THE JOURNAL OF PHYSICAL CHEMISTRY A
A JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

pubs.acs.org/JPCA
Article

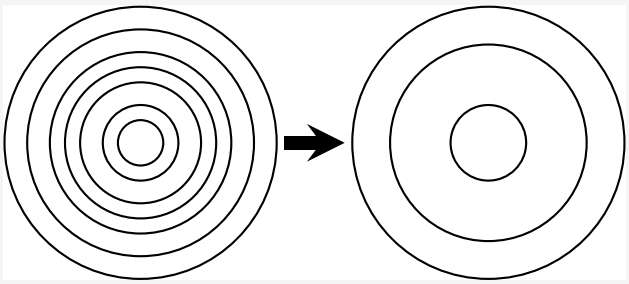
Economical Models for Electron Densities

Published as part of The Journal of Physical Chemistry A virtual special issue "Krishnan Raghavachari Festschrift".

Ellena K. G. Black and Peter M. W. Gill*

Cite This: J. Phys. Chem. A 2023, 127, 9346−9356
Read Online
ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: We discuss a new theoretical framework for modeling molecular electron densities. Our approach decomposes the total density into contributions from basis function products and then approximates each product using constrained least−squares approximation in a tailored local basis of functions with adjustable non−linear parameters. We show how to solve directly for the expansion coefficients and Lagrange multipliers and present an iterative method to optimize the non−linear parameters. Example products from the Dunning cc-pVTZ basis set are discussed.



1. INTRODUCTION

1.1. Background. The central tenet of density functional theory (DFT) is that, notwithstanding its conceptual simplicity, the electron density $D(\mathbf{r})$ in a system is a molecular property from which all others can be derived.[1] Consequently, its computation, representation, and visualization are among the most fundamental features of all modern electronic structure software packages. Our goal in this work is to develop a systematic scheme for efficiently constructing accurate approximations to $D(\mathbf{r})$.

In terms of a set of orthonormal orbitals $\psi_i(\mathbf{r})$, we can write the electron density as[2,3]

$$D(\mathbf{r}) = \sum_{i}^{n} \psi_i^*(\mathbf{r})\psi_i(\mathbf{r}) \tag{1}$$

and, if each of these orbitals is expanded in a one-electron basis set $\phi_\mu(\mathbf{r})$, we obtain

$$D(\mathbf{r}) = \sum_{\mu\nu}^{N} P_{\mu\nu}\phi_\mu(\mathbf{r})\phi_\nu(\mathbf{r}) \tag{2}$$

where $P_{\mu\nu}$ is a density matrix. If each basis function is a contracted Gaussian

$$\phi_\mu(\mathbf{r}) = Y_\mu(\mathbf{r}) \sum_{k=1}^{K_\mu} C_\mu^k \exp(-\alpha_\mu^k|\mathbf{r} - \mathbf{A}_\mu|^2) \tag{3}$$

where $Y_\mu(\mathbf{r})$ is the angular factor, C_μ^k is a contraction coefficient, α_μ^k is an exponent, and $\mathbf{A}_\mu$ is the Gaussian center, then the function products in (2) are

$$\phi_\mu(\mathbf{r})\phi_\nu(\mathbf{r}) = Y_\mu(\mathbf{r})Y_\nu(\mathbf{r})\rho(\mathbf{r}) \tag{4}$$

where we have introduced the Gaussian density

$$\begin{aligned}\rho(\mathbf{r}) &= \sum_{k=1}^{K_\mu}\sum_{l=1}^{K_\nu} C_\mu^k C_\nu^l \exp[-\alpha_\mu^k|\mathbf{r} - \mathbf{A}_\mu|^2] \exp[-\alpha_\nu^l|\mathbf{r} - \mathbf{A}_\nu|^2] \\ &= \sum_{k=1}^{K_\mu}\sum_{l=1}^{K_\nu} C_\mu^k C_\nu^l G_{\mu\nu}^{kl} \exp[-\alpha_{\mu\nu}^{kl}|\mathbf{r} - \mathbf{A}_{\mu\nu}^{kl}|^2]\end{aligned} \tag{5}$$

and its primitive prefactors, exponents and centers are, respectively,

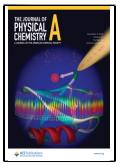
$$G_{\mu\nu}^{kl} = \exp\left[-\frac{|\mathbf{A}_\mu - \mathbf{A}_\nu|^2}{1/\alpha_\mu^k + 1/\alpha_\nu^l}\right] \tag{6}$$

$$\alpha_{\mu\nu}^{kl} = \alpha_\mu^k + \alpha_\nu^l \tag{7}$$

$$\mathbf{A}_{\mu\nu}^{kl} = \frac{\alpha_\mu^k \mathbf{A}_\mu + \alpha_\nu^l \mathbf{A}_\nu}{\alpha_\mu^k + \alpha_\nu^l} \tag{8}$$

The electrostatic energy between two products is given by the two-electron integral

Received: June 28, 2023
Revised: October 3, 2023
Accepted: October 4, 2023
Published: October 31, 2023



https://doi.org/10.1021/acs.jpca.3c04363
J. Phys. Chem. A 2023, 127, 9346−9356

$$(\mu\nu|\lambda\sigma) = \iint \phi_\mu(\mathbf{r}_1)\phi_\nu(\mathbf{r}_1)\frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|}\phi_\lambda(\mathbf{r}_2)\phi_\sigma(\mathbf{r}_2)\, d\mathbf{r}_1\, d\mathbf{r}_2 \quad (9)$$

and these integrals yield the Coulomb, exchange, and correlation energies.[4] The evaluation of the $(\mu\nu|\lambda\sigma)$ is one of the key bottlenecks of quantum chemistry, and research over many decades has led to an impressive array of algorithms for their computation, including the methods of Boys,[5] Dupuis et al.,[6] McMurchie and Davidson,[7] Obara and Saika,[8] Head-Gordon and Pople,[9] Hamilton and Schaefer,[10] Lindh et al.,[11] Gill and Pople,[12,13] Adams et al.,[14] Ishida,[15] Makowski,[16] and Komornicki and King.[17] Nonetheless, although some of these are very computationally efficient, all become expensive when the degrees of contraction $K_\mu K_\nu$ and $K_\lambda K_\sigma$ are large. For this reason, it was recognized long ago that it is desirable to construct approximations for products.

In the 20th century, a variety of methods for fitting function products in an auxiliary basis were introduced and we note the seminal contributions by Reeves and Fletcher,[18] O-ohata et al.,[19] Newton et al.,[20] Hehre et al.,[21] Stewart,[22] Billingsley and Bloor,[23] Baerends et al.,[24] Whitten,[25] Beebe and Linderberg,[26] Dunlap et al.,[27] Fortunelli and Salvetti,[28] Feyereisen et al.,[29] and Eichkorn et al.[30] Helpful discussions of the underlying theory were published subsequently by a number of workers.[31−37]

In the present paper, however, we eschew a global auxiliary basis in favor of approximating each product on its own locally generated basis. More specifically, we seek an approximation to the Gaussian density (5) that uses a few Gaussians whose coefficients, exponents and centers are explicitly optimized.

Throughout this paper, we will use italic symbols to denote scalars and bold (or double-struck) symbols to denote vectors or matrices.

1.2. Least-Squares Functional. We seek to approximate a given density $\rho(\mathbf{r})$ by a model $\chi(\mathbf{r})$ where

$$\rho(\mathbf{r}) = \mathbf{d}^{\mathrm{T}}\mathbf{a} \quad (10)$$

$$\chi(\mathbf{r}) = \mathbf{c}^{\mathrm{T}}\mathbf{b} \quad (11)$$

where $\mathbf{d} = (d_1, ..., d_n)$ is a list of positive coefficients, $\mathbf{c} = (c_1, ..., c_m)$ is a list of expansion coefficients, $\mathbf{a} = [a_1(\mathbf{r}), ..., a_n(\mathbf{r})]$ and $\mathbf{b} = [b_1(\mathbf{r}), ..., b_m(\mathbf{r})]$ are lists of normalized Gaussians

$$a_j(\mathbf{r}) = \left(\frac{1}{4\pi\alpha_j}\right)^{3/2} \exp\left[-\frac{x^2 + y^2 + (z - A_j)^2}{4\alpha_j}\right] \quad (12)$$

$$b_j(\mathbf{r}) = \left(\frac{1}{4\pi\beta_j}\right)^{3/2} \exp\left[-\frac{x^2 + y^2 + (z - B_j)^2}{4\beta_j}\right] \quad (13)$$

$\boldsymbol{\alpha} = (\alpha_1, ..., \alpha_n)$ and $\boldsymbol{\beta} = (\beta_1, ..., \beta_m)$ are lists of inverted exponents, and $\mathbf{A} = (A_1, ..., A_n)$ and $\mathbf{B} = (B_1, ..., B_m)$ are lists of Gaussian centers on the z axis. The assumption that all $d_i > 0$ sometimes requires that the basis set be reconstructed and this is discussed in Section 2.1. We assume that $m < n$, and we regard all of the c_j, β_j, and B_j as optimizable parameters.

For simplicity, we recast the problem in Fourier space and model $\hat{\rho}(\mathbf{k})$ by $\hat{\chi}(\mathbf{k})$ where

$$\hat{\rho}(\mathbf{k}) = \mathbf{d}^{\mathrm{T}}\hat{\mathbf{a}} \quad (14)$$

$$\hat{\chi}(\mathbf{k}) = \mathbf{c}^{\mathrm{T}}\hat{\mathbf{b}} \quad (15)$$

and the elements of $\hat{\mathbf{a}}$ and $\hat{\mathbf{b}}$ are the Fourier transforms

$$\hat{a}_j(\mathbf{k}) = \exp(-\alpha_j k^2)\exp(\mathrm{i}A_j k_z) \quad (16)$$

$$\hat{b}_j(\mathbf{k}) = \exp(-\beta_j k^2)\exp(\mathrm{i}B_j k_z) \quad (17)$$

Our models seek to minimize by least-squares the residual

$$\hat{R}(\mathbf{k}) = \hat{\rho}(\mathbf{k}) - \hat{\chi}(\mathbf{k}) \quad (18)$$

over all of space and, to achieve this, we introduce a Fourier weight function

$$\omega_p(k) = k^{2p-3}/(2\pi) \quad (19)$$

In general, our models must also satisfy l linear constraints on the expansion coefficients c_i. These are conveniently captured by the matrix equation

$$\mathbb{K}^{\mathrm{T}}\mathbf{c} = \mathbb{q} \quad (20)$$

where $\mathbb{K}$ is an $m \times l$ matrix and $\mathbb{q}$ is a vector of length l. For example, if we seek a model that conserves charge in an ss density, $\mathbb{K}$ must contain a column of ones and the corresponding element of $\mathbb{q}$ must hold $\hat{\rho}(\mathbf{0})$, i.e. the charge of $\rho(\mathbf{r})$. We therefore introduce a vector $\mathcal{L}$ of Lagrange multipliers and seek a stationary point of the functional

$$Z = \langle\hat{R}|\omega_p|\hat{R}\rangle + 2\mathcal{L}^{\mathrm{T}}(\mathbb{K}^{\mathrm{T}}\mathbf{c} - \mathbb{q}) \quad (21)$$

1.3. Weight Function. The parameter p in (19) determines the type of model that we form and we will write $L_p(m)$ to denote an m-Gaussian least-squares model with parameter p and charge conservation

$$\hat{\rho}(\mathbf{0}) = \hat{\chi}(\mathbf{0}) \quad (22)$$

constrained. If $p > 0$, the fit emphasizes high-frequency components of the density; if $p < 0$, it emphasizes low-frequency components. Historically, $p = 3/2$, $1/2$, or $-1/2$ have been advocated, but these are just three points on a continuum of possibilities.

$L_{3/2}(m)$ models minimize the square of the residual *density* $R(\mathbf{r})$ itself. Such models were used by Reeves and Fletcher[18] to perform approximate Slater orbital calculations, O-ohata et al.[19] to construct expansions of Slater functions, Newton et al.[20] in the PDDO approximation, Hehre et al.[21,22] to construct the STO-nG basis sets, Billingsley and Bloor[23] in the LEDO approximation and Baerends et al.[24] in their Hartree−Fock−

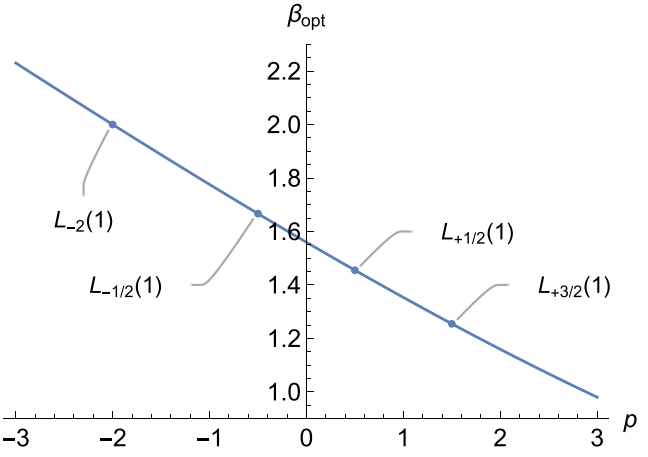


Figure 1. Optimal β for modeling a Slater function as the weight function $\omega_p(k)$ varies.

https://doi.org/10.1021/acs.jpca.3c04363
J. Phys. Chem. A 2023, 127, 9346−9356

Slater algorithm. Most of these authors used unconstrained fits, but Baerends et al. constrained charge conservation.

$L_{1/2}(m)$ models minimize the squared norm of the *electric field* of $R(\mathbf{r})$. Such models were advocated by Whitten[25] for modeling densities without charge conservation, used by Dunlap et al.[27] in his Coulomb fitting approach, developed further by Feyereisen et al.[29] and Eichkorn et al.[30] and analyzed in detail by Hall and Martin.[31]

$L_{-1/2}(m)$ models minimize the square of the *potential* value of $R(\mathbf{r})$. They were introduced by Fortunelli and Salvetti[28] and defined rigorously by Gill et al.[32]

The constraint (22) ensures that the integral in (21) exists when weight functions with $-2 < p \leq 0$ are used but, for $p \leq -2$, additional constraints must be applied. For certain isolated p values, for example $p = 0$, the weight function (19) is not positive definite and the functional (21) does not have a well-defined stationary point. However, if they are needed, such models can be obtained by taking limits in the neighborhood of p.

To illustrate this, we modeled a single Slater function by a single Gaussian function. In Fourier space, this corresponds to modeling $\hat{\rho}(\mathbf{k}) = (1 + k^2)^{-2}$ by $\hat{\chi}(\mathbf{k}) = \exp(-\beta k^2)$. By minimizing (21), one finds that the optimal β satisfies

$$(2\beta)^{p+1}\mathrm{U}(p+1, p, \beta) = 1 \tag{23}$$

where U is the Tricomi hypergeometric function.[38] Analysis of (23) reveals that

$$\beta_{\mathrm{opt}} = 2 + (4\ln 2 - 3)(p+2) + O(p+2)^2 \tag{24}$$

and Figure 1 shows that the optimal β varies smoothly and almost linearly over a wide domain.

1.4. Cost-Benefit Analysis. In a spatially extended system, most pairs of basis functions ϕ_μ and ϕ_ν are well separated, and the product $\phi_\mu(\mathbf{r})\phi_\nu(\mathbf{r})$ can therefore be neglected. It is easy to show that the number of *significant* $\phi_\mu(\mathbf{r})\phi_\nu(\mathbf{r})$ products grows only linearly with N and, thus, because the shell pair economizations are independent, the cost of economizing all significant shell pairs is $O(N)$. However, it is reasonable to ask whether this effort is worthwhile. There are two types of calculations where the cost is easy to justify.

First, calculations in which the electron density $D(\mathbf{r})$ will be evaluated at a large number K of points $\mathbf{r}$ as, for example, on a van der Waals surface. The cost of evaluating (2) at a point $\mathbf{r}$ is $O(N)$ and the total cost for all points is therefore $O(KN)$. In such situations, the economization cost will be more than compensated by the subsequent savings in density evaluations.

Second, calculations in which the shell pairs will be used to calculate two-electron integrals (9) as, for example, in a hybrid density functional calculation. The number of *significant* shell quartets, and therefore the number of significant ERIs, is $O(N^2)$. Thus, whereas the cost of economization scales linearly with N, the savings that accrue from the use of economized shell pairs to form ERIs scale quadratically with N and will more than compensate for the economization cost if N is sufficiently large.

2. METHODS

2.1. Basis Set Reconstruction. To facilitate the modeling process, we required all contraction coefficients to be positive. This is true of some basis functions (for example, the 5-fold contracted function on H atoms in the Dunning cc-pVTZ basis,[39] see Table 1) but it is not true of many basis functions (for example, the 10-fold contracted functions on C atoms in the same basis, see Table 2). However, it is often possible (at least, in shared-exponent basis sets such as those of Dunning[39] and Jensen[40]) to reconstruct the basis on each atom, creating all-positive basis functions by linearly combining the originals. Because the reconstructed basis set necessarily has the same span as the original, it is equivalent for quantum chemical purposes.

Table 1. cc-pVTZ Exponents and Coefficients[39,41] of an *s* Function on H

Primitive	Exponent	Coefficient
1	33.87	+0.006068
2	5.095	+0.045308
3	1.159	+0.202822
4	0.3258	+0.503903
5	0.1027	+0.383421

We formulate our problem as follows. Given a set V of linearly independent vectors $\{\mathbf{v}_1, ..., \mathbf{v}_n\} \in \mathbb{R}^\kappa$, we seek a set U of all-positive vectors $\{\mathbf{u}_1, ..., \mathbf{u}_n\}$ with $\mathrm{span}(U) = \mathrm{span}(V)$.

We define $\mathbf{v}_i^{\min}$ as the minimum (i.e., most negative) component of $\mathbf{v}_i$ and we then reorder the $\mathbf{v}_i$ so that $\mathbf{v}_1^{\min} > \mathbf{v}_2^{\min} > \cdots > \mathbf{v}_n^{\min}$. We then try to construct a *single* all-positive vector $\mathbf{u}_1$ from V by writing

$$\mathbf{u}_1 = \mathbf{v}_1 + \sum_{i=2}^{n} \gamma_i \mathbf{v}_i \tag{25}$$

and seeking the γ_i that maximize $\mathbf{u}_1^{\min}$. This maps onto the linear programming problem

$$\text{maximize } t \text{ subject to } v_{1j} + \sum_{i=2}^{n} \gamma_i v_{ij} \geq t \quad (j = 1, 2, \cdots, \kappa) \tag{26}$$

with the auxiliary variable t, and solved, for example, by converting to canonical form and applying the Simplex Method.[42] The optimized t is maximized $\mathbf{u}_1^{\min}$.

If the resulting $\mathbf{u}_1^{\min}$ value is positive, the original vectors are reconstructible. We then form

$$\mathbf{u}_i = \mathbf{v}_i + \delta_i \mathbf{u}_{i-1} \qquad (i = 2, ..., n) \tag{27}$$

where δ_i is the smallest value that yields $\mathbf{u}_i^{\min} = 0$, and normalize the resulting basis functions. This has the additional benefit of reducing the order of contraction by one even before modeling. We add the prefix "rec-" to the name of a reconstructed basis set.

It is important to note that the reconstructed basis functions are not orthogonal and, in some cases, can be almost parallel. As a consequence, we hope that, in the future, (a) improved reconstruction algorithms will be devised and (b) basis set developers will produce basis sets with all-positive coefficients, so that reconstruction is unnecessary.

If this method is applied to the 10-fold contracted functions on C in the cc-pVTZ basis set (Table 2), we find that $\mathbf{u}_1^{\min}$ is a concave piecewise linear function of γ_2 given by

$$\mathbf{u}_1^{\min} = \begin{cases} -0.008983 + 0.598684\gamma_2 & \gamma_2 < 0.0158885 \\ +0.000531 - 0.000113\gamma_2 & 0.0158885 \leq \gamma_2 < 2.03008 \\ +0.346129 - 0.170352\gamma_2 & 2.03008 \leq \gamma_2 \end{cases} \tag{28}$$

https://doi.org/10.1021/acs.jpca.3c04363
J. Phys. Chem. A 2023, 127, 9346−9356

Table 2. cc-pVTZ Exponents and Coefficients[39,41] of Two *s* Functions on C

		Coefficient			
		cc-pVTZ	rec-cc-pVTZ		
Primitive	Exponent	$\mathbf{v}_1$	$\mathbf{v}_2$	$\mathbf{u}_1$	$\mathbf{u}_2$
1	8236.	+0.000531	−0.000113	+0.000529	+0.000133
2	1235.	+0.004108	−0.000878	+0.004094	+0.001026
3	280.8	+0.021087	−0.004540	+0.021012	+0.005238
4	79.27	+0.081853	−0.018133	+0.081555	+0.019875
5	25.59	+0.234817	−0.055760	+0.233901	+0.053661
6	8.997	+0.434401	−0.126895	+0.432330	+0.077970
7	3.319	+0.346129	−0.170352	+0.343379	0
8	0.9059	+0.039378	+0.140382	+0.041603	+0.143342
9	0.3643	−0.008983	+0.598684	+0.000529	+0.533186
10	0.1285	+0.002385	+0.395389	+0.008666	+0.355805

which reaches its maximum value (+0.000529) at γ_2 = 0.0158885. We then form

$$\mathbf{u}_2 = \mathbf{v}_2 + 0.496042\mathbf{u}_1 \tag{29}$$

and renormalize both $\mathbf{u}_1$ and $\mathbf{u}_2$. The rec-cc-pVTZ basis functions are shown in Table 2 and we note that their overlap integral is S_{12} = 0.456.

2.2. Modeling Theory. *2.2.1. Linear Parameters.* The basic units of the optimization are the scalar

$$Z_0 = \langle\hat{\rho}|\omega_p|\hat{\rho}\rangle \tag{30}$$

the vector

$$\mathbb{f}_{00} = \langle\hat{b}_i|\omega_p|\hat{\rho}\rangle \tag{31}$$

and the matrix

$$\mathbb{F}_{00} = \langle\hat{b}_i|\omega_p|\hat{b}_j\rangle \tag{32}$$

From these, we can assemble the augmented vector

$$\mathbf{f}_{00} = \begin{pmatrix} \mathbb{f}_{00} \\ \mathbb{q} \end{pmatrix} \tag{33}$$

and the augmented matrix

$$\mathbf{F}_{00} = \begin{pmatrix} \mathbb{F}_{00} & \mathbb{K} \\ \mathbb{K}^{\mathrm{T}} & \mathbf{0} \end{pmatrix} \tag{34}$$

and then write (21) as

$$Z = Z_0 - 2\mathbf{f}_{00}^{\mathrm{T}}\mathbf{x} + \mathbf{x}^{\mathrm{T}}\mathbf{F}_{00}\mathbf{x} \tag{35}$$

where the coefficients and Lagrange multipliers have been conflated into the vector

$$\mathbf{x} = \begin{pmatrix} \mathbf{c} \\ \mathcal{L} \end{pmatrix} \tag{36}$$

Requiring that the gradient of (35) with respect to $\mathbf{x}$ vanish yields the optimal coefficients and Lagrange multipliers

$$\mathbf{x} = \mathbf{F}_{00}^{-1}\mathbf{f}_{00} \tag{37}$$

and the resulting minimal value

$$Z = Z_0 - \mathbf{f}_{00}^{\mathrm{T}}\mathbf{x} \tag{38}$$

For greatest numerical stability, (37) should be computed by solving a linear system.

Having thus solved for the linear parameters c_i and $\mathcal{L}_i$, we now have a functional (38) that depends only on the nonlinear parameters β_i and B_i. We now discuss the minimization of (38) with respect to those parameters.

2.2.2. Initial Guesses: Spherical Density. If $\rho(\mathbf{r})$ has spherical (K_h) symmetry, we choose all B_i = 0 and eschew further optimization. The exponents $\boldsymbol{\beta}$ are more challenging, and to construct an initial guess for these, we convert the modeling problem into a quadrature problem. First, we form the integral representation

$$\hat{\rho}(k) = \int_0^1 u^{\beta_0 k^2} w(u)\, \mathrm{d}u \tag{39}$$

where $\beta_0 > 0$ is a scale factor and an Inverse Mellin Transform[38] provides the weight function

$$w(u) = u^{-1}\mathcal{M}^{-1}\{\hat{\rho}(\sqrt{s/\beta_0})\}(u) \tag{40}$$

We then approximate (39) by the *m*-point Gauss−Christoffel quadrature formula[43]

$$\hat{\rho}(k) \approx \sum_{i=1}^{m} c_i u_i^{\beta_0 k^2} \tag{41}$$

where the weights $c_i > 0$ and roots $0 < u_i < 1$ are related to polynomials that are orthogonal on [0, 1] with respect to $w(u)$. In the Golub-Welsch method[44] (see Supporting Information), the c_i and u_i are found from the moments

$$\langle u^j\rangle = \int_0^1 u^j w(u)\, \mathrm{d}u = \hat{\rho}(\sqrt{j/\beta_0}) \quad (j = 0, 1, \ldots, 2m) \tag{42}$$

and, in order to sample $\hat{\rho}(k)$ well, we choose the arbitrary scale factor to be

$$\beta_0 = \sqrt{m}/\langle k^2\rangle = \sqrt{m}\left[\int_0^\infty \hat{\rho}(k)\, \mathrm{d}k\right]\Big/\left[\int_0^\infty k^2\hat{\rho}(k)\, \mathrm{d}k\right] \tag{43}$$

Finally, by substituting $u_i = \exp(-\beta_i/\beta_0)$ into (41), we obtain an approximation that we call the Q(m) model

$$\hat{\rho}(k) \approx \sum_{i=1}^{m} c_i \exp(-\beta_i k^2) \tag{44}$$

Although this model is less accurate than least-squares fitting, it may sometimes be useful in its own right. However, its low cost and reasonable accuracy make it a useful initial guess.

https://doi.org/10.1021/acs.jpca.3c04363
J. Phys. Chem. A 2023, 127, 9346−9356

There is an important caveat to the use of a quadrature scheme to generate an initial guess. There is no guarantee[45] that the Gauss−Christoffel roots u_i and weights c_i will exist or be satisfactory unless the weight function $w(u)$ in (39) is a non-negative function on $[0, 1]$. It was to ensure this that we assumed that all of the coefficients d_i in (10) are positive.

2.2.3. Initial Guesses: Cylindrical Density. Generating guesses for the exponents β_i and centers B_i when $\rho(\mathbf{r})$ has cylindrical ($C_{\infty v}$) symmetry is not difficult. In general, because of the primitive prefactors (6), the coefficients d_i of the n primitives in $\rho(\mathbf{r})$ range over many orders of magnitude and $\rho(\mathbf{r})$ is therefore dominated by the primitives with the largest absolute coefficients $|d_i|$. We exploit this by using the α_i and A_i of the m primitives with the largest $|d_i|$ as our guesses.

We start by ordering the α_i and A_i by $|d_i|$ value and take the first m primitives that are not already "covered" by a larger primitive. A primitive i is considered "covered" if its A_i value is within 0.25 au of the position A_j of a larger primitive j ($|d_j| > |d_i|$). This is because primitives that are close to each other tend to be represented by a single Gaussian in models, even if they both have large coefficients.

In the case of a tie for the mth largest "uncovered" primitive, there are two options. If it would not break another tie, the $(m-1)$th largest primitive is excluded, and the α_i and A_i of the mth and $(m+1)$th largest primitives are used. Otherwise, the average of the mth and $(m+1)$th largest primitives' quantities are taken.

Finally, if the guess formed this way has $k < m$ primitives, the $m-k$ largest previously excluded "covered" primitives are included as well.

2.2.4. Matrix Elements and Their Derivatives. For the basis functions (17), one can show from eq (6.631.1) of Gradshteyn and Ryzhik[46] that both the matrix (32) and its derivatives with respect to the nonlinear parameters (β_i and B_i) can be conveniently expressed in terms of the function

Table 3. Construction of an *m*-Gaussian Model for a Gaussian Density

Step	Description	Equations	Cost
1	Input ρ, m, p, $\mathbb{K}$, q		$O(1)$
	If ρ is spherical		
2a	Compute $\langle u^j \rangle$ and β_0	(42), (43)	$O(m)$
3a	Compute initial guess for β	(44)	$O(m^2)$
4a	Set B = 0		$O(m)$
	else if ρ is cylindrical		
2b	Sort primitives by $\|d_i\|$ value		$O(n)$
3b	Compute initial guess for β and B		$O(m)$
	end if		
5	Compute $\mathbf{f}_{00}$, $\mathbf{f}_{10}$, $\mathbf{f}_{01}$, $\mathbf{f}_{20}$, $\mathbf{f}_{11}$, $\mathbf{f}_{02}$	(33), (52)−(56)	$O(m)$
6	Compute $\mathbf{F}_{00}$, $\mathbf{F}_{10}$, $\mathbf{F}_{01}$, $\mathbf{F}_{20}$, $\mathbf{F}_{11}$, $\mathbf{F}_{02}$	(34), (52)−(56)	$O(m^2)$
7	Compute E_{10}, E_{01}, E_{20}, E_{11}, E_{02}	(59) − (63)	$O(m)$
8	Compute U_{10}, U_{01}, H_{20}, H_{11}, H_{02}	(59) − (63)	$O(m^3)$
9	Compute $\mathbf{x}$	(37)	$O(m^3)$
10	Compute $\mathbf{g}$	(57)	$O(m^2)$
11	Compute $\mathbf{H}$	(58)	$O(m^3)$
12	Compute $(\boldsymbol{\beta}, \mathbf{B}) \leftarrow (\boldsymbol{\beta}, \mathbf{B}) - \boldsymbol{\Delta}$	(69)	$O(m^3)$
13	If $\|\boldsymbol{\Delta}\| > 10^{-4}$ or **H is not positive definite**, go to Step 5		$O(m^3)$

$$\Phi_{s,t}(\zeta, R) = \frac{\Gamma(s)}{\zeta^s} \mathrm{M}\left[s, \frac{3}{2} + t, -\frac{R^2}{4\zeta}\right] \tag{45}$$

where Γ is the gamma function[38] and M is the Kummer hypergeometric function.[38] The matrix itself is given by

$$\mathbb{F}_{00} = \Phi_{p,0}(\beta_i + \beta_j, B_i - B_j) \tag{46}$$

and, using properties of the hypergeometric function, one finds the first derivatives

$$\frac{\partial}{\partial \zeta}\Phi_{p,0}(\zeta, R) = -\Phi_{p+1,0}(\zeta, R) \tag{47}$$

$$\frac{\partial}{\partial R}\Phi_{p,0}(\zeta, R) = -\frac{R}{3}\Phi_{p+1,1}(\zeta, R) \tag{48}$$

and the second derivatives

$$\frac{\partial^2}{\partial \zeta^2}\Phi_{p,0}(\zeta, R) = \Phi_{p+2,0}(\zeta, R) \tag{49}$$

$$\frac{\partial^2}{\partial \zeta \partial R}\Phi_{p,0}(\zeta, R) = \frac{R}{3}\Phi_{p+2,1}(\zeta, R) \tag{50}$$

$$\frac{\partial^2}{\partial R^2}\Phi_{p,0}(\zeta, R) = \frac{R^2}{15}\Phi_{p+2,2}(\zeta, R) - \frac{1}{3}\Phi_{p+1,1}(\zeta, R) \tag{51}$$

For half-integer values of p, all six of these can be expressed in terms of Gaussians and the error function.[38]

To construct the gradient of Z, we require the first derivative vectors and matrices

$$\mathbf{f}_{10} = \partial \mathbf{f}_{00}/\partial \beta_i \quad \mathbf{F}_{10} = \partial \mathbf{F}_{00}/\partial \beta_i \tag{52}$$

$$\mathbf{f}_{01} = \partial \mathbf{f}_{00}/\partial B_i \quad \mathbf{F}_{01} = \partial \mathbf{F}_{00}/\partial B_i \tag{53}$$

and, to construct its Hessian, we need the second derivative vectors and matrices

$$\mathbf{f}_{20} = \partial^2 \mathbf{f}_{00}/\partial \beta_i^2 \quad \mathbf{F}_{20} = \partial^2 \mathbf{F}_{00}/\partial \beta_i^2 \tag{54}$$

$$\mathbf{f}_{11} = \partial^2 \mathbf{f}_{00}/\partial \beta_i \partial B_i \quad \mathbf{F}_{11} = \partial^2 \mathbf{F}_{00}/\partial \beta_i \partial B_i \tag{55}$$

$$\mathbf{f}_{02} = \partial^2 \mathbf{f}_{00}/\partial B_i^2 \quad \mathbf{F}_{02} = \partial^2 \mathbf{F}_{00}/\partial B_i^2 (1 - \delta_{ij}) \tag{56}$$

Differentiating (38) shows the gradient and Hessian of Z with respect to $\boldsymbol{\beta}$ and $\mathbf{B}$ are

$$\tilde{\mathbf{g}} = 2\begin{pmatrix} [\mathbf{E}_{10}\mathbf{x}] \\ [\mathbf{E}_{01}\mathbf{x}] \end{pmatrix} \tag{57}$$

$$\tilde{\mathbf{H}} = 2\begin{pmatrix} [\tilde{\mathbf{H}}_{20}] & [\tilde{\mathbf{H}}_{11}] \\ [\tilde{\mathbf{H}}_{11}^{\mathrm{T}}] & [\mathbf{H}_{02}] \end{pmatrix} \tag{58}$$

where

$$\mathbf{E}_{10} = \mathrm{diag}(\mathbf{F}_{10}\mathbf{x} - \mathbf{f}_{10}) \quad \mathbf{U}_{10} = \mathbf{X}\mathbf{F}_{10} + \mathbf{E}_{10} \tag{59}$$

$$\mathbf{E}_{01} = \mathrm{diag}(\mathbf{F}_{01}\mathbf{x} - \mathbf{f}_{01}) \quad \mathbf{U}_{01} = \mathbf{X}\mathbf{F}_{01} + \mathbf{E}_{01} \tag{60}$$

$$\mathbf{E}_{20} = \mathrm{diag}(\mathbf{F}_{20}\mathbf{x} - \mathbf{f}_{20})$$
$$\tilde{\mathbf{H}}_{20} = \mathbf{X}(\mathbf{E}_{20} + \mathbf{F}_{20}\mathbf{X}) - \mathbf{U}_{10}\mathbf{F}_{00}^{-1}\mathbf{U}_{10}^{\mathrm{T}} \tag{61}$$

https://doi.org/10.1021/acs.jpca.3c04363
J. Phys. Chem. A 2023, 127, 9346−9356

Table 4. Models for the H($1s^2$) cc-pVTZ Density

	$Q(m)$		$L_{-1/2}(m)$		$L_{+1/2}(m)$		$L_{+3/2}(m)$	
m	λ_i	c_i	λ_i	c_i	λ_i	c_i	λ_i	c_i
1	+1.032	1.000	+0.876	1.000	+1.011	1.000	+1.159	1.000
	$\mathcal{E} = 1.8 \times 10^{-1}$		$\mathcal{E} = 1.7 \times 10^{-1}$		$\mathcal{E} = 1.8 \times 10^{-1}$		$\mathcal{E} = 2.1 \times 10^{-1}$	
2	+0.421	0.610	+0.315	0.572	+0.497	0.700	+0.719	0.824
	+1.827	0.390	+1.802	0.428	+2.083	0.300	+2.455	0.176
	$\mathcal{E} = 5.1 \times 10^{-2}$		$\mathcal{E} = 5.5 \times 10^{-2}$		$\mathcal{E} = 4.2 \times 10^{-2}$		$\mathcal{E} = 6.7 \times 10^{-2}$	
3	+0.004	0.262	+0.027	0.308	+0.219	0.456	+0.439	0.621
	+0.935	0.533	+1.102	0.563	+1.396	0.476	+1.741	0.348
	+2.250	0.204	+2.583	0.129	+2.974	0.068	+3.430	0.031
	$\mathcal{E} = 2.8 \times 10^{-2}$		$\mathcal{E} = 1.8 \times 10^{-2}$		$\mathcal{E} = 1.2 \times 10^{-2}$		$\mathcal{E} = 2.0 \times 10^{-2}$	
4	−0.174	0.161	−0.139	0.186	−0.036	0.249	+0.316	0.519
	+0.660	0.465	+0.765	0.521	+0.908	0.521	+1.478	0.412
	+1.564	0.295	+1.849	0.260	+2.027	0.208	+2.838	0.066
	+2.776	0.080	+3.361	0.032	+3.581	0.022	+4.621	0.003
	$\mathcal{E} = 1.2 \times 10^{-2}$		$\mathcal{E} = 5.4 \times 10^{-3}$		$\mathcal{E} = 4.5 \times 10^{-3}$		$\mathcal{E} = 1.2 \times 10^{-2}$	
5	−0.196	0.147	−0.165	0.169	−0.130	0.189	−0.057	0.228
	+0.550	0.325	+0.707	0.499	+0.758	0.506	+0.831	0.496
	+1.008	0.254	+1.709	0.278	+1.787	0.261	+1.854	0.238
	+1.833	0.233	+2.917	0.050	+3.074	0.042	+3.148	0.037
	+3.183	0.041	+4.455	0.004	+4.810	0.002	+4.924	0.002
	$\mathcal{E} = 5.7 \times 10^{-3}$		$\mathcal{E} = 1.6 \times 10^{-3}$		$\mathcal{E} = 6.6 \times 10^{-4}$		$\mathcal{E} = 7.1 \times 10^{-4}$	
6	−0.196	0.147	−0.190	0.152	−0.176	0.161	−0.135	0.183
	+0.539	0.305	+0.610	0.412	+0.660	0.459	+0.730	0.490
	+0.961	0.255	+1.225	0.211	+1.436	0.214	+1.684	0.254
	+1.738	0.219	+1.927	0.187	+2.043	0.132	+2.421	0.046
	+2.411	0.049	+3.147	0.036	+3.192	0.033	+3.268	0.026
	+3.456	0.024	+4.874	0.002	+4.936	0.002	+4.975	0.002
	$\mathcal{E} = 4.7 \times 10^{-3}$		$\mathcal{E} = 2.5 \times 10^{-4}$		$\mathcal{E} = 1.5 \times 10^{-4}$		$\mathcal{E} = 1.8 \times 10^{-4}$	

$$\mathbf{E}_{11} = \mathrm{diag}(\mathbf{F}_{11}\mathbf{x} - \mathbf{f}_{11})$$
$$\tilde{\mathbf{H}}_{11} = \mathbf{X}(\mathbf{E}_{11} - \mathbf{F}_{11}\mathbf{X}) - \mathbf{U}_{10}\mathbf{F}_{00}^{-1}\mathbf{U}_{01}^{\mathrm{T}} \quad (62)$$

$$\mathbf{E}_{02} = \mathrm{diag}(\mathbf{F}_{02}\mathbf{x} - \mathbf{f}_{02})$$
$$\mathbf{H}_{02} = \mathbf{X}(\mathbf{E}_{02} - \mathbf{F}_{02}\mathbf{X}) - \mathbf{U}_{01}\mathbf{F}_{00}^{-1}\mathbf{U}_{01}^{\mathrm{T}} \quad (63)$$

and $\mathbf{X} = \mathrm{diag}(\mathbf{x})$. The square brackets in (57) and (58) indicate that only the first m elements of the enclosed vector, or the first m-by-m block of the enclosed matrix, is retained.

It is well-known, however, that the optimal exponents of Gaussian expansions of exponential densities form a roughly geometrical sequence[35,47] and that the exponents typically span several orders of magnitude. We therefore choose to work with the log-exponents

$$\lambda_i = -\ln \beta_i \quad (64)$$

which are roughly evenly spaced and typically of the order of unity. By the chain rule, the elements of the gradient and Hessian of Z with respect to the λ_i are

$$[\mathbf{g}] = -\mathrm{diag}(\boldsymbol{\beta})[\tilde{\mathbf{g}}] \quad (65)$$

$$\mathbf{H}_{20} = -\mathrm{diag}([\mathbf{g}]) + \mathrm{diag}(\boldsymbol{\beta})[\tilde{\mathbf{H}}_{20}]\,\mathrm{diag}(\boldsymbol{\beta}) \quad (66)$$

$$\mathbf{H}_{11} = -\mathrm{diag}(\boldsymbol{\beta})[\tilde{\mathbf{H}}_{11}] \quad (67)$$

$$\mathbf{H} = 2\begin{pmatrix} \mathbf{H}_{20} & \mathbf{H}_{11} \\ \mathbf{H}_{11}^{\mathrm{T}} & [\mathbf{H}_{02}] \end{pmatrix} \quad (68)$$

2.2.5. Optimization Step. Given the gradient ($\mathbf{g}$) and Hessian ($\mathbf{H}$) of (38), we are in a position to find the optimal $\boldsymbol{\beta}$ and $\mathbf{B}$ via an iterative scheme. Although the Newton−Raphson method could be used, it is effective only if current $\boldsymbol{\beta}$ and $\mathbf{B}$ are near-optimal and can be disastrous otherwise. We prefer the slower (but more robust) approach taken by the Levenberg−Marquardt method,[48,49] wherein a small fraction $\sigma > 0$ of the identity matrix is added to the Hessian to form the shifted-Newton step

$$\boldsymbol{\Delta} = (\mathbf{H} + \sigma\mathbf{I})^{-1}\mathbf{g} \quad (69)$$

The shift σ should decay as the optimization proceeds and we have found that $\sigma = 10|\mathbf{g}|$ is effective for all the densities discussed below and hundreds of others that we have explored.

When optimizing $\boldsymbol{\beta}$ and $\mathbf{B}$, our goal is to reduce Z to within 1% of its minimum value. This is usually achieved when $\mathbf{H}$ is positive definite and $|\boldsymbol{\Delta}| < 10^{-4}$.

2.2.6. Algorithm. The inputs to our algorithm are the density $\rho(\mathbf{r})$, the number m of Gaussians in the desired model, the weight function parameter p, the constraint matrix $\mathbb{K}$ and the constraint vector $\mathbb{q}$. The way in which we generate initial guesses for the nonlinear ($\boldsymbol{\beta}$ and $\mathbf{B}$) parameters depends on whether the Gaussian density is spherical or cylindrical, but after that, the algorithm becomes the same. The overall algorithm, together with links to key equations, is shown in Table 3.

3. RESULTS AND DISCUSSION

3.1. Spherical Densities. *3.1.1. Preamble.* Armed with the necessary theory, we now consider how it can be applied in calculations using the Dunning cc-pVTZ basis set.[39] If the basis

https://doi.org/10.1021/acs.jpca.3c04363
J. Phys. Chem. A 2023, 127, 9346−9356

Table 5. Models for the rec-cc-pVTZ C(1s2s) Density (S = 0.456)

	$Q(m)$		$L_{-1/2}(m)$		$L_{+1/2}(m)$		$L_{+3/2}(m)$	
m	λ_i	c_i/S	λ_i	c_i/S	λ_i	c_i/S	λ_i	c_i/S
1	4.031	1.000	3.386	1.000	3.685	1.000	3.976	1.000
	$\mathcal{E} = 6.3 \times 10^{-1}$		$\mathcal{E} = 4.3 \times 10^{-1}$		$\mathcal{E} = 4.9 \times 10^{-1}$		$\mathcal{E} = 6.0 \times 10^{-1}$	
2	3.273	0.725	2.349	0.441	2.835	0.642	3.284	0.796
	5.098	0.275	4.445	0.559	4.954	0.358	5.454	0.204
	$\mathcal{E} = 1.5 \times 10^{-1}$		$\mathcal{E} = 1.6 \times 10^{-1}$		$\mathcal{E} = 1.4 \times 10^{-1}$		$\mathcal{E} = 2.1 \times 10^{-1}$	
3	2.855	0.524	1.423	0.117	2.334	0.382	2.872	0.615
	4.150	0.332	3.159	0.572	3.958	0.484	4.657	0.342
	5.561	0.144	5.036	0.312	5.699	0.134	6.423	0.043
	$\mathcal{E} = 6.1 \times 10^{-2}$		$\mathcal{E} = 1.1 \times 10^{-1}$		$\mathcal{E} = 5.1 \times 10^{-2}$		$\mathcal{E} = 8.7 \times 10^{-2}$	
4	2.625	0.416	0.854	0.046	1.708	0.155	2.574	0.467
	3.703	0.298	2.622	0.396	3.165	0.485	4.104	0.396
	4.715	0.216	4.027	0.419	4.646	0.307	5.504	0.127
	6.015	0.070	5.646	0.139	6.272	0.053	7.218	0.010
	$\mathcal{E} = 3.1 \times 10^{-2}$		$\mathcal{E} = 4.8 \times 10^{-2}$		$\mathcal{E} = 1.9 \times 10^{-2}$		$\mathcal{E} = 4.0 \times 10^{-2}$	
5	2.544	0.370	0.494	0.023	1.195	0.071	2.273	0.321
	3.369	0.167	2.117	0.187	2.750	0.403	3.596	0.412
	3.941	0.220	3.216	0.416	4.018	0.352	4.874	0.221
	4.877	0.190	4.554	0.306	5.279	0.156	6.180	0.044
	6.191	0.053	6.098	0.068	6.871	0.018	7.879	0.003
	$\mathcal{E} = 2.6 \times 10^{-2}$		$\mathcal{E} = 2.5 \times 10^{-2}$		$\mathcal{E} = 9.8 \times 10^{-3}$		$\mathcal{E} = 2.1 \times 10^{-2}$	
6	2.182	0.143	0.217	0.013	0.957	0.049	1.958	0.201
	2.768	0.281	1.602	0.081	2.534	0.318	3.208	0.416
	3.682	0.266	2.775	0.369	3.609	0.341	4.437	0.276
	4.457	0.156	3.924	0.331	4.732	0.225	5.576	0.093
	5.177	0.119	5.109	0.179	5.920	0.062	6.875	0.013
	6.424	0.035	6.624	0.028	7.503	0.006	8.583	0.001
	$\mathcal{E} = 1.7 \times 10^{-2}$		$\mathcal{E} = 1.1 \times 10^{-2}$		$\mathcal{E} = 3.7 \times 10^{-3}$		$\mathcal{E} = 1.2 \times 10^{-2}$	

functions are concentric, their density is spherical and has the Fourier transform (14)

$$\hat{\rho}(k) = \sum_{i=1}^{n} d_i \exp(-\alpha_i k^2) \tag{70}$$

and the associated weight function (40) is

$$w(u) = u^{-1} \sum_{i=1}^{n} d_i \delta(\ln u + \alpha_i/\beta_0) \tag{71}$$

where δ is the Dirac delta. If all $d_i > 0$, then $w(u)$ is non-negative and, from (43), we have

$$\beta_0 = 2\sqrt{m} \left[\sum_{i=1}^{n} d_i \alpha_i^{-1/2}\right] \Big/ \left[\sum_{i=1}^{n} d_i \alpha_i^{-3/2}\right] \tag{72}$$

To find the $Q(m)$, we need the moments

$$\hat{\rho}(\sqrt{j/\beta_0}) = \sum_{i=1}^{n} d_i \exp(-j\alpha_i/\beta_0) \tag{73}$$

and, to find the $L_p(m)$, we need the integrals

$$Z_0 = \sum_{i=1}^{n} \sum_{j=1}^{n} d_i d_j \Phi_{p,0}(\alpha_i + \alpha_j, 0) \tag{74}$$

$$\mathbb{f}_{00} = \sum_{j=1}^{n} d_j \Phi_{p,0}(\alpha_j + \beta_i, 0) \tag{75}$$

The $Q(m)$ and $L_p(m)$ models are constructed to mimic the original density $\rho(\mathbf{r})$ as well as possible over all space. It is also important, however, to assess the resulting models $\chi(\mathbf{r})$ by measuring the largest *pointwise* difference between $\rho(\mathbf{r})$ and $\chi(\mathbf{r})$. For a spherical density, it is natural to measure this through the maximum Jacobian-weighted error

$$\mathcal{E} = \max[4\pi r^2 |\rho(r) - \chi(r)|] \tag{76}$$

and we report this for each of the models described below.

3.1.2. H(1s1s) Density. Our first example comes from the 5-fold contracted basis function on an H atom in the cc-pVTZ basis set (see Table 1). The product of this function with itself yields a density with n = 15 distinct Gaussians and its $Q(m)$ and $L_p(m)$ models are listed in Table 4.

The exponents of the $L_p(m)$ change smoothly and predictably as p and m are varied. The $L_{-1/2}(m)$ are the most diffuse (i.e., their λ_i are most negative) and the $L_{+3/2}(m)$ are the most compact (i.e., their λ_i are the most positive). This is reasonable because, whereas the $L_{-1/2}(m)$ use a Fourier weight $1/(2\pi\, k^4)$ that emphasizes low-frequency components of the density and therefore its tail, the $L_{+3/2}(m)$ use a weight $1/(2\pi k^0)$ that places greater emphasis on the high-frequency components and therefore the region near the nuclear cusp. The $L_{+1/2}(m)$ exponents are remarkably close to the average of the analogous $L_{-1/2}(m)$ and $L_{+3/2}(m)$ exponents, thus generalizing the near-linear behavior seen in Figure 1.

The exponents of $Q(m)$ also change smoothly with m. Their most diffuse exponents are similar to those in $L_{+1/2}(m)$ when m is small and similar to those in $L_{-1/2}(m)$ when m is larger but the range $\lambda_m - \lambda_1$ of the $Q(m)$ exponents is smaller than in any of the least-squares models. For applications for which a quick and reasonably accurate approximation is required, the $Q(m)$ models appear useful.

https://doi.org/10.1021/acs.jpca.3c04363
J. Phys. Chem. A 2023, 127, 9346−9356

Table 6. $L_{-1/2}(m)$ Models for cc-pVTZ H(1s)H(1s) Densities with Various R

	R = 4.928, S = 0.100			R = 7.725, S = 0.010			R = 9.995, S = 0.001		
m	B_i	λ_i	c_i/S	B_i	λ_i	c_i/S	B_i	λ_i	c_i/S
1	+0.000	+0.015	1.000	+0.000	−0.239	1.000	+0.000	−0.262	1.000
	$\mathcal{E} = 3.6 \times 10^{-3}$			$\mathcal{E} = 2.2 \times 10^{-4}$			$\mathcal{E} = 1.0 \times 10^{-5}$		
2	−1.142	+0.282	0.500	−1.287	+0.019	0.500	−0.990	−0.125	0.500
	+1.142	+0.282	0.500	+1.287	+0.019	0.500	+0.990	−0.125	0.500
	$\mathcal{E} = 1.8 \times 10^{-3}$			$\mathcal{E} = 9.1 \times 10^{-5}$			$\mathcal{E} = 3.9 \times 10^{-6}$		
3	−1.382	+0.709	0.274	−2.320	+0.657	0.133	−2.743	+0.561	0.062
	+0.000	−0.182	0.452	+0.000	−0.165	0.734	+0.000	−0.193	0.875
	+1.382	+0.709	0.274	+2.320	+0.657	0.133	+2.743	+0.561	0.062
	$\mathcal{E} = 1.1 \times 10^{-3}$			$\mathcal{E} = 3.6 \times 10^{-5}$			$\mathcal{E} = 9.6 \times 10^{-7}$		
4	−1.763	+0.913	0.168	−2.323	+0.670	0.132	−2.744	+0.561	0.062
	+0.000	−0.163	0.510	+0.000	−0.184	0.717	+0.000	+3.312	−0.0001
	+0.000	+0.662	0.155	+0.000	+0.583	0.018	+0.000	−0.192	0.875
	+1.763	+0.913	0.168	+2.323	+0.670	0.132	+2.744	+0.561	0.062
	$\mathcal{E} = 8.0 \times 10^{-4}$			$\mathcal{E} = 3.6 \times 10^{-5}$			$\mathcal{E} = 9.6 \times 10^{-7}$		
5	−2.229	+1.761	0.030	−3.276	+1.640	0.013	−4.252	+1.696	0.002
	−1.137	+0.554	0.290	−1.982	+0.531	0.146	−2.607	+0.541	0.062
	+0.000	−0.273	0.361	+0.000	−0.198	0.683	+0.000	−0.196	0.870
	+1.137	+0.554	0.290	+1.982	+0.531	0.146	+2.607	+0.541	0.062
	+2.229	+1.761	0.030	+3.276	+1.640	0.013	+4.252	+1.696	0.002
	$\mathcal{E} = 3.1 \times 10^{-4}$			$\mathcal{E} = 1.1 \times 10^{-5}$			$\mathcal{E} = 1.8 \times 10^{-7}$		
6	−2.138	+1.915	0.025	−3.327	+1.785	0.010			
	−1.368	+0.596	0.220	−2.037	+0.550	0.144			
	+0.000	−0.186	0.445	+0.000	−0.195	0.689			
	+0.000	+0.872	0.064	+0.000	+0.886	0.002			
	+1.368	+0.596	0.220	+2.037	+0.550	0.144			
	+2.138	+1.915	0.025	+3.327	+1.785	0.010			
	$\mathcal{E} = 2.7 \times 10^{-4}$			$\mathcal{E} = 9.0 \times 10^{-6}$					

3.1.3. C(1s2s) Density. Our second example comes from the product of the two 10-fold contracted rec-cc-pVTZ basis functions on a C atom (see Table 2). Their product yields a density with n = 55 distinct Gaussians and its $Q(m)$ and $L_p(m)$ models are listed in Table 5.

3.1.4. Other Densities. We have applied our algorithm for spherical densities to all of the concentric *ss* products arising in the cc-pVDZ and cc-pVTZ basis sets of Dunning and the pc-1 and pc-2 basis set of Jensen, constructing $L_{-1/2}(m)$, $L_{1/2}(m)$ and $L_{3/2}(m)$ models with m = 1,..., 6.

The maximum error $\mathcal{E}$ does not show a strong dependence on p but we find that the smallest $\mathcal{E}$ is often obtained by $L_{3/2}(m)$ if m is small, and by $L_{1/2}(m)$ if m is larger.

In many cases, the number m of Gaussians required to accurately model a density with a large number n of Gaussians is surprisingly small. The explanation for this encouraging discovery appears to lie in the distribution of exponents $\alpha_{\mu\nu}^{kl}$ in a typical basis function product. We see from (7) that the $\alpha_{\mu\nu}^{kl}$ are simple sums of the α_{μ}^{k} and α_{ν}^{l} of the parent basis functions but, because the latter typically form a roughly geometrical sequence, the distribution of $\alpha_{\mu\nu}^{kl}$ is clustered. To illustrate this, consider a hypothetical basis function with α_{μ}^{k} = {1000, 100, 10, 1}. The product of this with itself generates a Gaussian density with $\alpha_{\mu\mu}^{kl}$ = {2000, 1100, 1010, 1001, 200, 110, 101, 20, 11, 2}, and although these 10 exponents are strictly distinct, the second, third and fourth Gaussians are similar (as are the sixth and seventh Gaussians) and can be modeled quite well by a single Gaussian.

3.2. Cylindrical Densities. *3.2.1. Preamble.* If two basis functions are on different centers, separated by a distance $R > 0$, their product yields a Gaussian density $\rho(\mathbf{r})$ with cylindrical symmetry and, to find the $L_p(m)$, we need

$$Z_0 = \sum_{i=1}^{n}\sum_{j=1}^{n} d_i d_j \Phi_{p,0}(\alpha_i + \alpha_j, A_i - A_j) \tag{77}$$

$$\mathbb{f}_{00} = \sum_{j=1}^{n} d_j \Phi_{p+q,0}(\alpha_j + \beta_i, A_j - B_i) \tag{78}$$

The overlap $S = \int \rho(\mathbf{r})\, d\mathbf{r}$ of the functions decays rapidly as R increases and it is interesting to compare the accuracies of the $L_p(m)$ models as a function of R.

The $L_p(m)$ models are constructed to mimic the original density $\rho(\mathbf{r})$ as well as possible over all space. It is also important, however, to assess the resulting models $\chi(\mathbf{r})$ by measuring the largest *pointwise* difference between $\rho(\mathbf{r})$ and $\chi(\mathbf{r})$. For a cylindrical density, it is natural to measure this through the maximum axial error

$$\mathcal{E} = \max[2\pi\sqrt{x^2 + y^2}\,|\rho(\mathbf{r}) - \chi(\mathbf{r})|] \tag{79}$$

and we report this for each of the models described below.

3.2.2. H(1s)H(1s) Density. Our first example is the product of the 5-fold contracted cc-pVTZ basis functions (Table 1) on two hydrogen atoms centered at (0, 0, $\pm R/2$). The density consists of 25 Gaussians, and $L_{-1/2}(m)$ models with m = 1 to 6 Gaussians are shown in Table 6.

We have considered three values of R, chosen so that the resulting overlap S is close to 0.1, 0.01, or 0.001. These correspond roughly to the most distant H atoms in the

https://doi.org/10.1021/acs.jpca.3c04363
J. Phys. Chem. A 2023, 127, 9346−9356

Table 7. $L_{-1/2}(m)$ Models for rec-cc-pVTZ C(2s)H(1s) Densities with Various R

	R = 4.669, S = 0.100			R = 7.305, S = 0.010			R = 9.446, S = 0.001		
m	B_i	λ_i	c_i/S	B_i	λ_i	c_i/S	B_i	λ_i	c_i/S
1	−0.285	+0.158	1.000	−0.593	−0.090	1.000	−0.712	−0.134	1.000
	$\mathcal{E}$ = 3.8 × 10^{-3}			$\mathcal{E}$ = 3.2 × 10^{-4}			$\mathcal{E}$ = 1.8 × 10^{-5}		
2	−1.179	+0.418	0.561	−1.779	+0.310	0.480	−2.135	+0.266	0.303
	+0.889	+0.378	0.439	+0.567	+0.025	0.520	−0.072	−0.082	0.697
	$\mathcal{E}$ = 3.1 × 10^{-3}			$\mathcal{E}$ = 1.2 × 10^{-4}			$\mathcal{E}$ = 4.6 × 10^{-6}		
3	−1.320	+0.761	0.344	−2.306	+0.758	0.196	−2.797	+0.677	0.120
	−0.289	−0.099	0.387	−0.437	−0.041	0.699	−0.537	−0.073	0.838
	+1.079	+0.733	0.270	+1.869	+0.662	0.105	+2.171	+0.595	0.041
	$\mathcal{E}$ = 2.9 × 10^{-3}			$\mathcal{E}$ = 1.1 × 10^{-4}			$\mathcal{E}$ = 3.0 × 10^{-6}		
4	−1.691	+0.948	0.197	−3.396	+2.405	0.006	−4.158	+1.924	0.003
	−0.274	−0.039	0.487	−2.197	+0.707	0.203	−2.710	+0.653	0.121
	−0.170	+0.776	0.171	−0.434	−0.052	0.681	−0.533	−0.076	0.834
	+1.529	+0.954	0.145	+1.813	+0.646	0.110	+2.136	+0.594	0.042
	$\mathcal{E}$ =2.7 × 10^{-3}			$\mathcal{E}$ =8.6 × 10^{-5}			$\mathcal{E}$ =2.4 × 10^{-6}		
5	−2.280	+3.202	0.008	−3.232	+1.878	0.012	−4.037	+1.686	0.005
	−1.345	+0.719	0.329	−2.068	+0.649	0.222	−2.675	+0.641	0.122
	−0.230	−0.152	0.345	−0.418	−0.080	0.640	−0.530	−0.078	0.831
	+0.652	+0.625	0.274	+1.487	+0.568	0.115	+2.061	+0.595	0.041
	+1.978	+1.429	0.044	+2.849	+1.438	0.011	+3.881	+1.783	0.001
	$\mathcal{E}$ = 1.6 × 10^{-3}			$\mathcal{E}$ = 9.4 × 10^{-5}			$\mathcal{E}$ = 2.5 × 10^{-6}		
6	−2.249	+3.354	0.007	−3.573	+3.933	0.002	−4.647	+3.942	0.0004
	−1.460	+0.796	0.257	−2.812	+1.213	0.032	−3.766	+1.398	0.007
	−0.318	−0.053	0.436	−1.959	+0.590	0.211	−2.646	+0.625	0.121
	+0.013	+0.918	0.093	−0.406	−0.083	0.631	−0.527	−0.078	0.829
	+1.131	+0.611	0.184	+1.523	+0.582	0.116	+2.052	+0.598	0.041
	+1.915	+1.868	0.024	+2.929	+1.587	0.009	+3.830	+1.672	0.001
	$\mathcal{E}$ = 1.5 × 10^{-3}			$\mathcal{E}$ = 3.5 × 10^{-5}			$\mathcal{E}$ = 8.7 × 10^{-7}		

equilibrium structures of ethane, propane and butane, respectively.

In the least-squares approximation of smooth functions by other smooth functions, pointwise convergence is expected to be exponential[50] and this seems to be observed here. As the number m of Gaussians in the model grows, the maximum error $\mathcal{E}$ decays rapidly. In all cases, the optimal centers B_i are symmetrically and more or less uniformly distributed on the z axis, the central Gaussian is the most diffuse (its λ_i is the most negative) and the Gaussians furthest from the origin are the least diffuse and have the smallest coefficients c_i.

It is particularly encouraging to note that, for $R \approx 10$, the 25-Gaussian density can be approximated very accurately ($\mathcal{E} < 1 \times 10^{-6}$) by a model with just three Gaussians, or accurately ($\mathcal{E} = 1 \times 10^{-5}$) by a model with just a single Gaussian!

The reported m = 4 model for $R \approx 10$ is a local minimum, with a Z value 0.5% larger than the global minimum. The $\mathcal{E}$ values are the same. It seems clear that adding a fourth Gaussian to the model provides negligible improvement, so its precise λ_i and c_i have little bearing on the overall accuracy. The fifth Gaussian is needed to noticeably lower both Z and $\mathcal{E}$, but this is unimportant since this density can be modeled so accurately with even a single Gaussian.

3.2.3. C(2s)H(1s) Density. Our second example is the product of the 10-fold contracted rec-cc-pVTZ 2s basis function on a C atom at (0, 0, −R/2) with the 5-fold contracted cc-pVTZ basis function on a H atom at (0, 0, +R/2) (see Tables 1 and 2). The product consists of 50 Gaussians, and $L_{-1/2}(m)$ models with m = 1 to 6 Gaussians are shown in Table 7.

As before, we have considered three values of R, chosen so that the resulting overlap S is close to 0.1, 0.01, or 0.001. These correspond roughly to the most separated C and H atoms in the equilibrium structures of ethane, propane and butane, respectively.

Although the maximum pointwise error $\mathcal{E}$ of the models almost always decreases as m grows, its decay is slower than for the H(1s)H(1s) densities above. More detailed investigation revealed that the pointwise error is usually largest (and oscillating) in a small region close to the carbon nucleus, but because of volume effects, this error contributes relatively little to the integral in (21) and is therefore only slowly reduced as more Gaussians are added to the model.

As was observed for the H(1s)H(1s) densities, the models in Table 7 are usually dominated by a diffuse Gaussian near the origin, flanked by tighter Gaussians with smaller coefficients, and the central dominance is even more marked in the large-R densities. It is clear from the $\mathcal{E}$ values that the 50-Gaussian densities arising from the overlap of well separated C(2s) and H(1s) functions can be very accurately modeled by models with very small numbers of Gaussians.

4. CONCLUSIONS

The electron density $D(\mathbf{r})$ is a fundamental quantity in all quantum chemical calculations and, in order that those calculations be efficient, it is desirable that $D(\mathbf{r})$ be represented as compactly as possible. Most of the successful previous work in this area projects $D(\mathbf{r})$ into a global auxiliary basis set but, in the present work, we have sought instead to approximate each

https://doi.org/10.1021/acs.jpca.3c04363
J. Phys. Chem. A 2023, 127, 9346−9356

basis function product that contributes to $D(\mathbf{r})$ in a small, local basis that is explicitly tailored to that product. This is achieved by performing a nonlinear optimization with a few adjustable parameters for each product, and we have presented expressions for all of the first and second derivatives required to use Newton's method or a similar second-order scheme.

We have applied our algorithm to many of the products that arise when Dunning's cc-pVnZ or Jensen's pc-n basis sets are used, and we find that, in many cases, products consisting of many primitive Gaussians can be accurately approximated by models with only a few, well-optimized Gaussians. We anticipate that when such models are used in quantum chemical calculations, the resulting errors will be small while the computational speed will be significantly enhanced.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jpca.3c04363.

The Golub-Welsch algorithm (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Peter M. W. Gill − *School of Chemistry, University of Sydney, Camperdown, NSW 2006, Australia;* orcid.org/0000-0003-1042-6331; Phone: +61 2 8627 9486; Email: p.gill@sydney.edu.au

Author

Ellena K. G. Black − *School of Chemistry, University of Sydney, Camperdown, NSW 2006, Australia*

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jpca.3c04363

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

E.K.G.B. thanks the University of Sydney for a Ph.D. scholarship and a Postgraduate Teaching Fellowship.

■ REFERENCES

(1) Hohenberg, P.; Kohn, W. Inhomogeneous electron gas. *Phys. Rev.* **1964**, *136*, B864−B871.

(2) Davidson, E. R. *Reduced density matrices in quantum chemistry*; Academic: New York, 1976; pp 1−2.

(3) Parr, R. G.; Yang, W. *Density-functional theory of atoms and molecules*; Clarendon Press: Oxford, U.K., 1989; pp 24−26.

(4) Szabo, A.; Ostlund, N. S. *Modern quantum chemistry*; McGraw-Hill: New York, 1989; pp 108−230.

(5) Boys, S. F. Electronic wave functions. 1. A general method of calculation for the stationary states of any molecular system. *Proc. R. Soc. London, Ser. A* **1950**, *200*, 542−554.

(6) Dupuis, M.; Rys, J.; King, H. F. Evaluation of molecular integrals over Gaussian basis functions. *J. Chem. Phys.* **1976**, *65*, 111−116.

(7) McMurchie, L. E.; Davidson, E. R. One- and two-electron integrals over Cartesian Gaussian functions. *J. Comput. Phys.* **1978**, *26*, 218−231.

(8) Obara, S.; Saika, A. Efficient recursive computation of molecular integrals over cartesian gaussian functions. *J. Chem. Phys.* **1986**, *84*, 3963−3974.

(9) Head-Gordon, M.; Pople, J. A. A method for two-electron gaussian integral and integral derivative evaluation using recurrence relations. *J. Chem. Phys.* **1988**, *89*, 5777−5786.

(10) Hamilton, T. P.; Schaefer, H. F., III New variations in two-electron integral evaluation in the context of direct SCF procedures. *Chem. Phys.* **1991**, *150*, 163−171.

(11) Lindh, R.; Ryu, U.; Liu, B. The reduced multiplication scheme of the Rys quadrature and new recurrence relations for auxiliary function based two-electron integral evaluation. *J. Chem. Phys.* **1991**, *95*, 5889−5897.

(12) Gill, P. M. W.; Pople, J. A. The PRISM algorithm for two-electron integrals. *Int. J. Quantum Chem.* **1991**, *40*, 753−772.

(13) Gill, P. M. W. Molecular integrals over gaussian basis functions. *Adv. Quantum Chem.* **1994**, *25*, 141−205.

(14) Adams, T. R.; Adamson, R. D.; Gill, P. M. W. A tensor approach to two-electron matrix elements. *J. Chem. Phys.* **1997**, *107*, 124−131.

(15) Ishida, K. Rigorous algorithm for the electron repulsion integral over the generally contracted solid harmonic Gaussian-type orbitals. *J. Chem. Phys.* **2000**, *113*, 7818−7829.

(16) Makowski, M. Simple yet powerful techniques for optimization of horizontal recursion steps in gaussian-type two-electron integral evaluation algorithms. *Int. J. Quantum Chem.* **2007**, *107*, 30−36.

(17) Komornicki, A.; King, H. F. A general formulation for the efficient evaluation of n-electron integrals over products of gaussian charge distributions with gaussian geminal functions. *J. Chem. Phys.* **2011**, *134*, 244115.

(18) Reeves, C. M.; Fletcher, R. Use of gaussian functions in the calculation of wavefunctions for small molecules. III. The orbital basis and its effect on valence. *J. Chem. Phys.* **1965**, *42*, 4073−4081.

(19) O-ohata, K.; Taketa, H.; Huzinaga, S. Gaussian expansions of atomic orbitals. *J. Phys. Soc. Jpn.* **1966**, *21*, 2306−2313.

(20) Newton, M. D.; Ostlund, N. S.; Pople, J. A. Projection of diatomic differential overlap: Least-squares projection of two-center distributions onto one-center functions. *J. Chem. Phys.* **1968**, *49*, 5192−5194.

(21) Hehre, W. J.; Stewart, R. F.; Pople, J. A. Self-consistent molecular-orbital methods. I. Use of gaussian expansions of Slater-type atomic orbitals. *J. Chem. Phys.* **1969**, *51*, 2657−2664.

(22) Stewart, R. F. Small gaussian expansions of atomic orbitals. *J. Chem. Phys.* **1969**, *50*, 2485−2495.

(23) Billingsley, F. P., II; Bloor, J. E. Limited expansion of diatomic overlap (LEDO): A near-accurate approximate ab initio LCAO MO method. I. Theory and preliminary investigations. *J. Chem. Phys.* **1971**, *55*, 5178−5190.

(24) Baerends, E. J.; Ellis, D. E.; Ros, P. Self-consistent molecular Hartree-Fock-Slater calculations. I. The computational procedure. *Chem. Phys.* **1973**, *2*, 41−51.

(25) Whitten, J. Coulombic potential energy integrals and approximations. *J. Chem. Phys.* **1973**, *58*, 4496−4501.

(26) Beebe, N. H. F.; Linderberg, J. Simplifications in the generation and transformation of two-electron integrals in molecular calculations. *Int. J. Quantum Chem.* **1977**, *12*, 683−705.

(27) Dunlap, B. I.; Connolly, J.; Sabin, J. On some approximations in applications of Xα theory. *J. Chem. Phys.* **1979**, *71*, 3396−3402.

(28) Fortunelli, A.; Salvetti, O. A simplified representation of the potential produced by gaussian charge distributions. *J. Comput. Chem.* **1991**, *12*, 36−41.

(29) Feyereisen, M.; Fitzgerald, G.; Komornicki, A. Use of approximate integrals in ab initio theory. An application in MP2 energy calculations. *Chem. Phys. Lett.* **1993**, *208*, 359−363.

(30) Eichkorn, K.; Treutler, O.; Öhm, H.; Häser, M.; Ahlrichs, R. Auxiliary basis sets to approximate Coulomb potentials. *Chem. Phys. Lett.* **1995**, *240*, 283−289.

(31) Hall, G. G.; Martin, D. Approximate electron densities for atoms and molecules. *Isr. J. Chem.* **1980**, *19*, 255−259.

(32) Gill, P. M. W.; Johnson, B. G.; Pople, J. A.; Taylor, S. W. Modeling the potential of a charge distribution. *J. Chem. Phys.* **1992**, *96*, 7178−7179.

(33) Weigend, F.; Kattannek, M.; Ahlrichs, R. Approximated electron repulsion integrals: Cholesky decomposition versus resolution of the identity methods. *J. Chem. Phys.* **2009**, *130*, 164106.

(34) Pedersen, T. B.; Aquilante, F.; Lindh, R. Density fitting with auxiliary basis sets from Cholesky decompositions. *Theor. Chem. Acc.* **2009**, *124*, 1−10.
(35) McKemmish, L. K.; Gill, P. M. W. Gaussian expansions of orbitals. *J. Chem. Theory Comput.* **2012**, *8*, 4891−4898.
(36) Kutzelnigg, W. Expansion of a wave function in a gaussian basis. I. Local versus global approximation. *Int. J. Quantum Chem.* **2013**, *113*, 203−217.
(37) García, V.; Zorrilla, D.; Sánchez-Márquez, J.; Fernández-Núnez, M. Software to obtain accurate gaussian expansions for a wide range of radial functions. *J. Mol. Model* **2017**, *23*, 165.
(38) Olver, F. W. J., Lozier, D. W., Boisvert, R. F., Clark, C. W., Eds. *NIST Handbook of Mathematical Functions*; Cambridge University Press: New York, 2010.
(39) Dunning, T. H. Gaussian basis sets for use in correlated molecular calculations. *J. Chem. Phys.* **1989**, *90*, 1007−1023.
(40) Jensen, F. Polarization consistent basis sets: Principles. *J. Chem. Phys.* **2001**, *115*, 9113−9125.
(41) Pritchard, B. P.; Altarawy, D.; Didier, B.; Gibson, T. D.; Windus, T. L. A new Basis Set Exchange: An open, up-to-date resource for the molecular sciences community. *J. Chem. Inf. Model.* **2019**, *59*, 4814−4820.
(42) Dantzig, G. B.; Thapa, M. N. *Linear Programming 1: Introduction*; Springer: New York, 1997; pp 63−97.
(43) Gautschi, W. Construction of Gauss-Christoffel quadrature formulas. *Math. Comp.* **1968**, *22*, 251−270.
(44) Golub, G. H.; Welsch, J. H. Calculation of Gauss quadrature rules. *Math. Comp.* **1969**, *23*, 221−230.
(45) Szegö, G. *Orthogonal polynomials*; American Mathematical Society: New York, 1959; pp 9, 111−158.
(46) Gradshteyn, I. S.; Ryzhik, I. M. In *Table of Integrals, Series, and Products*, 7th ed.; Jeffrey, A., Zwillinger, D., Eds.; Academic Press: London, 2007; pp 635−793.
(47) Bardo, R. D.; Ruedenberg, K. Even-tempered atomic orbitals. VI. Optimal orbital exponents and optimal contractions of Gaussian primitives for hydrogen, carbon, and oxygen in molecules. *J. Chem. Phys.* **1974**, *60*, 918−931.
(48) Levenberg, K. A method for the solution of certain non-linear problems in least squares. *Quart. Appl. Math.* **1944**, *2*, 164−168.
(49) Marquardt, D. W. An algorithm for least-squares estimation of nonlinear parameters. *J. Soc. Indust. Appl. Math.* **1963**, *11*, 431−441.
(50) Boyd, J. P. *Chebyshev and Fourier spectral methods*, 2nd ed.; Dover: New York, 2000; p 67.

https://doi.org/10.1021/acs.jpca.3c04363
J. Phys. Chem. A 2023, 127, 9346−9356