arXiv:2304.05505v1 [physics.comp-ph] 11 Apr 2023

通过张量超收缩实现更快速的固体精确交换

Adam Rettig, [†,‡] Joonho Lee, [*,¶] 和 Martin Head-Gordon[*,†,‡]

†加利福尼亚大学伯克利分校化学系，美国加利福尼亚州伯克利，*94720*

‡劳伦斯伯克利国家实验室化学科学部，美国加利福尼亚州伯克利，*94720*

¶哥伦比亚大学化学系，美国纽约州纽约，*10027*

电子邮箱：joonholee@g.harvard.edu；mhg@cchem.berkeley.edu

Abstract

由于精确交换的计算成本高昂，大型周期性体系的杂化密度泛函理论（DFT）计算仍难以实现。我们将张量超收缩（THC）（或插值可分离密度拟合）近似应用于采用高斯型轨道的周期性杂化密度泛函理论计算，以降低计算复杂度随基组函数数量（N）和$\mathbf{k}$点数量（N_k）的增长。此外，我们提出一种算法，通过 THC 仅拟合占据轨道乘积（即一组点， N_{ISDF}），以进一步减少计算时间和内存用量。该算法的计算成本随$\mathbf{k}$点数量线性增长，不显式依赖 N_{ISDF} 基组大小，且总体上随晶胞大小呈三次增长。对于中等大小的体系，该方法可显著加速计算并减少内存用量；对于大型体系，也能带来额外收益。使用 THC-oo-K 进行自洽计算时，可获得足够的精度

。我们以苯晶体为对象，在基组极限和热力学极限下进行示例性杂化密度泛函理论计算，以展示该算法的实用价值。

1 引言

过去几十年来，密度泛函理论（DFT）凭借其兼具令人满意的精度和相对较低的 $\mathcal{O}(N^3-N^5)$ 计算成本，主导了量子化学领域，其中 N 为基组函数的数量。在密度泛函理论中，杂化泛函包含 Hartree-Fock“精确交换”（见式（1）），

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

已成为几乎所有分子应用中精度最高、使用最广泛的泛函，[1–4] 与局域密度泛函相比，计算成本仅略有增加。不出所料，杂化密度泛函理论在当今针对分子开展的量子化学研究中占据了很大比重。

众所周知，杂化泛函在周期性边界条件（PBC）下也能显著提高某些体系的计算精度。[5–9] 尽管如此，由于计算效率较高，局域（或纯）密度泛函仍广泛用于 PBC 研究。周期性计算通常比典型的分子研究开销大得多，因为需要达到热力学极限（TDL）。借助平移对称性，可以通过对第一布里渊区进行数值积分（即对**k**点采样）来获得热力学极限结果。与等效的超胞方法相比，这种方法的计算复杂度更低。在局域 DFT 中，不同的**k** 点可以独立处理，因此计算复杂度会额外乘上**k** 点的数量， N_k。杂化密度泛函理论还会带来额外的开销，因为不同的**k** 点通过这些方式相互耦合。

周期性精确交换，如公式（2）所示。

$$K^k_{\mu\nu} = \sum_{\lambda\sigma k'} (\phi^k_\mu \phi^{k'}_\lambda | \phi^{k'}_\sigma \phi^k_\nu) P^{k'}_{\lambda\sigma}. \tag{2}$$

这会引入额外的计算复杂度因子 N_k^2，使得周期性精确交换的总体计算复杂度为 $\mathcal{O}(N_k^2 N^3)$，而局域泛函的计算复杂度为 $\mathcal{O}(N_k N^3)$。与等效的超胞计算相比，k 点计算可大幅降低局域泛函和杂化泛函的计算成本；按此处的记法，等效超胞计算的计算复杂度为 $\mathcal{O}(N_k^3 N^3)$。尽管 k 点计算能够加快计算速度，但对于周期性体系，杂化泛函的成本仍远高于局域泛函，额外开销为 N_k。此外，杂化泛函收敛至 TDL 的速度比局域泛函慢，因此需要进行规模更大的 k 点计算，进一步增加计算成本。

过去几十年来，分子体系的高效精确交换计算一直受到广泛研究；其中许多思路也已应用于周期性体系。对于足够大的体系，双电子积分张量中的非零元素数量仅随体系大小呈二次增长，这促使人们采用稀疏性感知方法，将有带隙体系中精确交换的标度降低至 $\mathcal{O}(N)$。[10–14] 在实践中，这一渐近区域难以达到，限制了此类方法的实用性。将双电子积分张量近似分解为低秩张量取得了更大的成功，由此发展出了伪谱方法[15]（以及与之密切相关的球链方法[16,17]）、乔列斯基分解[18,19]、恒等分解（RI）或密度拟合近似[20–22]，以及张量超收缩（THC）。[23,24]在这些方法中，RI 已成为精确交换和相关波函数理论中最常用的方法，这得益于其显著的加速效果，以及相对稳定且较低的每原子 $\sim 50\mu\mathrm{E}_H$ 误差（这种误差在能量差中通常也会相互抵消）。

许多周期性代码使用平面波作为基组，这与分子体系中使用的高斯型轨道 (GTO) 不同，因此难以将这些算法直接应用于周期性计算。自适应压缩交换 (ACE) 算法与为分子体系开发的 occ-RI-K 算法类似，已成为基于平面波的周期性量子化学代码中加速精确交换计算最常用的方法。[25] 最近，针对 THC 的插值可分离密度拟合 (ISDF)[26–28] 方法已被证明能够显著加速 Γ 点和**k**点平面波计算。[29–32]

我们通过布洛赫轨道框架使用高斯型轨道进行周期性计算，从而能够更直接地应用基于分子的精确交换算法。我们的晶体高斯型轨道基函数表示为：

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

其中 $\mathbf{R}$ 是直接晶格矢量，$\mathbf{k}$ 是晶体动量，而 $\phi_\mu(\mathbf{r})$ 是原子轨道。 $u_\mu^{\mathbf{k}}(\mathbf{r})$ 是布洛赫轨道的晶胞周期部分，沿晶格呈周期性（即 $u_\mu^{\mathbf{k}}(\mathbf{r}) = u_\mu^{\mathbf{k}}(\mathbf{r}+\mathbf{R})$），其表达式为：

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

最近，我们的团队展示了，将分子 occ-RI-K 算法[33] 扩展至基于 GTO 的周期性边界条件计算，可显著提升大基组计算的速度；对于所研究的体系，速度提升最高可达两个数量级。[34] 针对基于 GTO 和数值原子轨道（NAO）的 Γ点周期性边界条件计算，也已开发出 ISDF 方法。[35,36]

在本研究中，我们提出了一种利用 ISDF、适用于高斯型轨道基组中采用**k**点采样的周期性精确交换的 THC 算法。类似于平

面波中的相应实现，[30] 使用 ISDF 可以降低精确交换相对于**k**点数量的计算复杂度。此外，我们还将提出一种新的 ISDF方法，仅拟合占据轨道的乘积，正如平面波中的 ACE-ISDF 所做的那样，以进一步节省计算开销。我们展示了这些算法在计算复杂度方面的改进，并对苯晶体内聚能进行了示例性的杂化 DFT 研究。

2 理论

2.1 分子和 Γ点 ISDF 回顾

在分子和 Γ点周期性 ISDF 中，原子轨道的乘积可近似表示为一组插值向量 $\{\xi_P(r)\}$ 的加权和，其中权重为轨道在一组 N_{ISDF}插值点 $\{r_P\}$处的取值：

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

在上文中，我们用 $[nn]$ 上标标记插值向量，以表示这些函数拟合的是两个原子轨道的乘积。

上述插值点集合 $\{\mathbf{r}_P\}$ 可以通过多种方式选取；通常使用带列主元选取的 QR 分解[29] 或质心Voronoi镶嵌（CVT）k-均值算法[28] 。插值点的数量通过参数 $c^{[nn]}$ ISDF设定：

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

插值点选定后，插值向量 $\xi_P^{[nn]}$ (r)即可确定

通过最小二乘拟合[24,37] 来获得所需的轨道乘积:

$$\sum_{Q} \mathbf{S}_{PQ}^{[nn]} \xi_{Q}^{[nn]}(\mathbf{r}) = \sum_{\mu\nu} \phi_{\mu}(\mathbf{r})^{*} \phi_{\nu}(\mathbf{r}) \phi_{\nu}(\mathbf{r}_{P})^{*} \phi_{\mu}(\mathbf{r}_{P}) \tag{7}$$

其中，S 表示 ISDF 度量矩阵（不要将其与 AO 重叠矩阵混淆）：

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

随后，只需使用在一组插值点处计算的基组函数，以及插值向量之间的双电子积分，即可计算交换矩阵，M_{PQ}；插值向量的数量仅为 N^{2}_{ISDF}。

$$M_{PQ} = \int \int d\mathbf{r}_{1} d\mathbf{r}_{2} \frac{\xi_{P}(\mathbf{r}_{1}) \xi_{Q}(\mathbf{r}_{2})}{r_{12}} \tag{9}$$

M_{PQ}（见公式（9））可根据 GPW算法在倒易空间中计算。计算 M_{PQ} 的成本为 $\mathcal{O}$（$N_{g} N^{2}{}_{\mathrm{ISDF}}$），这是由于在倒易空间中对两个插值向量进行缩并所致。

对于大型体系，M_{PQ} 的计算成本可能非常高，但只需在计算开始时计算一次。随后，在每次迭代中，都可以根据公式（10）使用它来计算 $K_{\mu\nu}$。

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

最高效的做法是先将 $\phi_{\lambda}(\mathbf{r}_{P})\phi_{\sigma}(\mathbf{r}_{Q})^{*} P_{\lambda\sigma}$ 收缩，并与 M_{PQ}进行 Hadamard（逐元素）乘积。然后将这个中间结果与 $\phi_{\mu}(\mathbf{r}_{P})^{*}$ 收缩，再与 $\phi_{\nu}(\mathbf{r}_{Q})$收缩；这一步是瓶颈，成本为 $O(N N^{2}{}_{\mathrm{ISDF}})$。由于 $N << N_{g}$，每次迭代的成本远低于最初构造 M_{PQ}的成本。Sharma 等人表明，对于周期性 ϕ_{λ} 点计算，这种算法可以实现大幅

加速。[35]

2.2 周期性 k 点 ISDF

我们首先研究将上述提出的 Γ点算法扩展到**k**点基于高斯型轨道的计算。为此，我们仅将ISDF插值向量拟合到我们基组函数的晶胞周期部分， $u_{\mu}^{\mathbf{k}}(r)$：[30]

$$u_{\mu}^{\mathbf{k}}(\mathbf{r})^{*}u_{\nu}^{\mathbf{k}'}(\mathbf{r}) \approx \sum_{P} u_{\mu}^{\mathbf{k}}(\mathbf{r}_{P})^{*}u_{\nu}^{\mathbf{k}'}(\mathbf{r}_{P})\xi_{P}^{[nn]}(\mathbf{r}) \tag{11}$$

然后，通过加入布洛赫相位因子，即可近似表示基组函数的完整乘积:

$$\phi_{\mu}^{\mathbf{k}}(\mathbf{r})^{*}\phi_{\nu}^{\mathbf{k}'}(\mathbf{r}) \approx \sum_{P} u_{\mu}^{\mathbf{k}}(\mathbf{r}_{P})^{*}u_{\nu}^{\mathbf{k}'}(\mathbf{r}_{P})e^{i\mathbf{q}\cdot\mathbf{r}}\xi_{P}^{[nn]}(\mathbf{r}) \tag{12}$$

在上文中，我们定义了 $\mathbf{q} = \mathbf{k}' - \mathbf{k}$。请注意，我们的插值向量不依赖于 **k**-点。

插值函数再次通过最小二乘拟合确定，但现在我们必须包含对**k**点的求和:

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

S 矩阵同样包含对**k** 点的求和。随后可计算**k**点**K** 矩阵，如下:

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

我们已将基组函数中包含的布洛赫相位因子并入对

$\mathbf{M}$ 的定义中，因此它现在具有 $\mathbf{q}$ 依赖性：

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

然而，$\mathbf{M}$ 仅取决于$\mathbf{q}$，而非同时取决于$\mathbf{k}$ 和$\mathbf{k}'$。因此，唯一二电子积分的数量与$\mathbf{k}$点的数量呈线性关系。$\mathbf{M}$ 可通过 GPW 在 O($N_k N_{\mathrm{ISDF}}^2 N_g$) 时间内计算。然而，这个中间量只需计算一次，并可存储起来以供每次迭代重复使用。

直接计算式 (14) 会导致 O($N_k^2 N N_{\mathrm{ISDF}}^2$) 的计算复杂度。然而，我们将$\mathbf{k}'$写作$\mathbf{k} + \mathbf{q}$，以强调在将 $M_{\mathbf{q}}{}^{PQ}$与$u_\lambda^{\mathbf{k}+\mathbf{q}}(\mathbf{r}_P) u_\sigma^{\mathbf{k}+\mathbf{q}}(\mathbf{r}Q)^* P_{\lambda\sigma}^{\mathbf{k}+\mathbf{q}}$缩并时，对$\mathbf{q}$求和实际上是在$\mathbf{k}$空间中的卷积。利用快速傅里叶变换可在 O($N_k$log($N_k$)) 时间内完成卷积。利用这一技巧，我们可以将该算法（称为 THC-AO-K）的计算复杂度降低至每次迭代 O($N_k$log(N_k)$N N_{\mathrm{ISDF}}^2$)。因此，该算法的限速步骤是计算$\mathbf{M}$ ，因为 N_g的值很大。这样，在大型$\mathbf{k}$点计算中可显著节省计算量——交换项的计算复杂度可仅随体系大小呈立方增长，并随 $N_{\mathbf{k}}$呈线性增长。
THC-AO-K算法的概要见算法 1。

尽管该算法总体上呈立方级缩放，但计算$\mathbf{M}$ 矩阵需要较大的前因子。在分子体系中，这一前因子非常显著，以至于除非研究极大规模的体系，否则THC算法无法与缩放级别更高的方法展开竞争。[27] 在本研究中，我们将分析该算法的计时，以考察降低$\mathbf{k}$ 点缩放所带来的额外优势是否会使其在实用的体系规模下实现交叉。此外，还必须关注内存用量，因为存储 $\mathbf{M}$ 矩阵需要O($N_{\mathrm{ISDF}}^2 N_k$) 的空间；对于大型体系和大型基组而言，这很快会变得难以实际存储。

算法 1 THC-AO-K算法

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

通过如式（16）所示仅拟合占据轨道的乘积，可进一步减少计算时间和存储需求：

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

因此，我们在插值向量上标注 $[oo]$ 上标，以表示这些函数是针对两个占据轨道的乘积进行拟合的。这种方法与采用 ACE 程序的平面波 ISDF 代码中的做法类似，我们将这一算法称为 THC-oo-K。THC-oo-K 的出发点是，交换能 仅取决于占据轨道。当然，精确交换导致的轨道梯度贡献取决于虚轨道，但可以通过直接对 THC-oo-K 能量函数求导来计算。在 THC-oo-K 中，插值向量需要拟合的量少得多——其规模仅随占据轨道数增长（即与基组大小无关！）。因此，我们仿照公式 6 定义 ISDF 点的数量如下：

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

预期 $N^{[oo]}{}_{\mathrm{ISDF}} << N^{[nn]}{}_{\mathrm{ISDF}}$（其中 nn 指原子轨道）。随后，ISDF 交换能仍按前述方式计算：

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

尽管乍看之下，由于N_{ISDF} 的减少而带来的潜在内存和计算节省非常令人期待，但这种方法也会带来若干问题。首先，占据轨道在每次迭代中都会发生变化，因此需要在每次迭代中重新计算 $\xi_P(r)$，从而重新计算 $M_{\mathbf{q}}{}^{PQ}$。其次，由于 M 取决于占据轨道，交换能的轨道梯度不再像式（10）中所示的THC-AO-K情况那样简单。第一点表明，计算时间方面存在权衡——由于 N_{ISDF}的减少， M 的计算成本要低得多。然而，与THC-AO-K 只需计算一次相比，每次迭代都必须计算 M。随后，我们将研究在通常使用情况下，这种权衡对 THC-oo-K 是否有利。

第二点可以通过推导交换矩阵的 vo 块的解析表达式来解决，其形式为：

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

在上式中， $W_{\mathbf{q}}{}^{PQ} = \sum_{ij\ \mathbf{k}} u_i^{\mathbf{k}}(\mathbf{r}_P)^* u_j^{\mathbf{k}+\mathbf{q}}(\mathbf{r}_P) u_j^{\mathbf{k}+\mathbf{q}}(\mathbf{r}Q)^* u_i^{\mathbf{k}}(\mathbf{r}Q)$，其中对$\mathbf{k}$点的卷积可以再次使用快速傅里叶变换来完成。$\mathbf{M}$ 对占据轨道的依赖会在$\mathbf{K}$ 矩阵中引入另外两项，但由于在计算第一项时已经需要大部分相关部分，因此这些项相对容易计算。这些

额外项的影响是，据我们的经验，计算时间会增加一个 ~ 3 的额外因子。这些项可以采用与 THC-AO-K 类似的方式进行缩并，从而再次得到总体 O($N_k N_{\mathrm{ISDF}}^2 N_g$) 的标度。算法 2 总结了 THC-oo-K 算法。

在本研究中，我们将考察 THC-AO-K 和 THC-oo-K 的精度随 N_{ISDF} 的变化，并分析它们在各种体系规模下的相对计算时间。

算法 2 THC-oo-K 算法

$\{\mathbf{r}_P\} \leftarrow$ CVT k 均值
当 SCF 未收敛 **执行**
  计算 $\xi_P^{[oo]}(\mathbf{r})$
  $\xi_P^{[oo]}(\mathbf{G}) \leftarrow$ 快速傅里叶变换($\xi_P^{[oo]}(\mathbf{r})$)
  $M_{PQ}^{\mathbf{q}} \leftarrow \sum_{\mathbf{G}} \xi_P^{[oo]}(\mathbf{G}) V^{\mathbf{q}}(\mathbf{G}) \xi^{[oo]}{}_{\mathbf{Q}}(\mathbf{G})$
  $X_{PQ}^{\mathbf{k}} \leftarrow \sum_{iq} u_i^{\mathbf{k}+\mathbf{q}}(\mathbf{r}_P) M_{\mathbf{q}PQ} u_i^{\mathbf{k}+\mathbf{q}}(\mathbf{r}_Q)^*$(快速傅里叶变换卷积)
  计算 $\frac{\partial M}{\partial \theta_{ia}^{\mathbf{k}}}$
  $K_{ij}^{\mathbf{k}} \leftarrow \sum_{PQ} u_i^{\mathbf{k}}(\mathbf{r}_P)^* X_{PQ}^{\mathbf{k}} u_j^{\mathbf{k}}(\mathbf{r}_Q)$
  $K_{ia}^{\mathbf{k}} \leftarrow \sum_{PQ} u_i^{\mathbf{k}}(\mathbf{r}_P)^* X_{PQ}^{\mathbf{k}} u_a^{\mathbf{k}}(\mathbf{r}_Q) + \frac{\partial M}{\partial \theta_{ia}^{\mathbf{k}}}$
结束循环

3 计算细节

计算采用了 SZV-GTH、DZVP-GTH、TZV2P-GTH、QZV2P-GTH、unc-def2-QZVP[38] 基组，这些基组专为采用 Goedecker、Teter、Hutter (GTH) 赝势的周期性体系设计。所有 DFT 计算均使用 GTH-PBE 赝势[39,40]，而所有 Hartree–Fock 计算均使用 GTH-HF 赝势。

所有计算均采用高斯平面波（GPW）算法[41,42]来计算双电子积分。辅助基组使用了足够高的动能截断值,

以使下文中的所有计算完全收敛至 10^{-6} 哈特里。为处理精确交换的发散问题，我们采用马德隆修正。[43]

苯体系的计算采用反平衡校正进行，其中包括(2,2,2)超胞的基组函数。对于不同**k**点网格尺寸的内聚能，并通过以下方式外推至热力学极限：

$$E_{\mathrm{TDL}} = \frac{E_{N_{k_2}} N_{k_1}^{-1} - E_{N_{k_1}} N_{k_2}^{-1}}{N_{k_1}^{-1} - N_{k_2}^{-1}} \tag{20}$$

所有计算均在 Q-Chem 软件包中完成。[34,44]

4 结果与讨论

4.1 THC 精度

我们首先研究两种 THC 算法的精度，以确定c_{ISDF}可接受的取值。在分子体系中，研究发现，对于包括热化学和非共价相互作用在内的多个数据集，$c^{[nn]}{}_{\mathrm{ISDF}} = 20-40$足以重现 $50\ \mu\mathrm{E}_H$ 每个原子的RI 级精度。[27] 为了避免非变分塌缩到伪解（使用较小的 $c^{[nn]}{}_{\mathrm{ISDF}}$时发现），需要如此大的 $c^{[nn]}{}_{\mathrm{ISDF}}$ 。虽然我们没有在周期性边界条件计算中观察到任何非变分塌缩，但我们发现，正如后文所示，$c^{[nn]}{}_{\mathrm{ISDF}}$ 也需要采用类似的取值。然而，插值向量现在必须表示多个**k**点处的轨道对，因此我们预计，随着所使用的**k**点增加，该取值会略有增大，尽管最终应达到平台。[30] 此外，我们希望确定 $c^{[oo]}{}_{\mathrm{ISDF}}$的适当取值，从而能够对 THC-AO-K 和 THC-oo-K 的效率进行有意义的比较。

我们首先考察使用两种算法计算得到的 THC 能量的精度。图 1a 展示了在给定收敛 SCF 密度的情况下，对于金刚石（总计 8 个电子），采用 (2,2,2) **k** 点网格计算得到的 THC 能量的每原子绝对误差，涵盖了若干

不同基组。我们发现，除 SZV-GTH 外，THC-AO-K 均可使用$c^{[nn]}{}_{\mathrm{ISDF}} = 25$ 准确表示能量。SZV-GTH 是最小的基组；根据公式 6，对于较小的基组，相同的 $c^{[nn]}{}_{\mathrm{ISDF}}$ 参数会产生更少的 ISDF 点。因此，最小的有效 $c^{[nn]}{}_{\mathrm{ISDF}} = 25$ 值在一定程度上取决于基组大小。重要的是，我们发现 THC-AO-K 能量可以通过增加点数来系统地提高精度，因此若需要更准确的能量，可以增加所用的点数。我们确实观察到，随着 $c^{[nn]}{}_{\mathrm{ISDF}}$ 增大，误差会有一些波动——并非严格单调下降；这可能是由于用于选择 ISDF 点的 CVT 算法本身具有随机性。

在图 1b 中，我们展示了同一体系的轨道梯度均方根。由于我们使用的是经非 THC 算法收敛的轨道，如果不存在 ISDF 误差，我们预计该值应为零。我们发现，对于某些基组，在 $c^{[nn]}{}_{\mathrm{ISDF}} = 25$ 时，梯度的幅值仍为 $\sim 10^{-3}$，这表明由于 ISDF 分解误差，THC-AO-K 轨道与经非 THC 算法收敛的轨道略有不同。通过 $c^{[nn]}{}_{\mathrm{ISDF}}$ 参数，也可以系统地改善这一结果；随着使用的点数足够多，最终将收敛到精确解。与能量的讨论一致，我们发现，给定的 $c^{[nn]}{}_{\mathrm{ISDF}}$ 值通常会导致较小基组的 RMS 梯度更大（不过，由于前面提到的 ISDF 点选择问题，情况并非总是如此）。

我们在图 1c 和 1d 中展示了 THC-oo-K 的类似图表。可以看到，对于该算法，要达到可接受的误差，需要更大的$c^{[oo]}{}_{\mathrm{ISDF}} = 50$ 。不过需要指出的是，该算法的 ISDF 点数量由缩放关系决定N_{occ} (式 17)，因此尽管 c_{ISDF} 更高，其点数仍显著少于 THC-AO-K。因此，这意味着对于每种经过测试的基组，THC-oo-K 使用的 ISDF 点数量都相同！所以，对于大型基组，THC-oo-K 能够非常有效地紧凑表示 Hartree–Fock 交换能。另一方面，观察图 1d 中的梯度可见，THC-oo-K 轨道梯度显著

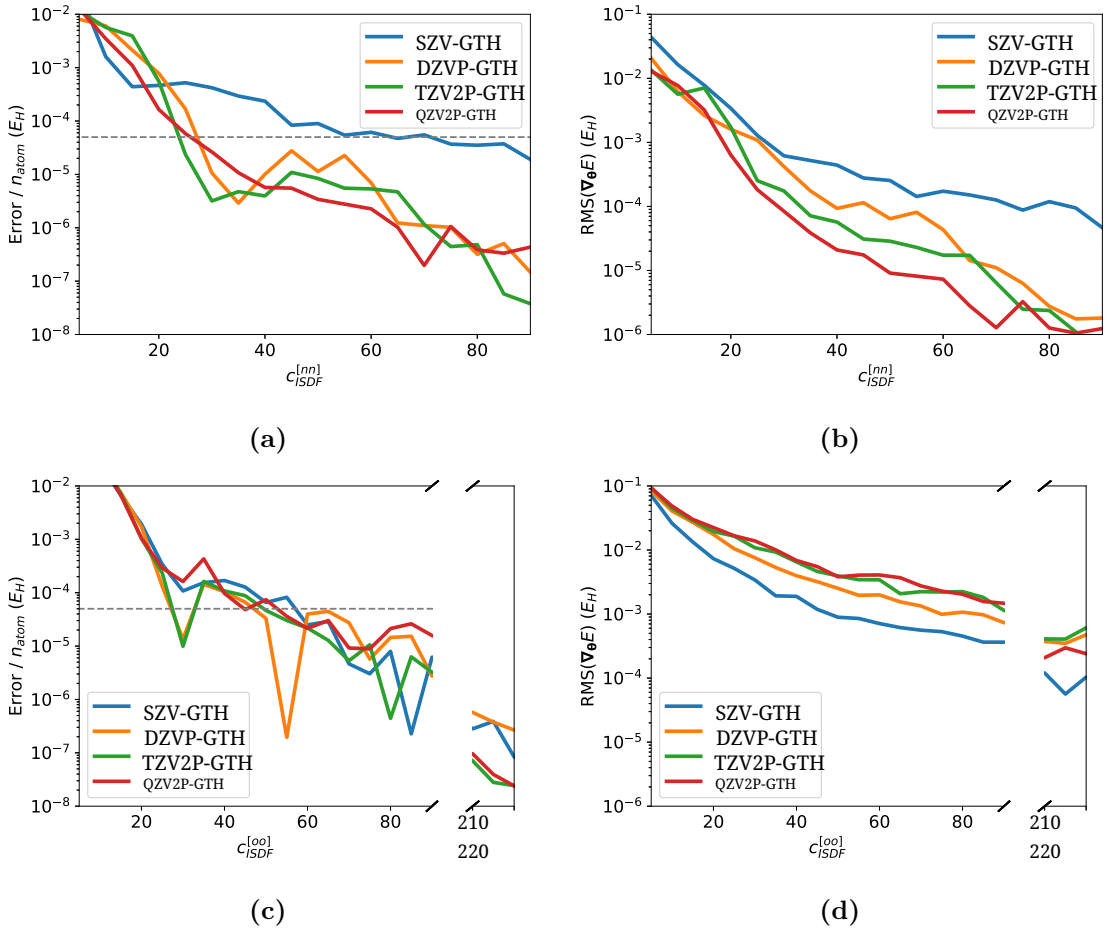
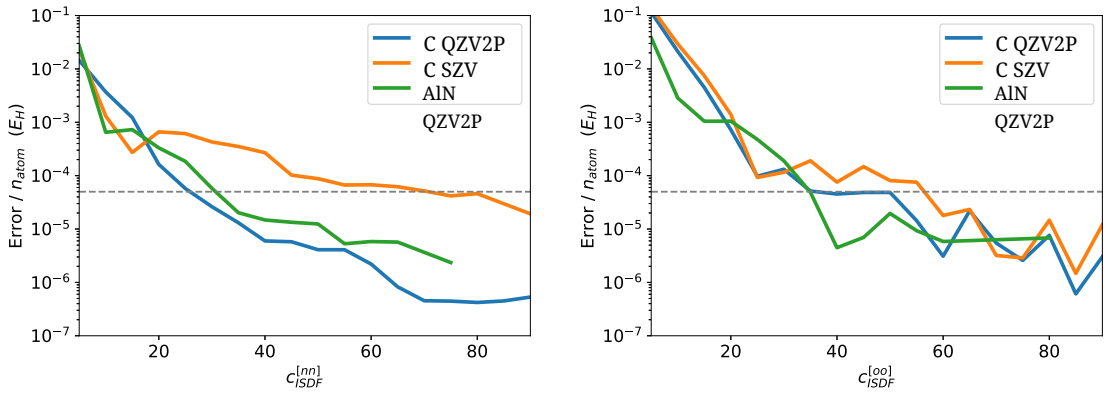


图 1： 对于金刚石，在使用不含 ISDF求积误差的收敛密度进行计算时，(2,2,2)k点网格下 THC-AO-K HF 能量 (a)、THC-AO-K轨道梯度 (b)、THC-oo-K HF能量 (c) 和 THC-oo-K轨道梯度 (d) 的误差。THC-AO-K 和 THC-oo-K 所使用的 ISDF 插值点数量分别通过公式 6 和 17 与参数 $c^{[nn]}$ ISDF 和 $c^{[oo]}$ ISDF 相关。

大于 THC-AO-K 的轨道梯度，并且随着 $c^{[oo]}$ ISDF 增加，收敛似乎相当缓慢。此外，它似乎在一定程度上取决于基组：较大的基组拥有多得多的自由度；SZV-GTH基组的梯度误差小得多，而 DZVP-GTH、TZV2P-GTH 和 QZV2P-GTH 的误差大致相近。梯度中的这种不准确性可能源于整个能量面上的 THC误差波动，并可能导致 SCF收敛问题。接下来，我们将研究通过 THC-K算法得到的自洽能量，以探讨这一问题。

接下来，我们研究了收敛密度下 HF 能量的误差——也就是说，我们使用 THC 算法优化轨道。在图2a中，我们展示了使用 THC-AO-K 得到的 HF 自洽总能量误差随 $c^{[nn]}_{\text{ISDF}}$ 变化的情况，涉及采用 QZV2P-GTH 和 SZV-GTH 基组的金刚石，以及采用 GTH-QZV2P 基组的 AlN（总计 16 个电子）；所有计算均使用 (2,2,2) k 点网格。可以看到，对于 QZV2P-GTH 体系，我们建议的 $c^{[nn]}_{\text{ISDF}} = 25$ 值表现相当不错，不过在 SZV-GTH 基组下，金刚石的误差确实明显更高。图2b给出了 THC-oo-K 的类似结果。尽管精确收敛密度下评估的轨道梯度可能存在显著偏差，从而引发担忧，THC-oo-K 算法仍能为这些体系提供足够准确的自洽解。最令人鼓舞的是，$c^{[oo]}_{\text{ISDF}}$ 参数显然可以在不同体系和基组之间通用。



(a) (b)

图2： 使用不同基组大小和 (2,2,2) k点网格，对金刚石和 AlN 采用 (a) THC-AO-K 与 c^{nn}_{ISDF}（见公式 6）以及 (b) THC-oo-K 与 c^{oo}_{ISDF}（见公式 17）计算的单位晶胞自洽 HF 能量的每原子绝对误差。虚线表示 50 μE_H/原子精度。

接下来，我们研究随着k点网格增大时的精度。我们的算法依赖于这样一个事实：为获得准确解所需的 ISDF 点数会随着k网格大小的增加而迅速趋于平台。在图 3 中，我们展示了使用THC-oo-K在QZV2P-GTH 基组中计算的金刚石自洽 HF 能量的绝对误差图，

几种不同的k点网格。我们看到，从 Γ 点到(2,2,2) k点网格时，所需的 ISDF 点数量急剧增加；显然，在多个k 点处表示占据轨道的乘积，比在单个k 点处困难得多。然而，进一步增加网格几乎没有影响。因此，对于所有k点网格尺寸， $c^{[oo]}{}_{\mathrm{ISDF}} = 50$ 值似乎都已足够，不过对于 Γ 点计算，显著更低的值可能就足够了。

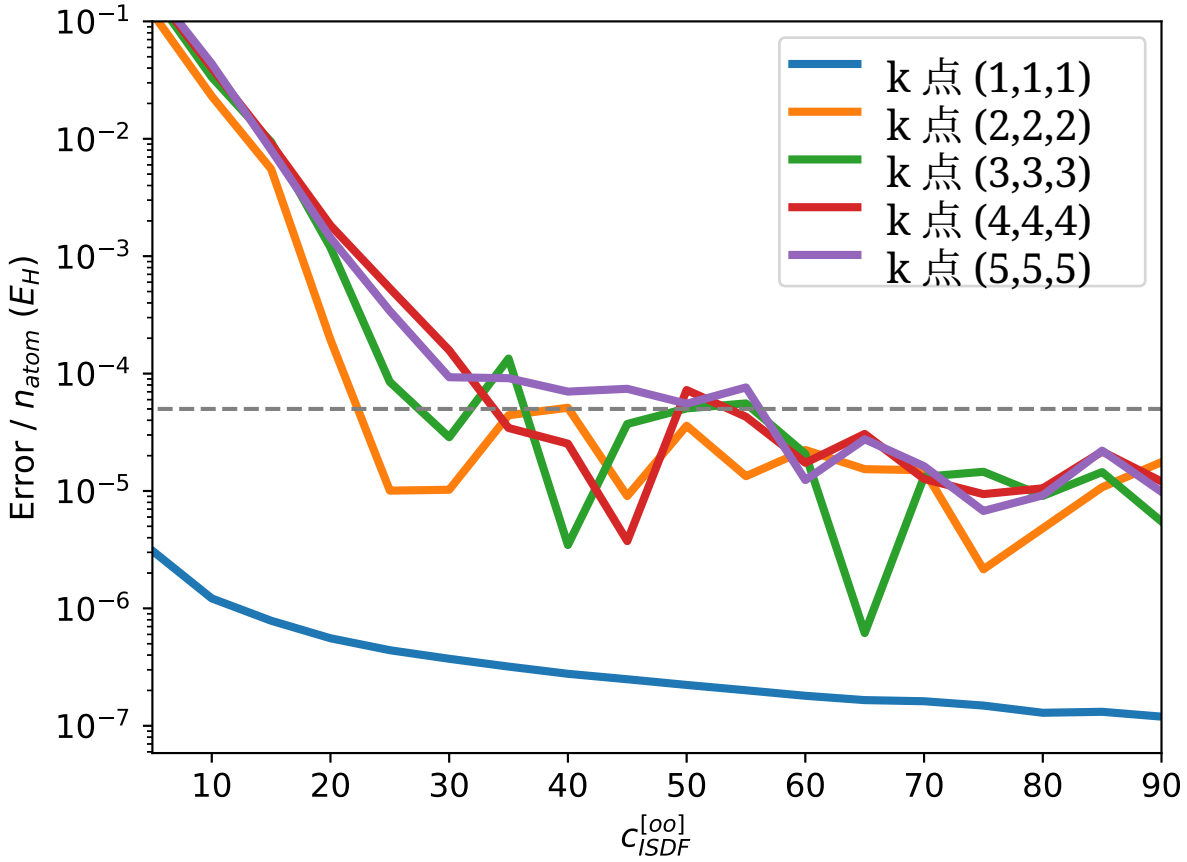


图 3： 使用 THC-oo-K 计算的 HF 能量绝对误差，随 c^{oo}_{ISDF} 变化；对象为采用 QZV2P-GTH 基组的金刚石，使用了多个k点网格。

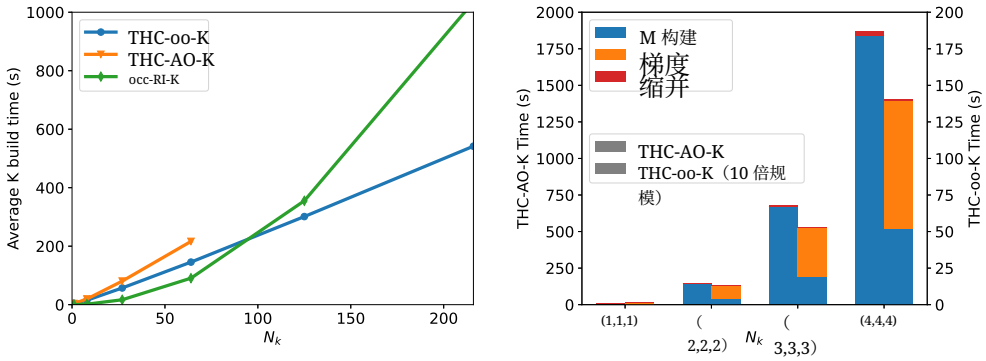
因此，我们暂时建议从数值 $c^{[nn]}{}_{\mathrm{ISDF}} = 25$、$c^{[oo]}{}_{\mathrm{ISDF}} = 50$ 开始。可以通过逐案增大这些参数并确保总体结果变化很小，来验证这样做是否足够。我们注意到，本研究分析了绝对能量，

这未考虑误差抵消对晶格能或带隙等可观测量的影响。误差抵消可能使 c_{ISDF} 参数的要求更宽松，从而减少所需的插值点数。此外，我们在此展示的是 HF 能量误差，其中包含 100% 的精确交换。使用仅包含少量精确交换或短程精确交换的杂化密度泛函时，THC 误差还会进一步缩小。事实上，我们在分子研究中观察到了这一点。[27] 因此，我们认为推荐的系数是保守的。

4.2 计算复杂度

我们发现，采用k点采样时，周期性 THC-AO-K 和 THC-oo-K 所需的插值点都比对应的分子体系更多。然而，研究也发现，除非体系非常大，否则即使是分子 THC 所需的插值点也过多，难以在计算上具有竞争力。因此，我们考察 K 构建时间，评估基于 ISDF 的算法对小型至中型周期性体系是否有用。在接下来的计算中，我们使用 $c^{[nn]}{}_{\mathrm{ISDF}} = 25$ 和 $c^{[oo]}{}_{\mathrm{ISDF}} = 55$。此外，假设进行 10 次迭代来计算 THC-AO-K 的平均迭代时间，因为该算法的大部分计算成本来自只需执行一次的 M 构建步骤；也就是说，我们将十分之一的 M 构建时间计入每次迭代的成本。如果所需迭代次数超过 10 次（即初始猜测较差），那么 THC-AO-K 将比此处所示更有优势，反之亦然。

在图 4a 中，我们绘制了 AlN 在 QZV2P-GTH 基组下，THC-AO-K 和 THC-oo-K 的 K 构建时间随k点数变化的关系。该图揭示了两个重要结论。首先，正如预期，得益于快速傅里叶变换卷积，THC-K算法的计算量随k点数线性增长。相比之下，occ-RI-K 的计算量随k点数呈二次增长。其次，THC-K算法的前置因子较大，因此在k点网格较小时，occ-RI-K 是更高效的方法。线性的k点。



(a) (b)

图4： (a) 每个循环的平均 K 构建时间；(b) AlN 的 QZV2P-GTH 基组中，THC 初始循环计时随k点数量变化的分解情况。

这种扩展性使THC-oo-K在减少大型k点计算的计算时间方面极为有效，并且该算法在(4,4,4)到(5,5,5)k点网格之间变得比occ-RI-K更高效。由于基组较大，因此 $N^{[nn]}_{\text{ISDF}}$也较大，THC-AO-K算法明显慢于另外两种算法，并且由于内存需求，无法在大于(4,4,4)的k点网格上运行。

图4b 展示了 AlN 体系中 THC-AO-K 和 THC-oo-K 首个循环计算时间的分解情况（其中包括完整的 THC-AO-K M 构建时间）。可以看到，THC-AO-K 的耗时主要来自最初的单次M 构建，因此按照式（14）收缩M 以形成K 的每次迭代成本可以忽略不计。THC-oo-K 的M 构建耗时显著更低（注意尺度相差 10 倍），但计算轨道梯度会使每次迭代的计算成本额外增加 2–3 倍。

我们还研究了 THC-K 的计算复杂度随基组大小的变化。此前我们发现，occ-RI-K 算法基本不受基组大小影响，因此在基组较大且k点数量较少的计算中非常有效。我们预计 THC-oo-K 也具有这一特性，因为如图2b所示， N_{ISDF} 与基组大小无关。因此，随着基组大小增加，我们预计 THC-oo-K 将比 THC-AO-K 更高效。在图5中，我们展示了平均 K 矩阵构建时间随

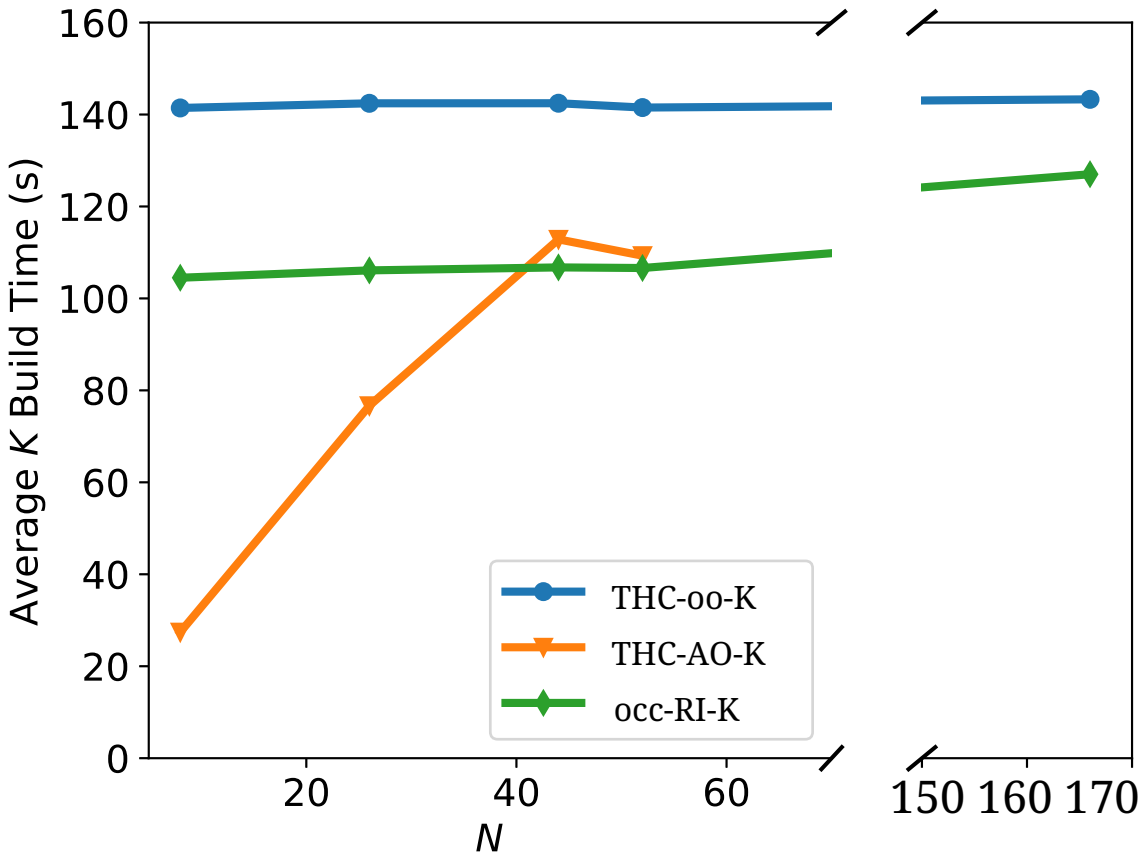


图5： 金刚石在 (7,7,7) k 点网格下，每周期平均 K 构建时间与基组函数数量的关系。

基组大小的变化，所研究的是采用 (7,7,7) k网格的金刚石体系。对于所用的前两种最小基组 SZV-GTH 和 DZVP-GTH，THC-AO-K 比 THC-oo-K 和 occ-RI-K 更高效；该算法的瓶颈在于预先计算 M 矩阵，而对于较小的基组，该矩阵相当小。对于 TZV2P-GTH 和 QZV2P-GTH，occ-RI-K 略优于 THC-AO-K，因为它不受基组大小影响。显然，只有在使用非常大的 unc-def2-QZVP-GTH 基组时，THC-oo-K 才能与 THC-AO-K 竞争。

最后，我们指出，与 occ-RI-K 相比，THC 算法可显著降低内存用量。对于这两种 THC 算法，我们仅存储 $\xi_P(\mathbf{r})$ 和 $M_{\mathbf{q}PQ}$，即不存储任何随网格和k点数量同时增长的量。这与

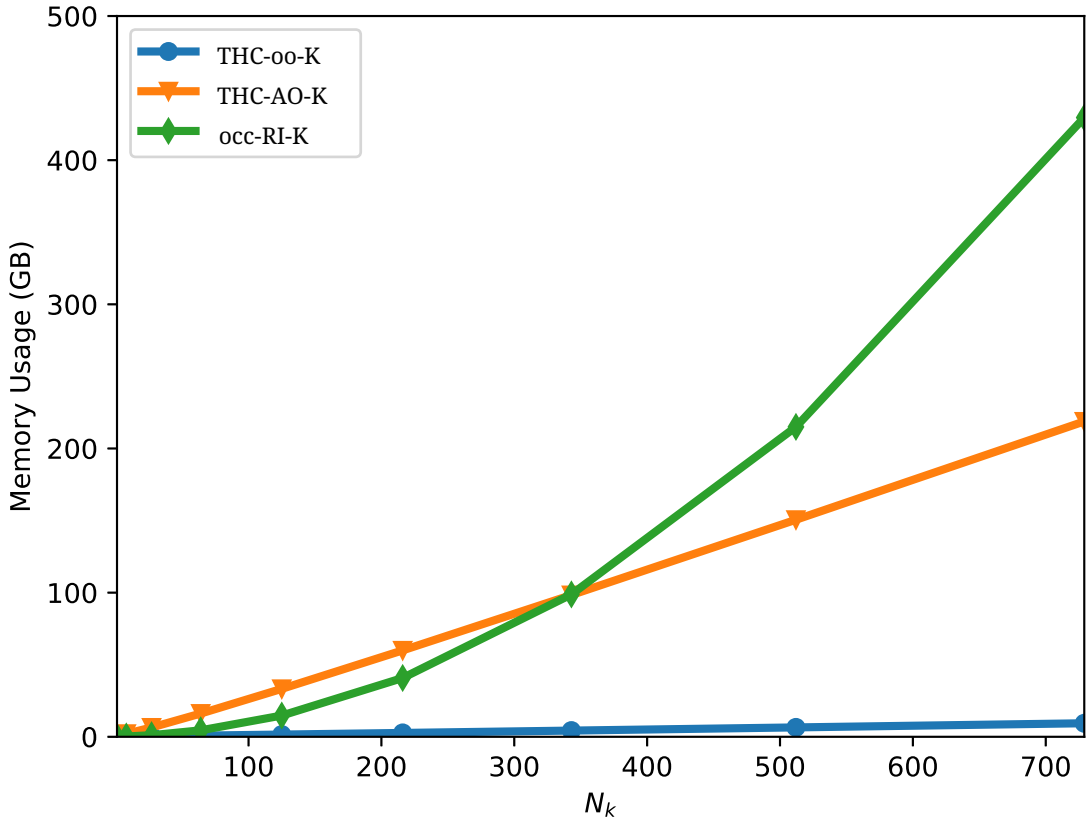


图 6： 使用 QZV2P-GTH 基组时，金刚石体系中 occ-RI-K、THC-AO-K 和 THC-oo-K 算法的内存用量随**k**点数量的变化。

occ-RI-K 算法默认将基组函数（在 N_k 中呈线性增长）和库仑核（在 N_k 中呈二次增长）存储在 Q-Chem 的网格上。采用积分直接法进行 occ-RI-K 将消除这一内存需求[34]，但会导致性能大幅下降，使其相较于 THC 算法的竞争力降低。我们在图 6 中对此进行了强调：绘制了 QZV2P-GTH 基组的金刚石体系中，各算法的内存需求随 k-点数量的变化关系。我们看到，对于小于 (6,6,6) 的 k-点网格，THC-AO-K 所需的大量 ISDF 点导致其内存用量高于 occ-RI-K；然而，N_k 中的线性扩展最终在大于该规模的 k-点网格上占据优势。由于 ISDF 点数量大幅减少，THC-oo-K 的内存用量较 THC-AO- K 显著降低。对于这个金刚石 QZV2P-GTH 示例，在使用 (9,9,9) k-点网格时，THC-oo-K 仅需 9 GB，而 occ-RI-K 则需要 430 GB！

现在，我们可以为本文研究的每种算法定义具体的适用场景。对于较小的k点网格，我们建议使用 occ-RI-K 算法，因为它比所有 THC 算法都快。对于较大的k点网格，如果使用较小的基组，我们建议使用 THC-AO-K；如果使用较大的基组，则建议使用 THC-oo-K。对于各种规模的体系和k点网格，如果推荐的算法需要占用过多内存，我们建议改用 THC-oo-K。它尤其适合在内存受限的计算设备（如图形处理器）上实现。

4.3 分子晶体

最后，我们通过计算苯晶体的内聚能，对 THC-K 的极限进行了严格测试。通过基于片段的方法开展的高阶波函数理论计算，可以获得该数值的准确理论估计。[45–47] 迄今为止，周期性计算相当有限，因为众所周知，HF 和纯密度泛函在结合方面表现不佳[48]，而更高阶的波函数理-

对于苯的周期性研究而言，成本极高。研究表明，采用色散修正时，局域密度泛函能够给出合理的结果，[49] 近期也已开展使用周期性 MP2 的研究。[50,51] 然而，受计算资源限制，这些周期性计算通常只能采用较小的k点网格或较小的基组，或者两者皆是。由于难以在热力学极限下获得精确交换能，仅有少数杂化泛函用于苯的测试[49, 52–55] 因此，我们决定利用 THC-oo-K 算法对用于计算内聚能的杂化泛函进行基准测试，以便与理论最佳估计值进行比较。

我们研究了经过色散修正的纯泛函B97M-rV[56,57] 以及杂化泛函PBE0-D3,[58,59] MN 15,[60] M06-2X,[61] SCAN0-D3,[62] ωB97X-rV泛函,[63] 以及 ωB97M-rV。[64] 我们尝试将结果收敛至完备基组（CBS）极限和热力学极限（TDL），其中包括反平衡校正（见第3节）。我们涵盖了同时采用D3 和rVV10 色散修正的杂化泛函，以及经过参数化以包含色散的经验泛函。我们注意到，MN15和 M06-2X 并未经过色散修正，但研究发现，它们在非共价相互作用分子数据集上的表现相当准确。[1] 对于不同大小的 k 点网格计算得到的内聚能见表 1。我们使用 QZV2P-GTH 基组进行了计算。

与理论最佳估计（TBE）的 -54.58kJ/mol相比，[45] 我们看到 DFT的性能存在很大差异。正如预期的那样，色散处理对于获得准确的内聚能至关重要。DFT-D3 色散处理给出了相当准确的结果，其中全局杂化泛函PBE0-D3 和SCAN0-D3 表现良好，误差分别为 3.4 和 2.0kJ/mol。经过参数化以隐式处理色散的MN15和 M06-2X给出了极为不同的结果；MN15 是所测试泛函中最准确的，误差为1.1kJ/mol，而 M06-2X 的表现显著差于其他所有泛函，误差达到13kJ/mol。使用 rVV10色散校正同样会产生不理想的结果；纯泛函 B97M-rV 的结合能过强了 9.0

kJ/mol，而长程分离杂化泛函 ωB97M-rV和 ωB97X-rV泛函是分子体系中最精确的泛函之一，[1] 其表现仅略好一些，误差分别为 7.2 和8.0 kJ/mol。这三个基于 VV10 的泛函持续过度结合，表明 rVV10 是其误差的主要来源。

我们数据中观察到的趋势与此前针对苯和 X23 数据集的色散修正研究结果一致。众所周知，DFT-D3 的表现相当出色，[49] 而 rVV10 在分子晶体中的失效此前被归因于缺乏屏蔽（即多体色散）效应。[65–67] 值得注意的是，通过比较 B97M-rV 与 ωB97M-rV，我们发现加入精确交换仅带来 1 kJ/mol 的小幅改进。

我们看到，对于一个 (3,3,3) k点网格，纯泛函基本上已经收敛，而杂化泛函与热力学极限的差距仍高达5kJ/mol。杂化泛函的收敛性在图7中得到了展示，其中收敛速率取决于泛函中精确交换所占的比例——这突显了周期性杂化密度泛函理论的难度。尽管如此，使用 $1/N_k$ 进行尺寸外推对该体系效果良好，使得可靠的杂化泛函计算成为可能。

表 1： 苯晶体内聚能（kJ/mol）的误差，作为多个密度泛函的函数，随k点网格大小变化。

N_k	B97M-rV	Hartree–Fock	PBE0-D3	MN15	M06-2X	SCAN0-D3	ωB97X-rV泛函	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

我们发现，对于苯晶体，加入精确交换可以小幅提高精度，但色散的处理重要得多。要评估杂化泛函用于周期性应用时的一般趋势，还需要进行更深入的分析。不过，分子[1]与周期性体系之间的差异表明，进一步开发适用于周期性应用的杂化泛函可能会有所助益。

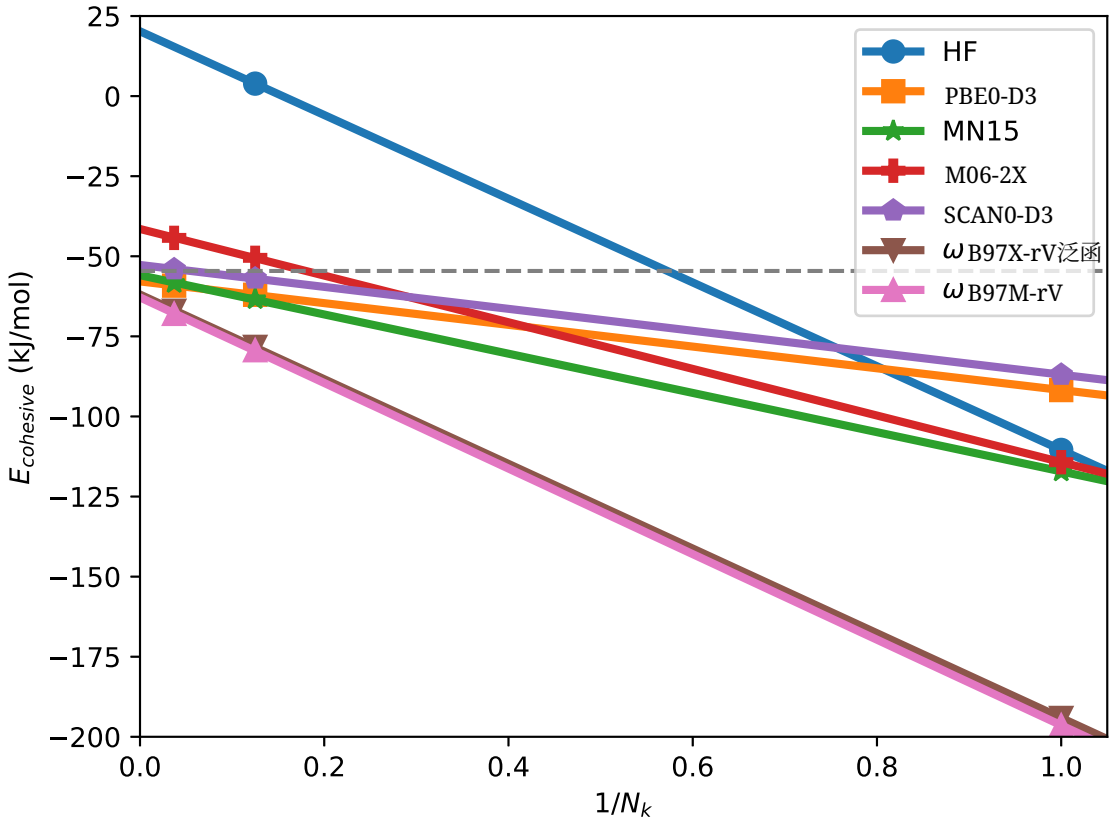


图 7： 苯晶体的内聚能随若干杂化泛函的**k**点采样而变化。虚线表示参考值。

5 结论

总而言之，我们提出了将 ISDF近似扩展用于采用高斯型轨道进行周期性k点计算中的精确交换。我们还提出了一种新算法 THC-oo-K：通过 ISDF方法仅拟合占据轨道的乘积，而计算能量只需要这一项。虽然这意味着 THC-oo-K轨道梯度（涉及占据-虚轨道乘积）的误差大于能量的误差，但对自洽优化后能量的影响仍然很小。我们证明，这些算法将精确交换的计算复杂度降低至随体系大小呈三次方增长、随k点数量呈线性增长；此外，THC-oo-K算法的 THC维度与基组大小无关。初步研究表明，对于大型k点网格，这些算法能显著节省计算开销；对于中等规模的k点网格，THC-oo-K相较于 occ-RI-K 也能节省计算开销，并在所有情况下大幅降低内存用量，同时总体精度仅略有牺牲。我们认为，这种成本降低将使采用杂化泛函开展周期性研究更加可行。我们通过一项苯的晶格能研究展示了这一点，该研究采用 QZV2P-GTH 基组，计算至 (3,3,3) k点网格。

通过 ISDF 仅拟合双电子积分张量的一侧，可进一步降低误差，这种方法称为稳健拟合THC。[35,68] 这使能量误差与 THC拟合误差呈二次关系，而非线性关系。我们对该方法在 THC-AO-K算法中的应用进行了初步研究。不过，我们发现，尽管稳健拟合确实大幅降低了误差，但由于必须存储 $V_P^q(r)$——由插值函数$\xi_P(r)$产生的势——内存需求显著增加，达到难以承受的程度。这使本已计算开销很大的 THC-AO-K 的内存用量增加了 N_k 倍；此外，GPW算法所需的昂贵快速傅里叶变换次数也增加了 N_k 倍，显著延长了计算时间。尽管如此，多节点

MPI 实现可能会将可用内存增加到足以使这一方法变得可行。[35] 对于 THC-oo-K 算法，无需在每次迭代中存储 $V_P^q(r)$，因此可以对这一量进行分批处理，从而避免存储其全部内容；然而，快速傅里叶变换数量的增加仍将导致计算成本显著上升。因此，稳健拟合 THC-oo-K算法是未来发展的一个有趣方向，但计算时间方面的权衡最终可能并不理想。

本文所有工作均采用 GPW算法，该算法仅适用于赝势计算。与更常用的平面波基组相比，使用高斯轨道进行周期性边界条件计算的一大优势是能够对芯轨道进行建模。研究这些算法的全电子版本将会很有意义。通过采用投影缀加波方法（PAW）、[69] 密度拟合、[70,71] 或类似方法，并像分子程序中那样从 Becke 网格中选取 ISDF 点，应该可以实现这一扩展。[27]

最后需要指出，这种方法为近似双电子积分提供了一个通用框架，其用途并不限于精确交换。相同的中间量也可用于加速计算 Møller–Plesset 理论、耦合簇理论及其他关联波函数方法所需的矩阵元。这一点已在分子 ISDF 中得到验证。[27,72,73] 将该算法扩展到关联波函数理论后，便可进行规模大得多的k点计算。

致谢：本研究由美国能源部科学办公室基础能源科学办公室主任办公室通过合同编号 DE-AC02-05CH11231 资助。A.R. 感谢国家科学基金会通过项目编号 DGE 1752814 提供资助。J.L. 感谢 David Reichman 的支持。

参考文献

(1) Mardirossian, N.; Head-Gordon, M. 计算化学中密度泛函理论的三十年：对 200 种密度泛函的概述与全面评估。*Mol. Phys.* **2017年**, *115*, 2315–2372.

(2) Goerigk, L.; Hansen, A.; Bauer, C.;Ehrlich,S.;Najibi,A.;Grimme, S. 通过用于通-用主族热化学、动力学和非共价相互作用的先进 GMTKN55 数据库，探讨密度泛函理论“动物园”。物理化学、化学物理 **2017年**, *19*,32184–32215。

(3) Dohm, S.; Hansen, A.; Steinmetz, M.; Grimme, S.; Checinskii, M. P. 真实闭壳层金属有机反应的综合热化学基准集。*J. Chem.Theory Comput.* **2018**, *14*, 2596–2608。

(4) Chan, B.; Gill, P.M.W.;Kimura, M.使用TMC151 数据集汇编评估过渡金属的 DFT 方法，并与主族化学的精度进行比较。 化学理论与计算杂志 **2019**, *15*,3610–3622。

(5) Muscat, J.; Wander, A.; Harrison, N. 关于利用杂化泛函理论预测带隙。《化学物理快报》 **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. 在计算硅的不同相及其缺陷的总能量时，对比屏蔽杂化密度泛函理论与扩散蒙特卡洛方法。《物理评论 *B*》 **2006**, *74*, 121102.

(7) De´ak, P.; Aradi, B.; Frauenheim, T.; Janz´en, E.; Gali, A. 通过 HSE06 范围分离杂化泛函获得准确的缺陷能级。《物理评论 *B*》**2010年**，*81*，153203。

(8) Komsa, H.-P.; Pasquarello, A. 评估杂化泛函在确定缺陷能级方面的准确性：As 反位缺陷在砷化镓中的应用。《物理评论 *B*》**2011年**，*84*，075207。

(9) Henderson, T. M.; Paier, J.; Scuseria, G. E. 使用 HSE屏蔽杂化泛函 精确处理固体。《固体物理学（*b*辑）》**2011年**, *248*, 767–774.

(10) Schwegler, E.; Challacombe, M. Hartree–Fock 交换矩阵的线性标度计算。《化学物理学杂志》1996年，105，2726–2734。

(11) Schwegler, E.; Challacombe, M.; Head-Gordon, M. Fock矩阵的线性标度计算。II. 交换积分和增量式Fock矩阵构建的严格界限。《化学物理学杂志》**1997年**, *106*, 9708–9717.

(12) Burant, J. C.; Scuseria, G. E.; Frisch,M.J. 一种用于大分子 Hartree–Fock交换计算的线性标度方法。《化学物理学杂志》 **1996年**,*105*, 8969–8972.

(13) Challacombe, M.; Schwegler, E. Fock矩阵的线性标度计算。《化学物理学杂志》1997年, 106, 5526–5536.

(14) Ochsenfeld, C.; White, C. A.; Head-Gordon, M. Hartree–Fock型交换矩阵的线性与次线性标度构造。《化学物理学杂志》**1998年**, *109*, 1663–1669.

(15)Friesner,R. A. 通过伪谱方法求解自洽场电子结构方程。《化学物理快报》《化学物理快报》 **1985年**，*116*，39–43。

(16) Neese, F.; Wennmohs, F.; Hansen, A.; Becker, U. 高效、近似且并行的 Hartree–Fock 与杂化密度泛函理论计算：一种用于 Hartree–Fock 交换的“球链”算法。《化学物理》**2009年**, *356*,98–109.

(17)Izs´ak, R.; Neese,F.一种重叠拟合球链交换方法。《化学物理学杂志》 **2011年**, *135*, 144105。

(18) Beebe, N. H.; Linderberg, J. 分子计算中双电子积分生成与变换的简化。《国际量子化学杂志》**1977**, *12*, 683–705。

(19)Røeggen, I.; Wisløff-Nilssen,E. 关于 Beebe-Linderberg 双电子积分近似。《化学物理快报》**1986**, *132*, 154–160。

(20) Baerends, E.; Ellis, D.; Ros, P. 自洽分子 Hartree—Fock—Slater 计算 I。计算程序。《化学物理》**1973**, *2*, 41–51。

(21) Whitten, J. L. 库仑势能积分及其近似方法。《化学物理学杂志》**1973**, *58*, 4496–4501。

(22) Jafri, J.; Whitten, J. 电子排斥积分近似与误差界限：分子应用。《化学物理学杂志》1974, 61, 2116–2121。

(23) Hohenstein, E. G.; Parrish, R. M.; Mart´ınez, T. J. 张量超收缩密度拟合。I. 四次标度的二阶和三阶莫勒-普莱塞微扰理论。《化学物理学杂志》**2012年**, *137*,044103.

(24) Parrish,R.M.;Hohenstein, E. G.; Mart´ınez,T.J.;Sherrill, C. D. 张量超收缩。II.最小二乘重归一化。《化学物理学杂志》**2012年**, *137*, 224106。

(25) Lin, L. 自适应压缩交换算符。《化学理论与计算杂志》**2016**, *12*, 2242–2249。

(26) Shenvi, N.; Van Aggelen, H.; Yang, Y.; Yang, W. 张量超收缩 ppRPA：将粒子-粒子随机相位近似的计算成本从 O (r 6) 降至 O (r 4)。《化学物理学杂志》**2014**，*141*，024119。

(27) Lee, J.; Lin, L.; Head-Gordon, M. 可系统改进的张量超收缩：将插值可分离密度拟合用于分子的精确交换、二阶和三阶 Møller–Plesset 微扰理论。《化学理论与计算杂志》**2019**，*16*，243–263。

(28) Dong, K.; Hu, W.; Lin, L. 通过质心 Voronoi 镶嵌实现插值可分离密度拟合，并应用于杂化泛函电子结构计算。《化学理论与计算杂志》**2018**，*14*，1311–1320。

(29) Hu, W.; Lin, L.; Yang, C. 用于加速杂化密度泛函计算的插值可分离密度拟合分解及其在硅中缺陷研究中的应用。《化学理论与计算杂志》**2017年**，*13*，5420–5431。

(30) Wu, K.; Qin, X.; Hu, W.; Yang, J. 通过低秩近似加速含 k 点采样的平面波杂化泛函计算。《化学理论与计算杂志》**2021**，*18*，206–218。

(31) Li, J.; Qin, X.; Wan, L.; Jiao, S.; Hu, W.; Yang, J. 基于复数 K-means 聚类的插值可分离密度拟合，用于采用平面波基组的大规模混合泛函从头算分子动力学。*arXiv* 预印本 *arXiv:2208.07731***2022**,

(32) Qin,X.;Hu,W.;Yang, J. 用于加速双电子积分的插值可分离密度拟合：理论视角。《化学理论与计算杂志》 **2023,**

(33) Manzer, S. F.; Horn, P. R.; Mardirossian, N.; Head-Gordon, M. 快速、准确地评估精确交换：occ-RI-K 算法。《化学物理学杂志》 **2015**, *143*, 024113.

(34) Lee, J.; Rettig, A.; Feng, X.; Epifanovsky, E.; Head-Gordon, M. 通过 occ-RI-K 更快地计算固体的精确交换：在简单固体的基组极限附近的组合优化范围分离杂化泛函中的应用。 *arXiv* 预印本 *arXiv:2207.09028* **2022**,

(35) Sharma, S.; White, A. F.; Beylkin, G. 使用高斯基组和稳健伪谱方法快速计算交换。*arXiv* 预印本 *arXiv:2207.04636* **2022**,

(36) Qin,X.;Liu,J.; Hu, W.; Yang, J.用于加速数值原子轨道中 Hartree–Fock交换计算的插值可分离密度拟合分解。《物理化学杂志 *A*》**2020**,*124*, 5664–5674。

(37) Matthews, D. A. 最小二乘张量超缩并中的网格优化与拟合改进。《化学理论与计算杂志》**2020**，*16*，1382–1385。

(38) Lee, J.; Feng, X.; Cunha, L. A.; Gonthier, J. F.; Epifanovsky, E.; Head-Gordon, M. 通过可迁移性逼近基于高斯轨道的周期性计算中的基组极限：纯密度泛函在简单半导体中的性能。《化学物理学杂志》 **2021年**，*155*，164102。

(39) Goedecker, S.; Teter, M.; Hutter, J. 可分离双空间高斯赝势。《物理评论 *B*》**1996年**，*54*，1703。

(40) Krack, M. 针对梯度校正交换-相关泛函优化的 H 至 Kr 赝势。《理论化学通讯》**2005**，*114*，145–152。

(41) Lippert, B. G.; PARRINELLO, J. H.; MICHELE, 一种高斯与平面波混合密度泛函方案。《分子物理学》**1997年**, *92*, 477–488.

(42) VandeVondele, J.; Krack, M.; Mohamed, F.; Parrinello, M.; Chassaing, T.; Hutter, J. Quickstep：采用高斯与平面波混合方法进行快速、精确的密度泛函计算。《计算物理通讯》**2005**, *167*, 103–128.

(43) Fraser, L. M.; Foulkes, W.; Rajagopal, G.; Needs, R.; Kenny, S.; Williamson, A. 周期性边界条件下均匀体系的量子蒙特卡罗计算中的有限尺寸效应与库仑相互作用。《物理评论 *B*》**1996年**, *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. 等。量子化学前沿软件：Q-Chem 5 软件包开发概览。《化学物理学杂志》**2021**, *155*, 084801.

(45) Yang, J.; Hu, W.; Usvyat, D.; Matthews, D.; Sch¨utz, M.; Chan, G. K.-L. 从头算确定晶态苯的晶格能，精度达到亚千焦/摩尔。《科学》**2014**, *345*, 640–643.

(46) Ringer, A. L.; Sherrill, C. D. 有机固体晶格能的第一性原理计算：苯晶体。《欧洲化学杂志》**2008**，*14*，2542–2547。

(47) Kennedy, M. R.; McDonald, A. R.; DePrince III, A. E.; Marshall, M. S.; Podeszwa, R.; Sherrill, C. D. 通讯：解析晶态苯晶格能中的三体贡献：耦合簇理论的基准结果。《化学物理学杂志》**2014**，*140*，121104。

(48) Civalleri, B.; Doll, K.; Zicovich-Wilson, C. 对结晶尿素的结构和内聚能进行从头算研究。《物理化学杂志 *B*》**2007**, *111*, 26–33.

(49) Moellmann, J.; Grimme, S. 对若干分子晶体的 DFT-D3 研究。《物理化学杂志 *C*》**2014**, *118*, 7615–7621.

(50) Bintrim, S. J.; Berkelbach, T. C.; Ye, H.-Z. 采用密度拟合的周期性体系积分直接 Hartree–Fock 方法和 Møller–Plesset 微扰理论：苯晶体中的应用。《化学理论与计算杂志》**2022**，*18*，5374–5381。

(51) Neumann, M. A.; Perrin, M.-A. 使用密度泛函理论计算和经验范德华校正对分子晶体进行能量排序。《物理化学杂志 *B*》**2005**，*109*，15531–15541。

(52) Civalleri, B.; Zicovich-Wilson, C. M.; Valenzano, L.; Ugliengo, P. 将结合经验色散项的 B3LYP（B3LYP-D*）应用于分子晶体。*Crys-tEngComm* **2008**，*10*，405–410。

(53) Loboda, O. A.; Dolgonos, G. A.; Boese, A. D. 通过基于片段的方法进行分子晶体的杂化密度泛函计算。《化学物理学杂志》**2018**，*149*，124104。

(54) Price, A. J.; Otero-de-la-Roza, A.; Johnson, E.R.经 XDM 校正的杂化DFT 与数值原子轨道相结合，以前所未有的准确度预测分子晶体晶格能。《化学科学》**2023**,

(55) Cutini, M.; Civalleri, B.; Corno, M.; Orlando, R.; Brandenburg, J. G.; Maschio, L.; Ugliengo, P. 对不同量子力学方法预测分子晶体的结构和内聚能的评估。《化学理论与计算杂志》 **2016**，*12*，3340–3352。

(56) Mardirossian, N.; Head-Gordon, M. 绘制元广义梯度

近似密度泛函的基因组图谱：寻找 B97M-V。《化学物理学杂志》**2015**，*142*，074111。

(57) Mardirossian, N.; Ruiz Pestana, L.; Womack, J. C.; Skylaris, C.-K.; Head-Gordon, T.; Head-Gordon, M. 在 B97M-V 密度泛函中使用 rVV10 非局域相关泛函：定义 B97M-rV 及相关泛函。《物理化学快报》**2017年**，*8*，35–40。

(58) Adamo, C.; Barone, V. 迈向无需可调参数的可靠密度泛函方法：PBE0 模型。《化学物理学杂志》**1999**，*110*，6158–6170。

(59) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. 针对 H-Pu 共 94 种元素的密度泛函色散校正（DFT-D）的一致且准确的从头算参数化。《化学物理学杂志》**2010**，*132*，154104。

(60) Haoyu, S. Y.; He, X.; Li, S. L.; Truhlar, D. G. MN15：一种 Kohn–Sham 全局杂化交换-相关密度泛函，在多参考和单参考体系及非共价相互作用方面具有广泛的准确性。《化学科学》**2016**，*7*，5032– 5051。

(61) Zhao, Y.; Truhlar, D. G. 用于主族热化学、热化学动力学、非共价相互作用、激发态和过渡元素的 M06 系列密度泛函：两个新泛函以及对四种 M06 类泛函和其他 12 种泛函的系统测试。《理论化学通讯》**2008**, *120*, 215–241.

(62)Hui,K.;Chai, J.-D. 基于 SCAN的混合密度泛函和双混合密度泛函：来自无拟合参数模型。《化学物理学杂志》2016，144，044114。

(63) Mardirossian, N.; Head-Gordon, M. ωB97X-V：一种通过适者生存策略设计的、含非局域相关的 10 参数、范围分离杂化广义梯度近似密度泛函,

Physical Chemistry Chemical Physics 物理化学化学物理学 **2014**，*16*，9904–9924。

(64) Mardirossian, N.; Head-Gordon,M. ω B97M-V：一种经过组合优化的、范围分离的混合型元-GGA 密度泛函，采用VV10 非局域相关。《化学物理学杂志》**2016**，*144*，214110。

(65) Hermann, J.; Tkatchenko, A. 范德华体系中的电子交换与相关：平衡半局域与非局域能量贡献。《化学理论与计算杂志》**2018**，*14*，1361–1369。

(66) St¨ohr,M.;Van Voorhis, T.;Tkatchenko, A. 电子结构计算中范德华相互作用建模的理论与实践。化学学会评论 **2019**,*48*, 4118–4154.

(67) Fabiano, E.; Cortona, P. 探索适用范围广泛的色散校正 GGA 泛函：TCA+ D3 和 RevTCA+ D3 在固态体系中的表现。计算材料科学**2023**，*216*，111826。

(68) Pierce, K.;Rishi,V.;Valeev, E. F. 张量网络的稳健近似：库仑相互作用的无网格张量分解应用。《化学理论与计算杂志》**2021**，*17*，2217–2230。

(69) Bl¨ochl，P. E. 投影缀加波方法。《物理评论 *B*》**1994**，*50*，17953。

(70) Sun, Q.; Berkelbach, T. C.; McClain,J.D.;Chan, G. K.-L. 用于周期性体系的高斯与平面波混合密度拟合。《化学物理学杂志》**2017年**，*147*，164119。

(71) Ye, H.-Z.; Berkelbach, T. C. 通过范围分离实现快速周期性高斯密度拟合。《化学物理学杂志》**2021年**，*154*，131104。

(72) Hohenstein, E. G.; Parrish, R. M.; Sherrill, C. D.; Martínez, T. J. 通讯：张量超收缩。III. 用于确定相关波函数的最小二乘张量超收缩。《化学物理学杂志》**2012年**, *137*, 221101。

(73) Lu, J.; Thicke, K. 使用插值可分离密度拟合计算 RPA 相关的立方标度算法。《计算物理学杂志》 2017年, 351,187–202.