Modern developments in Hartree-Fock theory: Fast methods for computing the Coulomb matrix. [a]

Matt Challacombe, Eric Schwegler and Jan Almlöf

Department of Chemistry
and
Minnesota Supercomputer Institute,
University of Minnesota, Minneapolis, MN 55415

Abstract

We describe a variety of new methods for accelerating construction of the Coulomb matrix that are based upon a simplified representation of the electron charge density. Central to these methods are the elimination of degeneracies, the reduction of complexity and a Hermite Gaussian representation. The utility of Hermite Gaussian-type functions as intermediates in the computation of the Coulomb matrix is stressed, and important consequences of this representation are pointed out, especially relationships to fast numerical methods.

Projection methods are reviewed and new results obtained with two-center potential-matched fits of the electron charge density are presented. Based on these results, and on simple arguments, we conclude that the use of a single projection method to approximate the density is inefficient and introduces errors that cannot be controlled systematically.

Hierarchical multipole methods and their quantum chemical adaptations are also reviewed. We introduce the fast Gauss transform to quantum chemistry and demonstrate its equivalence to a Cartesian multipole expansion in the asymptotic limit. A synthesis of the fast Gauss transform and a Barnes-Hut-like method is suggested and a first implementation involving only the Barnes-Hut method is described and shown to be highly competitive. In particular, calculations on large polypeptides and water clusters scale as favorably as $N^{1.6}$, where N is the number of basis functions. An important feature of this new method is that accuracy is not sacrificed for speed, and that errors may be systematically controlled with a single thresholding parameter.

[a] "Computational Chemistry: Review of Current Trends", pp. 53-107, Ed. J. Leczszynski, World Scientific, 1996.

Contents

1 Introduction **3**

2 Hermite Gaussian-type basis functions in quantum chemistry **7**
- 2.1 The McMurchie-Davidson algorithm 8
- 2.2 The density in an HGTF basis 9

3 The Cartesian multipole expansion **11**
- 3.1 HGTF-based ERIs in the multipole limit 12
- 3.2 Recurrence relations for computation of the Cartesian multipole tensor 13

4 Projection methods for approximation of the electron density **15**
- 4.1 Variational equations, constraints and measure 17
- 4.2 Two-center fits 19
  - Evaluation of the TCF Coulomb matrix 20
  - Results 22
  - Discussion 24

5 Fast methods in quantum chemistry **25**
- 5.1 Hierarchical methods for classical systems 26
- 5.2 Hierarchical methods in quantum chemistry 28
- 5.3 The fast Gauss transform 29
  - Multiresolution, wavelets and hierarchical multipole methods . 32
  - The FGT and secondary expansions 33
- 5.4 A quantum chemical tree code 34
  - Independent thresholding of distributions 35
  - Thresholding of multipolar interactions 37
  - Results 42
  - Discussion 43

6 Conclusion **43**

7 Bibliography **45**

1 Introduction

One of the most useful ideas in electronic structure theory is the LCAO concept, *i.e.* the notion that the molecular orbitals $\psi_i(\mathbf{r})$ of any molecule can be efficiently expanded in the atomic orbitals $\phi_a(\mathbf{r})$ of its constituent atoms as

$$\psi_i(\mathbf{r}) = \sum_a C_{ai}\phi_a(\mathbf{r}) \, . \tag{1}$$

Since a probability distribution is given by the square amplitude of the corresponding wavefunction, the LCAO expansion implies an expression for the electron density $\rho(\mathbf{r})$ in terms of atomic orbitals products;

$$\rho(\mathbf{r}) = \sum_{ab} P_{ab}\phi_a(\mathbf{r})\phi_b(\mathbf{r}) \, . \tag{2}$$

The introduction of an LCAO basis in the Hartree-Fock approximation leads to the Roothaan-Hall [1,2] or Pople-Nesbet [3] equations, and requires evaluation of Coulomb- and exchange-type interactions between such orbital products, expressed in terms of the electron-repulsion integrals (ERIs)

$$\left(\phi_a\phi_b\middle|\phi_c\phi_d\right) = \int d\mathbf{r} \int d\mathbf{r}' \, \phi_a\left(\mathbf{r}\right)\phi_b\left(\mathbf{r}\right) g(\mathbf{r},\mathbf{r})\phi_c\left(\mathbf{r}'\right)\phi_d\left(\mathbf{r}'\right) \; , \tag{3}$$

where $g(\mathbf{r},\mathbf{r}') = |\mathbf{r}-\mathbf{r}'|^{-1}$ is the Coulomb operator. The density in the LCAO approximation is an object with N^2 complexity (two different AO indices, each running over all N basis functions), and the computational work required to solve the Roothaan-Hall equations scales formally as N^4. This work is related to computation of the Coulomb matrix elements

$$J_{ab} = \sum_{cd} P_{cd}\left(\phi_a\phi_b\middle|\phi_c\phi_d\right) \; , \tag{4}$$

and elements of the exchange matrix,

$$K_{ab} = \sum_{cd} P_{cd}\left(\phi_a\phi_c\middle|\phi_b\phi_d\right) \; , \tag{5}$$

both involving contraction of four-center ERIs with the density matrix $\mathbf{P}$.

In the early days of quantum chemistry, the atomic orbitals in Eq. (1) were chosen to be exponential functions (ETOs) inspired by the exact eigenfunctions for the hydrogen atom (see *e.g.* Ref. [4] for exhaustive discussions of the subject). However, the N^4 ERIs that had to be evaluated through complicated, multi-dimensional quadrature severely limited the size of the systems amenable to study. For these reasons Boys introduced the Gaussian-type basis function (GTF) [5]. Because the Hartree-Fock density is known to possess a cusp at nuclei [6,7] and to decay exponentially at large distances [8,9,10], the GTF might at first seem inferior to the ETO. It can be argued that these advantages of ETOs are illusory [11,12], but we will not dwell further on the subject here. More important for the present discussion are the facts that

a) all modern applications of electronic structure theory to polyatomic systems use Cartesian Gaussian-type basis functions (CGTFs)

$$\phi_a(\mathbf{r}) = (r_x - A_x)^{l_a}(r_y - A_y)^{m_a}(r_z - A_z)^{n_a} \exp\left[-\zeta_a(\mathbf{r} - \mathbf{A})^2\right] \;, \quad (6)$$

b) CGTFs have been found to provide quite accurate LCAO expansions, permitting calculations whose results stand up in quantitative comparisons with experiment,

c) CGTFs allow the rapid evaluation of ERIs, making possible calculations on systems of chemical interest, and

d) the fast methods that we will discuss in the following are all consequences of the particular mathematical properties of the GTF representation, and would be impractical or impossible with an ETO basis.

For large systems a majority of the ERIs are small, and one would therefore expect that their computation can be avoided, provided that a quick and reliable estimate of their significance can be designed. Dyczmons [13] noted that if the total contribution to the energy from those small integrals is required to be less than some (small) tolerance, then the work associated with computing the remaining ERIs is proportional to $N^2 \log^2 N$ in the limit of large systems. Later, Ahlrichs [14] revised this estimate to $N^2 \log N$, and pointed out that if individual ERIs less than some tolerance are neglected, the work associated with computing the remaining ones scales as N^2.

The Roothaan-Hall equations are non-linear in the MO coefficients C_{ai} in Eq. (1), and their solution is therefore an iterative process requiring the ERIs in every iteration. Since the evaluation of ERIs is the time-critical part of the calculation, it was customary before 1980 to store the ERIs externally and to retrieve them for processing according to Eqs. (4) and (5). This, however, led to severe limitations in large calculations due to insufficient disk and I/O capacities. In the direct SCF approach [15] the integrals are recomputed as needed and then discarded, rather than re-cycled. While such an approach might at first seem wasteful, it has advantages aside from eliminating the storage bottleneck. For example, the prescreening of integrals can be based on their significance to the Fock matrix as in Eqs. (4) and (5), rather than on the magnitude of the ERI itself, leading to significantly improved efficiency in the screening.

The original direct SCF method made use of radial overlaps for the estimation of ERIs, allowing the thresholding of integral batches [15]. There have since been a number of more rigorous estimates proposed for the magnitude of an ERI [16,17,18] which have increased the efficiency of the direct method. The development of recurrence relations for the computation of ERIs [19−29] have extended the applicability of direct technologies by improving their efficiency to the point where the re-evaluation of integrals is competitive with their retrieval from disk. As a result of such improvements, very large direct Hartree-Fock calculations have been found to scale only slightly above the asymptotic N^2 limit [30,31].

It has been argued [31,32] that it is the calculation of ERIs significant to the Coulomb matrix which are responsible for the N^2 limitation, since exchange contributions typically show a more favorable screening behavior due to the structure of the density matrix. We will not belabor this issue here; instead we simply note that decoupling the computation of exchange and Coulomb matrices provides the freedom to pursue different specialized algorithms for the evaluation of each. This separation is required, or at least useful, in a number of advanced theoretical models such as DFT [33,34], HF/DFT hybrid methods [35,36] and the Møller-Plesset family of perturbation methods [37,38,39]. It may also facilitate computational developments, as for example parallel implementations on small-memory architectures [40]. In this article we consider specialized mathematical and algorithmic methods for computation of the Coulomb matrix, under the tacit assumption that efficient methods for computing the exchange interaction are available.

The $\mathcal{O}(N^2)$ limit of direct SCF methods may be explained in terms of the Gaussian product theorem (GPT). The GPT simply states that an uncontracted distribution $\rho_{ab} = \phi_a\phi_b$ can be expressed *exactly* as a finite (short) sum of CGTFs with exponent $\zeta_p = \zeta_a + \zeta_b$. For example, the product of two s-type Gaussians is

$$\begin{aligned} \exp\left[-\zeta_a(\mathbf{r}-\mathbf{A})^2\right] \times \exp\left[-\zeta_b(\mathbf{r}-\mathbf{B})^2\right] = \\ \exp\left[-\xi(\mathbf{A}-\mathbf{B})^2\right] \times \exp\left[-\zeta_p(\mathbf{r}-\mathbf{P})^2\right] \ , \end{aligned} \tag{7}$$

where $\xi = \zeta_a\zeta_b/\zeta_p$ and $\mathbf{P} = (\zeta_a\mathbf{A} + \zeta_b\mathbf{B})/\zeta_p$. As the system size increases the radial overlap $\exp\left[-\xi(\mathbf{A}-\mathbf{B})^2\right]$ falls exponentially with separation of the two centers A and B, and the number of distributions significant to the density, given by Eq. (2), approaches N from above. Likewise, the number of Coulomb matrix elements, $J_{ab} = \left(\rho_{ab}\middle|\rho\right)$, larger than some small threshold is also $\mathcal{O}(N)$. Therefore, computation of the Coulomb matrix by conventional direct methods is at best an $\mathcal{O}(N^2)$ process in the asymptotic limit.

In this chapter, we explore methods for calculation of the Coulomb matrix that accelerate convergence to the N^2 asymptote, as well as fast methods that allow penetration of this limit. Common to all these methods is a transformation of the electron charge density into a representation with reduced complexity. Although approaches involving grids and quadrature are popular and useful in many areas of electronic structure theory [32,41,42], they are outside the scope of this paper and will not be discussed here.

Central to our approach is a representation of the electron charge density in a basis of Hermite Gaussian-type functions (HGTFs). In Section 2, we review the use of HGTFs in quantum chemistry and discuss the elementary simplifications that can be obtained through transformation of the density into a HGTF representation. Hermite Gaussian-type functions are related to derivatives of the Dirac delta function, and in Section 3 we exploit this property to obtain a reduced form of HGTF-based ERIs in the asymptotic limit and to derive new recursion relations for computation of the Cartesian multipole tensor.

Projection methods may be used to approximate the density as a linear combination of atom-centered auxiliary basis functions;

$$\rho(\mathbf{r}) \approx \sum_c D_c\, \theta_c(\mathbf{r}). \tag{8}$$

Because the number of auxiliary basis functions is proportional to N, projection methods may yield a faster approach to the N^2 limit than standard direct methods. In Section 4, we review projection methods and describe our work involving potential-matched two-center fits of the density and atom-centered HGTF auxiliary functions.

Hermite Gaussian functions may be expressed in terms of other HGTFs with different exponents and with different centers using rapidly convergent series expansions. In the far-field, such translations are equivalent to the manipulations used in fast hierarchical multipole methods. Contrary to multipole translations, these Gaussian expansions are also convergent in the near-field and form the basis of the fast Gauss transform. In Section 5.1, we review classical hierarchical multipole methods and in Section 5.2 we look at adaptations of these methods to quantum chemistry. In Section 5.3, we introduce the fast Gauss transform, discuss its use in computing the near-field component of the Coulomb matrix, and point out its equivalence to multipole methods in the asymptotic limit. In Section 5.4 we discuss an implementation of hierarchical multipole methods in the *ab initio* electronic structure program SUPERMOLECULE[43] and report, for the first time, sub-N^2 scaling for computation of the Coulomb matrix.

2 Hermite Gaussian-type basis functions in quantum chemistry

Methods for the calculation of ERIs between Cartesian Gaussian-type distributions have been well developed in the last two decades. The McMurchie and Davidson (MD) algorithm[44,45,46] was one of the first methods to exploit the advantages of using HGTFs in the GPT expansion Eq. (7). In the MD algorithm, the products of CGTFs describing elementary charge distributions are expanded in HGTFs as

$$\rho_{ab}(\mathbf{r}) \equiv \phi_a(\mathbf{r})\,\phi_b(\mathbf{r}) = \sum_{L}^{l_a+l_b} \sum_{M}^{m_a+m_b} \sum_{N}^{n_a+n_b} e^{ab}_{LMN} \Lambda^{p}_{LMN}(\mathbf{r}) \ , \tag{9}$$

where the e^{ab}_{LMN} are HGTF expansion coefficients multiplied by radial overlaps and the HGTFs are defined by

$$\begin{aligned}
\Lambda^{p}_{LMN}(\mathbf{r}) &= \frac{\partial^L}{\partial P_x^L}\frac{\partial^M}{\partial P_y^M}\frac{\partial^N}{\partial P_z^N} \exp\left[-\zeta_p(\mathbf{r}-\mathbf{P})^2\right] \\
&= \zeta_p^{(L+M+N)/2} H_L\left[\zeta_p^{1/2}(x-P_x)\right] H_M\left[\zeta_p^{1/2}(y-P_y)\right] \\
&\qquad \times H_N\left[\zeta_p^{1/2}(z-P_z)\right] \exp\left[-\zeta_p(\mathbf{r}-\mathbf{P})^2\right] \ .
\end{aligned} \tag{10}$$

In Eq. (10), H_L is the L^{th} order Hermite polynomial

$$H_L(x) = \exp\left[x^2\right] \frac{\partial^L}{\partial x^L} \exp\left[-x^2\right] = \sum_{s=0}^{[L/2]} \frac{(-1)^s L!}{s!(L-2s)!}(2x)^{L-2s} \ . \tag{11}$$

The Hermite polynomials satisfy the relation

$$H_L(x) = 2xH_{L-1}(x) - 2(L-1)H_{L-2}(x) \ , \tag{12}$$

which forms the basis of the MD two-term recurrence relations for computation of the expansion coefficients e^{ab}_{LMN} as well as the MD three-term recurrence relations for calculation of HGTF-based ERIs.

2.1 The McMurchie-Davidson algorithm

By the GPT, the products of elementary CGTFs may be expressed as one-center distributions. In the MD algorithm, these one-center distributions are of HGTF form, rather than the CGTF expansions that were used in older algorithms. By analogy with Eq. (9), four-center CGTF-based ERIs may be transformed into linear combinations of the two-center HGTF-based ERIs

$$\left(\Lambda^{p}_{LMN}\middle|\Lambda^{q}_{L'M'N'}\right) = \int d\mathbf{r} \int d\mathbf{r}' \Lambda^{p}_{LMN}(\mathbf{r}) \left|\mathbf{r}-\mathbf{r}'\right|^{-1} \Lambda^{q}_{L'M'N'}(\mathbf{r}') \ . \tag{13}$$

The reduction to two-center ERIs is a direct consequence of the GPT. When HGTFs are used as a basis for the expansion, these ERIs may be further simplified by noting that $\partial\left|\mathbf{P}-\mathbf{Q}\right|/\partial P_i = -\partial\left|\mathbf{P}-\mathbf{Q}\right|/\partial Q_i$ and re-writing them as the one-center integrals

$$\begin{aligned}\left(\Lambda^{p}_{LMN}\middle|\Lambda^{q}_{L'M'N'}\right) &= (-1)^{L'+M'+N'} \left(\Lambda^{p}_{L+L',M+M',N+N'}\middle|\Lambda^{q}_{000}\right) \\ &= (-1)^{L'+M'+N'} u_{pq}\, R^{pq}_{L+L',M+M',N+N',0}\end{aligned} \tag{14}$$

where

$$R^{pq}_{lmn0} = \frac{\partial^l}{\partial PQ_x^l} \frac{\partial^m}{\partial PQ_y^m} \frac{\partial^n}{\partial PQ_z^n} F_0\left(t\right) \tag{15}$$

are "auxiliary functions",

$$u_{pq} = \frac{2\pi^{5/2}}{\zeta_p \zeta_q \sqrt{\zeta_p + \zeta_q}} \tag{16}$$

and PQ_x, PQ_y and PQ_z are the Cartesian components of $\mathbf{P}-\mathbf{Q}$. The auxiliary functions are related to the 0^{th} order reduced incomplete gamma function[47,48], which is

$$F_0(t) = \sqrt{\frac{\pi}{4t}}\, \mathrm{erf}\left(\sqrt{t}\right) = \int_0^1 e^{-tu^2} du \ , \tag{17}$$

with argument

$$t = \omega \left|\mathbf{P}-\mathbf{Q}\right|^2 \ , \tag{18}$$

where

$$\omega = \frac{\zeta_p \zeta_q}{\zeta_p + \zeta_q} \ . \tag{19}$$

Differentiation of $F_0(t)$ leads to higher order reduced incomplete gamma functions $F_{j+1}(t) = -F_j'(t)$, which we refer to hereafter simply as "gamma" functions.

For a class of HGTF-based ERIs with $L+M+N \leq \mathcal{L}_p$, $L'+M'+N' \leq \mathcal{L}_q$ and $\mathcal{L} = \mathcal{L}_p + \mathcal{L}_q$, the MD algorithm efficiently calculates all $\binom{\mathcal{L}+3}{3}$ auxiliary functions $R^{pq}_{l,m,n,0}$ with $l+m+n \leq \mathcal{L}$ from an initial set,

$$R^{pq}_{000j} = (-2\omega)^j F_j(t) \qquad (j = 0, 1, \ldots, \mathcal{L}), \tag{20}$$

using three-term recursion relations that derive from Eq. (12) [44,45,46] and which involve work that scales as $\binom{\mathcal{L}+3}{4}$.

Since the total number of HGTF-based ERIs grows as $\binom{\mathcal{L}_p+3}{3} \times \binom{\mathcal{L}_q+3}{3}$, the cost per HGTF-based ERI decreases with increasing $\mathcal{L}_p$ and $\mathcal{L}_q$. In a conventional calculation, this economy is offset by the costly back-transformation to a Cartesian representation. For example, the standard MD algorithm for evaluating the Coulomb matrix involves tightly nested loops with high operation counts inside the four-fold loop over atom centers, which are necessary to accomplish the transformation from a HGTF to a CGTF representation of the integrals;

$$\begin{aligned} J_{ab} = \sum_{cd} P_{cd}\, u_{pq} \sum_{L}^{l_a+l_b} \sum_{M}^{m_a+m_b} \sum_{N}^{n_a+n_b} e^{ab}_{LMN} \\ \times \sum_{L'}^{l_c+l_d} \sum_{M'}^{m_c+m_d} \sum_{N'}^{n_c+n_d} (-1)^{L'+M'+N'} e^{cd}_{L'M'N'} R^{pq}_{L+L',M+M',N+N'} \,. \end{aligned} \tag{21}$$

The bulk of the computational work is therefore done at the level of transforming, handling, and processing individual integrals in the LCAO basis rather than evaluating the primary quantities in the basis of expansion functions. Methods for improving the evaluation and processing of integrals should therefore focus on simplifying, reducing or eliminating transformations between CGTF and HGTF representations.

2.2 The density in an HGTF basis

Representation of the density in an HGTF basis goes back to the early work of Dunlap, Connoly, and Sabin [49] and of Mintmire [50], where Hermite Gaussians were used to expand the electron density in the Xα method. Saunders [51] studied the Mulliken partition of the charge density and its expansion in HGTFs. HGTFs have also been used by Wheatley [52,53] under the name Gaussian multipoles(!) to represent the charge density in the modeling of penetration effects between molecules. More recently, Ahmadi and Almlöf [54] studied the Coulomb matrix in an HGTF basis; our discussion closely follows their work, although we use a slightly different notation.

Instead of the usual approach where four-center ERIs are generated and then subsequently contracted with the density matrix as in Eq. (21), efficiency can be obtained by first transforming the density to an HGTF basis using the expansion coefficients e^{ab}_{LMN};

$$E^{q}_{L'M'N'} = \sum_{cd} P_{cd}\, e^{cd}_{L'M'N'} \ . \tag{22}$$

Asymptotically, this step requires effort and memory proportional to N and avoids the costly transformation from HGTF to CGTF inside the four-fold loop over atom centers. Representation of the density and Fock matrices in an HGTF basis also allows dramatic simplifications when the LCAO basis set shares exponents between CGTFs of different angular symmetry. Examples of such "family" basis sets include the Pople *sp*-type functions [55], the well-tempered basis sets of Huzinaga and co-workers [56−59], the even-tempered basis sets of Bardo and Ruedenberg [60], the geometrical basis sets of Clementi and Corongiu [61], and Wilson's universal basis set [62,63].

The expansion of CGTF distributions leads to HGTF families that are all related to a "mother" (s-type) Gaussian by differentiation. Using family LCAO basis sets is economical because *all* CGTF distributions are related to *all* HGTFs in the corresponding family by Eq. (9). The benefits obtained with a family LCAO basis set are two-fold. First of all, the redundancies associated with equivalent combinations of CGTF angular symmetries may be eliminated. For example, p-p and s-d products that share the same exponent and center correspond to the same family of HGTFs and their contributions to the density may be summed. Secondly, as the MD algorithm computes all R_{lmn} with $l+m+n \le \mathcal{L}$, two-center integrals with $l+m+n < \mathcal{L}$ are obtained at no additional cost.

Our algorithm for computing the Coulomb matrix in a HGTF basis involves the accumulation of contributions to the HGTF family p from all HGTF families q that constitute the density;

$$\mathcal{J}^{p}_{LMN} = \sum_{q} u_{pq} \sum_{L'M'N'} (-1)^{L'+M'+N'} E^{q}_{L'M'N'} R^{pq}_{L+L',M+M',N+N'} \ . \tag{23}$$

All elements of the Coulomb matrix belonging to the CGTF family of distributions p may then be evaluated as

$$J_{ab} = \mathbf{e}^{ab} \cdot \boldsymbol{\mathcal{J}}^{p} \,, \quad J_{a'b'} = \mathbf{e}^{a'b'} \cdot \boldsymbol{\mathcal{J}}^{p} \,, \quad \dots \qquad (ab,\, a'b',\, \dots \in p) \,, \tag{24}$$

where we have conveniently arrayed elements of the vectors $\mathbf{e}^{ab}$ and $\boldsymbol{\mathcal{J}}^{p}$ using the canonical index

$$LMN = \binom{L+M+N+2}{3} + \binom{M+N+1}{2} + N + 1 \ . \tag{25}$$

The benefits of using a family LCAO basis should now be obvious. As the work associated with Eq. (22) and Eq. (24) takes place outside the four-fold loop over basis functions, the limiting step is the evaluation of Eq. (23), which is proportional to the square of the number of HGTF families. If family LCAO basis sets are employed, the total number of HGTF families can be reduced dramatically.

3 The Cartesian multipole expansion

In large molecular systems, the dominant component of the two-electron energy is due to classical electrostatics. When penetration between the two electron densities ρ_A and ρ_B is negligible, the electrostatic energy of their interaction

$$E_{\text{el}}^{AB} = \int d\mathbf{r} \int d\mathbf{r}' \rho_A(\mathbf{r}) \left|\mathbf{r}-\mathbf{r}'\right|^{-1} \rho_B(\mathbf{r}') \tag{26}$$

may be approximated accurately by developing $\left|\mathbf{r}-\mathbf{r}'\right|^{-1}$ in a multipole expansion. The basic mathematical apparatus for evaluating such expansions has been well established for some time, involving the development of $\left|\mathbf{r}-\mathbf{r}'\right|^{-1}$ in either spherical harmonics [64,65] or a Taylor series in Cartesian coordinates [66]. Taylor series expansion of $\left|\mathbf{r}-\mathbf{r}'\right|^{-1}$ about both $\mathbf{A}$ and $\mathbf{B}$ leads to

$$E_{\text{el}}^{AB} = \sum_{LMN} \sum_{L'M'N'} (-1)^{L'+M'+N'} Q_{LMN}^{A} T_{L+L',M+M',N+N'}^{AB} Q_{L'M'N'}^{B} \ , \tag{27}$$

where

$$Q_{LMN}^{A} = \frac{1}{L!M!N!} \int d\mathbf{r}\, x^L y^M z^N \rho_A(\mathbf{r}) \tag{28}$$

is an unmodified Cartesian multipole moment of rank $\ell_a = L+M+N$. In Eq. (27) and throughout, we shall assume that the multipole expansion is truncated at order $\mathcal{L} = \mathcal{L}_a + \mathcal{L}_b$, where $\mathcal{L}_a$ is the order of the series expansion about $\mathbf{A}$, and $\mathcal{L}_b$ is the order of the series expansion about $\mathbf{B}$.

Alternatively, Eq. (27) can be re-summed and expressed in terms of Buckingham's traceless multipole moments [66]

$$M_{LMN}^{A} = (-1)^{\ell_a} \frac{1}{\ell_a!} \int d\mathbf{r}\, \rho_A(\mathbf{r}) \ \left|\mathbf{r}\right|^{2\ell_a+1} \frac{\partial^{\ell_a}}{\partial x^L \partial y^M \partial z^N} \frac{1}{\left|\mathbf{r}\right|} \quad (\ell_a = L+M+N) \ , \tag{29}$$

which may be used in Eq. (27) by replacing Q_{LMN}^{A} with $2^{\ell_a} \ell_a! M_{LMN}^{A} / (2\ell_a)!$. In either case, elements of the Cartesian multipole tensor are

$$T_{lmn}^{AB} = \frac{\partial^{\ell}}{\partial AB_x^l \, \partial AB_y^m \, \partial AB_z^n} \frac{1}{\left|\mathbf{A}-\mathbf{B}\right|} \qquad (\ell = l+m+n) \, , \tag{30}$$

where AB_x, AB_y and AB_z are the Cartesian components of $\mathbf{A}-\mathbf{B}$.

In 1982, Cipriani and Silvi [67] deduced the expression

$$T^{AB}_{lmn} = \frac{(-1)^\ell l!m!n!}{2^\ell|\mathbf{A}-\mathbf{B}|^{\ell+1}} \sum_{s=0}^{[l/2]} \sum_{t=0}^{[m/2]} \sum_{u=0}^{[n/2]} \Bigg[\frac{(-1)^{s+t+u}\,(2\ell-2s-2t-2u)!}{s!t!u!(l-2s)!(m-2t)!(n-2u)!(\ell-s-t-u)!} \times \left(\frac{AB_x}{|\mathbf{A}-\mathbf{B}|}\right)^{l-2s} \left(\frac{AB_y}{|\mathbf{A}-\mathbf{B}|}\right)^{m-2t} \left(\frac{AB_z}{|\mathbf{A}-\mathbf{B}|}\right)^{n-2u} \Bigg] , \tag{31}$$

which is currently the basis for computing Cartesian multipole interactions through arbitrary order [68], and involves $\mathcal{O}(\mathcal{L}^6)$ work.

3.1 HGTF-based ERIs in the multipole limit

In the multipole limit, HGTFs are related to derivatives of the Dirac delta function by

$$\delta^p_{LMN}(\mathbf{r}) = \frac{\partial^L}{\partial P_x^L}\frac{\partial^M}{\partial P_y^M}\frac{\partial^N}{\partial P_z^N}\,\delta(\mathbf{r}-\mathbf{P}) = \lim_{\zeta_p\to\infty}\left(\frac{\zeta_p}{\pi}\right)^{3/2}\Lambda^p_{LMN}(\mathbf{r}) \ . \tag{32}$$

The derivatives δ^p_{LMN} obey the identity [69]

$$\int d\mathbf{r}\, f(\mathbf{r})\ \delta^p_{LMN}(\mathbf{r}) = (-1)^{L+M+N}\frac{\partial^L}{\partial P_x^L}\frac{\partial^M}{\partial P_y^M}\frac{\partial^N}{\partial P_z^N}\ f(\mathbf{P}) \ . \tag{33}$$

Substituting Eqs. (32) and (33) into Eq. (14) and taking the far-field (multipole) limit, the following relation between HGTF-based ERIs and elements of the Cartesian multipole tensor (T-tensor) can readily be obtained:

$$\left(\Lambda^p_{LMN}\middle|\Lambda^q_{L'M'N'}\right) = (-1)^{L'+M'+N'}\left(\frac{\pi}{\zeta_p}\right)^{3/2} T^{PQ}_{L+L',M+M',N+N'}\left(\frac{\pi}{\zeta_q}\right)^{3/2} . \tag{34}$$

In Section 2.2, we discussed the benefits of working in an HGTF basis. The introduction of the multipole limit would at first not seem important, since the evaluation of integrals over HGTFs is not significantly slower than evaluation of elements of the Cartesian multipole tensor (see the following Section). However, closer inspection of Eq. (34) reveals that in the asymptotic regime exponents are coupled only by elements of the T-tensor, which in turn depends only on the centers P and Q. We may therefore sum all moments corresponding to center Q,

$$\varrho^{Q}_{L'M'N'} = \sum_{q} \left(\frac{\pi}{\zeta_q}\right)^{3/2} \sum_{cd} P_{cd}\, e^{cd}_{L'M'N'} \ , \tag{35}$$

independently and accumulate the potential at center P due to the entire density as

$$\mathcal{J}^{P}_{LMN} = \sum_{Q} \sum_{L'M'N'} (-1)^{L'+M'+N'} \varrho^{Q}_{L'M'N'} T^{PQ}_{L+L',M+M',N+N'} \ . \tag{36}$$

All elements of the Coulomb matrix with distributions centered at P may then be computed as

$$J_{ab} = \varrho^{p} \mathbf{e}^{ab} \cdot \boldsymbol{\mathcal{J}}^{P} \ , \quad J_{a'b'} = \varrho^{p} \mathbf{e}^{a'b'} \cdot \boldsymbol{\mathcal{J}}^{P} \ , \quad \ldots \qquad (ab,\, a'b',\, \ldots \in p) \ , \tag{37}$$

where $\varrho^{p} = (\pi/\zeta_p)^{3/2}$. As expected, this formalism is very similar to the relations obtained in Section 2.2 for the evaluation of the Coulomb matrix in a HGTF basis.

The multipole approximation allows summation of moments corresponding to different exponents but the same center. However, the advantages gained by the straightforward application of the multipole approximation are limited as multiple exponents will only rarely correspond to the same center, *e.g.* in the case of one-center distributions centered at nuclei or distributions located at the mid-point between equivalent nuclei. More important savings can be realized if the multipole moments are translated to common centers and combined; such an approach forms the basis of the hierarchical multipole methods discussed in Section 5.

3.2 Recurrence relations for computation of the Cartesian multipole tensor

Equation (34) identifies elements of the Cartesian multipole tensor within the MD 3-term recursion relations; in particular, by inspection of Eqs. (14), (20) and (34) the correspondence

$$T^{PQ}_{000j} \Longleftrightarrow (\zeta_p \zeta_q)^{3/2} \pi^{-3} u_{pq} R^{pq}_{000j} \tag{38}$$

is established. Substituting the multipole limit of the gamma functions,

$$F_j\,(t) = \int_0^{\infty} u^{2m} e^{-tu^2} du = \frac{\pi^{1/2}(2j-1)!!}{2^{(j+1)} t^{(j+1/2)}} \qquad (t \to \infty) \ , \tag{39}$$

into Eq. (38), we find that, from an initial set of auxiliary tensor elements

$$T^{PQ}_{000j} = (-1)^j (2j-1)!! \, |\mathbf{P}-\mathbf{Q}|^{-(2j+1)} \qquad (j = 0, 1, \ldots, \mathcal{L}) \ , \tag{40}$$

all $\binom{\mathcal{L}+3}{3}$ elements of the Cartesian multipole tensor T_{lmn0} with $l+m+n \leq \mathcal{L}$ may be calculated with the 3-term recursion relations

$$\begin{aligned}
T_{l,0,0,j} &= PQ_x T_{l-1,0,0,j+1} + (l-1) T_{l-2,0,0,j+1} \ , & (41) \\
T_{l,m,0,j} &= PQ_y T_{l,m-1,0,j+1} + (m-1) T_{l,m-2,0,j+1} \ , & (42) \\
T_{l,m,n,j} &= PQ_z T_{l,m,n-1,j+1} + (n-1) T_{l,m,n-2,j+1} \ , & (43)
\end{aligned}$$

where for convenience we have dropped the superscript PQ. The appropriate control structures for implementation of Eqs. (41–43) are outlined in Ref. [44].

Hermite functions are also intermediates in the van der Velde algorithm [70,71,46] for the calculation of CGTF-based ERIs. Both the van der Velde algorithm and the Cipriani and Silvi formula, Eq. (31), are based on the explicit form of the Hermite polynomials, Eq. (11), rather than the recursion relations given in Eq. (12). Therefore, there is a direct correspondence in the multipole limit between the van der Velde algorithm and the Cipriani and Silvi formula.

Computing elements of the Cartesian multipole tensor with Eqs. (41–43) is superior to current methods [68] based on Eq. (31) because:

a) Evaluation of all Cartesian multipole tensor elements through order $\mathcal{L}$ scales as $\mathcal{L}^4$, as compared to $\mathcal{L}^6$ with the Cipriani and Silvi equation.

b) Storage of intermediate values is not required.

c) The dominant portion of the work may be performed in vector loops of *average* length $\frac{1}{4}\binom{\mathcal{L}+3}{3}$, resulting in enhanced performance on vector computers and RISC-based platforms alike [72].

We have implemented two approaches to calculation of the Cartesian multipole tensor (T-tensor). The first approach involves three control structures corresponding to Eqs. (41–43). As in standard implementations of the MD algorithm, our first method involves an innermost loop, which for fixed $\mathcal{L}_p$ and $\mathcal{L}_q$ allows vectorization over lists of PQ_x, PQ_y, and PQ_z. We will refer to this method as a "list-directed" build.

In many applications it may not be possible to build long lists, particularly for large $\mathcal{L}$. We have developed a second approach to address this issue which allows vectorization over the indices LMN in the calculation of individual T-tensors. In this approach, which we refer to as "index collapsing" [73], the nested loops over the indices L, M, and N in the three control structures are first

Table 1: Performance and timings for computation and contraction of the Cartesian multipole tensor with $\mathcal{L}_a = 0$. For the list directed method vectors of length 128 were used. All timings are single processor Cray C-90 cpu seconds.

	List Directed		Index Collapsing	
$\mathcal{L}_b$	MFLOPs	cpu sec	MFLOPs	cpu sec
4	304.5	1.1e-6	46.9	1.5e-5
6	304.7	2.7e-6	93.2	2.2e-5
8	329.8	6.0e-6	180.7	3.1e-5
10	334.5	1.2e-5	205.5	4.4e-5
12	338.0	2.0e-5	255.8	6.3e-5
14	342.4	3.2e-5	309.7	8.6e-5

coalesced [74] to provide an outer loop over j and a single inner loop representing the recursion over L, M, and N. The three indices of the inner loop are then collapsed into a single index LMN to provide long vector lengths and reduced memory conflicts. This approach involves vector loops with average length equal to $\frac{1}{4}\binom{\mathcal{L}+3}{3}$.

In addition to the calculation of the Cartesian multipole tensor, it is important to perform its contraction with the multipole moments efficiently; using $\mathbf{P}-\mathbf{Q}$ and $(-1)^{L+M+N}$ or $\mathbf{Q}-\mathbf{P}$ and $(-1)^{L'+M'+N'}$ in Eq. (27) we place the longest contraction in the innermost loop. A long vector loop over the collapsed index $L+L'$, $M+M'$, and $N+N'$ may be then used as before.

Timings and execution rates of our two implementations are listed in Table 1. The increase in MFLOPs seen in the list directed method is due to the increasing dominance of the T-tensor build over the multipole contraction step. Currently, the contraction step operates at $\sim$ 200 MFLOPs, while the T-tensor build runs at $\sim$ 420 MFLOPs. For large $\mathcal{L}$, the index collapsing method is only three times slower than the optimal list directed method. Complete details of these developments and more extensive timings will appear elsewhere [72].

4 Projection methods for approximation of the electron density

Projection methods were originally developed for the approximation of two-electron integrals by Boys and Shavitt [75], Mulliken [76], Ruedenberg [77], Löwdin [78], Cizek [79], Harris and Rein [80]. We refer the reader to Ref. [81] for a complete review of the early literature, and begin our discussion with the projection of the differential diatomic overlap (PDDO) method of Newton *et al.* [82,83,84] and the limited expansion of diatomic overlap (LEDO) method due to Billingsley and Bloor [81]. The PDDO and LEDO methods were developed to simplify and economize MO based methods by projecting two-center distributions onto auxiliary functions centered on the two progenital atoms of that distribution;

$$\rho_{ab} \approx \widetilde{\rho}_{ab} = \mathbf{d}^a \cdot \boldsymbol{\theta}_A + \mathbf{d}^b \cdot \boldsymbol{\theta}_B \ . \tag{44}$$

In Eq. (44), $\mathbf{d}^a$ is a vector of expansion coefficients corresponding to the "a part" of distribution ρ_{ab} and the list of auxiliary functions $\boldsymbol{\theta}_A$ centered on atom A.

The most important distinction between the PDDO and the LEDO methods is that the PDDO method uses a least square fit involving overlap integrals (the L_2 norm) and the LEDO method employs the physically more appropriate Coulomb norm attributed to the early work of Harris and Rein [80]. Okniński [85] noted that apart from this difference, the two approaches are just generalizations of the same variational principle. These issues are discussed further in Section 4.1.

The auxiliary basis functions used in the PDDO and LEDO methods consist of the original one-center LCAO distributions supplemented with additional functions of higher angular symmetry. The LEDO method was abandoned by Billingsley due to ill-conditioned matrix equations and difficulties associated with approximation of tight two-center distributions [86]. Baerends, Ellis and Ros [87] extended the PDDO method to approximation of the two-center component of the charge density, but found that both accuracy and economy could be obtained using optimized auxiliary basis functions. Generalizations of this approach, in the sense of Okniński [85], will be referred to as two-center fits (TCFs) in the following.

Dunlap, Connolly and Sabin [88] later gained significant improvements in accuracy with the introduction of global expansion methods (GEMs) together with the Coulomb norm. A global expansion is one in which atom-centered auxiliary functions on every center can contribute to the representation of every distribution;

$$\rho_{ab} \approx \widetilde{\rho}_{ab} = \mathbf{d}^{ab} \cdot \boldsymbol{\theta} \,, \tag{45}$$

where $\boldsymbol{\theta} = \{\boldsymbol{\theta}_1, \ldots \boldsymbol{\theta}_{N_{\mathrm{ats}}}\}$ is a list of *all* atom-centered auxiliary functions used in the calculation (N_{ats} being the total number of atoms in the system), and $\mathbf{d}^{ab}$ is the corresponding list of expansion coefficients. The auxiliary basis functions originally used by Dunlap, Connolly and Sabin were obtained by augmenting the basis with functions of higher angular symmetry, and by doubling the original LCAO exponents, in order to obtain an accurate fit for same-center distributions. These authors also noted that minimizing residuals of the density in the Coulomb norm is equivalent to a least squares fit of the electric field.

In 1980, Campbell, Hillier and Saunders [89] introduced the core-valence approximation (CVA),

$$\rho_{cv} \approx \widetilde{\rho}_{cv} = \mathbf{d}^{cv} \cdot \boldsymbol{\theta}_{C} \,, \tag{46}$$

in which tight distributions arising from the product of a core (tight) and valence (diffuse) basis function are expanded about $\mathbf{C}$ using an auxiliary basis of CGTFs with higher angular symmetry functions involving the LCAO core exponent. Expansion coefficients in the CVA are determined by requiring equivalence of low order multipole moments. The multiplicative integral approximation (MIA) introduced by Van Alsenoy [90] extended the applicability of the CVA by introducing a more flexible one-center auxiliary basis. The MIA approximates distributions involving one tight and one diffuse basis function with a least squares fit and constraints requiring conservation of monopole and dipole moments. Although the CVA and the MIA are not appropriate for approximation of the total density, significant speedups may be obtained that are comparable to those achieved by the use of pseudopotentials.

GEMs have gained wide popularity for fitting the electron charge density in density functional theory [91,92,93], as well as for approximation of four-center ERIs in MO theory [94–97]. They have been especially useful in molecular orbital treatments of electron correlation as they simplify AO to MO transformations [98,99]. Several implementations [100,94,96] of correlated methods have utilized the fact that quantities with at most three AO indices appear in these global methods, and their transformation to a MO basis thus scales no worse than N^4.

Two common criticisms of GEMs are that the choice of auxiliary basis set is non-trivial and that reliable error estimates are not available [32]. However, Eichkorn *et al.* [93] have developed contracted auxiliary basis sets for use in GEMs that incur errors in the total energy which are often less than 0.1 mE_{h} per atom. These results suggest that the prevailing pessimism regarding the accuracy obtainable with GEMs may soon be dispelled.

4.1 Variational equations, constraints and measure

A projection method is defined by an auxiliary basis and an operator $\widehat{\mathcal{O}}$ that provides a measure between exact and approximate distributions. The goal of the fitting process is to reduce a residual function Δ through minimization of the quantity $(\Delta|\mathcal{O}|\Delta)$ for a given operator $\widehat{\mathcal{O}}$. The fitting can be made for the total density using $\Delta = \rho - \widetilde{\rho}$, or for each individual density distribution with $\Delta = \rho_{ab} - \widetilde{\rho}_{ab}$. The former case corresponds to the GEMs, and the latter to the TCF. Solving the extremal equation $\delta\,(\Delta|\mathcal{O}|\Delta) = 0$ yields expansion coefficients for the TCF

$$\mathbf{d}^{ab} = \left(\boldsymbol{\theta}_{AB} \left| \hat{\mathcal{O}} \right| \boldsymbol{\theta}_{AB}\right)^{-1} \cdot \left(\boldsymbol{\theta}_{AB} \left| \hat{\mathcal{O}} \right| \rho_{ab}\right) \ , \tag{47}$$

or equivalently the GEM expansion coefficients

$$\mathbf{D} = \left(\boldsymbol{\theta} \left| \hat{\mathcal{O}} \right| \boldsymbol{\theta}\right)^{-1} \cdot \left(\boldsymbol{\theta} \left| \hat{\mathcal{O}} \right| \rho\right) \ . \tag{48}$$

In Eq. (48), $\mathbf{D}$ is a vector of expansion coefficients corresponding to the list $\boldsymbol{\theta}$ of auxiliary functions on every center; $\widetilde{\rho} = \mathbf{D} \cdot \boldsymbol{\theta}$. With the TCF, the approximate two-center component of the density may be constructed iteratively by accumulating projections from each unique pair of atoms as

$$\mathbf{D}^{2c} \leftarrow \mathbf{D}^{2c} + \sum_{ab} P_{ab} \mathbf{d}^{ab} \qquad (a \in A, b \in B, A \neq B) \ . \tag{49}$$

In practice, solving equation Eq. (47) or Eq. (48) is computationally intensive because the matrices of two-center integrals are typically ill-conditioned due to linear dependencies. An eigenvector decomposition usually has to be performed and small eigenvalues projected out; this approach is essentially a singular value decomposition (SVD) of a square matrix [81,101]. In the case of global expansion methods the SVD requires $\mathcal{O}(N_{\text{ats}}^3 N_{\text{aux}}^3)$ work and memory (N_{ats} is the number of atoms, N_{aux}^3 the number of auxiliary basis functions) while with the TCF $\mathcal{O}(N_{\text{ats}} N_{\text{aux}}^3)$ work and $\mathcal{O}(N_{\text{aux}}^3)$ memory is required.

When fitting the density, or its two-center component, there may be first-order properties of the original "exact" density that are important to conserve. For instance, it is often desirable to have the approximate density conserve the total charge of the exact one, a criterion that is met if

$$\int d\mathbf{r} \Delta(\mathbf{r}) = 0. \tag{50}$$

In addition to charge conservation, constraining the density to yield correct multipole moments is often useful [89,90]. These and other criteria can easily be enforced by the method of Lagrange multipliers.

In addition to selecting the constraints, the expansion basis set, and choosing between a local or a global fit, one also has a choice in the operator $\widehat{\mathcal{O}}$. Clearly, that choice will affect the accuracy of the expansion. Choosing $\widehat{\mathcal{O}} = \delta\left(|\mathbf{r} - \mathbf{r}'|\right)$ results in an approximate density having optimal overlap with the exact one, and is equivalent to a fit of the density in a conventional least squares sense; we shall refer to such an approach as the S approximation.

A better choice for $\widehat{\mathcal{O}}$ is the Coulomb operator $\widehat{\mathcal{O}} = |\mathbf{r}-\mathbf{r}'|^{-1}$, leading to an expansion which we refer to as the V approximation. The V approximation is equivalent to a least squares fit of the electric field [88]. In the context of approximate ERIs, the V approximation typically gives better results than the S approximation because it specifically minimizes the error in a more appropriate norm [88,95].

An alternative measure has been introduced by Gill *et al.* [102] that employs the distance operator $\widehat{D} = |\mathbf{r}-\mathbf{r}'|$. Constraining the charge as in Eq. (50) and choosing $\widehat{\mathcal{O}} = \widehat{D}$, yields an approximate density optimal with respect to the electric potential

$$V_q(\mathbf{r}) = \int d\mathbf{r}' \, |\mathbf{r}-\mathbf{r}'|^{-1} \, \rho_q(\mathbf{r}') \ . \tag{51}$$

This approach, which we shall refer to as the D approximation, is expected to yield more accurate ERIs as the error to first order

$$(\delta\rho_p|\delta\rho_q) \approx \int d\mathbf{r} \, \rho_p(\mathbf{r})\delta V_q(\mathbf{r}) + \int d\mathbf{r} \, \rho_q(\mathbf{r})\delta V_p(\mathbf{r}) \tag{52}$$

involves one less integration over an approximate quantity than with the V approximation.

4.2 Two-center fits

In both the GEM and TCF, the number of auxiliary basis functions is proportional (3–10 times) to the number of LCAO basis functions. Solving the SVD associated with the GEM expansion coefficients given by Eq. (48) therefore requires $\mathcal{O}(N^3)$ work and memory, which may be prohibitive for very large systems. Also, errors due to the necessary elimination of near-zero eigenvalues will grow linearly with system size. Perhaps most importantly, the work required to both fit the density and to evaluate the Coulomb matrix is $\mathcal{O}(N^2)$. While fast methods (see Section 5) may be used to penetrate the N^2 limit for Coulomb matrix evaluation, there is no known way to combine such methods with GEMs in fitting the density; accordingly GEMs are not expected to be competitive for very large systems.

In the following, we consider TCFs as an alternative to GEMs for approximating the density with atom-centered auxiliary functions. The TCF methods have several computational advantages over GEMs: First of all, the computation of integrals required for the fitting scales linearly with thresholding; if distributions are of negligible magnitude, there is no need to project them onto auxiliary functions. Secondly, solving the TCF SVDs involves $\mathcal{O}(N_{\rm ats}N_{\rm aux}^3)$ work and requires only $N_{\rm aux}^3$ memory. Finally, by solving only small SVDs, the accumulation of errors due to the elimination of small eigenvalues can be reduced.

Evaluation of the TCF Coulomb matrix

One-center CGTF distributions and the portions of the density they represent may be expanded exactly in atom-centered HGTFs with the two-term McMurchie-Davidson recursion relations [45]. Although these one-center distributions could also be projected onto an auxiliary set of HGTFs, the error resulting from such an approximation is unjustified, in so far as same-center distributions describe the majority of the electron density, and because computations involving two-center distributions will asymptotically become dominant.

We write one-center distributions as

$$\rho_{aa'}(\mathbf{r}) = \mathbf{e}^{aa'} \cdot \boldsymbol{\vartheta}_A(\mathbf{r}) \qquad (a, a' \in A) \ , \tag{53}$$

where $\mathbf{e}^{aa'}$ is a vector of canonically ordered MD expansion coefficients and $\boldsymbol{\vartheta}_A$ is the corresponding vector of atom-centered HGTFs. The portion of the density described by one-center distributions is

$$\rho^{1c}(\mathbf{r}) = \sum_A \mathbf{E}_A \cdot \boldsymbol{\vartheta}_A(\mathbf{r}) \ , \tag{54}$$

where

$$\mathbf{E}_A = \sum_{a,a' \in A} P_{aa'} \, \mathbf{e}^{aa'} \tag{55}$$

is a vector of MD expansion coefficients contracted with the density matrix.

In the TCF expansion, the remainder of the density is approximated by projecting two-center distributions onto auxiliary basis functions centered on the progenital atoms;

$$\rho_{ab}(\mathbf{r}) \approx \widetilde{\rho}_{ab}(\mathbf{r}) = \mathbf{d}^a \cdot \boldsymbol{\theta}_A(\mathbf{r}) + \mathbf{d}^b \cdot \boldsymbol{\theta}_B(\mathbf{r}) \qquad (a \in A, \ b \in B) \ . \tag{56}$$

In Eq. (56), $\mathbf{d}^a$ and $\mathbf{d}^b$ are vectors of expansion coefficients corresponding to the atom-centered HGTFs $\boldsymbol{\theta}_A$ and $\boldsymbol{\theta}_B$, and are obtained by the projection methods described in Section 4.1. Note that with a global approximation, the two-term sum in Eq. (56) would be replaced by one extending over all centers, while with a MIA- or CVA-like approximation the sum would contain only one term. The two-center component of the electron density,

$$\rho^{2c}(\mathbf{r}) = \sum_{A \neq B} \sum_{a \in A} \sum_{b \in B} P_{ab} \phi_a(\mathbf{r}) \phi_b(\mathbf{r}) \ , \tag{57}$$

may be approximated as

$$\widetilde{\rho}^{2c}(\mathbf{r}) = \mathbf{D}^{2c} \cdot \boldsymbol{\theta}(\mathbf{r}) \tag{58}$$

where $\mathbf{D}^{2c}$ is obtained iteratively as described in Section 4.1.

In the approximation of two-center densities with atom-centered auxiliary functions, matrix elements of the Coulomb operator fall into two categories, those with exact one-center ⟨BRA| s;

$$\widetilde{J}_{aa'} = \sum_{C} \left[\mathbf{e}^{aa'} \cdot \left(\boldsymbol{\vartheta}_{A} \middle| \boldsymbol{\vartheta}_{C} \right) \cdot \mathbf{E}_{C} + \mathbf{e}^{aa'} \cdot \left(\boldsymbol{\vartheta}_{A} \middle| \boldsymbol{\theta}_{C} \right) \cdot \mathbf{D}_{C}^{2c} \right] \qquad (a, a' \in A) \ , \tag{59}$$

and those with approximate two-center ⟨BRA| s;

$$\begin{aligned} \widetilde{J}_{ab} = \sum_{C} \Big[& \mathbf{d}^{a} \cdot \left(\boldsymbol{\theta}_{A} \middle| \boldsymbol{\vartheta}_{C} \right) \cdot \mathbf{E}_{C} + \mathbf{d}^{b} \cdot \left(\boldsymbol{\theta}_{B} \middle| \boldsymbol{\vartheta}_{C} \right) \cdot \mathbf{E}_{C} \\ & + \mathbf{d}^{a} \cdot \left(\boldsymbol{\theta}_{A} \middle| \boldsymbol{\theta}_{C} \right) \cdot \mathbf{D}_{C}^{2c} + \mathbf{d}^{b} \cdot \left(\boldsymbol{\theta}_{B} \middle| \boldsymbol{\theta}_{C} \right) \cdot \mathbf{D}_{C}^{2c} \Big] \quad (a \in A,\, b \in B;\, A \neq B) \ . \end{aligned} \tag{60}$$

The TCF approximation has been implemented in our *ab initio* program SUPERMOLECULE[43]. We use Eq. (47) for approximation of two-center densities and Eqs. (59) and (60) for the approximate Coulomb matrix. The auxiliary basis sets are even-tempered and uncontracted, with exponents

$$\zeta_{A,k} = \zeta_{A}^{\min} \left(\frac{\zeta_{A}^{\max}}{\zeta_{A}^{\min}} \right)^{k/(N_s - 1)} \qquad (k = 0,\, 1,\, \ldots,\, N_s - 1) \ , \tag{61}$$

where N_s is the number of s-type auxiliary functions and $\zeta_{A}^{\max}$ and $\zeta_{A}^{\min}$ are empirically determined. We employ the notation $(\ell_1, \ell_2, \ldots, \ell_{N_s})$ for auxiliary basis sets, where ℓ_k is the maximum ℓ-value of the HGTF family k with exponent ζ_k.

All two- and three-center integrals are calculated using the McMurchie-Davidson (MD)[45] algorithm. Klopper and Röhse[103] have developed formulae for the two-electron integrals involving the $\widehat{D}$ operator in the context of explicitly correlated wavefunctions. Here, we have used a different formulation. Integrals involving $\widehat{D}$ and arbitrary HGTFs may be calculated using the MD algorithm with v_{pq} instead of u_{pq} and G_{000j} instead of R_{000j}, where

$$G_{000j} = \frac{(-2\omega)^j}{2j - 1} \left[(2j - 1 - 2t) F_j(t) - e^{-t} \right] \tag{62}$$

and

$$v_{pq} = \frac{\pi^{5/2} \left(\zeta_p + \zeta_q \right)^{1/2}}{\zeta_p^2 \zeta_q^2} \ . \tag{63}$$

Note that the $\widehat{D}$ operator does not define a proper metric, but that the other choices of $\widehat{O}$ that we have considered do. Accordingly, the D approximation requires the inversion (singular value decomposition) of the matrix $(\boldsymbol{\theta}|\widehat{D}|\boldsymbol{\theta})$ which is not positive-definite, and for which nothing can be expected with regard to the magnitude or sign of the eigenvalues. Furthermore, in the far-field limit

$$\left(\Lambda_{000}^{p}\left|\widehat{D}\right|\Lambda_{000}^{q}\right) = \frac{\pi^3}{(\zeta_p\zeta_q)^{3/2}}\,|\mathbf{P}-\mathbf{Q}| \qquad (|\mathbf{P}-\mathbf{Q}|\to\infty), \tag{64}$$

so that matrix elements actually grow with separation. This property makes the D approximation poorly suited for use in GEMs applied to large systems as it is certain to result in very ill-conditioned matrices. On the other hand, with a TCF, large separations may be avoided by thresholding small, insignificant distributions.

Solving the SVD for Eq. (47) in the V approximation, we discard all eigenvalues less than 10^{-14}, while with the D approximation we project out all eigenvalues with magnitude less than 2×10^{-8}.

In the following we will employ the mean deviation error,

$$\delta E_{\mathrm{MD}} = (1/2)\sum_{ab}\left|P_{ab}\left(\widetilde{J}_{ab}-J_{ab}\right)\right|, \tag{65}$$

which is a measure of auxiliary basis set quality.

Results

For the test molecule NCOH, auxiliary basis sets were optimized in the D approximation by the following method: Beginning with an auxiliary set denoted (6,6,6,6,6,6,6,6,6) in the nomenclature described above, δE_{MD} was minimized by varying $\zeta_{\min}$ and $\zeta_{\max}$ for each of the 3-21G, 6-31G and 6-31G** LCAO basis sets. We found the values $\zeta_{\min}=0.3$ and $\zeta_{\max}=66.0$ to be nearly optimal for all of the basis sets under consideration, although the minima are in general very shallow. For each original LCAO basis we performed an iterative refinement of the original 9×6 auxiliary basis sets by lowering the index ℓ_k corresponding to the smallest increase in the mean deviation error δE_{MD}, until a threshold is reached; here we have chosen 0.5 and 0.1 mE_{h}. The results of this refinement procedure are listed in Table 2.

With the LCAO[TCF] basis sets listed in Table 2, the total energies of the $\phi=0°$ and $\phi=180°$ glycine rotamers were calculated at geometries taken from Ref. [104]. The total energies for each rotamer, together with their differences, are listed in Table 3.

Table 2: The auxiliary HGTF basis sets with $\zeta_{\min} = 0.3$ and $\zeta_{\max} = 66.0$ obtained through iterative refinement and the corresponding mean deviation errors for CONH in the D approximation.

LCAO[TCF] Basis set	Heavy	Hydrogen	δE_{MD} (mE_h)
3-21G[A,D]	3 3 5 5 2 0 0 0 0	0 0 5 3 0 0 0 0 0	0.500
6-31G[A,D]	0 5 6 4 0 0 0 0 0	0 0 4 2 0 0 0 0 0	0.500
6-31G**[A,D]	0 5 6 4 0 0 0 0 0	0 0 6 2 0 0 0 0 0	0.500
3-21G[B,D]	4 5 6 5 3 2 0 0 0	0 2 6 3 2 0 0 0 0	0.098
6-31G[B,D]	5 5 6 5 3 0 0 0 0	0 4 6 5 0 0 0 0 0	0.099
6-31G**[B,D]	5 6 6 5 0 5 0 0 0	1 5 6 5 0 0 0 0 0	0.099

Table 3: Comparison of the Hartree-Fock LCAO and LCAO[TCF] results: Total and relative energies of two rotational isomers of glycine.[a]

	Energies in Hartree		
Basis set[b]	$E(\phi = 0°)^c$	$E(\phi = 180°)^c$	$E(0°) - E(180°)$
3-21G[A,V]	-281.249 183	-281.245 831	0.003 352
3-21G[A,D]	-281.247 584	-281.244 434	0.003 150
3-21G[B,V]	-281.247 388	-281.244 409	0.002 979
3-21G[B,D]	-281.247 304	-281.244 311	0.002 993
3-21G	-281.247 393	-281.244 434	0.002 959
6-31G[A,V]	-282.717 512	-282.713 412	0.004 100
6-31G[A,D]	-282.695 860	-282.691 707	0.004 153
6-31G[B,V]	-282.696 920	-282.693 705	0.003 215
6-31G[B,D]	-282.696 905	-282.693 697	0.003 208
6-31G	-282.696 834	-282.693 580	0.003 254
6-31G**[A,V]	-282.847 088	-282.843 430	0.003 658
6-31G**[A,D]	-282.843 345	-282.840 297	0.003 048
6-31G**[B,V]	-282.844 180	-282.841 648	0.002 532
6-31G**[B,D]	-282.844 432	-282.841 702	0.002 768
6-31G**	-282.844 469	-282.841 672	0.002 797

[a] Geometries taken from Ref. [104].
[b] Auxiliary basis sets described in text and listed in Table 2.
[c] ϕ is the N-C-C=O torsion angle.

Table 4: Comparison of the Hartree-Fock LCAO and LCAO[TCF] optimized geometries and corresponding total energies of benzene.

Basis set[a]	R_{CC}	R_{CH}	E_{HF}
3-21G[A,D]	2.614	2.027	-229.414 980
3-21G	2.616	2.026	-229.419 446
6-31G**[A,D]	2.615	2.034	-230.713 328
6-31G**	2.619	2.033	-230.713 860

[a] Auxiliary basis sets described in text and listed in Table 2.

We next calculated the geometries of benzene using the 3-21G[A,D] and 6-31G**[A,D] approximations and finite differences. The results of these calculations are presented in Table 4. With the V approximation, we were unable to obtain sufficient convergence of the density matrix to allow the numerically stable calculation of gradients and Hessian by finite differences, even when the better "B" auxiliary basis sets were used. This is presumably a result of a small but significant density "leakage", due to the fact that our approximate density was not constrained to conserve charge.

While optimization of our auxiliary basis sets certainly involves multiple minima and is molecule specific, our crude procedure outlines gross features which involve large-ℓ HGTFs with diffuse exponents. This conclusion is verified by the reasonable performance of our optimized auxiliary basis sets in the calculation of glycine total energies and benzene geometries.

The data presented in Table 3 indicate a superior accuracy of the D approximation with the small "A" auxiliary basis sets, while for the larger "B" auxiliary basis set the advantages are not well defined. One might assume that this is a result of using the D approximation in the optimization of auxiliary basis sets. However, optimization of auxiliary basis sets with the V approximation merely leads to larger auxiliary basis sets, and does not place different emphasis on their structure.

Discussion

We have reconsidered TCFs of the charge density and their utility for efficiently calculating the Coulomb matrix in extended systems. Historically, the LEDO approximation was considered untenable due to ill-conditioned matrix equations and difficulties associated with the approximation of tight distributions [86]. In our re-examination, we have considered auxiliary basis sets specifically optimized to reduce the error in TCF approximations. Based on our evaluation of auxiliary basis set requirements (see Table 2) we conclude, contrary to Billingsley and Trindle [86], that TCF methods are impractical for approximation of the entire two-center density because of the severe demands placed on HGTFs with diffuse exponents. Specifically, high-ℓ HGTFs with diffuse exponents are required to approximate distributions located between widely separated centers. These demands will increase for extended systems, being limited only by elements of the density matrix coupling diffuse functions and slowly decaying radial overlaps.

Global methods are capable of approximating diffuse distributions with less demand on the auxiliary basis set than TCFs. However, GEMs applied to extended systems perform excessive amounts of work, attempting to represent tight distributions with expansions in diffuse auxiliary functions and vice versa. Clearly, the shortcomings of TCFs and GEMs are directly related to the translation of distributions further than their extents allow, and to the expansion of distributions using auxiliary functions of very different extent. In other words, errors and inefficiencies are related to expansions of source distributions in inappropriate auxiliary functions. By inappropriate we mean auxiliary functions that

a) have exponents that differ greatly from the exponent of the source distribution, ζ, and

b) are at a distance Δ from the source distribution such that $\Delta\sqrt{\zeta} \gg 1$.

Based on these criteria, *we conclude that no single projection method is capable of efficiently approximating the entire density.* However, projection methods that approximate specific portions of the density corresponding to a given length scale, such as the CVA [89] or MIA [90], are certainly possible and are likely to be both efficient and accurate.

5 Fast methods in quantum chemistry

The simulation of a system of classical particles interacting via two-body forces is an N-body problem, where the direct computation of forces and energies requires $\mathcal{O}(N^2)$ work. A variety of fast hierarchical methods that significantly reduce the complexity of such N-body problems have been introduced [105–116] A common feature of these methods is that they subdivide space into a fixed hierarchy of cubes (cells), and rely on the same tree-like data structure. Starting with the root (parent), which contains all particles, space is recursively subdived into smaller cells (children). The complexity of the computation is reduced by approximating the distributions of particles within a cell as a series (multipole) expansion that converges rapidly in the far-field. Computations involving the hierarchy of cells are effective because the series expansions become increasingly more accurate with separation; larger cells may be used as the interaction distance increases.

Hierarchical methods may also be used to compute far-field components of the Coulomb matrix in quantum chemistry. However, unlike the classical N-body problem, quantum chemical methods involve continuous distributions. Integrals in the far-field regime (see Section 3) can be efficiently evaluated with multipole methods, while short-range interactions must be set apart and computed with methods that are valid in the near-field. The problem is complicated by the fact that the near/far distinction depends on the size of $t = \omega\,|\mathbf{P} - \mathbf{Q}|$, rather than on the separation $|\mathbf{P} - \mathbf{Q}|$ as in classical systems. However, working with continuous distributions also has its advantages; using the fast Gauss transform (FGT), which we introduce in Section 5.3, it is possible to represent cell distributions with series expansions that are rapidly convergent both locally and in the far-field. If a large number of distributions occupy a cell, their representation with the FGT may significantly reduce the complexity of computing both near- and far-field contributions to the Coulomb matrix.

5.1 Hierarchical methods for classical systems

Hierarchical multipole methods may be classified according to the type of multipole interactions computed. The two most common methods are the Barnes-Hut (BH) [105–111] and the fast multipole method (FMM) [112–116]. Both the original BH and FMM methods employ a BH tree, the leaves of which are cells ideally containing a single particle. The BH method uses low order Cartesian multipole approximations, while the FMM has traditionally relied on high order spherical tensors.

The BH method computes particle-cell multipole interactions, and employs a multipole acceptability criteria (MAC) [110,111] to determine their validity; failure to meet the MAC results in recursion down the (inverted) tree to smaller cells, which may or may not meet the MAC. Ultimately, the algorithm reaches the lowest level (leaves) and computes all remaining interactions directly.

The FMM is characterized by interactions between cells wherein a far-field multipole expansion representing the potential of one cell is transformed into a locally convergent Taylor series expansion in another, which is well separated from the first. Following these cell-cell interactions, the local Taylor expansions are translated down the tree, accumulating far-field potentials from all well separated cells. At the lowest level, the interaction between a particle and all well separated cells is evaluated with the Taylor series coefficients. Finally, remaining near-field interactions are evaluated directly.

Central to the FMM is the "well-separated" index (WSI), which for a given expansion order specifies the minimum number of intervening cells that guarantee convergence of all far-field expansions. Though rigorous, the FMM definition of what constitutes well-separated cells is based on a worst case error analysis of *all* the particles in a cell interacting with *all* the particles in another. On the other hand, the use of MACs in BH methods may lead to significantly tighter error estimates. While there are well known cases where MACs such as the BH opening criteria can lead to exploding galaxies and other (numerical) catastrophes [110], rigorous MACs suitable for use with Cartesian multipoles have recently been developed [111].

Although the FMM has traditionally made use of high-order expansions in spherical harmonics, several Cartesian variants, called cell multipole methods (CMM), have also been developed [117–120]. Like the BH methods, the CMMs make use of low order Cartesian multipole expansions. Also, CMMs do not employ rigorous error bounds, but rely on a WSI equivalent to the BH opening criteria [117].

An important concept in the discussion of a new hierarchical methods is the computational complexity, *i.e.* the asymptotic dependence of cpu time on system size. The BH algorithm is usually cited as $\mathcal{O}(N \log N)$ and the FMM as $\mathcal{O}(N)$. It is generally accepted that for three–dimensional systems the BH methods break even sooner than the FMM [108,121], while the FMM is superior for very large systems. A key factor in the derivation of the aforementioned complexity estimates is the assumption of homogeneity [122], under which all N particles are partitioned at a tree depth of $\log_8 N$. However, Aluru *et al.* [123] have argued that the complexity of both the BH and the FMM are distribution dependent; in the most favorable case they scale as $\mathcal{O}(N \log^4 N)$ and at worst they are unbounded. These authors argue that the FMM and BH method become unbounded when the BH tree involves long paths in which every cell contains the same set of particles. Such paths are the result of inhomogeneities, and may be arbitrarily large if the BH tree is extended to resolve all particles. On the other hand, if the BH tree is truncated before the particle distribution is well resolved, a substantial $\mathcal{O}(N^2)$ amount of work may be left to perform directly.

We disagree with the claim that the FMM involves a BH tree extending to the finest level of particle resolution and is therefore potentially unbounded. The FMM originally proposed by Greengard [122] involves recursion to a depth of $\log_8 N$, while most modern implementations of the FMM treat the BH tree depth as a tuning parameter [116,117,124]. Aluru *et al.* [123] have in fact developed distribution independent versions of the BH method and the FMM, although both methods are at best $\mathcal{O}(N \log^2 N)$.

5.2 Hierarchical methods in quantum chemistry

Hierarchical methods may be used also in quantum chemistry to compute the Coulomb matrix. Asymptotically, only $\mathcal{O}(N)$ distributions are significant to the density, and so hierarchical methods may, in principle, be used to compute all $\mathcal{O}(N)$ significant elements of the Coulomb matrix in $\mathcal{O}(N \log^2 N)$ time (using the best case estimates in [123]).

The Q-Chem group has developed the continuous fast multipole method (CFMM) [125], which is an adaptation of the Greengard FMM that accounts for the various distributions with a modified WSI. The CFMM WSI specifies the number of intervening boxes which certify both the absence of penetration effects and convergence of the multipole expansion. For a given cell, the CFMM allows the summation of multipole moments only for distributions with identical WSIs. As moments are passed up the tree (to larger cells) the WSI is halved and the multipole moments of more distributions may be summed. The CFMM employs a truncated BH tree, obtained by subdividing the parent cell so as to obtain children in which the number of particles are approximately independent of the total number [116,125]. The CFMM employs conventional integral methods for direct computations and a simple 10^{-10} cutoff to eliminate insignificant distributions.

Kutteh and Nichols have implemented the generalized cell multipole method (GCMM) [126] in the NWChem electronic structure program. The GCMM is based on the CMM, but is capable of handling moments higher than monopole [119,120]. The GCMM implementation in NWChem uses a density that has been fit to atom-centered auxiliary basis functions with the GEM described in Section 4, and computes the near-field integrals numerically by solving Poisson's equation, as discussed by Termath and Handy [32]. Like the CFMM, the GCMM also employs a truncated BH tree, which is determined by performance and accuracy issues.

As is evident from the preceding discussions, hierarchical methods for the simulation of classical point charges are still rapidly evolving. Often, implementations are highly platform and application dependent [111]. The incorporation of hierarchical methods into electronic structure programs is certain to add extra layers of complexity. This observation is underscored by the fact that, to date, no accounts of actual SCF calculations employing fast methods have been published, despite the introduction of these methods more than ten years ago.

The following is a list of issues that are likely to effect quantum chemical adaptations of classical hierarchical methods:

a) Distributions with large exponents cluster tightly about nuclear centers while diffuse distributions are spaced (more or less) uniformly. Thus, for large systems and good quality basis sets, *e.g.* with the largest exponents $\sim \mathcal{O}(10^4)$, generating a BH tree to partition all distributions leads to unbounded BH trees. Alternatively, truncating the BH tree prematurely can lead to cells that contain $\mathcal{O}(N)$ tight distributions and $\mathcal{O}(N^2)$ direct work.

b) Although the number of significant diffuse distributions making up the density scale as $\mathcal{O}(N)$, this asymptotic value is reached only slowly. In addition, applicability of the multipole approximation is limited by the most diffuse distribution in a cell. Thus, if only distributions with similar extents are combined the multipole approximation may be used more often, but with a concomitant increase in complexity.

c) With the exception of the GFMM, hierarchical multipole methods were designed for simulations involving point charges, for which the direct, near-field part of the computation is trivial. On the other hand, quantum chemical methods often employ basis functions of up to f-type and involve integrals that are non-trivial. It is therefore important that introduction of cell-based hierarchical methods do not impinge on the efficiency of computing near-field contributions.

5.3 The fast Gauss transform

The fast Gauss transform (FGT) [127,128,129] was originally developed to accelerate the numerical solution of the heat equation involving a Gaussian kernel. Like the BH and FMM, the FGT is a clustering technology that relies on rapidly convergent series to reduce complexity. Although originally employed as a device to accelerate quadrature, our adaptation preserves the integrity associated with the analytic evaluation of two-electron integrals.

Starting with the generating function of the Hermite polynomials,

$$\exp\left(2xt - t^2\right) = \sum_{k=0}^{\infty} \frac{t^k}{k!} H_k(x) \,, \tag{66}$$

we obtain the translation

$$\exp\left[-\zeta(r-q)^2\right] = \exp\left[-\zeta(r-c)^2\right] \sum_{k=0}^{\infty} \frac{(q-c)^k}{k!} \zeta^{k/2} H_k\left[\zeta^{1/2}(r-c)\right] \,, \tag{67}$$

which is both the Taylor and the Hermite expansion of a one dimensional s-type Gaussian. *The key feature of Eq. (67) is that this expansion converges rapidly both locally and in the far-field.* This expansion is illustrated in Figure 1.

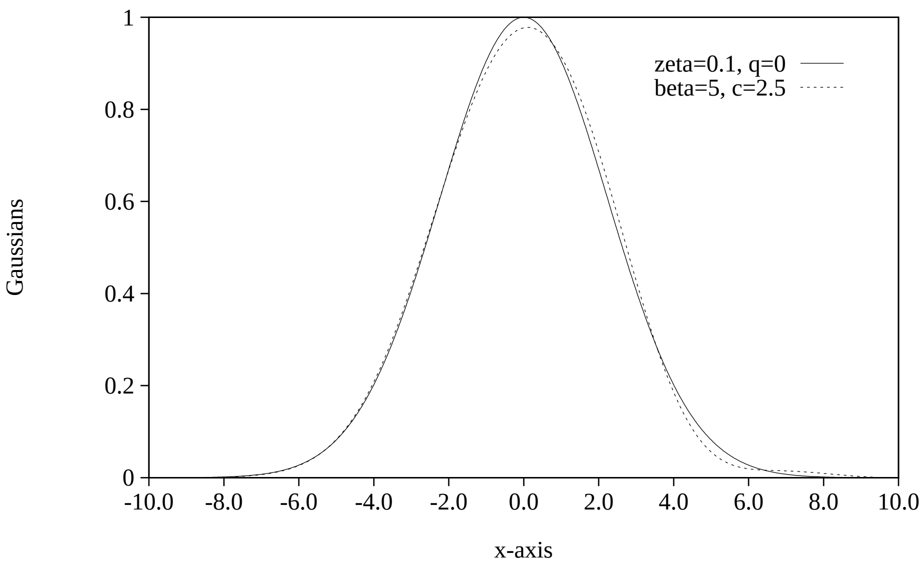

Figure 1: An unnormalized one-dimensional s-type Gaussian with $\zeta = 0.1$ centered at $q = 0$ and its expansion about $c = 2.5$ with $\beta = 5$.

Generalizing Eq. (67) to three dimensions, an s-type HGTF at $\mathbf{Q}$ may be expanded about $\mathbf{C}$ as

$$\Lambda^q_{000}(\mathbf{r}) = \sum_{uvw}{}' \left[(Q_x - C_x)^u (Q_y - C_y)^v (Q_z - C_z)^w \frac{\Lambda^c_{uvw}(\mathbf{r})}{u!v!w!}\right] + \epsilon^{q\to c}_{(0,\beta)}(\mathbf{r}) \,, \quad (68)$$

where the prime indicates the restricted summation $u+v+w \le \beta$ and $\epsilon^{q\to c}_{(0,\beta)}(\mathbf{r})$ is the truncation error. Similarly, we can expand arbitrary HGTFs as

$$\begin{aligned} \Lambda^q_{lmn}(\mathbf{r}) \quad &= \quad \sum_{uvw}{}' \left[(Q_x - C_x)^u (Q_y - C_y)^v (Q_z - C_z)^w \frac{\Lambda^c_{(l+u,m+v,n+w)}(\mathbf{r})}{u!v!w!}\right] \\ &\quad + \epsilon^{q\to c}_{(lmn,\beta)}(\mathbf{r}) \,. \end{aligned} \quad (69)$$

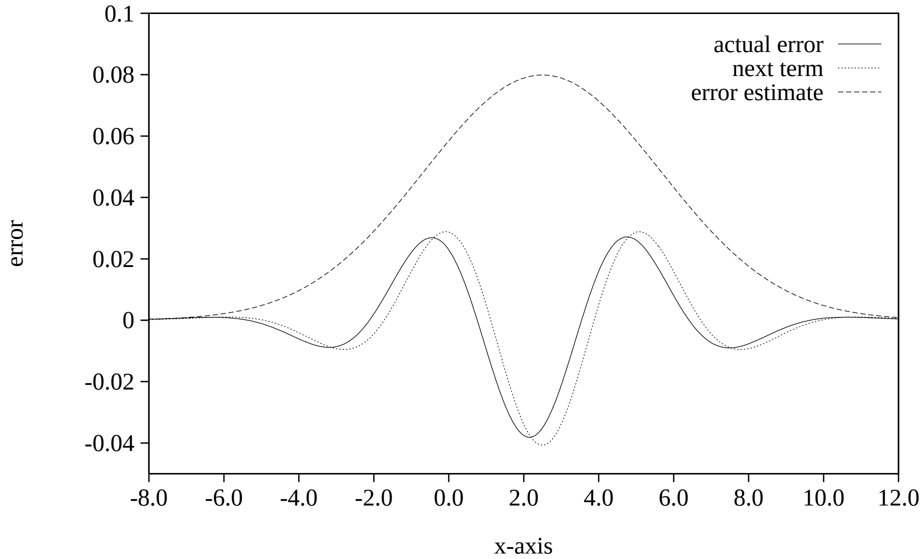

Figure 2: The actual error, the error estimate given by Eq. (73) and the next ($\beta + 1 = 6$) term corresponding to the expansion depicted in Figure 1.

Using Cramer's inequality [130]

$$H_\beta(x) \le K 2^{\beta/2} \sqrt{\beta!}\, e^{x^2/2} \qquad (K < 1.09) \tag{70}$$

and Stirling's approximation,

$$\beta! \approx \sqrt{2\pi\beta}\, (\beta/e)^\beta\,, \tag{71}$$

we obtain

$$\epsilon^{q\to c}_{(0,\beta)}(\mathbf{r}) < \frac{K^3}{[2\pi(\beta+1)/3]^{3/4}} \left(\frac{6e\,\zeta_q\,\Delta^2}{\beta+1}\right)^{\frac{\beta+1}{2}} \exp\left(-\frac{\zeta_q}{2}|\mathbf{r}-\mathbf{C}|^2\right)\,. \tag{72}$$

Equation (72) is an estimate of the error resulting from translation of a s-type Gaussian a distance less than Δ using an expansion truncated at order $u+v+w = \beta$. Equivalent error estimates may be obtained for higher-ℓ HGTFs, and show that such HGTFs are more difficult to translate. The quality of $\epsilon^{q\to c}_{(0,\beta)}$

can be gauged by Figure 2, where we show the error of the one-dimensional expansion in Figure 1, the next term in that expansion, and the error estimate

$$\epsilon_{(0,\beta)}^{q\to c}(x) < \frac{K}{[2\pi(\beta+1)]^{1/4}} \left(\frac{2e\,\zeta_q\,\Delta^2}{\beta+1}\right)^{\frac{\beta+1}{2}} \exp\left[-\frac{\zeta_q}{2}(x-c)^2\right], \tag{73}$$

which is the one-dimensional analog of Eq. (72).

The McMurchie-Davidson algorithm evaluates all integrals $(\Lambda^p_{LMN}|\Lambda^q_{L'M'N'})$ with $L+M+N \leq \mathcal{L}_p$ and $\ell_q = L'+M'+N' \leq \mathcal{L}_q$ in $\binom{\mathcal{L}_p+\mathcal{L}_q+3}{4}$ cpu time. If, for a given expansion length β and corresponding cell width 2Δ, $N_{\rm cell}$ ℓ_q-type HGTFs occupy a cell and

$$\binom{\mathcal{L}_p+\beta+3}{4} \ll \binom{\mathcal{L}_p+\mathcal{L}_q+3}{4} \times N_{\rm cell}\,, \tag{74}$$

then a significant reduction in cpu time may be realized by employing the rapidly convergent expansions given by Eqs. (68) and (69); this is our quantum chemical adaptation of the fast Gauss transform (FGT).

Multiresolution, wavelets and hierarchical multipole methods

For a given error threshold and expansion length, distributions with similar extents may be safely translated similar distances. Therefore, the application of the FGT suggests resolution of the density into distributions with similar extents and the association of each resolvent with a different sized cell.

Decomposition of the density into characteristic length scales has much in common with wavelet theory and multiresolution analysis [131,132]. By analogy (but not in the strict sense) we call this decomposition multiresolution. Hermite Gaussian functions are often called wavelets [133] or quasi-wavelets [134], and have many features common to wavelets. Wavelets are families of functions derived from a single function by dilation and translation, as are HGTFs. Indeed p-type functions are often referred to as Mexican-Hat wavelets [135]. Wavelets, and by association HGTFs, are connected with a variety of fast numerical methods [136] at a fundamental level, and may present important opportunities for future research.

For large systems, the diffuse component of the density involves distributions that persist even for widely separated centers. Likewise, very tight distributions also persist for widely separated centers, but cluster near nuclei. Also, integrals involving diffuse distributions enter the far-field regime slowly. Therefore, both large cells corresponding to small exponents and small cells (near nuclei) corresponding to large exponents may encompass sufficient complexity to warrant use of the FGT. Clearly, the use of the FGT is justified for very tight distributions, for which it should be possible to contain all $\mathcal{O}(N)$ of them in a few cells surrounding each center.

The FGT is based on the equivalence of Taylor's series and the Hermite expansion of a Gaussian. As HGTFs of higher angular symmetry correspond to derivatives of the Dirac delta function (see Section 3.1), the FGT reduces to a Cartesian multipole expansion in the far-field. This result, together with the aforementioned considerations suggest a synthesis of FGT and BH-like methods. A natural approach is to use FGT cells as nodes on a truncated BH tree, with the smallest cells corresponding to the largest exponents.

The FGT and secondary expansions

The FGT is certainly an interesting approach to the calculation of near-field contributions to the Coulomb matrix; it admits the systematic control of error and can yield large savings for well populated cells. The efficiency of the FGT can be increased dramatically with secondary expansions about new exponents, clustering distributions with closely related exponents, thereby increasing cell populations. One approach is to perform one-center projections of HGTFs onto auxiliary functions with the same angular symmetries. For an s-type function, such an expansion reads

$$\exp\left[-\zeta_q(\mathbf{r}-\mathbf{Q})^2\right] = \sum_i c_i \exp\left[-\alpha_i(\mathbf{r}-\mathbf{Q})^2\right] \ , \tag{75}$$

where the set of expansion exponents $\{\alpha_i\}$ has been chosen to at least span the level of resolution that ζ_q belongs to (as described above). The expansion coefficients c_i may be determined once and for all by projection using any of the operators and constraints described in Section 4.

Another approach is to Taylor expand distributions with similar extents about a nearby exponent α. For an s-type function such an expansion leads to the series

$$\exp\left[-\zeta_q(\mathbf{r}-\mathbf{Q})^2\right] = \sum_{l=0}^{n} \frac{(-\delta)^l}{l!}(\mathbf{r}-\mathbf{Q})^{2l} \exp\left[-\alpha(\mathbf{r}-\mathbf{Q})^2\right] \ , \tag{76}$$

which is rapidly convergent if $\delta = \zeta_q - \alpha$ is small. Yet a third approach is to perform one-center projections of HGTFs onto auxiliary functions with the same *and higher* angular symmetries, as for example

$$\exp\left[-\zeta_q(\mathbf{r}-\mathbf{Q})^2\right] = \sum_{l=0}^{n} \sum_i c_{l,i}(\mathbf{r}-\mathbf{Q})^{2l} \exp\left[-\alpha_i(\mathbf{r}-\mathbf{Q})^2\right] \ . \tag{77}$$

Unlike the projection methods considered in Section 4, the expansions (76-77) are rapidly convergent as the auxiliary function exponents $\{\alpha_i\}$ either span the source exponent ζ_q or are very close to it. These expansions involve CGTF functions that may be easily translated to another center with the FGT and re-summed, thereby increasing cell populations. The best choice of one-center expansion is not obvious and is complicated by the fact that higher-ℓ HGTFs are more difficult to translate with the FGT. There is clearly a trade-off between the number of exponents α_i and the value of n in Eq. (77), and the balance between the two can be tuned.

Secondary expansions are certain to be useful in at least two circumstances: First of all, tight distributions may be re-expanded about a "core" exponent as in the CVA [89] or MIA [90]. Secondly, all diffuse distributions may be expanded in just a few exponents. As the diffuse exponents also correspond to large cells, we expect that very large populations may be accumulated. This is significant as the diffuse component of the density is the most difficult to treat with multipole methods.

Using the third approach, Eq. (77), it appears possible to expand the entire density with fewer than 20 exponents and $n \leq 4$. If this preliminary observation should indeed turn out to be reliable, it may lead to a dramatic reduction of work as the number of different exponents in a typical calculation is usually of the order of several hundred, and a transformation to the auxiliary set $\{\alpha_i\}$ can reduce that number by more than an order of magnitude. However, there is a complicated interplay between such reductions and the increased work associated with translating the resulting high-ℓ HGTFs. This interplay is mediated by cell populations, which in turn are system and basis set dependent. Clearly, determining optimum secondary expansions for use with the FGT will require significant experimentation.

5.4 A quantum chemical tree code

We have recently developed a quantum chemical tree code, QCTC, based loosely on the Barnes-Hut hierarchical algorithm [105]. Although it does not yet perform the FGT to compute near-field integrals, QCTC has some interesting features which we list here:

a) When referencing the density or Coulomb matrix elements, the outer most loops are over atom pairs and the exponents of primitive distributions.

b) $\langle$BRA$|$ s and $|$KET$\rangle$ s (see below) are subject to independent thresholding at the level of primitive distributions.

c) Parent cells are created from their children by fusion; for a given level of resolution the smallest cell is a tuning parameter.

d) Elements of the density are stored in an HGTF basis and referenced with cell-based pointers.

e) The Cartesian multipole tensor is computed through arbitrary order with the methods described in Section 3.2.

f) Both $\langle$BRA$|$ and $|$KET$\rangle$ multipole moments are screened rank by rank as described in below.

We have implemented QCTC in the *ab initio* electronic structure program SUPERMOLECULE v1.1 [43]. In the spirit of other direct approaches, QCTC does not store any integrals, and recomputes all intermediates with every cycle. Also, as explained in the following two Sections, the accuracy of QCTC is controlled by a single thresholding parameter. Currently, QCTC incurs the lowest execution times for medium sized systems with a tree depth of 4, cells at the finest level of spatial resolution equal to 3.3 au and cell multipole expansions through order $\mathcal{L}_c = 5$.

Independent thresholding of distributions

Conventional direct methods control truncation errors in the Fock matrix by testing on the (estimated) product between two-electron integrals and their corresponding elements of the density matrix. However, in order to assign elements of the density to cells one must know *a priori* which distributions are significant. Also, thresholding at the level of primitive (or contracted) integrals introduces an element of overhead which may eventually dominate the entire calculation. It is therefore critical to gauge the significance of a $\langle$BRA$|$ primitive based on the total density rather than on the various contributions of standard integrals.

We wish to compute all elements J_{ab} of the Coulomb matrix to within a given accuracy while independently thresholding contributions of primitive distributions to the density and to the $\langle$BRA$|$ basis functions. To first order, two sources of error ϵ_1 and ϵ_2 are introduced by the independent thresholding of $\langle$BRA$|$ and $|$KET$\rangle$ distributions;

$$J_{ab} \approx \left(\rho_{ab} + \epsilon_1 \middle| \rho + \epsilon_2\right) . \tag{78}$$

Our objective is to limit the error in J_{ab} to within $\mathcal{O}(\texttt{thresh})$, where `thresh` is a thresholding parameter.

In the following, lower case Roman letters refer to basis functions, the corresponding Greek symbols refer to primitive functions contributing to those basis functions, and upper case Roman letters indicate the basis function centers.

Beginning with the familiar Schwartz inequality [16,137,138],

$$\left(\rho_{ab}\middle|\epsilon_2\right) \le \left(\epsilon_2\middle|\epsilon_2\right)^{1/2}\left(\rho_{ab}\middle|\rho_{ab}\right)^{1/2}, \tag{79}$$

we assume that the average component of ϵ_2 due to neglect of the primitive distribution $\rho_{\gamma\delta}$ is

$$N_{\rm prim}^{-1}\left(\epsilon_2\middle|\epsilon_2\right)^{1/2} < \max\left[|c_\gamma c_\delta P_{cd}|\left(\rho_{\gamma\delta}\middle|\rho_{\gamma\delta}\right)^{1/2}\right] \qquad \forall\,\gamma\in c, \delta\in d\,. \tag{80}$$

To arrive at Eq. (80), we have assumed that asymptotically there are $\mathcal{O}(N_{\rm prim})$ small distributions that are possibly significant to the density. Making the additional assumption that the largest contributions to the density derive from s-type functions, we arrive at the simple criterion for neglect of primitive contributions to the density involving the exponent $\zeta_q = \zeta_\gamma + \zeta_\delta$ and center Q;

$$\zeta_q^{-5/4}\exp\left[-\frac{\zeta_\gamma\zeta_\delta}{\zeta_\gamma+\zeta_\delta}(\mathbf{C}-\mathbf{D})^2\right]\max(|c_\gamma c_\delta P_{cd}|\,) \le \tau_1 \tag{81}$$

where the max is taken for all primitive values $\gamma\in c$ and $\delta\in d$, and where

$$\tau_1 = \texttt{thresh}\times\left\{2\pi^{5/2}N_{\rm prim}\max\left[|c_\alpha c_\beta|(\zeta_\alpha+\zeta_\beta)^{-5/4}\right]\right\}^{-1} \qquad \forall\,\alpha,\beta\,. \tag{82}$$

With family basis sets (see Section 2.2), a single distribution exponent may correspond to multiple basis functions of different angular symmetry; we therefore consider the maximum contraction coefficients and density matrix elements corresponding to *all* primitives that involve exponent ζ_q and center Q.

The neglect of $\langle$BRA$|$ primitives introduces the error ϵ_1, which from Holder's inequality [18,139] obeys the bound

$$\left(\epsilon_1\middle|\rho\right) \le N_{\rm el}\int d\mathbf{r}\,|\mathbf{r}|^{-1}\epsilon_1(\mathbf{r}). \tag{83}$$

Again assuming that the largest errors result from the neglect of s-type primitives one arrives at the simple expression

$$\frac{2\pi|c_\alpha c_\beta|}{\zeta_p}\exp\left[-\frac{\zeta_\alpha\zeta_\beta}{\zeta_\alpha+\zeta_\beta}(\mathbf{A}-\mathbf{B})^2\right] \le \tau_2 = \frac{\texttt{thresh}}{N_{\rm el}} \tag{84}$$

for neglect of $\langle$BRA$|$ primitives involving the exponent $\zeta_p = \zeta_\alpha + \zeta_\beta$ and center P.

In an incremental Fock build, $N_{\rm el}$ in Eq. (84) can be replaced with $N'_{\rm el}$, corresponding to the density matrix correction $\mathbf{\Delta P}$;

$$N'_{\rm el} = \sum_{ab} \Delta P_{ab} S_{ab} \ , \tag{85}$$

where S_{ab} are elements of the overlap matrix. As $\mathbf{\Delta P}$ becomes smaller, so too does $N'_{\rm el}$, and τ_2 may be relaxed accordingly.

Thresholding of multipolar interactions

After establishing that an integral is in the asymptotic far-field regime (see Section 3), it may be computed as

$$\left(\rho_{\alpha\beta}\middle|\rho_c\right) = \sum_{LMN} (-1)^{L+M+N} \varrho^p_{LMN} \sum_{L'M'N'} T^{CP}_{L+L',M+M',N+N'} \varrho^C_{L'M'N'} \ , \tag{86}$$

where ϱ^p_{LMN} is a rank $\ell_p = L+M+N$ multipole moment corresponding to the primitive distribution $\rho_{\alpha\beta}$ centered at $\mathbf{P}$, and $\varrho^C_{L'M'N'}$ is a rank $\ell_c = L'+M'+N'$ multipole moment corresponding to the density in cell c, centered at $\mathbf{C}$. As all elements of the rank ℓ Cartesian multipole tensor satisfy the bound

$$\left|T^{CP}_{lmn}\right| \le \ell! \, |\mathbf{C}-\mathbf{P}|^{-(\ell+1)} \qquad (\ell = l+m+n) \ , \tag{87}$$

we can determine not only admissibility of the multipole approximation, but we can threshold both $\langle$BRA$|$ and $|$KET$\rangle$ moments on the fly with the simple criterion

$$\varrho^p_{\ell_p} \varrho^C_{\ell_c} (\ell_p+\ell_c)! \, |\mathbf{C}-\mathbf{P}|^{-(\ell_p+\ell_c+1)} \le \frac{\texttt{thresh}}{n_{\rm cell}} \ , \tag{88}$$

where $\varrho^p_{\ell_p}$ and $\varrho^C_{\ell_c}$ are order of magnitude estimates, and $n_{\rm cell}$ is the number of occupied cells at a given level of spatial resolution. Thus, if $|\mathbf{C}-\mathbf{P}|$ is large, it may be sufficient to restrict the expansion in Eq.(86) to *e.g.* monopole-monopole interactions only, even if the original primitive $\langle$BRA$|$ distribution contains high-ℓ functions. Implicit in Eq. (88) is the assumption that the estimates $\varrho^p_{\ell_p}$ and $\varrho^C_{\ell_c}$ decay monotonically with rank; to guarantee this, these estimates are computed as the backwards sums of the maximum multipole moments in each rank.

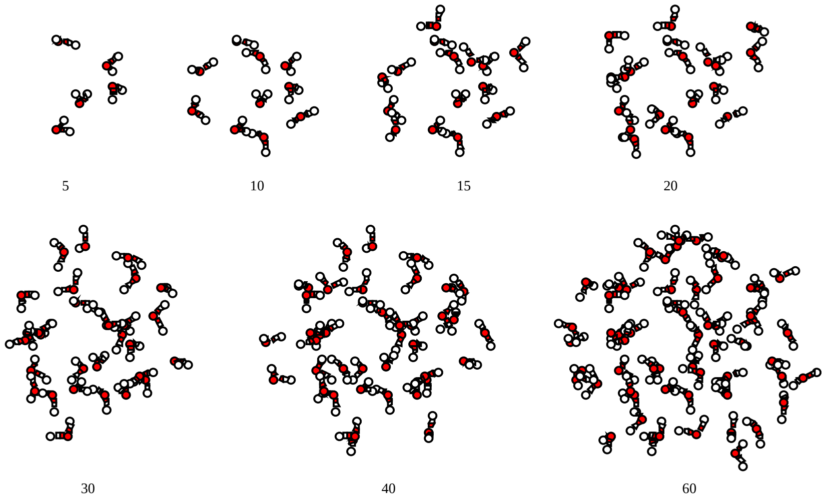

Figure 3: A sequence of water clusters taken from an equilibrated molecular dynamics simulation corresponding to 25 °C and 1 atm. All water molecules have the fixed geometry $\angle_{\mathrm{HOH}} = 104.52°$ and $R_{\mathrm{OH}} = 0.957\mathring{A}$

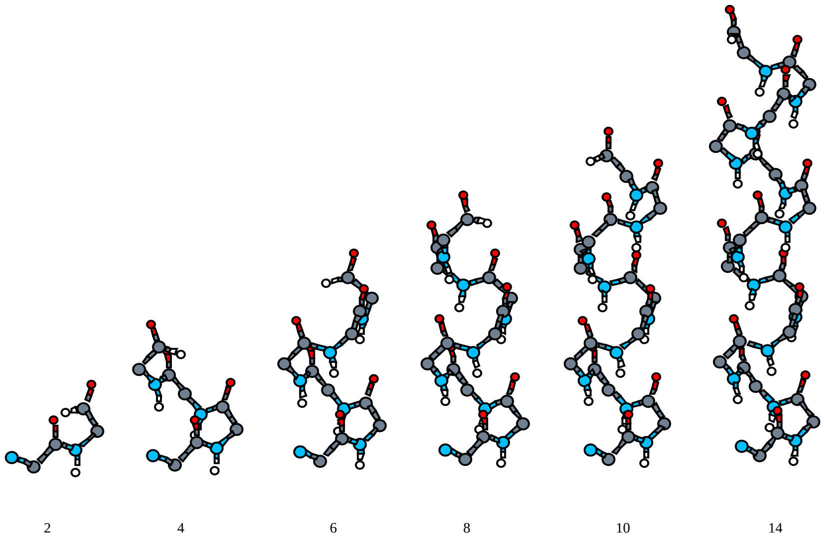

Figure 4: A sequence of polyglycine α-helices. Coordinates were obtained with the Insight II program[140].

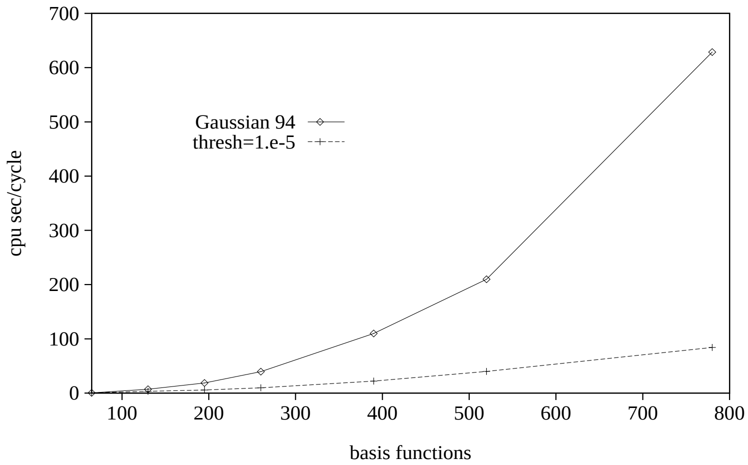

Figure 5: Average cpu times spent in LINK502 and average times for one QCTC build of the Coulomb matrix using **thresh** $= 10^{-5}$. Each data point corresponds to one structure in the sequence of water clusters shown in Figure 3.

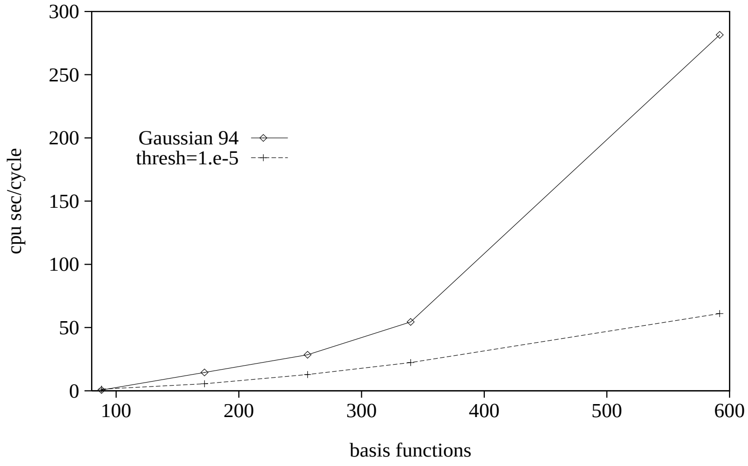

Figure 6: Average cpu times spent in LINK502 and average timings for one QCTC build of the Coulomb matrix using **thresh** $= 10^{-5}$. Each data point corresponds to one structure in the sequence of polyglycine α-helices shown in Figure 4.

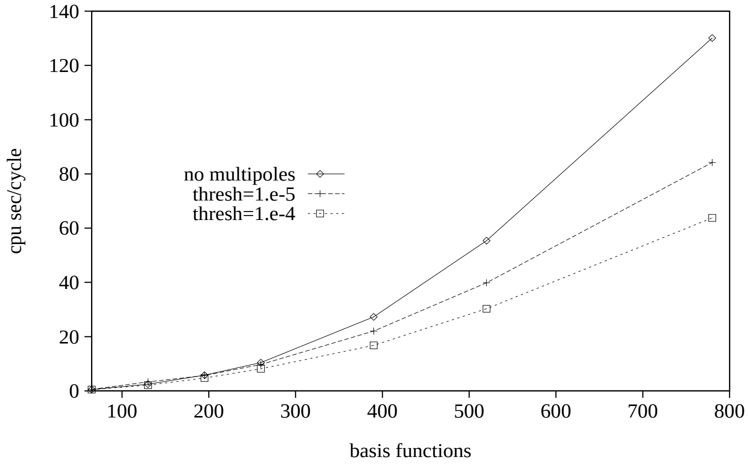

Figure 7: Average QCTC timings for the sequence of water clusters using **thresh** $= 10^{-5}$, **thresh** $= 10^{-4}$, and **thresh** $= 10^{-5}$ but no multipoles.

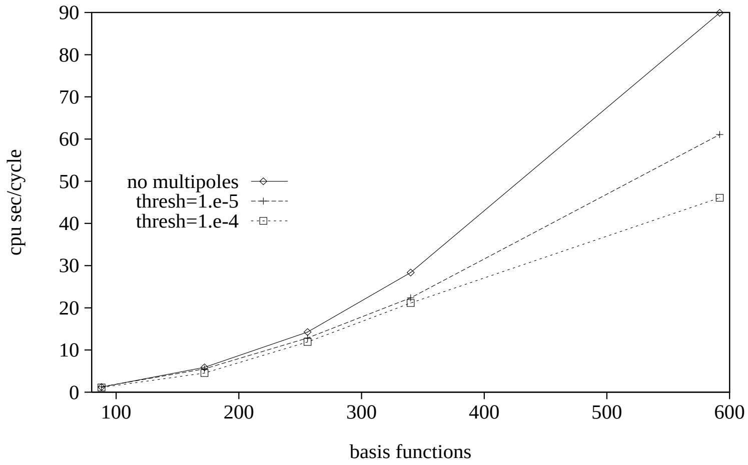

Figure 8: Average QCTC timings for the sequence of α-helices using **thresh** $= 10^{-5}$, **thresh** $= 10^{-4}$, and **thresh** $= 10^{-5}$ but no multipoles.

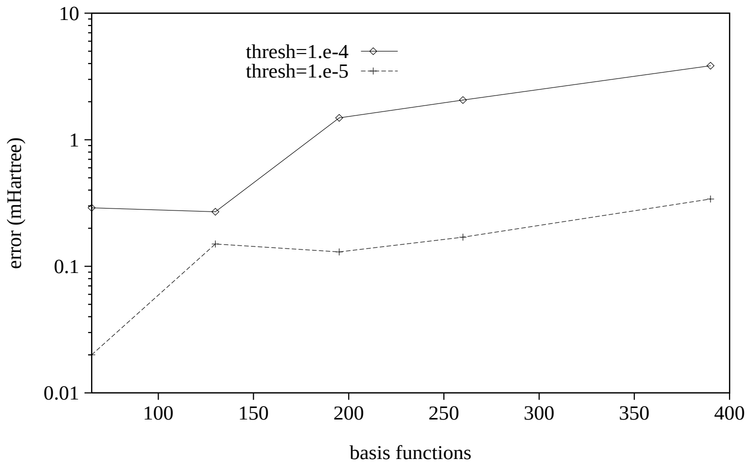

Figure 9: Errors in the converged HF/3-21G total energies corresponding to water clusters 5, 10, 15, 20 and 30 shown in Figure 3. The corresponding timings are shown in Figure 7

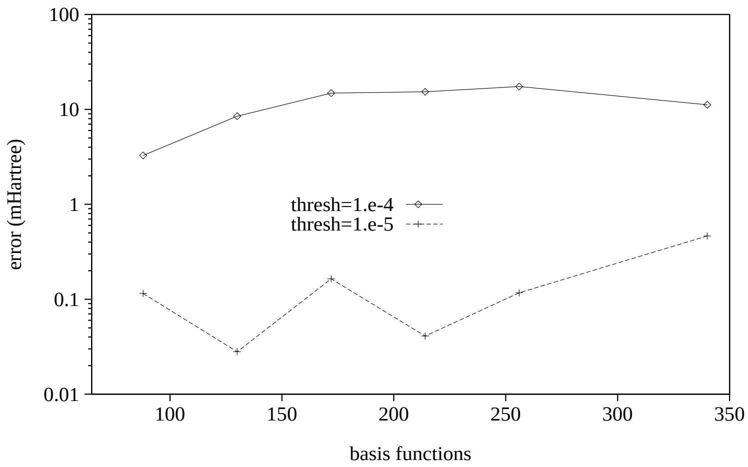

Figure 10: Errors in the converged HF/3-21G total energies corresponding to helices 2, 4, 6 and 8 shown in Figure 4. The corresponding timings are shown in Figure 8.

Results

Although QCTC is still in rapid development, it may be interesting to present some results to show the capability of the method. We have performed HF/3-21G calculations with SUPERMOLECULE/QCTC and GAUSSIAN94[141] on the sequence of water clusters shown in Figure 3 and the sequence of polyglycine α-helices shown in Figure 4. In Figure 5, GAUSSIAN94 and QCTC cpu times per SCF cycle are shown for the water clusters. Figure 6 shows comparable cpu times for the sequence of polyglycine α-helices.

The timings shown in Figures 5 and 6 are average cpu times for building the Coulomb matrix with QCTC and the average cpu time spent computing the Fock matrix and solving the SCF equations in GAUSSIAN94. The latter operations represent approximately twice the work performed by QCTC. All times were obtained with a single Cray C-90 processor and all GAUSSIAN94 timings refer to time spent in LINK502 using the SCF=(Direct,NoIncFock) options. To facilitate comparison we have not used incremental Fock builds in the QCTC calculations, although doing so should result in significant acceleration. Timings corresponding to helices 10,12, and 14, as well as water clusters 40 and 60, were computed as averages of the first three SCF cycles.

The QCTC times for building the Coulomb matrix with $\texttt{thresh} = 10^{-5}$, $\texttt{thresh} = 10^{-4}$ and with $\texttt{thresh} = 10^{-5}$ but using no multipole approximations are shown for the sequence of water clusters in Figure 7, and for the sequence of α-helices in Figure 8.

Although not yet approaching the asymptotic limit, it is interesting to compare the scaling of the methods. In Table 5, we list the approximate power dependence x obtained by fitting the cpu times for water clusters 20, 30, 40 and 60 and for helices 8, 10 and 14 to $a + bN^x$.

Table 5: Power dependence of GAUSSIAN94 and QCTC.

Method	Power Dependence	
	Water Clusters	α-helices
GAUSSIAN94	2.5	2.8
QCTC: no multipoles	2.3	2.2
QCTC: $\texttt{thresh} = 10^{-5}$	2.0	1.9
QCTC: $\texttt{thresh} = 10^{-4}$	1.9	1.6

Errors, computed as differences between total energies obtained with GAUSSIAN94 and SUPERMOLECULE/QCTC, are shown in Figures 9 and 10 for subsets of the water cluster and α-helix sequences, respectively. Note that the reported

GAUSSIAN94 cpu times could be decreased by relaxing the integral screening threshold, although probably not by more than 10-20%. The GAUSSIAN94 timings that we have reported correspond to the default accuracy equal to 10^{-5} Hartree, which is the level of accuracy obtained with SUPERMOLECULE/QCTC using no multipoles and $\texttt{thresh} = 10^{-5}$.

Discussion

We report, for the first time, sub-N^2 scaling for computation of the Coulomb matrix. Importantly, the efficiency of our method does *not* depend critically on sacrifices in accuracy, and whatever trade-off there is between the two is easily controlled with a single parameter. While these results are encouraging, more significant gains are yet to be realized, both from simple grooming and by major algorithmic developments.

While the current version of QCTC does not employ the FGT to compute near-field contributions to the Coulomb matrix, our current method is very competitive. This is due to the independent thresholding of $\langle\mathrm{BRA}|$ s and $|\mathrm{KET}\rangle$ s (see Section 5.4), and working in a HGTF representation (see Section 2.2). It has been argued that the rigorous Schwartz inequality is to be preferred to thresholds based on radial overlaps because it is a tighter bound. However, thresholding on the radial overlaps of primitive families as we have done excludes significant portions of work that is routinely performed by standard electronic structure programs. The additional work performed by these codes results from the use of the Schwartz inequality at the level of contracted quartets, which links the evaluation of all primitive integrals in the contracted quartet to those involving the most diffuse distributions.

Another reason for the competitive performance of our direct algorithm is the efficient implementation of the gamma functions, $F_m(t)$, which dominate the evaluation of R_{lmn} for small $\mathcal{L}$. These gamma functions are computed using three-term Chebychev interpolation [142–146], with expansion coefficients that have been computed and tabulated using Mathematica [147] and `ChebyshevRules` [146].

6 Conclusion

We have described a variety of new methods for accelerating construction of the Coulomb matrix, based upon a simplified representation of the electron charge density. Central to these methods are the elimination of degeneracies, the reduction of complexity and the use of fast expansion methods made possible by a Hermite Gaussian representation. Sub-N^2 cpu dependencies for assembly of the Coulomb matrix in realistic, real-chemistry applications have been reported for the first time, and the errors in the calculated SCF energies are shown to be under control.

An interesting aspect of our fast HGTF-based methods is that they are independent of contraction length. This is an added benefit as working with an uncontracted basis yields lower (more accurate) total energies. Another feature is that these methods may be easily extended to the computation of other integrals including derivatives of the Coulomb matrix and analytic integrals of the form

$$X_{ab} = \sum_{cd} P_{cd} \left(\rho_{ab} \middle| \widehat{x} \middle| \rho_{cd}\right) \ , \tag{89}$$

where $\widehat{x}$ is a generic operator. Because the fast HGTF-based methods retain the convenience of standard CGTF-based two-electron integrals, we foresee no difficulties integrating them into the well developed methods of molecular orbital theory.

While the methods developed here are quite promising, we have by no means explored all possible advantages to be gained by working in an HGTF basis. For example, the FGT may be used to expand both $\langle \textsc{bra}|$ and $|\textsc{ket}\rangle$ distributions, an approach which is the near-field analog of the fast multipole method. Also, the integral representation of HGTFs [148]

$$\zeta^{n/2} H_n\left(\sqrt{\zeta} x\right) e^{-\zeta x^2} = (-1)^{[n/2]} \zeta^{n/2} 2^{n+1} \pi^{-1/2} \times \begin{cases} \int_0^\infty e^{-t^2} t^n \cos\left(2\sqrt{\zeta} x t\right) dt & n \text{ even} \\ \int_0^\infty e^{-t^2} t^n \sin\left(2\sqrt{\zeta} x t\right) dt & n \text{ odd,} \end{cases} \tag{90}$$

as well as connections with wavelet theory [133,134] and related methods for the compression of dense matrices[136] provide intriguing avenues for future research.

The advent of sub-N^2 methods for assembly of the Coulomb matrix has brought very large scale applications closer to reality. The achievements reported here are useful in several ways: the improvements by themselves already allow larger applications than before, but more importantly, the further erosion of the dreaded "N^4 bottleneck" is making future progress easier to achieve as hardware and software development continues. While the problem of evaluating exchange interactions has not been addressed in the present work, there exist a multitude of alternatives to the conventional evaluation of Hartree-Fock exchange that will complement the techniques described here. Finally, it must also be stressed that while the current discussion has dealt exclusively with Coulombic interaction in Hartree-Fock theory, it is fully possible to implement similar techniques in direct, correlated methods. Especially within the

Møller-Plesset family of perturbation methods, we showed some time ago how the energy expressions can be rewritten in alternative forms that resemble the Hartree-Fock expressions [37,38,39], and how Coulomb- and exchange-like contributions can be identified. Combining these techniques with the ones discussed here can make very large systems accessible not only to SCF calculations, but also to correlated methods.

Acknowledgments

This work was supported by the Minnesota Supercomputer Institute, and by the National Science Foundation with grant CHE-9223782. Matt Challacombe acknowledges a postdoctoral fellowship in computational science from the NSF with matching support from MSI. Thanks are also due to Dr. John Mertz of Cray Research Inc. for supplying the geometries of representative water clusters.

7 Bibliography

1. C.C.J. Roothaan. New developments in molecular orbital theory. *Rev. Mod. Phys.*, **23**, 69–89 (1951).
2. G. G. Hall. The molecular orbital theory of chemical valency. VIII. A method of calculating ionization potentials. *Proc. Roy. Soc.*, **A 205**, 541–552 (1951).
3. J. A. Pople and R. K. Nesbet. Self-consistent orbitals for radicals. *J. Chem. Phys.*, **22**, 571–574 (1954).
4. ETO multicenter molecular integrals (Proceedings of the first international conference held at Florida A and M University). C. A. Weatherford and H. W. Jones, eds., D. Reidel, Dordrecht (1982).
5. S. F. Boys. Electronic wave functions. I. A general method of calculation for the stationary states of any molecular system. *Proc. Roy. Soc.*, **A 200**, 542–554 (1950).
6. T. Kato. On the eigenfunctions of many-particle systems in quantum mechanics. *Commun. Pure Appl. Math.*, **10**, 151–177 (1957).
7. E. Steiner. Charge densities in atoms. *J. Chem. Phys.*, **39**, 689–695 (1963).
8. N. C. Handy, M. T. Marron, and H. J. Silverstone. Long-range behavior of Hartree-Fock orbitals. *Phys. Rev.*, **180**, 45–48 (1968).
9. M. Levy and R. G. Parr. Long-range behavior of natural orbitals and electron density. *J. Chem. Phys.*, **64**, 2707–2708 (1976).

10. T. Hoffman-Ostenhof, M. Hoffman-Ostenhof, and R. Ahlrichs. "Schrödinger inequalities" and asymptotic behavior of many-electron densities. *Phys. Rev.*, **A 18**, 328–328 (1978).
11. J. Almlöf. Methods of modern Hartree-Fock theory. In *Lecture Notes in Chemistry*, B. Roos, Ed. Springer (1994) pp. 1-90.
12. D. Hankinson and J. Almlöf. A comparative evaluation of Gaussian- and Slater-type basis functions. (to be publ.).
13. V. Dyczmons. No N^4-dependence in the calculation of large molecules. *Theoret. Chim. Acta*, **28**, 307–310 (1973).
14. R. Ahlrichs. Methods for efficient evaluation of integrals for Gaussian type basis sets. *Theoret. Chim. Acta*, **33**, 157–167 (1974).
15. J. Almlöf, K. Faegri, and K. Korsell. Principles for a direct SCF approach to LCAO-MO *ab-initio* calculations. *J. Comput. Chem.*, **3**, 385–399 (1982).
16. M. Häser and R. Ahlrichs. Improvements on the Direct SCF method. *J. Comput. Chem.*, **10**, 104–111 (1989).
17. X. Jiancheng and J. Shouping. The rapid estimate of a Gauss integral. *Intern. J. Quantum Chem.*, **39**, 123–130 (1991).
18. P. M. W. Gill, B. G. Johnson, and J. A. Pople. A simple yet powerful upper bound for Coulomb integrals. *Chem. Phys. Letters*, **217**, 65–68 (1994).
19. S. Obara and A. Saika. Efficient recursive computation of molecular integrals over Cartesian Gaussian functions. *J. Chem. Phys.*, **84**, 3963–3974 (1985).
20. M. Head-Gordon and J. A. Pople. A method for two-electron Gaussian integral and integral derivative evaluation using recurrence relations. *J. Chem. Phys.*, **89**, 5777–5786 (1988).
21. P. M. W. Gill, M. Head-Gordon, and J. A. Pople. An efficient algorithm for the generation of two-electron repulsion integrals over Gaussian basis functions. *Intern. J. Quantum Chem.*, **S 23**, 269–280 (1989).
22. T. P. Hamilton and H. F. Schaefer. New variations in two-electron integral evaluation in the context of direct SCF procedures. *Chem. Phys.*, **150**, 163–171 (1991).
23. T. P. Hamilton and H. F. Schaefer. The use of special coordinate axes in direct and semi-direct implementations of second-order perturbation theory, including the derivation of a horizontal recurrence relation. *Can. J. Chem.*, **70**, 416–420 (1992).
24. P. M. W. Gill and J. A. Pople. The PRISM algorithm for two-electron integrals. *Intern. J. Quantum Chem.*, **40**, 753–772 (1991).

25. I. Panas. Two-electron integrals and integral derivatives revisited. *Chem. Phys. Letters*, **184**, 86–92 (1991).
26. U. Ryu, Y. S. Lee, and R. Lindh. An efficient method of implementing the horizontal recurrence relation in the evaluation of electron repulsion integrals using Cartesian Gaussian functions. *Chem. Phys. Letters*, **185**, 562–568 (1991).
27. R. Lindh. The reduced multiplication scheme of the Rys-Gauss quadrature for 1st order integral derivatives. *Theoret. Chim. Acta*, **85**, 423–440 (1993).
28. S. Ten-no and S. Iwata. An efficient algorithm for electron repulsion integrals over contracted Gaussian-type functions. *Chem. Phys. Letters*, **211**, 259–264 (1993).
29. B. G. Johnson, P. M. W. Gill, and J. A. Pople. The efficient transformation of $(m0|n0)$ to $(ab|cd)$ two-electron repulsion integrals. *Chem. Phys. Letters*, **206**, 229–238 (1993).
30. D. L. Strout and G. E. Scuseria. A quantitative study of the scaling properties of the Hartree-Fock method. *J. Chem. Phys.*, **102**, 8448–8452 (1995).
31. I. Panas, J. Almlöf, and M. W. Feyereisen. *Ab initio* methods for large systems. *Intern. J. Quantum Chem.*, **40**, 797–807 (1991).
32. V. Termath and N. C. Handy. A Kohn-Sham method involving the direct determination of the Coulomb potential on a numerical grid. *Chem. Phys. Letters*, **230**, 17–24 (1994).
33. R. G. Parr and W. Yang. Density-Functional Theory of Atoms and Molecules. Oxford University Press, Oxford (1989).
34. E. S. Kryachko and E. V. Ludeña. Energy Density Functional Theory of Many Electron Systems. Kluwer Academic, Boston (1990).
35. B. G. Johnson, P. M. W. Gill, and J. A. Pople. The performance of a family of density functional methods. *J. Chem. Phys.*, **98**, 5612–5626 (1993).
36. A. D. Becke. Density-functional thermochemistry. III. The role of exact exchange. *J. Chem. Phys.*, **98**, 5648–5652 (1993).
37. J. Almlöf. Elimination of energy denominators in Møller-Plesset perturbation theory by a Laplace transform approach. *Chem. Phys. Letters*, **181**, 319–320 (1991).
38. M. Häser and J. Almlöf. Laplace transform techniques in Møller-Plesset perturbation theory. *J. Chem. Phys.*, **96**, 489–494 (1992).
39. M. Häser. Møller-Plesset (MP2) perturbation theory for large molecules. *Theoret. Chim. Acta*, **87**, 147–173 (1993).

40. A. L. Sargent, J. Almlöf, and M. W. Feyereisen. Massively parallel algorithms for electronic structure calculations in quantum chemistry. *SIAM News*, **26(2)**, (1993).
41. B. H. Greeley, T. V. Russo, D. T. Mainz, R. A. Friesner, J. M. Langlois, W. A. Goddard, R. E. Donnelly, and M. N. Ringnalda. New pseudospectral algorithms for electronic structure calculations: Length scale separation and analytical two-electron integral corrections. *J. Chem. Phys.*, **101**, 4028–4041 (1994).
42. K. Iyer, M. Merrick, and T. Beck. Application of a distributed nucleus approximation in grid based minimization of the Kohn-Sham energy functional. *J. Chem. Phys.*, **103**, 227–233 (1995).
43. J. Almlöf and M. W. Feyereisen. SUPERMOLECULE, a program for *ab initio* electronic structure calculations. University of Minnesota (1993).
44. L. E. McMurchie. Ph.D. thesis, University of Seattle (1977).
45. L. E. McMurchie and E. R. Davidson. One- and two-electron integrals over Cartesian Gaussian functions. *J. Comp. Phys.*, **26**, 218–231 (1978).
46. V. R. Saunders. Molecular integrals for Gaussian type functions. In *Methods in Computational Molecular Physics,* NATO ASI Ser. C. G.H.F. Diercksen and S. Wilson, eds., pp. 1–36. D. Reidel Publishing Company, Boston (1983).
47. I. Shavitt. The Gaussian function in calculations of statistical mechanics and quantum mechanics. *Meth. Comp. Phys.*, **2**, 1–45 (1963).
48. P. J. Davis. Gamma function and related functions. In *Handbook of Mathematical Functions, 9th edition*, M. Abramowitz and I. A. Stegun, eds., pp. 260–262. Dover Publications Inc., New York (1987).
49. B. I. Dunlap, J. W. D. Connolly, and J. R. Sabin. On the applicability of the LCAO-Xα methods to molecules containing transition metal atoms: the nickel atom and nickel hydride. *Intern. J. Quantum Chem.*, **S 11**, 81–87 (1977).
50. J. W. Mintmire. LCAO methods in the Xα formalism: A comparison of results for the F_2 molecule. *Intern. J. Quantum Chem.*, **S 13**, 163–168 (1979).
51. V. R. Saunders, C. Freyria-Fava, L. Salasco R. Dovesi, and C. Roetti. On the electrostatic potential in crystalline systems where the charge density is expanded in Gaussian functions. *Mol. Phys.*, **77**, 629–665 (1992).
52. R. J. Wheatley. Gaussian multipole functions for describing molecular charge distributions. *Mol. Phys.*, **79**, 597–610 (1993).
53. R. J. Wheatley and J. B. O. Mitchell. Gaussian multipoles in practice: Electrostatic energies for intermolecular potentials. *J. Comput. Chem.*,

15, 1187–1198 (1994).
54. G. R. Ahmadi and J. Almlöf. The Coulomb operator in a Gaussian product basis. *Chem. Phys. Letters*, in press.
55. W. J. Hehre, L. Radom, P. von Schleyer, and J. A. Pople. *Ab Initio* Molecular Orbital Theory. J. Wiley and Sons, New York, first edition (1986).
56. S. Huzinaga. Basis sets for molecular calculations. *Comp. Phys. Rept.*, **2**, 279–339 (1985).
57. S. Huzinaga, M. Klobukowski, and H. Tatewaki. The well-tempered GTF basis sets and their applications in the SCF calculations on N_2, CO, Na_2, and P_2. *Can. J. Chem.*, **63**, 1812–1828 (1985).
58. S. Huzinaga and B. Miguel. A comparison of the geometrical sequence formula and the well-tempered formulas for generating GTO basis orbital exponents. *Chem. Phys. Letters*, **175**, 289–291 (1990).
59. S. Huzinaga and M. Klobukowski. Well-tempered Gaussian basis sets for the calculation of matrix Hartree-Fock wavefunctions. *Chem. Phys. Letters*, **212**, 260–264 (1993).
60. R. Bardo and K. Ruedenberg. Even-tempered atomic orbitals. VI. Optimal orbital exponents and optimal orbital contractions of Gaussian primitives for hydrogen, carbon and oxygen in molecules. *J. Chem. Phys.*, **60**, 918–931 (1974).
61. E. Clementi and G. Corongiou. Geometrical basis sets for molecular calculations. *Chem. Phys. Letters*, **90**, 359–363 (1982).
62. S. Wilson. Systematic sequences of even-tempered Gaussian primitives in electron correlation studies using many-body perturbation theory. *Theoret. Chim. Acta*, **57**, 53–61 (1980).
63. S. Wilson. Universal systematic sequences of even-tempered Gaussian primitive functions in electronic correlation studies. *Theoret. Chim. Acta*, **58**, 31–40 (1980).
64. H. Eyring, J. Walter, and G. E. Kimball. Quantum Chemistry. J. Wiley and Sons, New York, first edition (1944).
65. J. O. Hirschfelder, C. F. Curtiss, and R. B. Bird. Molecular Theory of Gases and Liquids. J. Wiley and Sons, New York, first edition (1944).
66. A. D. Buckingham. Basic theory of intermolecular forces: Applications to small molecules. In *Intermolecular Interactions: From Diatomics to Biopolymers*, B. Pullman, ed., pp. 2–62. J. Wiley and Sons, New York (1987).
67. J. Cipriani and B. Silvi. Cartesian expressions for electric multipole moment operators. *Mol. Phys.*, **45**, 259–272 (1982).

68. C. E. Dykstra. Efficient calculation of electrically based intermolecular potentials of weakly bonded clusters. *J. Comput. Chem.*, **9**, 476–487 (1988).
69. C. Cohen-Tannoudji, B. Diu, and F. Laloë. Quantum Mechanics, Vol. II. J. Wiley and Sons, New York (1977). See Appendix II for a useful discussion of the Dirac delta function and its derivatives.
70. G. A. van der Velde. Ph.D. thesis, University of Groningen (1974).
71. D. Hegarty and G. A. van der Velde. Integral evaluation algorithms and their implementation. *Intern. J. Quantum Chem.*, **23**, 1135–1153 (1983).
72. E. Schwegler, M. Challacombe, and J. Almlöf. (to be publ.).
73. F. E. Allen and J. Cocke. A catalogue of optimizing transformations. In *Design and Optimization of Compilers*, R. Rustin, ed., pp. 1–30. Prentice-Hall, New Jersey (1971).
74. C. D. Polychronopoulos. A compiler transformation for parallel machines. In *Proceedings of the International Conference on Parallel Processing*, S. K. Sahni, ed., pp. 235–242. Pennsylvania State University Press, Pennsylvania (1987).
75. S. F. Boys and I. Shavitt. Technical Report WIS-AF-13, University of Wisconsin (1959).
76. R. S. Mulliken. Quelques aspects de la théorie des orbitales moléculaires. *J. Chim. Phys.*, **46**, 497–542 (1946).
77. K. Ruedenberg. On the three- and four-center integrals in molecular quantum mechanics. *J. Chem. Phys.*, **19**, 1433–1434 (1951).
78. P. O. Löwdin. Approximate formulas for many-center integrals in the theory of molecules and crystals. *J. Chem. Phys.*, **21**, 374–375 (1953).
79. J. Cizek. A new method of approximative calculation of polycentric integrals used in quantum mechanical study of molecular structure. *Mol. Phys.*, **6**, 19–31 (1963).
80. F. E. Harris and R. Rein. Integral approximations for molecular orbital theory. *Theoret. Chim. Acta*, **6**, 73–82 (1966).
81. F. Billingsley and J. Bloor. Limited expansion of diatomic overlap (LEDO): A near-accurate approximate *ab initio* LCAO MO method. I. Theory and preliminary investigations. *J. Chem. Phys.*, **55**, 5178–5190 (1971).
82. M. D. Newton, N. S. Ostlund, and J. A. Pople. Projection of diatomic differential overlap: Least-squares projection of two-center distributions onto one-center functions. *J. Phys. Chem.*, **49**, 5192–5194 (1968).
83. M. D. Newton. Self-consistent molecular-orbital methods. II. Projection of diatomic differential overlap (PDDO). *J. Phys. Chem.*, **51**, 3917–3926

(1969).
84. M. D. Newton, W. A. Lathan, W. J. Hehre, and J. A. Pople. Self-consistent molecular-orbital methods. III. Comparison of Gaussian expansion and PDDO methods using minimal STO basis sets. *J. Phys. Chem.*, **51**, 3927–3932 (1969).
85. A. Okniński. Generalized treatment of the approximations for the two-center charge distributions. *J. Chem. Phys.*, **60**, 4098–4099 (1974).
86. F. Billingsley and C. Trindle. Mixed basis scheme for the rapid computation of molecular electronic energy in the SCF-LCAO-(STO)-MO formalism, and its application to the lithium superoxide molecule. *J. Phys. Chem.*, **76**, 2995–2999 (1972).
87. E. Baerends, D. Ellis, and P. Ros. Self-consistent molecular Hartree-Fock-Slater calculations. I. The computational procedure. *Chem. Phys.*, **2**, 41–51 (1973).
88. B. Dunlap, J. Connolly, and J. Sabin. On some approximations in applications of Xα theory. *J. Chem. Phys.*, **71**, 3396–3402 (1979).
89. J. C. Campbell, I. H. Hillier, and Saunders. A systematic classification of atomic integrals in *ab initio* molecular orbital calculations. I. The core valence approximation molecular orbital method. *J. Chem. Phys.*, **74**, 3971–3976 (1981).
90. C. Van Alsenoy. *Ab initio* calculations on large molecules: The multiplicative integral approximation. *J. Comput. Chem.*, **9**, 620–626 (1988).
91. J. Andzelm and E. Wimmer. Density functional Gaussian-type-orbital approach to molecular geometries, vibrations, and reaction energies. *J. Chem. Phys.*, **96**, 1280–1303 (1992).
92. A. Komornicki and G. Fitzgerald. Molecular gradients and Hessians implemented in density functional theory. *J. Chem. Phys.*, **98**, 1398–1421 (1993).
93. K. Eichkorn, O. Treutler, H. Öhm, M. Häser, and R. Ahlrichs. Auxiliary basis sets to approximate Coulomb potentials. *Chem. Phys. Letters*, **240**, 283–289 (1995).
94. M. Feyereisen, G. Fitzgerald, and A. Komornicki. Use of approximate integrals in *ab initio* theory. An application in MP2 energy calculations. *Chem. Phys. Letters*, **208**, 359–363 (1993).
95. O. Vahtras, J. Almlöf, and M. W. Feyereisen. Integral approximations for LCAO-SCF calculations. *Chem. Phys. Letters*, **213**, 514–518 (1993).
96. A. P. Rendell and T. J. Lee. Coupled-cluster theory employing approximate integrals: An approach to avoid the input/output and storage bottlenecks. *J. Chem. Phys.*, **101**, 400–408 (1994).

97. S. Ten-no and S. Iwata. Three-center expansion of electron repulsion integrals with linear combination of atomic electron distributions. *Chem. Phys. Letters*, **240**, 578–584 (1995).
98. N. Beebe and J. Linderberg. Simplifications in the generation and transformation of two-electron integrals in molecular calculations. *Intern. J. Quantum Chem.*, **12**, 683–705 (1977).
99. N. Beebe and J. Linderberg. A different approach to integrals and integral transformations, or, do we really need all the integrals? In *Numerical Algorithms in Chemistry: Algebraic Methods*, C. Moler and I. Shavitt, eds., pp. 158–161. Lawerence Berkeley Laboratory, Berkeley, CA (1978).
100. I. Røeggen and E. Wisløff-Nielsen. On the Beebe-Linderberg two-electron integral approximation. *Chem. Phys. Letters*, **132**, 154–160 (1986).
101. W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. Numerical Recipies in FORTRAN. Cambridge University Press, Port Chester, New York (1992). (see especially pp. 53–54).
102. P. M. W. Gill, B. G. Johnson, J. A. Pople, and S. W. Taylor. Modeling the potential of a charge distribution. *J. Chem. Phys.*, **96**, 7178–7179 (1992).
103. W. Klopper and R. Röhse. Computation of some new two-electron Gaussian integrals. *Theoret. Chim. Acta*, **83**, 441–453 (1992).
104. K. Siam, V. Klimkowski, J. Ewbank, C. Van Alsenoy, and L. Schäfer. *Ab initio* studies of structural features not easily amenable to experiment. *Theochem*, **110**, 171–182 (1984).
105. J. Barnes and P. Hut. A hierarchical $\mathcal{O}(NlogN)$ force-calculation algorithm. *Nature*, **324**, 446–449 (1986).
106. L. Hernquist. Performance characteristics of tree codes. *Ap. J. Suppl.*, **64**, 715–734 (1987).
107. J. E. Barnes and P. Hut. Error analysis of a tree code. *Ap. J. Suppl.*, **70**, 389–417 (1989).
108. M. S. Warren and J. K. Salmon. A parallel hashed oct-tree N-body algorithm. In *Proceedings SUPERCOMPUTING '93*, pp. 12–21. Los Alamitos, CA (1993). IEEE Computer Society Press.
109. S. Pfalzner and P. Gibbon. A 3D hierarchical tree code for dense plasma simulation. *Comp. Phys. Comm.*, **79**, 24–38 (1994).
110. J. K. Salmon. Fast parallel tree codes for gravitational and fluid dynamical N-body problems. *Intern. J. Super. Appl.*, **8**, 129–142 (1994).
111. M. S. Warren and J. K. Salmon. A portable parallel particle program. *Comp. Phys. Comm.*, **87**, 266–290 (1995).

112. L. Greengard and V. Rokhlin. A fast algorithm for particle simulations. *J. Comp. Phys.*, **73**, 325–348 (1987).
113. K. E. Schmidt and M. A. Lee. Implementing the fast multipole method in three dimensions. *J. Stat. Phys.*, **63**, 1223–1235 (1991).
114. H.G. Petersen, D. Soelvason, J. W. Perram, and E. R. Smith. The very fast multipole method. *J. Chem. Phys.*, **101**, 8870–8876 (1994).
115. H.G. Petersen, E. R. Smith, and D. Soelvason. Error estimates for the fast multipole method. II. The three-dimensional case. *Proc. Roy. Soc.*, **A 448**, 401–418 (1994).
116. C. A. White and M. Head-Gordon. Derivation and efficient implementation of the fast multipole method. *J. Chem. Phys.*, **101**, 6593–6605 (1994).
117. H. Q. Ding, N. Karasawa, and W. A. Goddard III. Atomic level simulations on a million particles: The cell multipole method for Coulomb and London nonbond interactions. *J. Chem. Phys.*, **97**, 4309–4315 (1992).
118. H.Q. Ding, N. Karasawa, and W. A. Goddard III. The reduced cell multipole method for Coulomb interactions in periodic systems with million-atom unit cells. *Chem. Phys. Letters*, **196**, 6–10 (1992).
119. R. Kutteh and J. B. Nicholas. Efficient dipole iteration in polarizable charged systems using the cell multipole method and application to polarizable water. *Comp. Phys. Comm.*, **86**, 227–235 (1995).
120. R. Kutteh and J. B. Nicholas. Implementing the cell multipole method for dipolar and charged dipolar systems. *Comp. Phys. Comm.*, **86**, 236–254 (1995).
121. A. Y. Grama, V. Kumar, and A. Sameh. Scalable parallel formulations of the Barnes-Hut method for N-body simulations. In *Proceedings SUPERCOMPUTING '94*, pp. 439–448, Los Alamitos, CA (1994). IEEE Computer Society Press.
122. L. Greengard and V. Rokhlin. The numerical solution of the N-body problem. *Comp. Phys.*, **4**, 142–152 (1990).
123. S. Aluru, G. M. Prabhu, and J. Gustafson. Truly distribution-independent algorithms for the N-body problem. In *Proceedings SUPERCOMPUTING '94*, pp. 420–428, Los Alamitos, CA (1994). IEEE Computer Society Press.
124. A. Windemuth. Advanced Algorithms for Molecular Dynamics Simulation: The Program PMD. ACS Books (1995).
125. C. A. White, B. G. Johnson, P. M. W. Gill, and M. Head-Gordon. The continuous fast multipole method. *Chem. Phys. Letters*, **230**, 8–16 (1994).

126. R. Kutteh and J. B. Nicholas. A generalized fast multipole approach for Hartree-Fock and density functional computations. *Chem. Phys. Letters*, **238**, 173–179 (1995).
127. L. Greengard and J. Strain. The fast Gauss transform. *J. Sci. Stat. Comput.*, **12**, 79–94 (1991).
128. J. Strain. The fast Gauss transform with variable scales. *J. Sci. Stat. Comput.*, **12**, 1131–1139 (1991).
129. J. Strain. Fast adaptive methods for the free-space heat equation. *J. Sci. Comput.*, **15**, 185–206 (1994).
130. E. Hille. A class of reciprocal functions. *Ann. Math.*, **27**, 427–443 (1926).
131. I. Daubechies. Ten Lectures on Wavelets. SIAM Press, Philidelphia (1992).
132. C. K. Chui. Wavelets - A Tutorial in Theory and Applications. Academic, New York (1992).
133. L. R. LoConte, R. Merletti, and G. V. Sandri. Hermite expansions of compact support waveforms: Applications to myoelectric signals. *IEEE Trans. Bio. Eng.*, **41**, 1147–1159 (1994).
134. K. Sabetfakhri and L.P.B. Katehi. On the application of quasi-wavelet expansions to open dielectric waveguide problems. In *Proceedings of IEEE Antennas and Propagation Society International Symposium*, pp. 439–448, New York, NY (1994). IEEE.
135. S. Palavajjhala, R.L. Motard, and B. Joseph. Computational Aspects of Wavelet Transforms. In *Computational Aspects of Wavelet Transforms*, R. L. Motard and B. Joseph, eds., p. 33. Kluwer Academic, Boston (1994).
136. G. Beylkin, R. Coifman, and V. Rokhlin. Fast wavelet transforms and numerical algorithms I. E-print (1995). Many related e-prints are available under "wavelets and numerical analysis" at http://www.mathsoft.com/wavelets.html.
137. J. L. Whitten. Coulombic potential energy integrals and approximations. *J. Chem. Phys.*, **58**, 4496–4501 (1973).
138. J. D. Power and R. M. Pitzer. Inequalities for electron repulsion integrals. *Chem. Phys. Letters*, **24**, 478–483 (1974).
139. S. R. Gadre and R. K. Pathak. Bounds to atomic and molecular energy functionals. *Adv. Quantum Chem.*, **22**, 211–300 (1991).
140. Biosym Technologies. *InsightII User Guide* (1992).
141. M. J. Frisch, G. W. Trucks, H. B. Schlegel, P. M. W. Gill, B. G. Johnson, M. A. Robb, J. R. Cheeseman, T. Keith, G. A. Petersson, J. A. Montgomery, K. Raghavachari, M. A. Al-Laham, V. G. Zakrzewski, J. V.

Ortiz, J. B. Foresman, C. Y. Peng, P. Y. Ayala, W. Chen, M. W. Wong, J. L. Andres, E. S. Replogle, R. Gomperts, R. L. Martin, D. J. Fox, J. S. Binkley, D. J. Defrees, J. Baker, J. P. Stewart, M. Head-Gordon, C. Gonzalez, and J. A. Pople. *Gaussian 94, Revision B.3*. GAUSSIAN (1994).

142. L. L. Shipman and R. E. Christoffersen. High speed evaluation of $F_0(x)$. *Comp. Phys. Comm.*, **2**, 201–206 (1971).
143. S. T. Elbert and E. R. Davidson. Evaluation of electron repulsion integrals over Gaussian lobe basis functions. *J. Comp. Phys.*, **16**, 391–395 (1974).
144. P. Gill, B. Johnson, and J. Pople. Two-electron repulsion integrals over Gaussian s functions. *Intern. J. Quantum Chem.*, **40**, 745–752 (1991).
145. S. Yahiro and Y. Gondo. A vector processing algorithm of auxiliary integral evaluation for two-electron Gaussian integrals. *J. Comput. Chem.*, **13**, 12–16 (1992).
146. M. Challacombe. Ph.D. thesis, Florida State University (1994).
147. S. Wolfram. *Mathematica, A System for Doing Mathematics by Computer*. Addison-Wesley, Redwood City, CA, second edition (1991).
148. G. Szego. *Orthogonal Polynomials*. American Mathematical Society, Providence, fourth edition (1975).