathbf{q}}(\mathbf{r}_P) M_{PQ}^{\mathbf{q}} u_i^{\mathbf{k}+\mathbf{q}}(\mathbf{r}_Q)^*$ (FFT Convolve)
    Compute $\frac{\partial M}{\partial \theta_{ia}^{\mathbf{k}}}$
    $K_{ij}^{\mathbf{k}} \leftarrow \sum_{PQ} u_i^{\mathbf{k}}(\mathbf{r}_P)^* X_{PQ}^{\mathbf{k}} u_j^{\mathbf{k}}(\mathbf{r}_Q)$
    $K_{ia}^{\mathbf{k}} \leftarrow \sum_{PQ} u_i^{\mathbf{k}}(\mathbf{r}_P)^* X_{PQ}^{\mathbf{k}} u_a^{\mathbf{k}}(\mathbf{r}_Q) + \frac{\partial M}{\partial \theta_{ia}^{\mathbf{k}}}$
end while

---

3 Computational Details

Calculations were performed using the SZV-GTH, DZVP-GTH, TZV2P-GTH, QZV2P-GTH, unc-def2-QZVP[38] basis sets designed for use in periodic systems with the Goedecker, Teter, Hutter (GTH) pseudopotentials. The GTH-PBE pseudopotential[39,40] was used for all DFT calculations, while the GTH-HF pseudopotential was used for all HF calculations.

All calculations were performed using the Gaussian Plane Wave (GPW) algorithm[41,42] for computing 2-electron integrals. A sufficient kinetic energy cutoff for the auxiliary basis

was used to fully converge all computations below 10^{-6} Hartrees. To treat the divergence of the exact exchange, we use a Madelung correction.[43]

Calculations for the benzene system were performed with a counterpoise correction including basis functions for a (2,2,2) supercell. Cohesive energies for varying **k**-point mesh sizes and extrapolated to the thermodynamic limit via:

$$E_{\text{TDL}} = \frac{E_{N_{k_2}} N_{k_1}^{-1} - E_{N_{k_1}} N_{k_2}^{-1}}{N_{k_1}^{-1} - N_{k_2}^{-1}} \qquad (20)$$

All calculations were performed in the Q-Chem software package.[34,44]

4 Results and Discussion

4.1 THC Accuracy

We first investigate the accuracy of both THC algorithms to determine acceptable values of c_{ISDF}. In the molecular case, it was found that $c_{\text{ISDF}}^{[nn]} = 20 - 40$ was sufficient to reproduce the RI-level of accuracy of 50 μE_H per atom for several datasets including thermochemistry and noncovalent interactions.[27] Such a large $c_{\text{ISDF}}^{[nn]}$ was needed to avoid non-variational collapse to spurious solutions (found with smaller $c_{\text{ISDF}}^{[nn]}$). While we did not observe any non-variational collapse in PBC calculations, we found that a similar value is needed for $c_{\text{ISDF}}^{[nn]}$ as we will see. The interpolation vectors must now represent orbital pairs at multiple **k**-points however so we expect this value to increase slightly as more **k**-points are used, although it should eventually plateau.[30] Additionally, we seek to determine an appropriate value for $c_{\text{ISDF}}^{[oo]}$, so that the efficiency of THC-AO-K and THC-oo-K can be meaningfully compared.

We first investigate the accuracy of the THC energy computed for both algorithms. In Figure 1a we show the absolute error per atom in the THC energy computed for diamond (8 electrons total) with a (2,2,2) **k**-point mesh given the converged SCF density for several

different basis sets. We see that THC-AO-K can accurately represent the energy using $c_{\text{ISDF}}^{[nn]} = 25$ for all bases other than SZV-GTH. SZV-GTH is the *smallest* basis and from Eq. 6, we see that the same $c_{\text{ISDF}}^{[nn]}$ parameter gives rise to smaller numbers of ISDF points for smaller basis sets. The smallest useful $c_{\text{ISDF}}^{[nn]} = 25$ value will therefore depend somewhat on basis set size. Importantly, we see that the THC-AO-K energy is systematically improvable, so one may increase the number of points used if a more accurate energy is desired. We do observe a some fluctuations in the error as $c_{\text{ISDF}}^{[nn]}$ is increased - the error is not strictly monotonically decreasing; this is likely due to the randomness inherent to the CVT algorithm used to select ISDF points.

In Figure 1b, we show the root mean square of the orbital gradient for the same system. As we are using non-THC converged orbitals, we expect this value to be zero if there were no ISDF error. We see that at $c_{\text{ISDF}}^{[nn]} = 25$, the gradient still has a magnitude of $\sim 10^{-3}$ for some basis sets, suggesting that the THC-AO-K orbitals are slightly different than the converged non-THC algorithms, due to the ISDF factorization error. This is also systematically improvable through the $c_{\text{ISDF}}^{[nn]}$ parameter and will eventually converge to the exact solution once enough points are used. Consistent with the discussion of the energies, we see that typically (though not always due to ISDF point selection as already mentioned) a given value of $c_{\text{ISDF}}^{[nn]}$ leads to larger RMS gradients for the smaller basis sets.

We show similar plots for THC-oo-K in figure 1c and 1d. We see that a larger value of $c_{\text{ISDF}}^{[oo]} = 50$ is needed to reach an acceptable error for this algorithm. It should, however, be pointed out that the number of ISDF points for this algorithm is determined by scaling N_{occ} (Eq. 17), and this leads to significantly fewer points than THC-AO-K despite the higher c_{ISDF}. Accordingly, this means that the number of ISDF points used in THC-oo-K for each of the basis sets tested is the same! THC-oo-K is, therefore, highly effective at compactly representing the HF exchange energy for large basis sets. On the other hand, looking at the gradient in figure 1d, we see that the THC-oo-K orbital gradient is significantly
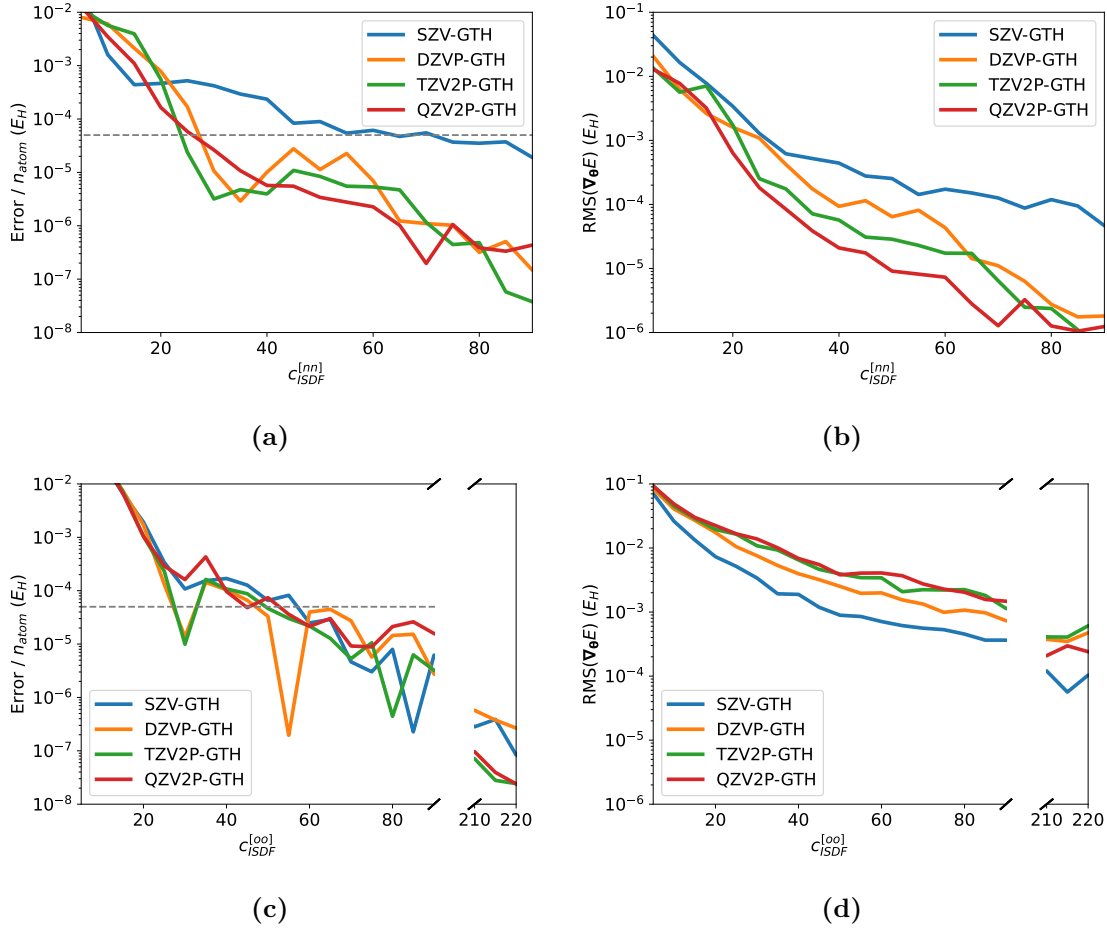
Figure 1: Error in the THC-AO-K HF energy (a), THC-AO-K orbital gradient (b), THC-oo-K HF energy (c), THC-oo-K orbital gradient (d), for diamond with a (2,2,2) **k**-point mesh, given the converged density computed without the ISDF quadrature error. The numbers of ISDF interpolation points used for THC-AO-K and THC-oo-K are related to the parameters $c_{\text{ISDF}}^{[nn]}$ and $c_{\text{ISDF}}^{[oo]}$ via Eqs. 6 and 17 respectively.

larger than that of THC-AO-K and the convergence with increasing $c_{\text{ISDF}}^{[oo]}$ seems to be quite slow. Moreover, there seems to be a degree of basis set dependence as a larger basis set has many more degrees of freedom; the SZV-GTH basis has a much smaller error in the gradient, DZVP-GTH, TZV2P-GTH, and QZV2P-GTH all have roughly similar errors. This inaccuracy in the gradient could be due to fluctuations in the THC error over the entire energy surface and could potentially lead to SCF convergence issues. We will next investigate self-consistent energies obtained via THC-K algorithms to address this.

14

We next investigated the error in the HF energy for the converged density - i.e. we optimize orbitals using the THC algorithms. In figure 2a, we show the error in the HF self-consistent total energy using THC-AO-K as a function of $c_{\text{ISDF}}^{[nn]}$ for diamond in the QZV2P-GTH and SZV-GTH bases, as well as AlN (16 electrons total) in the GTH-QZV2P basis, all with a (2,2,2) **k**-point mesh. We see that our suggested value of $c_{\text{ISDF}}^{[nn]} = 25$ holds up pretty well for the QZV2P-GTH systems, although diamond in the SZV-GTH basis does have a noticeably higher error in this case. A similar plot for THC-oo-K can be seen in figure 2b. We see that despite possible concerns arising from significant deviations in the orbital gradient when evaluated with exact converged density, the THC-oo-K algorithm also yields adequately accurate self-consistent solutions for these systems. Most encouragingly, the $c_{\text{ISDF}}^{[oo]}$ parameter is evidently transferable across different systems and basis sets.
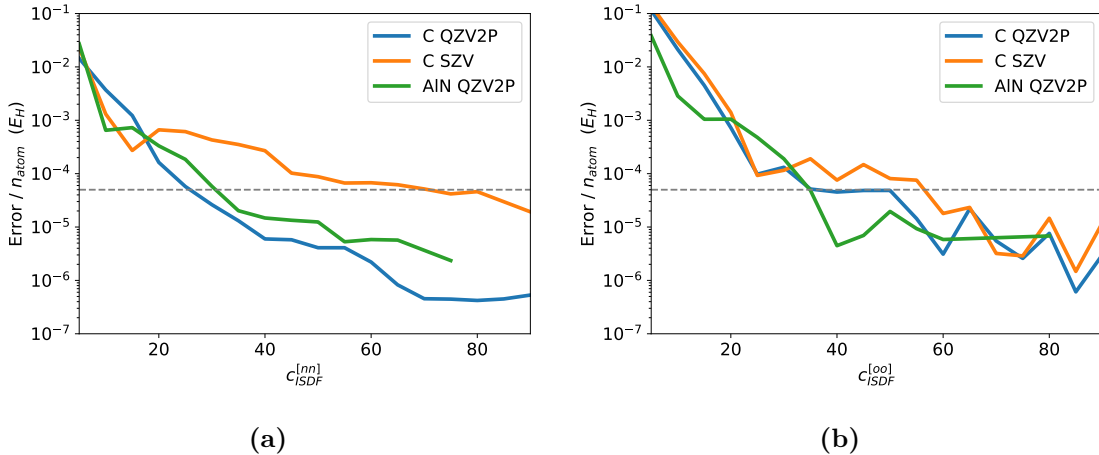

(a) (b)

Figure 2: Absolute Error per atom in the unit cell in the self-consistent HF energy using (a) THC-AO-K vs c_{ISDF}^{nn} (see Eq. 6) and (b) THC-oo-K vs. c_{ISDF}^{oo} (see Eq. 17) for both diamond and AlN using varying basis set sizes and a (2,2,2) **k**-point mesh. The dashed line represents 50 μE_H / atom accuracy.

We next investigate the accuracy as the **k**-point mesh is increased. Our algorithms rely on the fact that the number of ISDF points required to yield an accurate solution will plateau quickly with the **k**-mesh size. In Figure 3, we show a plot of the absolute error in the self-consistent HF energy of diamond in the QZV2P-GTH basis computed with THC-oo-K for

several different **k**-point meshes. We see that there is a very steep increase in the number of ISDF points required when going from gamma point to a (2,2,2) **k**-point mesh; evidently, it is much more difficult to represent products of occupied orbitals at multiple **k** points than at a single **k** point. However, increasing the mesh further has very little effect. It therefore appears that the $c_{\text{ISDF}}^{[oo]} = 50$ value is adequate for all **k**-point mesh sizes, although a significantly lower value may likely be sufficient for Γ-point calculations.
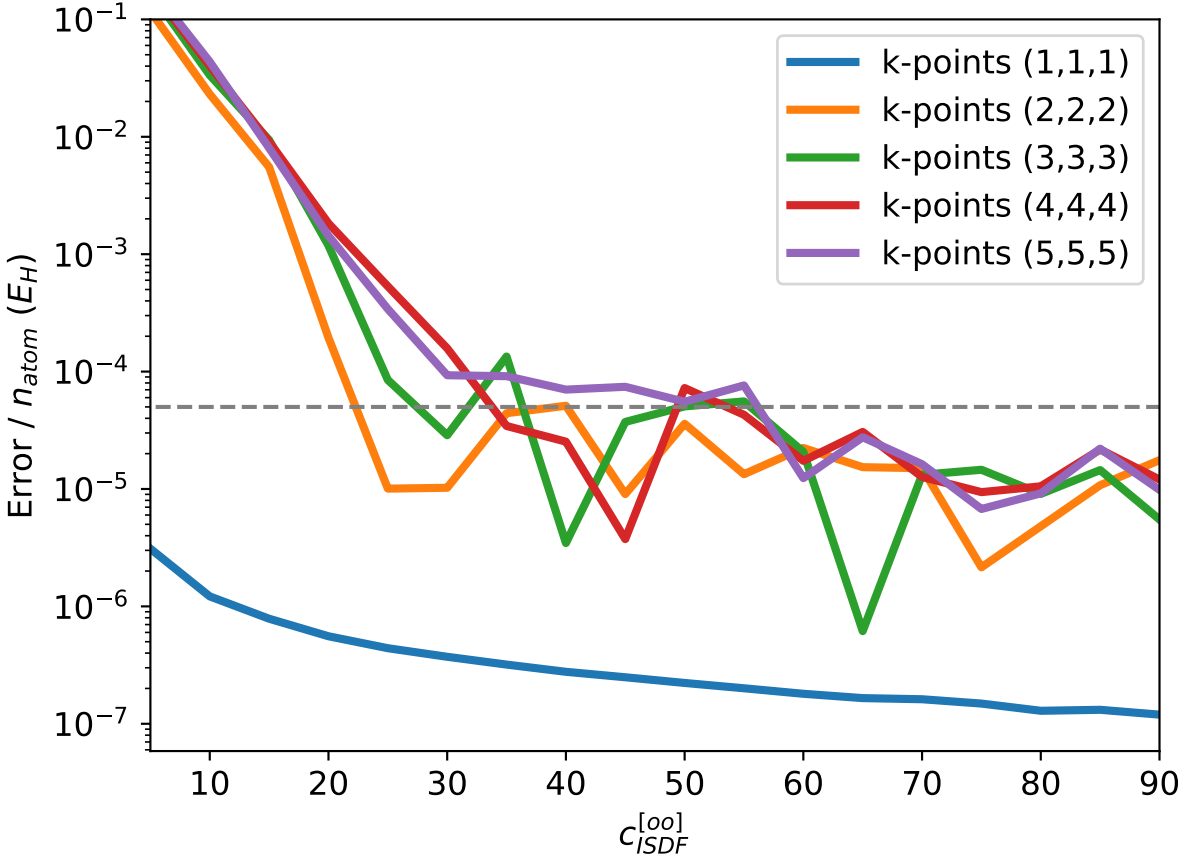

Figure 3: Absolute Error in the HF energy using THC-oo-K, as a function of c_{ISDF}^{oo} for diamond in the QZV2P-GTH basis with several **k**-point meshes.

We therefore tentatively recommend starting with values $c_{\text{ISDF}}^{[nn]} = 25$, $c_{\text{ISDF}}^{[oo]} = 50$. One may verify this is adequate on a case-by-case basis by increasing the parameters and ensuring little variation in overall results. We note that this study analyzed the absolute energies,

16

which leaves out the effect of error cancellation in observables such as the lattice energy or band gaps. Error cancellation could lead to more leniency in the c_{ISDF} parameters, necessitating fewer interpolation points. Additionally, we showed errors in the HF energy here, which includes 100% exact exchange. When using a hybrid density functional with a small fraction of exact exchange or short-range exact exchange, the THC errors will therefore be further scaled down. Indeed, this was observed in our molecular study.[27] We therefore believe that our recommended coefficients are conservative.

4.2 Computational scaling

We found that both periodic THC-AO-K and THC-oo-K require more interpolation points than their molecular counterparts when using **k**-point sampling. However, even molecular THC was seen to require too many interpolation points to be computationally competitive except for very large systems. We therefore investigate the K build time for both THC-AO-K and THC-oo-K to see if ISDF-based algorithms are useful for small to medium-sized periodic systems. In the following calculations, we use $c_{\text{ISDF}}^{[nn]} = 25$ and $c_{\text{ISDF}}^{[oo]} = 55$. Additionally, we will compute the average THC-AO-K iteration time by assuming 10 iterations, as the majority of the computational cost of this algorithm is the M build step which is performed only once, i.e. we add one-tenth of the M build time into the per iteration cost. If more than 10 iterations are required (i.e., the initial guess was poor), THC-AO-K will be more favorable than shown here, and vice versa.

In Fig. 4a, we plot the K build time for AlN in the QZV2P-GTH basis as a function of the number of **k**-points for both THC-AO-K and THC-oo-K. We see two important conclusions from this plot. First, as anticipated, the THC-K algorithms are linear with the number of **k**-points due to the FFT convolution. By contrast, occ-RI-K exhibits quadratic scaling with number of **k**-points. Second, the THC-K algorithms have a large prefactor which causes occ-RI-K to be the more efficient approach for small **k**-point meshes. The linear **k**-point

17

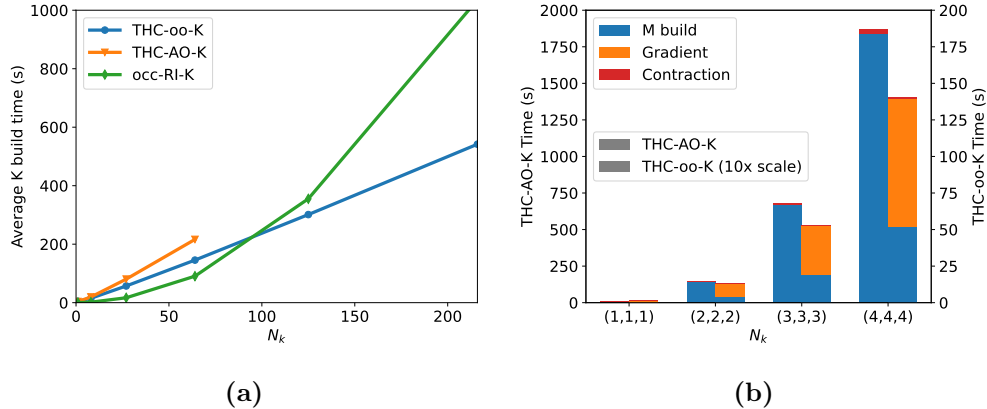**Figure 4:** (a) Average K build time per cycle and (b) breakdown of the THC initial cycle timings vs the number of **k**-points for AlN in the QZV2P-GTH basis.

scaling makes THC-oo-K extremely effective at reducing compute time for large **k**-point calculations, and the algorithm becomes more efficient than occ-RI-K between a (4,4,4) and (5,5,5) **k**-point mesh. Due to the large basis set and therefore large $N_{\text{ISDF}}^{[nn]}$, the THC-AO-K algorithm is significantly slower than both other algorithms, and could not be run for **k**-point meshes larger than (4,4,4) due to memory requirements.

We present a breakdown of the THC-AO-K and THC-oo-K first cycle computation time (the full THC-AO-K **M** build time is included) for the AlN system in Fig. 4b. We see that the THC-AO-K time is dominated by the initial single **M** build, making the per iteration cost to contract of **M** to form **K** according to Eq. (14) negligible. The THC-oo-K **M** is significantly cheaper (note the 10x scale difference) to form, but the computation of the orbital gradient adds an extra factor of 2-3 to the per-iteration compute cost.

We additionally investigated the computational scaling of THC-K with basis set size. We previously saw the occ-RI-K algorithm was largely independent of basis set size, making it very effective for large basis calculations (with small numbers of **k**-points). We expect THC-oo-K to have this same property, as N_{ISDF} is independent of the basis size, shown by fig. 2b. We, therefore, expect THC-oo-K to become more efficient than THC-AO-K as the basis set size increases. In figure 5, we show the average K matrix build time as a function of
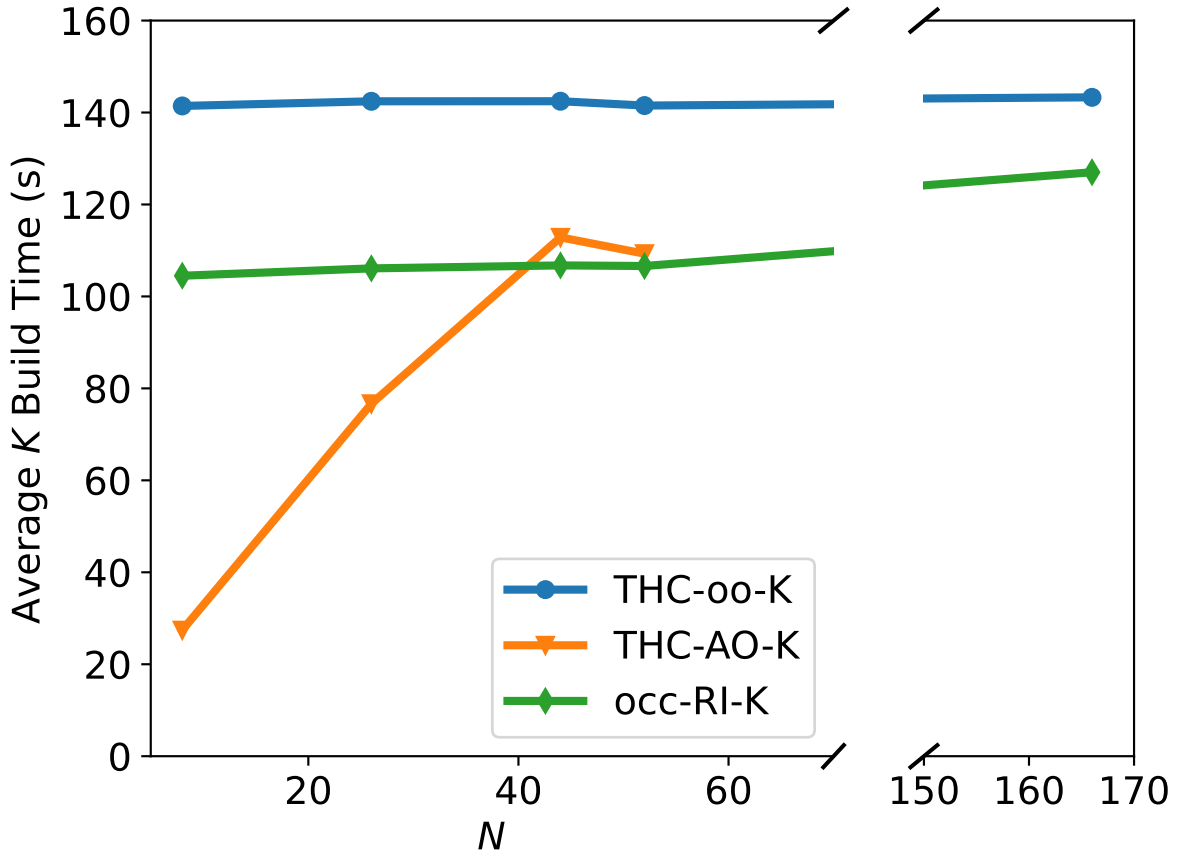
18

Figure 5: Average K build time per cycle vs the number of basis functions for diamond with a (7,7,7) **k**-point mesh.

the basis set size for the diamond system with a (7,7,7) **k**-mesh. THC-AO-K is more efficient than THC-oo-K and occ-RI-K for the smallest two basis sets used, SZV-GTH and DZVP-GTH; the bottleneck of this algorithm is precomputing the M matrix, which is fairly small for smaller basis set sizes. occ-RI-K is slightly better than THC-AO-K for TZV2P-GTH and QZV2P-GTH as it is independent of basis size. THC-oo-K evidently does not become competitive with THC-AO-K until the very large unc-def2-QZVP-GTH basis set is used.

Finally, we note that the THC algorithms can offer significantly reduced memory usage compared to occ-RI-K. For both THC algorithms, we store only $\xi_P(\mathbf{r})$ and $M_{PQ}^{\mathbf{q}}$, i.e. no quantities that scale with both the grid and number of **k**-points. This is in contrast to
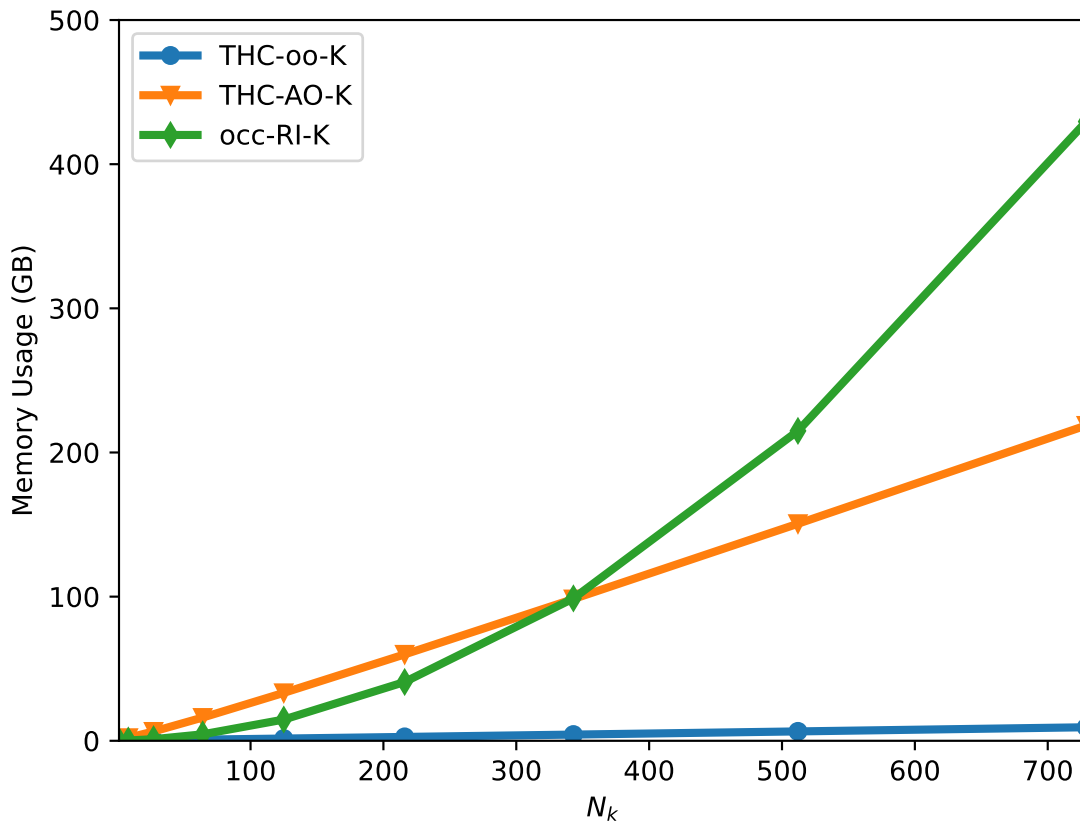
Figure 6: Memory usage for the occ-RI-K, THC-AO-K, and THC-oo-K algorithms as a function of the number of **k**-points for the diamond system using the QZV2P-GTH basis.

the occ-RI-K algorithm which stores the basis functions (linear in N_k) and coulomb kernel (quadratic in N_k) on the grid by our default in Q-Chem. Using an integral direct approach to occ-RI-K will eliminate this memory requirement[34] but results in much slower performance, making it less competitive with the THC algorithms. We highlight this in figure 6 by plotting the memory requirements for each algorithm for the diamond system in the QZV2P-GTH basis as a function of the number of **k**-points. We see that the large number of ISDF points required for THC-AO-K leads to larger memory consumption than occ-RI-K for **k**-point meshes smaller than (6,6,6), however, the linear scaling in N_k eventually wins out at **k**-point meshes larger than this. THC-oo-K offers a huge reduction in memory usage over THC-AO-K due to a large reduction in the number of ISDF points. For this diamond QZV2P-GTH example, with a (9,9,9) **k**-point mesh, THC-oo-K requires only 9 GB compared to 430 GB for occ-RI-K!

We can now define specific use cases for each algorithm studied here. For small **k**-point meshes, we recommend using the occ-RI-K algorithm as it will be faster than all THC algorithms. For large **k**-point meshes we recommend using THC-AO-K if using a small basis set and THC-oo-K for large basis sets. Across all system sizes and **k**-point mesh sizes, if the suggested algorithm requires too much memory, we recommend using THC-oo-K instead. It is particularly well-suited for implementation on memory-constrained computing devices, such as graphical processing units.

4.3 Molecular Crystals

Finally, we perform an illustrative computation of the cohesive energy of the benzene crystal as a rigorous test of the limits of THC-K. Accurate theoretical estimates of this value are available via high-order wavefunction theory calculations performed via fragment-based approaches.[45–47] Periodic calculations have thus far been fairly limited as HF and pure density functionals are known to fail for binding purposes[48] and higher order wavefunction theo-

ries are very expensive for periodic studies of benzene. Local density functionals have been shown to yield reasonable results when dispersion corrections are utilized,[49] and recent studies using periodic MP2 have been performed.[50,51] However, these periodic calculations are generally limited to small **k**-point meshes or small basis sets, or both due to computational constraints. Only a few hybrid functionals have been tested for benzene[49,52–55] due to the computational difficulty of obtaining exact exchange energies in the thermodynamic limit. We, therefore, decided to utilize the THC-oo-K algorithm to benchmark hybrid functional for cohesive energies in order to compare with theoretical best estimates.

We investigated the dispersion-corrected pure functional B97M-rV[56,57] and the hybrid functionals PBE0-D3,[58,59] MN15,[60] M06-2X,[61] SCAN0-D3,[62] ωB97X-rV,[63] and ωB97M-rV.[64] We attempted to converge our results to both the complete basis set (CBS) limit and the TDL, including counterpoise corrections (see Sec. 3). We cover hybrid functionals utilizing both the D3 and rVV10 dispersion corrections and empirical functionals parameterized to include dispersion. We note that MN15 and M06-2X are not dispersion-corrected but their performance on molecular data sets of non-covalent interaction was found to be quite accurate.[1] The cohesive energies computed for varying sizes of **k**-point meshes can be seen in table 1. We performed our calculations using the QZV2P-GTH basis.

Compared to the theoretical best estimate (TBE) of -54.58 kJ/mol,[45] we see a great variance in the performance of DFT. As expected, the dispersion treatment is paramount to obtaining accurate, cohesive energies. The DFT-D3 treatment of dispersion yields quite accurate results, with the global hybrid functionals PBE0-D3 and SCAN0-D3 performing well, giving errors of 3.4 and 2.0 kJ/mol, respectively. MN15 and M06-2X, parameterized to implicitly treat dispersion, yield extremely different results; MN15 is the most accurate of the functionals tested with an error of 1.1 kJ/mol while M06-2X performs substantially worse than all other functionals, giving an error of 13 kJ/mol. Use of the rVV10 dispersion correction yields unsatisfactory results as well; the pure functional B97M-rV overbinds by 9.0

kJ/mol while the range separated hybrids ωB97M-rV and ωB97X-rV, which are among the most accurate functionals for molecular systems,[1] perform only slightly better with errors of 7.2 and 8.0 kJ/mol respectively. The consistent overbinding of these three VV10-based functionals suggests that rVV10 is the main source of their errors.

The trends seen in our data are in line with previous studies on dispersion corrections for benzene and the X23 dataset. DFT-D3 is known to perform quite well,[49] while the failure of rVV10 for molecular crystals has previously been attributed to the lack of screening (i.e., many-body dispersion) effects.[65–67] Notably, by comparing B97M-rV to ωB97M-rV, we find that the addition of exact exchange makes a minor improvement of 1 kJ/mol.

We see that with a (3,3,3) **k**-point mesh, pure functionals are essentially converged, whereas the hybrid functionals are still up to 5 kJ/mol off from the TDL. The convergence of the hybrid functionals is demonstrated in figure 7, where the convergence rate is dependent on the fraction of exact exchange in the functional - highlighting the difficulty of periodic hybrid DFT. Nonetheless, the size extrapolation using $1/N_k$ works well for this system, making reliable hybrid functional calculations accessible.

Table 1: Error in the benzene crystal cohesive energy in kJ/mol for several density functionals as a function of the **k**-point mesh size.

N_k	B97M-rV	HF	PBE0-D3	MN15	M06-2X	SCAN0-D3	ωB97X-rV	ωB97M-rV
1^3	-10.97	-55.87	-37.18	-62.59	-59.70	-32.38	-139.54	-141.48
2^3	-8.29	58.48	-7.44	-8.81	4.09	-2.33	-23.70	-24.74
3^3	-8.98		-4.59	-3.76	10.33	0.54	-12.42	-13.40
TDL	-8.98	74.81	-3.38	-1.63	12.95	1.75	-7.68	-8.62

We see that for the benzene crystal, including exact exchange can lead to modest improvements in accuracy, but the treatment of dispersion is much more important. More in-depth analysis is required to evaluate general trends of hybrid functionals for periodic applications. Still, the discrepancy between molecular[1] and periodic cases suggests that further functional development may be beneficial for hybrids applied to periodic applications.
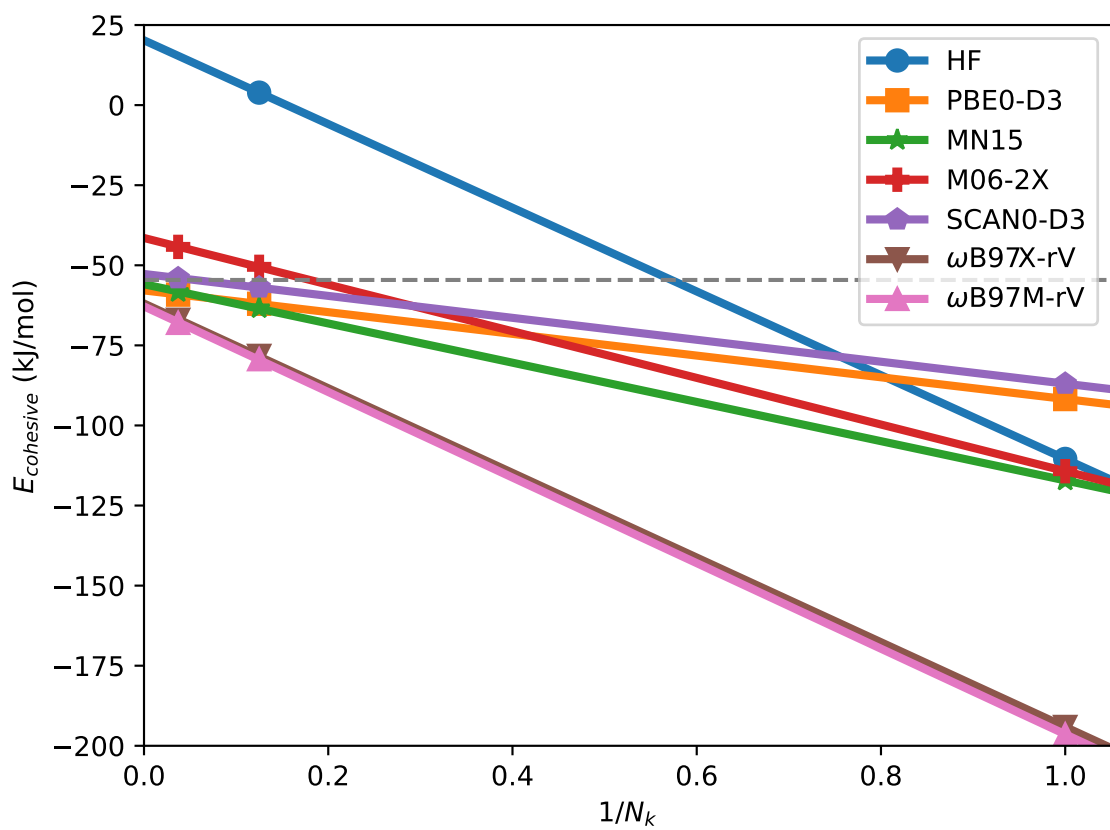
Figure 7: Cohesive energy of the benzene crystal as a function of **k**-point sampling for several hybrid functionals. The dashed line indicates the reference value.

5   Conclusions

To summarize, we have presented an extension of the ISDF approximation for exact exchange for use with periodic **k**-point calculations using GTOs. We additionally presented a new algorithm, THC-oo-K, by fitting only the product of occupied orbitals via the ISDF approach, which is all that is needed for the energy. While this means that errors in THC-oo-K orbital gradients (which involve occupied-virtual products) are larger than for the energy, the effect on the self-consistently optimized energy is still small. We have shown that these algorithms reduce the computational scaling for exact exchange to cubic in system size and linear with the number of **k**-points; the THC-oo-K algorithm has the additional advantage of THC dimension scaling independently of basis set size. Initial investigation showed that these algorithms provide substantial computational savings for large **k**-point meshes; THC-oo-K additionally provides computational savings over occ-RI-K at even medium-sized **k**-point meshes and huge reductions in memory usage in all cases, at only a minor compromise in overall accuracy. We believe this cost reduction will make periodic studies using hybrid functionals more feasible, which we illustrated via a study of the benzene lattice energy, computed with the QZV2P-GTH basis up to a (3,3,3) **k**-point mesh.

Further reduction in error may be obtained by fitting only one side of the two-electron integral tensor via ISDF, termed robust fit THC.[35,68] This reduces the error in the energy to quadratic in the THC fit error rather than linear. We performed preliminary studies of this approach for the THC-AO-K algorithm. Still, we found that while the robust fit did reduce error substantially, it significantly increased memory requirements to an intractable degree due to the requirement of storing $V_P^q(r)$, the potential due to interpolating function $\xi_P(r)$ which increases the memory usage of the already expensive THC-AO-K by a factor of N_k; additionally, the number of expensive FFTs for the GPW algorithm is increased by a factor of N_k, significantly increasing the compute time. Nonetheless, a multi-node

MPI implementation may increase the available memory enough to make this approach tractable.[35] For the THC-oo-K algorithm, there is no need to store $V_P^q(r)$ each iteration, so it is possible to batch over this quantity and avoid storing the entire thing; the increase in the number of FFTs will however still lead to a significant increase in compute cost. A robust fit THC-oo-K algorithm is therefore an interesting direction for future development, but the trade-off in compute time may turn out to be unfavorable.

All work presented here was used with the GPW algorithm, which only applies to pseudopotential calculations. A big advantage of Gaussian orbitals in PBC calculations compared to the more popular plane waves basis is the ability to model core orbitals. It would be interesting to investigate all-electron versions of these algorithms. This extension should be possible using the projector augmented-wave method (PAW),[69] density fitting,[70,71] or similar approaches, by selecting ISDF points from Becke grids as is done for molecular codes.[27]

Finally, we note that this approach offers a general framework for approximating the two-electron integrals and its usefulness is not limited to exact exchange. The same intermediates may be used to speed up the computation of matrix elements necessary for Møller-Plesset theory, coupled cluster theory, and other correlated wavefunction methods. This has already proved successful for molecular ISDF.[27,72,73] The extension of this algorithm to correlated wavefunction theory would allow for much larger **k**-point calculations.

References

(1) 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.

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

(3) Dohm, S.; Hansen, A.; Steinmetz, M.; Grimme, S.; Checinskii, M. P. Comprehensive thermochemical benchmark set of realistic closed-shell metal organic reactions. *J. Chem. Theory Comput.* **2018**, *14*, 2596–2608.

(4) Chan, B.; Gill, P. M. W.; Kimura, M. Assessment of DFT methods for transition metals with the TMC151 compilation of data sets and comparison with accuracies for main-group chemistry. *J. Chem. Theory Comput.* **2019**, *15*, 3610–3622.

(5) Muscat, J.; Wander, A.; Harrison, N. On the prediction of band gaps from hybrid functional theory. *Chemical Physics Letters* **2001**, *342*, 397–401.

(6) Batista, E. R.; Heyd, J.; Hennig, R. G.; Uberuaga, B. P.; Martin, R. L.; Scuseria, G. E.; Umrigar, C.; Wilkins, J. W. Comparison of screened hybrid density functional theory to diffusion Monte Carlo in calculations of total energies of silicon phases and defects. *Physical Review B* **2006**, *74*, 121102.

(7) Deák, P.; Aradi, B.; Frauenheim, T.; Janzén, E.; Gali, A. Accurate defect levels obtained from the HSE06 range-separated hybrid functional. *Physical Review B* **2010**, *81*, 153203.

(8) Komsa, H.-P.; Pasquarello, A. Assessing the accuracy of hybrid functionals in the determination of defect levels: Application to the As antisite in GaAs. *Physical Review B* **2011**, *84*, 075207.

(9) Henderson, T. M.; Paier, J.; Scuseria, G. E. Accurate treatment of solids with the HSE screened hybrid. *physica status solidi (b)* **2011**, *248*, 767–774.

(10) Schwegler, E.; Challacombe, M. Linear scaling computation of the Hartree–Fock exchange matrix. *The Journal of chemical physics* **1996**, *105*, 2726–2734.

(11) Schwegler, E.; Challacombe, M.; Head-Gordon, M. Linear scaling computation of the Fock matrix. II. Rigorous bounds on exchange integrals and incremental Fock build. *The Journal of chemical physics* **1997**, *106*, 9708–9717.

(12) Burant, J. C.; Scuseria, G. E.; Frisch, M. J. A linear scaling method for Hartree–Fock exchange calculations of large molecules. *The Journal of chemical physics* **1996**, *105*, 8969–8972.

(13) Challacombe, M.; Schwegler, E. Linear scaling computation of the Fock matrix. *The Journal of chemical physics* **1997**, *106*, 5526–5536.

(14) Ochsenfeld, C.; White, C. A.; Head-Gordon, M. Linear and sublinear scaling formation of Hartree–Fock-type exchange matrices. *The Journal of chemical physics* **1998**, *109*, 1663–1669.

(15) Friesner, R. A. Solution of self-consistent field electronic structure equations by a pseudospectral method. *Chemical physics letters* **1985**, *116*, 39–43.

(16) Neese, F.; Wennmohs, F.; Hansen, A.; Becker, U. Efficient, approximate and parallel Hartree–Fock and hybrid DFT calculations. A 'chain-of-spheres' algorithm for the Hartree–Fock exchange. *Chemical Physics* **2009**, *356*, 98–109.

(17) Izsák, R.; Neese, F. An overlap fitted chain of spheres exchange method. *The Journal of chemical physics* **2011**, *135*, 144105.

(18) Beebe, N. H.; Linderberg, J. Simplifications in the generation and transformation of two-electron integrals in molecular calculations. *International Journal of Quantum Chemistry* **1977**, *12*, 683–705.

(19) Røeggen, I.; Wisløff-Nilssen, E. On the Beebe-Linderberg two-electron integral approximation. *Chemical physics letters* **1986**, *132*, 154–160.

(20) Baerends, E.; Ellis, D.; Ros, P. Self-consistent molecular Hartree—Fock—Slater calculations I. The computational procedure. *Chemical Physics* **1973**, *2*, 41–51.

(21) Whitten, J. L. Coulombic potential energy integrals and approximations. *The Journal of Chemical Physics* **1973**, *58*, 4496–4501.

(22) Jafri, J.; Whitten, J. Electron repulsion integral approximations and error bounds: Molecular applications. *The Journal of Chemical Physics* **1974**, *61*, 2116–2121.

(23) Hohenstein, E. G.; Parrish, R. M.; Martínez, T. J. Tensor hypercontraction density fitting. I. Quartic scaling second-and third-order Møller-Plesset perturbation theory. *The Journal of chemical physics* **2012**, *137*, 044103.

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

(25) Lin, L. Adaptively compressed exchange operator. *Journal of chemical theory and computation* **2016**, *12*, 2242–2249.

(26) Shenvi, N.; Van Aggelen, H.; Yang, Y.; Yang, W. Tensor hypercontracted ppRPA: Reducing the cost of the particle-particle random phase approximation from O (r 6) to O (r 4). *The Journal of Chemical Physics* **2014**, *141*, 024119.

(27) 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. *Journal of chemical theory and computation* **2019**, *16*, 243–263.

(28) Dong, K.; Hu, W.; Lin, L. Interpolative separable density fitting through centroidal voronoi tessellation with applications to hybrid functional electronic structure calculations. *Journal of Chemical Theory and Computation* **2018**, *14*, 1311–1320.

(29) Hu, W.; Lin, L.; Yang, C. Interpolative separable density fitting decomposition for accelerating hybrid density functional calculations with applications to defects in silicon. *Journal of Chemical Theory and Computation* **2017**, *13*, 5420–5431.

(30) Wu, K.; Qin, X.; Hu, W.; Yang, J. Low-rank approximations accelerated plane-wave hybrid functional calculations with k-point sampling. *Journal of Chemical Theory and Computation* **2021**, *18*, 206–218.

(31) Li, J.; Qin, X.; Wan, L.; Jiao, S.; Hu, W.; Yang, J. Complex-Valued K-means Clustering for Interpolative Separable Density Fitting to Large-Scale Hybrid Functional Ab Initio Molecular Dynamics with Plane-Wave Basis Sets. *arXiv preprint arXiv:2208.07731* **2022**,

(32) Qin, X.; Hu, W.; Yang, J. Interpolative Separable Density Fitting for Accelerating Two-Electron Integrals: A Theoretical Perspective. *Journal of Chemical Theory and Computation* **2023**,

(33) Manzer, S. F.; Horn, P. R.; Mardirossian, N.; Head-Gordon, M. Fast, accurate evaluation of exact exchange: The occ-RI-K algorithm. *J. Chem. Phys.* **2015**, *143*, 024113.

(34) Lee, J.; Rettig, A.; Feng, X.; Epifanovsky, E.; Head-Gordon, M. Faster Exact Exchange for Solids via occ-RI-K: Application to Combinatorially Optimized Range-Separated Hybrid Functionals for Simple Solids Near the Basis Set Limit. *arXiv preprint arXiv:2207.09028* **2022**,

(35) Sharma, S.; White, A. F.; Beylkin, G. Fast exchange with Gaussian basis set using robust pseudospectral method. *arXiv preprint arXiv:2207.04636* **2022**,

(36) Qin, X.; Liu, J.; Hu, W.; Yang, J. Interpolative separable density fitting decomposition for accelerating Hartree–Fock exchange calculations within numerical atomic orbitals. *The Journal of Physical Chemistry A* **2020**, *124*, 5664–5674.

(37) Matthews, D. A. Improved grid optimization and fitting in least squares tensor hyper-contraction. *Journal of chemical theory and computation* **2020**, *16*, 1382–1385.

(38) Lee, J.; Feng, X.; Cunha, L. A.; Gonthier, J. F.; Epifanovsky, E.; Head-Gordon, M. Approaching the basis set limit in Gaussian-orbital-based periodic calculations with transferability: Performance of pure density functionals for simple semiconductors. *J. Chem. Phys.* **2021**, *155*, 164102.

(39) Goedecker, S.; Teter, M.; Hutter, J. Separable dual-space Gaussian pseudopotentials. *Physical Review B* **1996**, *54*, 1703.

(40) Krack, M. Pseudopotentials for H to Kr optimized for gradient-corrected exchange-correlation functionals. *Theoretical Chemistry Accounts* **2005**, *114*, 145–152.

(41) Lippert, B. G.; PARRINELLO, J. H.; MICHELE, A hybrid Gaussian and plane wave density functional scheme. *Molecular Physics* **1997**, *92*, 477–488.

(42) VandeVondele, J.; Krack, M.; Mohamed, F.; Parrinello, M.; Chassaing, T.; Hutter, J. Quickstep: Fast and accurate density functional calculations using a mixed Gaussian and plane waves approach. *Computer Physics Communications* **2005**, *167*, 103–128.

(43) Fraser, L. M.; Foulkes, W.; Rajagopal, G.; Needs, R.; Kenny, S.; Williamson, A. Finite-size effects and Coulomb interactions in quantum Monte Carlo calculations for homogeneous systems with periodic boundary conditions. *Physical Review B* **1996**, *53*, 1814.

(44) Epifanovsky, E.; Gilbert, A. T.; Feng, X.; Lee, J.; Mao, Y.; Mardirossian, N.; Pokhilko, P.; White, A. F.; Coons, M. P.; Dempwolff, A. L., et al. Software for the frontiers of quantum chemistry: An overview of developments in the Q-Chem 5 package. *J. Chem. Phys.* **2021**, *155*, 084801.

(45) Yang, J.; Hu, W.; Usvyat, D.; Matthews, D.; Schütz, M.; Chan, G. K.-L. Ab initio determination of the crystalline benzene lattice energy to sub-kilojoule/mole accuracy. *Science* **2014**, *345*, 640–643.

(46) Ringer, A. L.; Sherrill, C. D. First principles computation of lattice energies of organic solids: The benzene crystal. *Chemistry–A European Journal* **2008**, *14*, 2542–2547.

(47) Kennedy, M. R.; McDonald, A. R.; DePrince III, A. E.; Marshall, M. S.; Podeszwa, R.; Sherrill, C. D. Communication: Resolving the three-body contribution to the lattice energy of crystalline benzene: Benchmark results from coupled-cluster theory. *The Journal of chemical physics* **2014**, *140*, 121104.

(48) Civalleri, B.; Doll, K.; Zicovich-Wilson, C. Ab initio investigation of structure and cohesive energy of crystalline urea. *The Journal of Physical Chemistry B* **2007**, *111*, 26–33.

(49) Moellmann, J.; Grimme, S. DFT-D3 study of some molecular crystals. *The Journal of Physical Chemistry C* **2014**, *118*, 7615–7621.

(50) Bintrim, S. J.; Berkelbach, T. C.; Ye, H.-Z. Integral-Direct Hartree–Fock and Møller–Plesset Perturbation Theory for Periodic Systems with Density Fitting: Application to the Benzene Crystal. *Journal of Chemical Theory and Computation* **2022**, *18*, 5374–5381.

(51) Neumann, M. A.; Perrin, M.-A. Energy ranking of molecular crystals using density functional theory calculations and an empirical van der Waals correction. *The Journal of Physical Chemistry B* **2005**, *109*, 15531–15541.

(52) Civalleri, B.; Zicovich-Wilson, C. M.; Valenzano, L.; Ugliengo, P. B3LYP augmented with an empirical dispersion term (B3LYP-D*) as applied to molecular crystals. *CrystEngComm* **2008**, *10*, 405–410.

(53) Loboda, O. A.; Dolgonos, G. A.; Boese, A. D. Towards hybrid density functional calculations of molecular crystals via fragment-based methods. *The Journal of Chemical Physics* **2018**, *149*, 124104.

(54) Price, A. J.; Otero-de-la Roza, A.; Johnson, E. R. XDM-corrected hybrid DFT with numerical atomic orbitals predicts molecular crystal lattice energies with unprecedented accuracy. *Chemical Science* **2023**,

(55) Cutini, M.; Civalleri, B.; Corno, M.; Orlando, R.; Brandenburg, J. G.; Maschio, L.; Ugliengo, P. Assessment of different quantum mechanical methods for the prediction of structure and cohesive energy of molecular crystals. *Journal of chemical theory and computation* **2016**, *12*, 3340–3352.

(56) Mardirossian, N.; Head-Gordon, M. Mapping the genome of meta-generalized gradient

approximation density functionals: The search for B97M-V. *The Journal of chemical physics* **2015**, *142*, 074111.

(57) Mardirossian, N.; Ruiz Pestana, L.; Womack, J. C.; Skylaris, C.-K.; Head-Gordon, T.; Head-Gordon, M. Use of the rVV10 nonlocal correlation functional in the B97M-V density functional: Defining B97M-rV and related functionals. *The journal of physical chemistry letters* **2017**, *8*, 35–40.

(58) Adamo, C.; Barone, V. Toward reliable density functional methods without adjustable parameters: The PBE0 model. *The Journal of chemical physics* **1999**, *110*, 6158–6170.

(59) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *The Journal of chemical physics* **2010**, *132*, 154104.

(60) Haoyu, S. Y.; He, X.; Li, S. L.; Truhlar, D. G. MN15: A Kohn–Sham global-hybrid exchange–correlation density functional with broad accuracy for multi-reference and single-reference systems and noncovalent interactions. *Chemical science* **2016**, *7*, 5032–5051.

(61) Zhao, Y.; Truhlar, D. G. The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: two new functionals and systematic testing of four M06-class functionals and 12 other functionals. *Theoretical chemistry accounts* **2008**, *120*, 215–241.

(62) Hui, K.; Chai, J.-D. SCAN-based hybrid and double-hybrid density functionals from models without fitted parameters. *The Journal of chemical physics* **2016**, *144*, 044114.

(63) Mardirossian, N.; Head-Gordon, M. ωB97X-V: A 10-parameter, range-separated hybrid, generalized gradient approximation density functional with nonlocal correlation, de-

signed by a survival-of-the-fittest strategy. *Physical Chemistry Chemical Physics* **2014**, *16*, 9904–9924.

(64) Mardirossian, N.; Head-Gordon, M. ω B97M-V: A combinatorially optimized, range-separated hybrid, meta-GGA density functional with VV10 nonlocal correlation. *The Journal of chemical physics* **2016**, *144*, 214110.

(65) Hermann, J.; Tkatchenko, A. Electronic exchange and correlation in van der Waals systems: Balancing semilocal and nonlocal energy contributions. *Journal of chemical theory and computation* **2018**, *14*, 1361–1369.

(66) Stöhr, M.; Van Voorhis, T.; Tkatchenko, A. Theory and practice of modeling van der Waals interactions in electronic-structure calculations. *Chemical Society Reviews* **2019**, *48*, 4118–4154.

(67) Fabiano, E.; Cortona, P. Seeking widely applicable dispersion-corrected GGA functionals: The performances of TCA+ D3 and RevTCA+ D3 on solid-state systems. *Computational Materials Science* **2023**, *216*, 111826.

(68) Pierce, K.; Rishi, V.; Valeev, E. F. Robust Approximation of Tensor Networks: Application to Grid-Free Tensor Factorization of the Coulomb Interaction. *J. Chem. Theory Comput.* **2021**, *17*, 2217–2230.

(69) Blöchl, P. E. Projector augmented-wave method. *Physical review B* **1994**, *50*, 17953.

(70) Sun, Q.; Berkelbach, T. C.; McClain, J. D.; Chan, G. K.-L. Gaussian and plane-wave mixed density fitting for periodic systems. *The Journal of chemical physics* **2017**, *147*, 164119.

(71) Ye, H.-Z.; Berkelbach, T. C. Fast periodic Gaussian density fitting by range separation. *The Journal of Chemical Physics* **2021**, *154*, 131104.

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

(73) Lu, J.; Thicke, K. Cubic scaling algorithms for RPA correlation using interpolative separable density fitting. *Journal of Computational Physics* **2017**, *351*, 187–202.