PHYSICAL REVIEW E **108**, 065308 (2023)

Explicitly correlated Gaussians for high-precision variational calculations of S^e, P^e, and D^e states of quantum systems: An efficient algorithm

Toreniyaz Shomenov[*] and Sergiy Bubin[†]
Department of Physics, Nazarbayev University, Astana 010000, Kazakhstan

(Received 17 July 2023; revised 23 October 2023; accepted 10 November 2023; published 22 December 2023)

In this work we consider an efficient algorithm for variational calculations of quantum few-particle systems in S, P, and D states of the even parity using all-particle explicitly correlated Gaussian (ECG) basis sets. We primarily focus on the description of states where the dominant configuration contains either two particles in p states or a single particle in a d state (all other particles are in s states). The basis functions we consider are products of spherically symmetric ECGs and bipolar harmonics. We introduced a scheme for deriving expressions for matrix elements of the overlap, kinetic and potential energy, as well as their derivatives with respect to the nonlinear parameters of the Gaussians. This allowed us to improve the efficiency of numerical calculations of the matrix elements (which is the most critical part of any code that uses ECGs) by one to two orders of magnitude compared to previous implementations. We provide a complete set of formulas for all basic matrix elements using the formalism of matrix differential calculus and discuss some technical details relevant to their efficient implementation. Lastly, we report a few example calculations of the ${}^2D^e$ states of the Li atom, ${}^4P^e$ and ${}^2D^e$ states of the B atom, and ${}^3P^e$ states of the C atom.

DOI: 10.1103/PhysRevE.108.065308

I. INTRODUCTION

Obtaining nearly exact solutions to the Schrödinger equation for a system of particles with some nontrivial interparticle interaction, e.g., Coulomb forces, is a challenging task. Many, if not most, approaches that target ultrahigh accuracy in the calculations of such systems are based on the wave-function *Ansätze* that explicitly depend on the distances between particles—explicitly correlated functions. The first one dates back to the pioneering work of Hylleraas on the helium atom [1]. Some benchmark variational calculations of two-electron systems that employed the Hylleraas-type basis functions or their modified variants have reported truly remarkable results, where the estimated accuracy of the obtained energy eigenvalues exceeds 20 or even 40 decimal figures [2–5]. However, the use of the Hylleraas-type basis becomes impractical for systems with more than three electrons (or four particles in total) due to the difficulties of evaluating necessary matrix elements in the analytic form. This has led to the adoption of simpler forms of all-particle explicitly correlated basis sets—explicitly correlated Gaussians (ECGs)—which contain a quadratic form of all interparticle distances in the exponent [6,7]. The ECG basis functions demonstrate great flexibility and performance in many applications that involve quantum few-body systems of various nature, ranging from the nuclear structure and excitonic complexes to the electronic structure of small atoms and molecules [8–13].

Calculations of S states ($L = 0$) of few-particle systems require the use of spherically symmetric basis functions, such as simple all-particle ECGs [14]. In order to study states with nonzero angular momenta, it is necessary to multiply the spherically symmetric ECGs by certain angular factors that can be conveniently expressed in terms of Cartesian coordinates [15–17]. In this work we focus on ECGs with prefactors that are suitable for describing states with the following dominant particle configurations:

(a) Two particles are in p states ($l_i = 1$, where l_i are individual orbital angular momenta) while all others are in s states ($l_i = 0$). The resulting multiparticle states can be of S^e, P^e, or D^e symmetry.

(b) A single particle is in a d state while all others are in s states. This yields multiparticle states of D^e symmetry.

The superscript e here denotes the even parity. ECGs of this type have been previously used in calculations of small atoms [18–24]. A somewhat related EGG approach involving Gaussians with single-particle angular prefactors for arbitrary angular momenta of individual particles was also considered in Ref. [25].

However, the applications of previously implemented algorithms were essentially limited to lithium and beryllium atoms only. Reaching sufficiently high accuracy for the ground state energy of carbon atom [13] using existing computer hardware was already difficult and required prohibitively long calculations. The main reason for this is the numerical complexity of the formulas for the matrix elements of the Hamiltonian and overlap. It should also be noted that using ECGs in high-accuracy calculations of quantum few-body systems requires extensive optimization of nonlinear parameters and generation of large basis sets. This is partly due to the behavior of ECGs at short interparticle distances (Gaussian basis functions do not satisfy the Kato cusp condition [26]) and their rapid decay at long distances. Therefore, developing new approaches that

---

[*] toreniyaz.shomenov@nu.edu.kz

[†] sergiy.bubin@nu.edu.kz

can facilitate much faster computation of matrix elements and better optimization of nonlinear parameters is very critical for practical applications.

In this work, we introduce an efficient algorithm for computing the matrix elements of the Hamiltonian, overlap, and their gradients with respect to the nonlinear parameters of ECGs. This is achieved by using a different approach in deriving the formulas for the matrix elements with the products of spherically symmetric Gaussians and bipolar harmonics. It results in a significant reduction of the numerical complexity of computing the matrix elements and amounts to an overall speedup of about one to two orders of magnitude in comparison with previous implementations.

This article is organized in the following way. First, we introduce the notation scheme we adopt. Second, we consider the Hamiltonian for an arbitrary N-particle system in the laboratory coordinate frame and the transformation that excludes the center-of-mass motion. Then we introduce the basis functions employed in this work and discuss their properties. After that we provide the derivation of the matrix elements for the Hamiltonian, overlap, analytic energy gradient, and other useful quantities with these basis functions. Lastly, we present a few numerical results that demonstrate the performance of the approach.

II. NOTATIONS

In this article we derive and present analytic expressions in general matrix form and make extensive use of the formalism of matrix differential calculus [27]. In the matrix formulation, however, we encounter several types of vectors and matrices whose dimensionality varies with the number of particles, number of degrees of freedom per particle, choice of coordinates, and the length of the variational basis. In order to avoid any confusion with vector and/or matrix expressions, in this section we introduce the notation scheme used in the article.

In the very beginning let us specify a few key integer constants that define the dimensionalities and reserve certain symbols to denote them:

(i) ν denotes the number of degrees of freedom per particle. In this article we will only deal with particles moving in three dimensions, $\nu \equiv 3$.

(ii) N denotes the total number of particles in the system. $\nu N = 3N$ gives the total number of degrees of freedom in the system.

(iii) $n = N - 1$ denotes the effective number of (pseudo)particles after separating the motion of the center of mass. $\nu n = 3n$ is the number of degrees of freedom corresponding to the relative (or internal) motion of the particles.

(iv) $\mathcal{N}$ denotes the number of ECG basis functions used to expand the trial variational wave function. This integer defines the dimension of the Hilbert space spanned by the basis functions and the size of the Hamiltonian and overlap matrices.

To distinguish between different types of vectors and matrices we will adopt notations that use different alphabets, cases, and font type. The intention is to indicate what type of matrix, vector, or scalar is represented by a specific character. There may be some exceptions to the rules listed below (e.g., they do not apply to the symbols N, n, and $\mathcal{N}$ we reserved above). However, the exceptions are obvious, and should not cause any confusion.

(i) α, β, γ, etc.: Lowercase greek letters denote scalars.

(ii) $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$, $\boldsymbol{\gamma}$, etc.: Lowercase greek letters written in bold font denote 3-component vectors. The individual components of these vectors, if needed, can be labeled by indices (1, 2, 3 or x, y, z), for example, $\boldsymbol{\alpha}_x$ or $\boldsymbol{\alpha}_1$.

(iii) a, b, c, etc.: Lowercase latin letters are used to denote n-component vectors, where n is the number of pseudoparticles. Here, again, the components can be labeled with an index; for instance, a_i is the ith component.

(iv) $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}$, etc.: Lowercase latin letters written in bold font are used for $3n$-component vectors. Note that $\boldsymbol{a}_i$ denotes not a single component but a 3-component vector.

(v) a, b, c, etc.: Lowercase latin letters in sans-serif font denote N-component vectors, where N is the number of particles.

(vi) $\boldsymbol{\mathsf{a}}$, $\boldsymbol{\mathsf{b}}$, $\boldsymbol{\mathsf{c}}$, etc.: Lowercase latin letters in bold sans-serif font denote $3N$-component vectors.

(vii) $\boldsymbol{\Lambda}$, $\boldsymbol{\Xi}$, $\boldsymbol{\Omega}$, etc.: Uppercase bold greek letters are used for 3×3 matrices.

(viii) A, B, C, etc.: Uppercase latin letters are used for $n \times n$ matrices.

(ix) $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$, etc.: Uppercase latin letters written in bold font are used for $3n \times 3n$ matrices.

(x) $\mathtt{a}$, $\mathtt{b}$, $\mathtt{c}$, etc.: Lowercase latin letters in typewriter font are used to denote $\mathcal{N}$-component vectors in the Hilbert space of basis functions.

(xi) $\mathtt{A}$, $\mathtt{B}$, $\mathtt{C}$, etc.: Uppercase latin letters in typewriter font are used to denote $\mathcal{N} \times \mathcal{N}$ matrices.

(xii) $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$, etc.: Calligraphic font is used to denote operators, e.g., the Hamiltonian, and operators related to particle permutations such as antisymmetrizers, symmetrizers, pair permutations, etc.

(xiii) $\mathbf{I}$, I, $\boldsymbol{I}$, $\mathtt{I}$: Variants of the letter I are used to denote an identity matrix. Specifically, $\mathbf{I}$ (bold capital greek iota) is the 3×3 identity matrix, I (capital italic i) is the $n \times n$ identity matrix, $\boldsymbol{I}$ is the $3n \times 3n$ identity matrix, and $\mathtt{I}$ is the $\mathcal{N} \times \mathcal{N}$ identity matrix.

The symbols defined below are used to denote certain common operations on scalars, matrices, and vectors:

(i) $\operatorname{tr} A$, $\operatorname{tr} \boldsymbol{A}$, etc., stand for the trace of a matrix.

(ii) $|\boldsymbol{\Lambda}|$, $|A|$, $|\boldsymbol{A}|$, etc.: The vertical bars stand for the determinant of a matrix. However, if the object in between the vertical bars is a vector (e.g., $|\boldsymbol{\alpha}|$, $|a|$) or scalar, then the bars stand for the magnitude of the vector or scalar, respectively.

(iii) $\boldsymbol{\Lambda}^{-1}, A^{-1}, \boldsymbol{A}^{-1}$, etc., stand for the inverse of a matrix.

(iv) $\boldsymbol{\alpha}'$, a', A', $\boldsymbol{A}'$, etc.: The prime symbol stands for the transpose of a matrix or vector.

(v) $a \otimes \boldsymbol{\alpha}$, $A \otimes B$, $\boldsymbol{A} \otimes \boldsymbol{B}$, etc.: The symbol $\otimes$ denotes the Kronecker product. When the Kronecker product applies to vectors they are treated as rectangular matrices.

(vi) The common rules of matrix multiplication are assumed. Vectors are treated as rectangular matrices. For instance, if A and B are $n \times n$ matrices then $C = AB$ is also an $n \times n$ matrix. Similarly, the $a'b$ product of two n-component vectors is a scalar, the ab' product of two n-component vectors is an $n \times n$ matrix, etc.

We will also use two operations, vec and vech, that transform matrices into vectors. They both stack the columns of a matrix, one underneath the other. The difference between them is that vech takes only the part of each column below the diagonal, including the diagonal element. Hence, vec transforms an $n \times n$ matrix into an n^2-component vector, while vech transforms the same matrix into an $[n(n+1)/2]$-component vector. For example, if $n = 3$ and A is a 3×3 matrix, then

$$\operatorname{vec} A = \begin{pmatrix} A_{11} \\ A_{12} \\ A_{13} \\ A_{21} \\ A_{22} \\ A_{23} \\ A_{31} \\ A_{32} \\ A_{33} \end{pmatrix}, \qquad \operatorname{vech} A = \begin{pmatrix} A_{11} \\ A_{12} \\ A_{13} \\ A_{22} \\ A_{23} \\ A_{33} \end{pmatrix}. \tag{1}$$

III. HAMILTONIAN

We consider a nonrelativistic Coulombic system of N particles. Let r_i, m_i, and q_i denote the position vector, mass, and charge of the ith particle in the laboratory frame. In atomic units, the Hamiltonian of the system reads

$$\mathcal{H} = -\sum_{i=1}^{N} \frac{1}{2\mathsf{m}_i} \nabla^2_{\mathsf{r}_i} + \sum_{i=1}^{N} \sum_{j>i}^{N} \frac{\mathsf{q}_i \mathsf{q}_j}{\mathsf{r}_{ij}}, \tag{2}$$

where $\mathsf{r}_{ij} = |\mathsf{r}_j - \mathsf{r}_i|$ is the distance between the ith and jth particles, and ∇_{r_i} is the gradient with respect to the position of the ith particle. Unless the mass of at least one particle in Eq. (2) is infinitely heavy, this Hamiltonian does have a discrete spectrum. The discrete (or bound) states are embedded in a continuum that results from the free motion of the system as a whole. Indeed, Hamiltonian (2) is invariant with respect to an arbitrary translation of all particle positions by the same displacement vector. Thus, in order to target these bound states in calculations, one must first separate the motion of the center-of-mass from the laboratory-frame Hamiltonian (2). This can be done in various ways. In this article we adopt perhaps the most natural and straightforward approach in which the heaviest particle of the system (e.g., atomic nucleus) is placed at the origin of the new, internal, reference frame (see Fig. 1). Without any loss of generality we can assume that this reference particle is particle 1. Then the new internal coordinates of other particles are defined relative to the reference particle. The transformation from N laboratory-frame position vectors r_i to the set of coordinates composed of $n = N-1$ internal frame position vectors $\boldsymbol{r}_i$ and the position of the center of mass $\boldsymbol{r}_N$ is given by

$$\begin{aligned} \boldsymbol{r}_i &= \mathsf{r}_{i+1} - \mathsf{r}_1, \qquad i = 1, \dots, n, \\ \boldsymbol{r}_N &= \mathsf{r}_{\text{cm}} = \sum_{i=1}^{N} \frac{\mathsf{m}_i \mathsf{r}_i}{\mathsf{m}_{\text{tot}}}, \end{aligned} \tag{3}$$

where $\mathsf{m}_{\text{tot}} = \sum_{i=1}^{N} \mathsf{m}_i$ is the total mass of the system.

For further consideration it is convenient to introduce the notations for charges and masses with a shifted index: $m_i = \mathsf{m}_{i+1}$, $q_i = \mathsf{q}_{i+1}$, where $i = 0, \dots, n$. In these notations m_0 and q_0 are the mass and charge of the reference particle (particle 1), respectively.

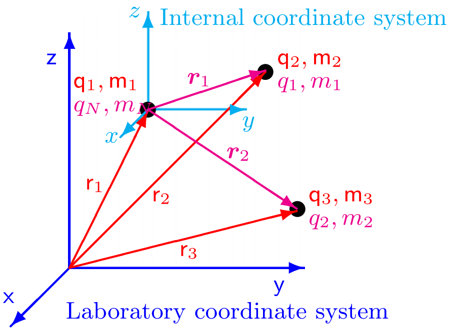

FIG. 1. Relation between the laboratory-frame coordinates and the internal coordinates, where the origin is placed at the first particle

In the new coordinates $\{\boldsymbol{r}, \boldsymbol{r}_N\}$ the laboratory-frame Hamiltonian separates into two independent terms,

$$\mathcal{H} = \mathcal{H}_{\text{int}} + \mathcal{H}_{\text{cm}}, \tag{4}$$

that correspond to the intrinsic motion of particles in the system and free motion of the center of mass:

$$\begin{aligned} \mathcal{H}_{\text{int}} = &-\sum_{i=1}^{n} \frac{1}{2\mu_i} \nabla^2_{\boldsymbol{r}_i} - \sum_{i=1}^{n} \sum_{i \neq j}^{n} \frac{1}{2m_0} \nabla'_{\boldsymbol{r}_i} \nabla_{\boldsymbol{r}_j} \\ &+ \sum_{i=1}^{n} \frac{q_0 q_i}{r_i} + \sum_{i=1}^{n} \sum_{j>i}^{n} \frac{q_i q_j}{r_{ij}}, \end{aligned} \tag{5}$$

$$\mathcal{H}_{\text{cm}} = -\frac{1}{2\mathsf{m}_{\text{tot}}} \nabla^2_{\boldsymbol{r}_N}. \tag{6}$$

In the above expressions $\mu_i = m_0 m_i/(m_0 + m_i)$ are reduced masses and $r_{ij} = |\boldsymbol{r}_j - \boldsymbol{r}_i|$.

The motions of the pseudoparticles are described by their kinetic energy terms, $-\sum_{i=1}^{n} \nabla^2_{\boldsymbol{r}_i}/(2\mu_i)$, and they are coupled through the so-called mass-polarization term, $-\sum_{i=1}^{n} \sum_{i \neq j}^{n} \nabla'_{\boldsymbol{r}_i} \nabla_{\boldsymbol{r}_j}/(2m_0)$, as well as through the Coulombic interactions. We use the term "pseudoparticles" because, despite the fact that the charges of the particles described by the internal Hamiltonian are actual particle charges, their masses are reduced masses rather than actual particle masses. Lastly, we will rewrite the internal Hamiltonian in the matrix form as it is more convenient for further analysis and derivations. This is done by combining the n position vectors and gradients into $3n$-component vectors as follows:

$$\boldsymbol{r} = \begin{pmatrix} \boldsymbol{r}_1 \\ \boldsymbol{r}_2 \\ \dots \\ \boldsymbol{r}_n \end{pmatrix}, \qquad \nabla_{\boldsymbol{r}} = \begin{pmatrix} \nabla_{\boldsymbol{r}_1} \\ \nabla_{\boldsymbol{r}_2} \\ \dots \\ \nabla_{\boldsymbol{r}_n} \end{pmatrix}. \tag{7}$$

Then the internal Hamiltonian can be cast in the matrix form as

$$\mathcal{H}_{\text{int}} = -\nabla'_{\boldsymbol{r}} \boldsymbol{M} \nabla_{\boldsymbol{r}} + \sum_{i=1}^{n} \frac{q_0 q_i}{r_i} + \sum_{i=1}^{n} \sum_{j>i}^{n} \frac{q_i q_j}{r_{ij}}, \tag{8}$$

where $\boldsymbol{M} = M \otimes \mathbf{I}$ is the Kronecker product of $n \times n$ matrix M and 3×3 identity matrix $\mathbf{I}$. The elements of a symmetric matrix M are given by

$$M_{ij} = \begin{cases} \frac{1}{2\mu_i}, & i = j \\ \frac{1}{2m_0}, & i \neq j. \end{cases} \tag{9}$$

IV. BASIS FUNCTIONS

The spatial part of ECG basis functions suitable for expanding the wave function of a system that contains only particles with zero angular momenta should be spherically symmetric (i.e., rotationally invariant). They consist of a radial exponential part only and can be written in the following compact matrix form [14]:

$$\phi_k = \exp[-\boldsymbol{r}'(A_k \otimes \mathbf{I})\boldsymbol{r}]. \tag{10}$$

The argument of the exponent contains the Kronecker product of an $n \times n$ real matrix A_k, and 3×3 identity matrix $\mathbf{I}$. The matrix A_k stores nonlinear variational parameters. The product $\boldsymbol{r}'(A_k \otimes \mathbf{I})\boldsymbol{r}$ defines a certain quadratic form in terms of pseudoparticle coordinates. In fact, there is an infinite number of possible choices for matrix A_k that yield the same quadratic form. In order to remove this arbitrariness and simplify calculations, we will always assume that matrix A_k is symmetric, in which case there is one-to-one correspondence between A_k and the quadratic form. Subscript k is used to indicate that, in general, matrix A_k is unique for each basis function. Because bound state wave functions must be square integrable, it is necessary to require that matrix A_k is positive definite. The positive definiteness is automatically satisfied if matrix A_k is represented in a Cholesky-factored form, $A_k = L_k L_k'$, where L_k is a lower triangular matrix. This allows unconstrained variations of the elements of L_k in the range $[-\infty, \infty]$ when optimizing the nonlinear parameters of the basis function using the variational method.

In order to describe few-particle systems with nonzero total orbital angular momenta, simple ECGs (10) must be modified by including a suitable angular factor in front of the spherically symmetric Gaussian. This angular factor can be obtained by means of the standard rules of addition of angular momentum states of individual particles. For multiparticle states with the total orbital angular momentum quantum number L and its projection on the z axis, M, that correspond to dominant single-particle configurations where only two particles (say, particles i and j) have nonzero orbital angular momenta, the angular factors θ_{LM} are proportional to the bipolar harmonics:

$$\theta_{LM}(\boldsymbol{r}_i, \boldsymbol{r}_j) = r_i^{l_i} r_j^{l_j} \sum_{\substack{m_i, m_j \\ M = m_i + m_j}} C^{LM}_{l_i m_i l_j m_j} Y_{l_i m_i}(\boldsymbol{r}_i) Y_{l_j m_j}(\boldsymbol{r}_j). \tag{11}$$

Here l_i and l_j are the orbital angular momentum quantum numbers of particles i and j, respectively, Y_{lm} are the usual spherical harmonics, and $C^{LM}_{l_i m_i l_j m_j}$ are the Clebsch-Gordan coefficients. Using the Cartesian form of the spherical harmonics and the table of the Clebsch-Gordan coefficients, one can obtain the angular factor θ_{LM} for an arbitrary value of the total orbital angular momentum number and its projection on the z axis. For example, in the case of $L = 1$, $M = 0$, even total parity, and $l_i = l_j = 1$, the angular factor comes out to be

$$\theta_{10} = \frac{3i}{4\pi\sqrt{2}}(x_i y_j - x_j y_i), \tag{12}$$

while the corresponding ECG basis function is

$$\phi_k = (x_{i_k} y_{j_k} - x_{j_k} y_{i_k}) \exp[-\boldsymbol{r}'(A_k \otimes \mathbf{I})\boldsymbol{r}]. \tag{13}$$

In the above expression, integer indices i_k and j_k, which can range from 1 to n, define a specific set of two particles that have nonzero orbital angular momenta. Since the exact wave function is a mixture of states with different single-particle configurations, these indices can take different values for different basis functions (hence i and j have a subscript k that relates them to a specific basis function). In general, indices i_k and j_k can be considered integer variational parameters in calculations and their optimal choice for each individual basis function should be dictated by the energy minimization.

For $L = 2$, $M = 0$, and even parity, we can consider two distinct cases. The first one is when $l_i = l_j = 1$. The corresponding basis function has the following form:

$$\phi_k = (x_{i_k} x_{j_k} + y_{i_k} y_{j_k} - 2 z_{i_k} z_{j_k}) \exp[-\boldsymbol{r}'(A_k \otimes \mathbf{I})\boldsymbol{r}]. \tag{14}$$

The second case is when $l_i = 2$ and all other individual orbital angular momenta are zeros. The corresponding basis function is

$$\phi_k = \left(x_{i_k}^2 + y_{i_k}^2 - 2 z_{i_k}^2\right) \exp[-\boldsymbol{r}'(A_k \otimes \mathbf{I})\boldsymbol{r}]. \tag{15}$$

Lastly, for $L = 0$, $M = 0$, even parity, and $l_i = l_j = 1$ the corresponding basis function also acquires a factor, but it is rotationally invariant:

$$\phi_k = (x_{i_k} x_{j_k} + y_{i_k} y_{j_k} + z_{i_k} z_{j_k}) \exp[-\boldsymbol{r}'(A_k \otimes \mathbf{I})\boldsymbol{r}]. \tag{16}$$

It should be noted that, in principle, states with $L = 0$ that result from coupling individual orbital angular momenta $l_i = 1$ and $l_j = 1$ can be expanded in terms of simpler basis functions (10). However, the presence of a factor that has a node at the origin may provide certain practical advantages for basis functions (16). In general, for states where two particles have nonzero individual orbital angular momenta, basis functions (16) should yield faster convergence when the basis size is increased in comparison with the case of using functions (10).

As mentioned previously, indices i_k and j_k in Eqs. (13)–(16) can range from 1 to n. The only restriction occurs for basis functions (13), where the value of i_k must not be equal to j_k as in that case the function automatically becomes a null function. When dealing with sets of identical particles, the range of possible values for i_k and j_k may be reduced to avoid redundancy that occurs due to nondistinguishability of particles. However, such a reduction is not at all required.

In this article we report expressions for the matrix elements of the Hamiltonian, overlap, and the analytic energy gradient with basis functions (13)–(16). These basis functions are suitable for high-accuracy variational calculations of the ground and excited states of many different few-particle systems. While our primary focus in this work is on few-electron atoms and ions, they can also be used for other types of quantum few-body systems, such as few-exciton complexes in materials, systems composed of or containing exotic particles (e.g., positrons, muons), nuclei, small molecules, etc.

In order to derive matrix elements with basis functions (13)–(16) in matrix form it is convenient to represent them as a sum of two or three general terms that read

$$\begin{aligned}\phi_k &= [\boldsymbol{v}_k'\boldsymbol{r}][\boldsymbol{w}_k'\boldsymbol{r}]\exp[-\boldsymbol{r}'\boldsymbol{A}_k\boldsymbol{r}]\\ &= [(v_k\otimes\boldsymbol{\epsilon}^v)'\boldsymbol{r}][(w_k\otimes\boldsymbol{\epsilon}^w)'\boldsymbol{r}]\exp[-\boldsymbol{r}'(A_k\otimes\mathbf{I})\boldsymbol{r}]. \end{aligned}\tag{17}$$

In this expression, $3n$-component sparse vectors $\boldsymbol{v}_k$ and $\boldsymbol{w}_k$ define a specific product of Cartesian coordinates (e.g., $x_{i_k}y_{j_k}$). The sparsity of $\boldsymbol{v}_k$ and $\boldsymbol{w}_k$ here means that each of them has only a single nonzero component. These $3n$-component vectors can be represented as Kronecker products, $v_k\otimes\boldsymbol{\epsilon}^v$ and $w_k\otimes\boldsymbol{\epsilon}^w$, that contain the corresponding n-component sparse vectors v_k and w_k. The only nonzero element of v_k is the i_kth element and it equals unity. Similarly, the only nonzero element of w_k is the j_kth element that also equals unity. Each of the two three-dimensional (3D) unit vectors, $\boldsymbol{\epsilon}^v$ and $\boldsymbol{\epsilon}^w$, also contains just a single nonzero element (equal to one) that controls which Cartesian coordinate of a pseudoparticle (x, y, or z) should appear in the prefactor of the Gaussian. For example, if the system of interest has three particles ($n = N - 1 = 2$) and the needed prefactor is $[\boldsymbol{v}_k'\boldsymbol{r}][\boldsymbol{w}_k'\boldsymbol{r}] = x_2y_1$, then

$$i_k = 2,\qquad \boldsymbol{\epsilon}^v = \boldsymbol{\epsilon}^x,\qquad j_k = 1,\qquad \boldsymbol{\epsilon}^w = \boldsymbol{\epsilon}^y,$$

and the corresponding sparse vectors are

$$\boldsymbol{v}_k = v_k\otimes\boldsymbol{\epsilon}^v = \begin{pmatrix}0\\1\end{pmatrix}\otimes\begin{pmatrix}1\\0\\0\end{pmatrix} = \begin{pmatrix}0\\0\\0\\1\\0\\0\end{pmatrix},$$

$$\boldsymbol{w}_k = w_k\otimes\boldsymbol{\epsilon}^w = \begin{pmatrix}1\\0\end{pmatrix}\otimes\begin{pmatrix}0\\1\\0\end{pmatrix} = \begin{pmatrix}0\\1\\0\\0\\0\\0\end{pmatrix}.$$

For the derivation of matrix elements it is convenient to introduce the following generating function:

$$\varphi_k = \exp[-\boldsymbol{r}'(A_k\otimes\mathbf{I})\boldsymbol{r} + \alpha_k\boldsymbol{v}_k'\boldsymbol{r} + \beta_k\boldsymbol{w}_k'\boldsymbol{r}].\tag{18}$$

The mixed second derivative of φ_k with respect to parameters α_k and β_k generates ϕ_k:

$$\phi_k = \frac{\partial}{\partial\alpha_k}\frac{\partial}{\partial\beta_k}\exp[-\boldsymbol{r}'(A_k\otimes\mathbf{I})\boldsymbol{r} + \alpha_k\boldsymbol{v}_k'\boldsymbol{r} + \beta_k\boldsymbol{w}_k'\boldsymbol{r}]\bigg|_{\alpha_k,\beta_k=0}.\tag{19}$$

V. PERMUTATIONAL SYMMETRY

The Hamiltonian of a system containing k identical particles is invariant with respect to the permutations of these particles, i.e.,

$$[\mathcal{H}, \mathcal{X}_i] = 0,\tag{20}$$

where $\mathcal{X}_i$ is an arbitrary element (permutation) of the finite symmetric group S_k that has order $k!$. This implies the simultaneous construction of both Hamiltonian and symmetric group eigenfunctions. It is accomplished by projecting the spatial basis functions onto the relevant irreducible representations of the symmetric group. Because the nonrelativistic Hamiltonian (8) does not depend on the spin of particles, we can eliminate the spin variables from further consideration. In this spin-free formalism, the Young operators for the irreducible representations of the symmetric group are constructed from their corresponding Young tableaux [28,29]. For a system of k identical particles, the Young tableaux are generated from a series of k-connected empty boxes that are called Young frames. For example, in the case of a system containing four identical particles, they may look as follows:

$$\begin{array}{|c|c|c|c|}\hline \ & \ & \ & \ \\\hline\end{array},\quad \begin{array}{|c|c|c|}\hline \ & \ & \ \\\hline \multicolumn{1}{|c|}{\ } \\\cline{1-1}\end{array},\quad \begin{array}{|c|c|}\hline \ & \ \\\hline \ & \ \\\hline\end{array},\quad \begin{array}{|c|c|}\hline \ & \ \\\hline \multicolumn{1}{|c|}{\ } \\\cline{1-1} \multicolumn{1}{|c|}{\ } \\\cline{1-1}\end{array},\quad \begin{array}{|c|}\hline \ \\\hline \ \\\hline \ \\\hline \ \\\hline\end{array}.\tag{21}$$

The shape of the suitable Young frame depends on whether the identical particles involved are fermions or bosons, their spin, and the total spin of the system. For the case of fermions with spin $1/2$, the Young frame for the spatial wave function must contain no more than two columns, while for fermions with spin $3/2$, the maximum number of columns is four. The number of columns for bosons is not limited, but the number of rows is. These are derived from the general rule of constructing the Young frames [30]. The present work primarily focuses on electrons, which are spin-$1/2$ particles. To create a Young frame for such systems with the total spin quantum number S, we first need to calculate the symmetry quantum number, $p = k/2 - S$. Then a partition is written as $\mu = [2^p\,1^{k-2p}]$. The Young frame then consists of two boxes in the first p rows and one box in the remaining $k - 2p$ rows.

A Young tableau is generated by filling a Young frame with numbers labeling the identical particles (e.g., from 1 to k). A specific order in which those numbers are placed in the boxes is not important; e.g., they can increase when going from left to right and from top to bottom. For example, let us consider the carbon atom in a triplet spin state. In this case $k = 6$ (the number of electrons) and for the total spin quantum number $S = 1$ (triplet multiplicity) we have $p = 2$. Based on that, we can write the partition, $\mu = [2^2\,1^2]$. If we assume that the labels of the particles range from 2 to 7 (the first particle can be a nucleus) then the corresponding Young tableau is filled as follows:

$$\begin{array}{|c|c|}\hline 2 & 3 \\\hline 4 & 5 \\\hline \multicolumn{1}{|c|}{6} \\\cline{1-1} \multicolumn{1}{|c|}{7} \\\cline{1-1}\end{array}.\tag{22}$$

Having the Young tableau, the Young symmetry projector can now be constructed in the following way:

$$\mathcal{Y} = \mathcal{S}\mathcal{A},\tag{23}$$

where $\mathcal{S}$ and $\mathcal{A}$ are symmetrization and antisymmetrization operators. $\mathcal{S}$ is defined as a product of symmetrizers over particles in each row of the Young tableaux, while $\mathcal{A}$ is a product of antisymmetrizers over particles in each column. The symmetrizers and antisymmetrizers are easily expressed using the transpositions (pair permutations of particles), which we will denote $\mathcal{P}_{ij}$. For instance, the Young tableau in Eq. (22) yields the following Young operator:

$$\mathcal{Y} = \mathcal{S}_{2,3}\mathcal{S}_{4,5}\mathcal{A}_{2,4,6,7}\mathcal{A}_{3,5}, \tag{24}$$

where subscripts list the particles over which the symmetrization or antisymmetrization must be carried out. Note that because each symmetrizer applies to a different set of particles, the order in which symmetrizers are written is arbitrary. Similarly, the mutual placement of antisymmetrizers is also arbitrary. Expressed in terms of transpositions the symmetrizers and antisymmetrizers in expression (24) are

$$\begin{aligned}
\mathcal{S}_{2,3} &= (\mathcal{I}+\mathcal{P}_{23}),\\
\mathcal{S}_{4,5} &= (\mathcal{I}+\mathcal{P}_{45}),\\
\mathcal{A}_{2,4,6,7} &= (\mathcal{I}-\mathcal{P}_{24})(\mathcal{I}-\mathcal{P}_{26}-\mathcal{P}_{46})(\mathcal{I}-\mathcal{P}_{27}-\mathcal{P}_{47}-\mathcal{P}_{67}),\\
\mathcal{A}_{3,5} &= (\mathcal{I}-\mathcal{P}_{35}),
\end{aligned}$$

where $\mathcal{I}$ is the identity operator and for the sake of simplicity we dropped all common normalization factors.

Spatial matrix elements of spin-independent operators evaluated with symmetry-projected spin-free basis functions (projected by applying operator $\mathcal{Y}$) are equivalent to the matrix elements with the corresponding full (spatial + spin) basis states after the summation over spin variables.

It is worth noting that in many cases the operators whose expectation values need to be evaluated in variational calculations are not only Hermitian but also commute with all transpositions in symmetry projector (23). Consequently, computations of matrix elements for such operators can always be limited to the scenario where the projection needs to be applied to the ket basis functions only. This is because

$$\langle\mathcal{Y}\phi_k|\mathcal{O}|\mathcal{Y}\phi_l\rangle = \langle\phi_k|\mathcal{O}|\mathcal{Y}^\dagger\mathcal{Y}\phi_l\rangle, \tag{25}$$

where $\mathcal{O}$ is some operator that commutes with all transpositions involving identical particles and $\mathcal{Y}^\dagger = \mathcal{A}^\dagger\mathcal{S}^\dagger = \mathcal{A}\mathcal{S}$ is the Hermitian conjugate of $\mathcal{Y}$. The operator $\mathcal{Y}^\dagger\mathcal{Y} \propto \mathcal{A}\mathcal{S}\mathcal{A}$ can be simplified by means of multiplying the transpositions and reducing identical terms. Such automatic simplification can be implemented in the computer program that does variational calculations. With this, the number of independent terms due to the permutational symmetry that need to be summed when evaluating a single matrix element of $\mathcal{O}$ reduces to $k!$.

When the system of interest contains multiple sets of identical particles (e.g., electrons and protons, or electrons and positrons), the elimination of spin from explicit consideration can be achieved by sequentially applying the Young symmetry projectors for each set of particles.

When the summation over permutations of identical particles is carried out in the course of evaluating matrix elements it is necessary to know how these permutations transform basis functions in the ket, $|\phi_l\rangle$. Apparently, the functional form of the basis functions is not altered. What changes is the nonlinear variational parameters ($\boldsymbol{A}_l$, $\boldsymbol{v}_l$, and $\boldsymbol{w}_l$). When an arbitrary permutational operator $\mathcal{P}$ acts on the laboratory-frame coordinates $\mathbf{r}$ of the real particles, these coordinates undergo a linear transformation. The transformation of internal coordinates $\boldsymbol{r}$ (e.g., pseudoparticle coordinates) as a result of permutations of the real particles is also linear. This allows to represent the transformation of the internal coordinates by means of the permutation matrix $\boldsymbol{P} = P \otimes \mathbf{I}$. If we act on basis function (17) with $\mathcal{P}$, we obtain

$$\begin{aligned}
\mathcal{P}\phi_l &= (\boldsymbol{v}_l)'(\boldsymbol{P}\boldsymbol{r})(\boldsymbol{w}_l)'(\boldsymbol{P}\boldsymbol{r})\exp[-(\boldsymbol{P}\boldsymbol{r})'\boldsymbol{A}_l(\boldsymbol{P}\boldsymbol{r})]\\
&= (\boldsymbol{P}'\boldsymbol{v}_l)'\boldsymbol{r}(\boldsymbol{P}'\boldsymbol{w}_l)'\boldsymbol{r}\exp[-\boldsymbol{r}'(\boldsymbol{P}'\boldsymbol{A}_l\boldsymbol{P})\boldsymbol{r}].
\end{aligned} \tag{26}$$

This expression shows how the nonlinear parameters of the Gaussians, i.e., matrices A_l and the corresponding sparse vectors v_l and w_l, are transformed. In the derivation of matrix elements in the subsequent sections, we will make an explicit assumption that the ket basis function may be transformed by some permutations of particles. This implies that the matrix A_l and the vectors v_l and w_l presented in all expressions should actually be replaced with $P'A_lP$, $P'v_l$, and $P'w_l$, respectively. However, for the sake of brevity, we will avoid including P's explicitly in the expressions.

VI. MATRIX ELEMENTS

The derivation of matrix elements will be carried out using the formalism of matrix differential calculus [27]. To begin, let us first list some useful relations for differentials involving matrices. For an arbitrary square matrix X, the differentials of its determinant, trace, and inverse are given by

$$\begin{aligned}
\mathrm{d}\,|X| &= |X|\,\mathrm{tr}[X^{-1}\,\mathrm{d}\,X],\\
\mathrm{d}\,\mathrm{tr}\,[X] &= \mathrm{tr}\,[\mathrm{dX}],\\
\mathrm{d}(X^{-1}) &= -X^{-1}(\mathrm{dX})X^{-1}.
\end{aligned} \tag{27}$$

We will also use the value of the following n-dimensional Gaussian integral,

$$\int_{-\infty}^{\infty}\exp[-x'Ax + y'x]\,\mathrm{d}\,x = \frac{\pi^{n/2}}{|A|^{1/2}}\exp\left[\frac{1}{4}y'A^{-1}y\right], \tag{28}$$

where x and y are n-component vectors (the first one contains n variables $x_1, \ldots, x_n$, while the latter is composed of constants $y_1, \ldots, y_n$), and A is an $n \times n$ positive definite symmetric matrix.

A. Overlap integral

The overlap integral between generating functions φ_k given by expression (18) is easily obtained using formula (28),

$$\begin{aligned}
\langle\varphi_k|\varphi_l\rangle &= \int\exp[-\boldsymbol{r}'\boldsymbol{A}_{kl}\boldsymbol{r} + (\alpha_k\boldsymbol{v}_k+\beta_k\boldsymbol{w}_k+\alpha_l\boldsymbol{v}_l+\beta_l\boldsymbol{w}_l)'\boldsymbol{r}]d\boldsymbol{r}\\
&= \frac{\pi^{3n/2}}{|\boldsymbol{A}_{kl}|^{1/2}}\exp\bigg[\frac{1}{4}(\alpha_k\boldsymbol{v}_k+\beta_k\boldsymbol{w}_k+\alpha_l\boldsymbol{v}_l+\beta_l\boldsymbol{w}_l)'\boldsymbol{A}_{kl}^{-1}\\
&\quad\times(\alpha_k\boldsymbol{v}_k+\beta_k\boldsymbol{w}_k+\alpha_l\boldsymbol{v}_l+\beta_l\boldsymbol{w}_l)\bigg],
\end{aligned} \tag{29}$$

where $\boldsymbol{A}_{kl} = \boldsymbol{A}_k + \boldsymbol{A}_l$. Note that the determinant of a $3n \times 3n$ matrix $\boldsymbol{A}_{kl}$ can be expressed through the determinant of an $n \times n$ matrix A_{kl}:

$$|\boldsymbol{A}_{kl}| = |A_{kl} \otimes \mathbf{I}| = |A_{kl}|^3. \tag{30}$$

The expression for the overlap integral with basis functions (17) will acquire somewhat different forms depending on the specific combination of the Cartesian coordinates in the Gaussian prefactors defined in vectors v_k and w_k. Three distinct cases can be identified and we list them separately below.

1. Case $(x_{i_k}x_{j_k}|x_{i_l}x_{j_l})$

Let us start with the scenario where $\boldsymbol{\epsilon}^v = \boldsymbol{\epsilon}^w$. For definiteness we can consider $\boldsymbol{\epsilon}^v = \boldsymbol{\epsilon}^w = \boldsymbol{\epsilon}^x$. Due to the rotational symmetry of the simple Gaussians (10) the integral will be the same in the case of $\boldsymbol{\epsilon}^v = \boldsymbol{\epsilon}^w = \boldsymbol{\epsilon}^y$ or $\boldsymbol{\epsilon}^v = \boldsymbol{\epsilon}^w = \boldsymbol{\epsilon}^z$. Using the following relation,

$$(a' \otimes \boldsymbol{\epsilon}^{x'})(A_{kl} \otimes \mathbf{I})(a \otimes \boldsymbol{\epsilon}^x) = a'A_{kl}a \otimes \boldsymbol{\epsilon}^{x'}\mathbf{I}\boldsymbol{\epsilon}^x = a'A_{kl}a, \tag{31}$$

where a is some vector, we can reduce the size of matrix $\boldsymbol{A}_{kl}$ and vectors in Eq. (29) from $3n$ to n:

$$\langle\varphi_k|\varphi_l\rangle = \frac{\pi^{3n/2}}{|A_{kl}|^{3/2}} \exp\left[\frac{1}{4}(\alpha_k v_k + \beta_k w_k + \alpha_l v_l + \beta_l w_l)' \times A_{kl}^{-1}(\alpha_k v_k + \beta_k w_k + \alpha_l v_l + \beta_l w_l)\right]. \tag{32}$$

Then the overlap integral between basis functions (17) is obtained by means of taking four derivatives, which yields

$$\begin{aligned}\mathsf{S}_{kl} &= \frac{\partial}{\partial\beta_l}\frac{\partial}{\partial\beta_k}\frac{\partial}{\partial\alpha_l}\frac{\partial}{\partial\alpha_k}\langle\varphi_k|\varphi_l\rangle\bigg|_{\alpha_k,\alpha_l,\beta_k,\beta_l=0} \\ &= \frac{\pi^{3n/2}}{|A_{kl}|^{3/2}}\frac{1}{4}\Big(\big[v_k'A_{kl}^{-1}v_l\big]\big[w_k'A_{kl}^{-1}w_l\big] \\ &\quad + \big[v_k'A_{kl}^{-1}w_k\big]\big[v_l'A_{kl}^{-1}w_l\big] + \big[v_k'A_{kl}^{-1}w_l\big]\big[w_k'A_{kl}^{-1}v_l\big]\Big).\end{aligned} \tag{33}$$

For brevity, let us introduce the following quantities that define vector-matrix-vector products:

$$\begin{aligned}\gamma_1 &= v_k'A_{kl}^{-1}v_l, \\ \gamma_2 &= w_k'A_{kl}^{-1}w_l, \\ \gamma_3 &= v_k'A_{kl}^{-1}w_k, \\ \gamma_4 &= v_l'A_{kl}^{-1}w_l, \\ \gamma_5 &= v_k'A_{kl}^{-1}w_l, \\ \gamma_6 &= w_k'A_{kl}^{-1}v_l.\end{aligned} \tag{34}$$

With that we can rewrite the expression for the overlap integral as

$$\mathsf{S}_{kl} = \frac{1}{4}\frac{\pi^{3n/2}}{|A_{kl}|^{3/2}}\gamma, \tag{35}$$

where

$$\gamma = \gamma_1\gamma_2 + \gamma_3\gamma_4 + \gamma_5\gamma_6. \tag{36}$$

2. Case $(x_{i_k}y_{j_k}|y_{i_l}x_{j_l})$

Because $\boldsymbol{\epsilon}^{x'}\mathbf{I}\boldsymbol{\epsilon}^y = \boldsymbol{\epsilon}^{y'}\mathbf{I}\boldsymbol{\epsilon}^x = 0$, expression (32) becomes

$$\begin{aligned}\langle\varphi_k|\varphi_l\rangle = \frac{\pi^{3n/2}}{|A_{kl}|^{3/2}} &\exp\left[\frac{1}{4}(\alpha_k v_k + \alpha_l v_l)'A_{kl}^{-1}(\alpha_k v_k + \alpha_l v_l)\right] \\ &\times \exp\left[\frac{1}{4}(\beta_k w_k + \beta_l w_l)'A_{kl}^{-1}(\beta_k w_k + \beta_l w_l)\right].\end{aligned} \tag{37}$$

Performing the same four differentiations as in the previous case gives

$$\mathsf{S}_{kl} = \frac{1}{4}\frac{\pi^{3n/2}}{|A_{kl}|^{3/2}}\gamma_1\gamma_2. \tag{38}$$

3. Case $(x_{i_k}x_{j_k}|y_{i_l}y_{j_l})$

The last distinct case is when both coordinates in the prefactor of the bra are x coordinates, while in the ket both coordinates are y coordinates. The result of taking four derivatives is then

$$\mathsf{S}_{kl} = \frac{1}{4}\frac{\pi^{3n/2}}{|A_{kl}|^{3/2}}\gamma_3\gamma_4. \tag{39}$$

At this point it should be clear that the expressions and derivation of matrix elements in all three distinct cases is very similar. For this reason, in the following sections for other matrix elements, such as those of the kinetic energy, potential energy, and the analytic gradient, we will only present the final expressions for the three cases rather than repeat almost identical derivations three times.

B. Kinetic energy

1. Case $(x_{i_k}x_{j_k}|x_{i_l}x_{j_l})$

For matrix elements of the kinetic energy and some other quantities we will need the value of an auxiliary integral containing a quadratic form with an arbitrary $3n \times 3n$ symmetric matrix $\boldsymbol{X}$:

$$\begin{aligned}\langle\phi_k|\boldsymbol{r}'\boldsymbol{X}\boldsymbol{r}|\phi_l\rangle &= -\frac{\partial}{\partial\beta}\langle\phi_k|\exp[-\beta\boldsymbol{r}'\boldsymbol{X}\boldsymbol{r}]|\phi_l\rangle\bigg|_{\beta=0} = -\frac{\partial}{\partial\beta}\frac{\pi^{3n/2}}{|A_{kl}+\beta X|^{3/2}}\Big([v_k'(A_{kl}+\beta X)^{-1}v_l][w_k'(A_{kl}+\beta X)^{-1}w_l] \\ &\quad + [v_k'(A_{kl}+\beta X)^{-1}w_k][v_l'(A_{kl}+\beta X)^{-1}w_l] + [v_k'(A_{kl}+\beta X)^{-1}w_l][w_k'(A_{kl}+\beta X)^{-1}v_l]\Big)\bigg|_{\beta=0},\end{aligned} \tag{40}$$

where $\boldsymbol{X} = X \otimes \mathbf{I}_3$. After taking the derivative using relations (27) we obtain

$$\begin{aligned}\langle\phi_k|\boldsymbol{r}'\boldsymbol{X}\boldsymbol{r}|\phi_l\rangle &= \frac{3}{2}\,\mathrm{tr}\left[A_{kl}^{-1}X\right]\mathsf{S}_{kl} + \frac{\pi^{3n/2}}{4|A_{kl}|^{3/2}}\Big(\big[v_k'A_{kl}^{-1}XA_{kl}^{-1}v_l\big]\big[w_k'A_{kl}^{-1}w_l\big] + \big[v_k'A_{kl}^{-1}v_l\big]\big[w_k'A_{kl}^{-1}XA_{kl}^{-1}w_l\big] + \big[v_k'A_{kl}^{-1}XA_{kl}^{-1}w_k\big] \\ &\quad \times \big[v_l'A_{kl}^{-1}w_l\big] + \big[v_k'A_{kl}^{-1}w_k\big]\big[v_l'A_{kl}^{-1}XA_{kl}^{-1}w_l\big] + \big[v_k'A_{kl}^{-1}XA_{kl}^{-1}w_l\big]\big[w_k'A_{kl}^{-1}v_l\big] + \big[v_k'A_{kl}^{-1}w_l\big]\big[w_k'A_{kl}^{-1}XA_{kl}^{-1}v_l\big]\Big).\end{aligned} \tag{41}$$

The evaluation of the kinetic energy matrix elements is easiest when we take into account the anti-Hermiticity of the gradient operators so that $\boldsymbol{\nabla}_{\boldsymbol{r}}$ is applied to both the bra and ket generating functions (18). Given the action of $\boldsymbol{\nabla}_{\boldsymbol{r}}$ on φ_k, namely,

$$\boldsymbol{\nabla}_{\boldsymbol{r}}\varphi_k = [-2\boldsymbol{A}_k\boldsymbol{r} + \alpha_k\boldsymbol{v}_k + \beta_k\boldsymbol{w}_k]\varphi_k, \tag{42}$$

then we obtain

$$\begin{aligned}\langle\boldsymbol{\nabla}_{\boldsymbol{r}}\varphi_k|\boldsymbol{M}|\boldsymbol{\nabla}_{\boldsymbol{r}}\varphi_l\rangle &= \langle(-2\boldsymbol{A}_k\boldsymbol{r} + \alpha_k\boldsymbol{v}_k + \beta_k\boldsymbol{w}_k)\varphi_k|\boldsymbol{M}|(-2\boldsymbol{A}_l\boldsymbol{r} + \alpha_l\boldsymbol{v}_l + \beta_l\boldsymbol{w}_l)\varphi_l\rangle = 4\langle\varphi_k|\boldsymbol{r}'\boldsymbol{A}_k'\boldsymbol{M}\boldsymbol{A}_l\boldsymbol{r}|\varphi_l\rangle - 2\alpha_l\langle\varphi_k|\boldsymbol{r}'\boldsymbol{A}_k'\boldsymbol{M}\boldsymbol{v}_l|\varphi_l\rangle\\ &\quad - 2\beta_l\langle\varphi_k|\boldsymbol{r}'\boldsymbol{A}_k'\boldsymbol{M}\boldsymbol{w}_l|\varphi_l\rangle - 2\alpha_k\langle\varphi_k|\boldsymbol{v}_k'\boldsymbol{M}\boldsymbol{A}_l\boldsymbol{r}|\varphi_l\rangle + \alpha_k\alpha_l\boldsymbol{v}_k'\boldsymbol{M}\boldsymbol{v}_l\langle\varphi_k|\varphi_l\rangle + \alpha_k\beta_l\boldsymbol{v}_k'\boldsymbol{M}\boldsymbol{w}_l\langle\varphi_k|\varphi_l\rangle\\ &\quad - 2\beta_k\langle\varphi_k|\boldsymbol{w}_k'\boldsymbol{M}\boldsymbol{A}_l\boldsymbol{r}|\varphi_l\rangle + \beta_k\alpha_l\boldsymbol{w}_k'\boldsymbol{M}\boldsymbol{v}_l\langle\varphi_k|\varphi_l\rangle + \beta_k\beta_l\boldsymbol{w}_k'\boldsymbol{M}\boldsymbol{w}_l\langle\varphi_k|\varphi_l\rangle.\end{aligned} \tag{43}$$

As usual, matrix elements of the kinetic energy operator with basis functions (17) are obtained by taking four derivatives of the above expression. In this process we end up with three types of integrals. The first one, which originates from the first term on the right-hand side of Eq. (43), contains a quadratic form. We can make the matrix defining this quadratic form symmetric and then use expression (41) substituting $\boldsymbol{X} = (\boldsymbol{A}_k'\boldsymbol{M}\boldsymbol{A}_l + \boldsymbol{A}_l'\boldsymbol{M}\boldsymbol{A}_k)/2$. The second type of integral originates from the second, third, fourth, and seventh terms on the right-hand side of Eq. (43). These integrals are similar to the overlap integral (35) after certain substitutions of matrices. Lastly, the third type of integral happens to be similar to the overlap integral of Gaussians with single Cartesian prefactors derived in Ref. [15]. After some lengthy but straightforward algebraic manipulation we obtain the final expression for the kinetic energy matrix elements:

$$\mathtt{T}_{kl} = \mathtt{S}_{kl}\left(6\tau + \frac{4\zeta}{\gamma}\right), \tag{44}$$

where

$$\begin{aligned}\tau &= \mathrm{tr}\left[A_{kl}^{-1}A_lMA_k'\right],\\ \zeta &= (\gamma_1\eta_1 + \gamma_2\eta_2 - \gamma_3\eta_3 - \gamma_4\eta_4 + \gamma_5\eta_5 + \gamma_6\eta_6),\\ \eta_1 &= \left(w_k'A_{kl}^{-1}A_l'MA_kA_{kl}^{-1}w_l\right),\\ \eta_2 &= \left(v_k'A_{kl}^{-1}A_l'MA_kA_{kl}^{-1}v_l\right),\\ \eta_3 &= \left(v_l'A_{kl}^{-1}A_k'MA_kA_{kl}^{-1}w_l\right),\\ \eta_4 &= \left(v_k'A_{kl}^{-1}A_l'MA_lA_{kl}^{-1}w_k\right),\\ \eta_5 &= \left(w_k'A_{kl}^{-1}A_l'MA_kA_{kl}^{-1}v_l\right),\\ \eta_6 &= \left(v_k'A_{kl}^{-1}A_l'MA_kA_{kl}^{-1}w_l\right).\end{aligned} \tag{45}$$

2. *Case* $(\boldsymbol{x}_{i_k}\boldsymbol{y}_{j_k}|\boldsymbol{y}_{i_l}\boldsymbol{x}_{j_l})$

Following the same recipe as in the previous case, we obtain

$$\mathtt{T}_{kl} = \mathtt{S}_{kl}\left(6\tau + 4\frac{\eta_2}{\gamma_1} + 4\frac{\eta_1}{\gamma_2}\right). \tag{46}$$

3. *Case* $(\boldsymbol{x}_{i_k}\boldsymbol{x}_{j_k}|\boldsymbol{y}_{i_l}\boldsymbol{y}_{j_l})$

Similarly, in this case we get

$$\mathtt{T}_{kl} = \mathtt{S}_{kl}\left(6\tau - 4\frac{\eta_4}{\gamma_3} - 4\frac{\eta_3}{\gamma_4}\right). \tag{47}$$

C. Potential energy

1. *Case* $(\boldsymbol{x}_{i_k}\boldsymbol{x}_{j_k}|\boldsymbol{x}_{i_l}\boldsymbol{x}_{j_l})$

First, let us note that the squares of the interparticle distances can be represented as quadratic forms:

$$r_i^2 = \boldsymbol{r}'\boldsymbol{J}_{ii}\boldsymbol{r}, \quad r_{ij}^2 = \boldsymbol{r}'\boldsymbol{J}_{ij}\boldsymbol{r}. \tag{48}$$

Here matrix $\boldsymbol{J}_{ij} = J_{ij} \otimes \mathbf{I}$ is defined as

$$J_{ij} = \begin{cases} E_{ii} & \text{if } i = j\\ E_{ii} + E_{jj} - E_{ij} - E_{ji} & \text{if } i \neq j, \end{cases} \tag{49}$$

where all the elements of matrix E_{ij} are zeros except the element in the ith row and jth column, which is equal to one. Now, using the overlap matrix elements (35), we can transform the Coulomb integrals $\langle\phi_k|1/r_i|\phi_l\rangle$ and $\langle\phi_k|1/r_{ij}|\phi_l\rangle$ as follows:

$$\begin{aligned}\mathtt{R}_{kl}^{ij} = \langle\phi_k|\frac{1}{r_{ij}}|\phi_l\rangle &= \frac{2}{\sqrt{\pi}}\int_0^\infty\langle\phi_k|\exp[-\beta^2\boldsymbol{r}'\boldsymbol{J}_{ij}\boldsymbol{r}]|\phi_l\rangle\mathrm{d}\beta\\ &= \frac{\pi^{\frac{3n-1}{2}}}{2}\int_0^\infty\frac{1}{|A_{kl} + \beta^2J_{ij}|^{3/2}}\big([v_k'(A_{kl}+\beta^2J_{ij})^{-1}v_l]\\ &\quad\times [w_k'(A_{kl}+\beta^2J_{ij})^{-1}w_l] + [v_k'(A_{kl}+\beta^2J_{ij})^{-1}w_k]\\ &\quad\times [v_l'(A_{kl}+\beta^2J_{ij})^{-1}w_l] + [v_k'(A_{kl}+\beta^2J_{ij})^{-1}w_l]\\ &\quad\times [w_k'(A_{kl}+\beta^2J_{ij})^{-1}v_l]\big)\mathrm{d}\beta.\end{aligned} \tag{50}$$

To solve this integral, we need to simplify the terms enclosed in square brackets, e.g., the terms such as $v_k'(A_{kl} + \beta^2J_{ij})^{-1}v_l$. This can be done by using the Sherman-Morrison formula, e.g.,

$$\begin{aligned}v_k'(A_{kl}+\beta^2J_{ij})^{-1}v_l &= \mathrm{tr}[(A_{kl}+\beta^2J_{ij})^{-1}v_lv_k']\\ &= \mathrm{tr}[A_{kl}^{-1}v_lv_k'] - \frac{\beta^2\,\mathrm{tr}\left[A_{kl}^{-1}J_{ij}A_{kl}^{-1}v_lv_k'\right]}{1 + \beta^2\,\mathrm{tr}\left[A_{kl}^{-1}J_{ij}\right]}.\end{aligned} \tag{51}$$

After making such substitutions in Eq. (50), taking the integral, and doing some algebraic manipulations, we obtain the final expression for the Coulomb integrals,

$$\mathtt{R}_{kl}^{ij} = \frac{\pi^{(3n-1)/2}}{2|A_{kl}|^{3/2}\omega^{1/2}}\left(\gamma - \frac{\kappa}{3\omega} + \frac{\rho}{5\omega^2}\right), \tag{52}$$

where we introduced the following quantities:

$$\begin{aligned}\rho_1 &= \left(w_k'A_{kl}^{-1}J_{ij}A_{kl}^{-1}w_l\right),\\ \rho_2 &= \left(v_k'A_{kl}^{-1}J_{ij}A_{kl}^{-1}v_l\right),\end{aligned}$$

$$\begin{aligned} \rho_3 &= \left(v_l' A_{kl}^{-1} J_{ij} A_{kl}^{-1} w_l\right), \\ \rho_4 &= \left(v_k' A_{kl}^{-1} J_{ij} A_{kl}^{-1} w_k\right), \\ \rho_5 &= \left(w_k' A_{kl}^{-1} J_{ij} A_{kl}^{-1} v_l\right), \\ \rho_6 &= \left(v_k' A_{kl}^{-1} J_{ij} A_{kl}^{-1} w_l\right), \\ \omega &= \mathrm{tr}\left[A_{kl}^{-1} J_{ij}\right], \\ \rho &= [\rho_1\rho_2 + \rho_3\rho_4 + \rho_5\rho_6], \\ \kappa &= [\gamma_1\rho_1 + \gamma_2\rho_2 + \gamma_3\rho_3 + \gamma_4\rho_4 + \gamma_5\rho_5 + \gamma_6\rho_6]. \end{aligned} \tag{53}$$

The expression for the complete potential energy matrix elements is given by the sum of $\mathtt{R}_{kl}^{i}$ and $\mathtt{R}_{kl}^{ij}$:

$$\mathtt{V}_{kl} = \sum_{i=1}^{n} q_0 q_i \mathtt{R}_{kl}^{i} + \sum_{i<j}^{n} q_i q_j \mathtt{R}_{kl}^{ij}. \tag{54}$$

2. Case $(x_{i_k} y_{j_k} | y_{i_l} x_{j_l})$

In this case, the expression for $\mathtt{R}_{kl}^{ij}$ is similar to that in Eq. (52) except that we need to substitute

$$\gamma \to \gamma_1\gamma_2, \quad \kappa \to \gamma_1\rho_1 + \gamma_2\rho_2, \quad \rho \to \rho_1\rho_2. \tag{55}$$

3. Case $(x_{i_k} x_{j_k} | y_{i_l} y_{j_l})$

Similarly, in this case in Eq. (52) we need to substitute

$$\gamma \to \gamma_3\gamma_4, \quad \kappa \to \gamma_3\rho_3 + \gamma_4\rho_4, \quad \rho \to \rho_3\rho_4. \tag{56}$$

D. Matrix elements with the Dirac δ function and related quantities

The expectation values of the 3D Dirac δ function with an argument in the form of various linear combinations of interparticle coordinates $\boldsymbol{r}_i$ occurs in many practical applications. For example, the particle contact density, $\langle\delta(\boldsymbol{r}_{ij})\rangle$, which is defined as the probability density of finding particles i and j in the same point in space, is used in the calculations of relativistic corrections and particle-antiparticle annihilation rates. Let us consider the case of an arbitrary linear combination of $\boldsymbol{r}_i'$s in the Dirac δ function:

$$\delta(a_1\boldsymbol{r}_1 + a_2\boldsymbol{r}_2 + \cdots + a_n\boldsymbol{r}_n - \boldsymbol{\xi}) = \delta((a\otimes\mathbf{I})'\boldsymbol{r} - \boldsymbol{\xi}), \tag{57}$$

where a and $\boldsymbol{\xi}$ are some arbitrary n-component and 3-component real vectors, respectively. For the derivation of matrix elements it is convenient to use the following representation of the δ function:

$$\delta((a\otimes\mathbf{I})'\boldsymbol{r} - \boldsymbol{\xi}) = \lim_{\beta\to\infty}\left(\frac{\beta^{3/2}}{\pi^{3/2}}\exp\left[-\beta\left((a\otimes\mathbf{I})'\boldsymbol{r} - \boldsymbol{\xi}\right)^2\right]\right). \tag{58}$$

First let us compute the matrix elements of the Dirac δ function with generating functions (18):

$$\langle\varphi_k|\delta((a\otimes\mathbf{I})'\boldsymbol{r} - \boldsymbol{\xi})|\varphi_l\rangle = \lim_{\beta\to\infty}\left(\frac{\beta^{3/2}}{\pi^{3/2}}\exp[-\beta\boldsymbol{\xi}^2]\langle\varphi_k|\exp[-\beta\boldsymbol{r}'(aa'\otimes\mathbf{I})\boldsymbol{r} + 2\beta((a\otimes\mathbf{I})\boldsymbol{\xi})'\boldsymbol{r}]|\varphi_l\rangle\right). \tag{59}$$

Here the matrix element inside the limit is similar to the overlap integral in Eq. (29):

$$\begin{aligned} \langle\varphi_k|\delta((a\otimes\mathbf{I})'\boldsymbol{r} - \boldsymbol{\xi})|\varphi_l\rangle = \lim_{\beta\to\infty}\left(\frac{\beta}{\pi}\right)^{3/2}\exp[-\beta\boldsymbol{\xi}^2]\frac{\pi^{3n/2}}{|A_{kl}+\beta aa'|^{3/2}}\exp\Bigg[\frac{1}{4}(\alpha_k\boldsymbol{v}_k + \beta_k\boldsymbol{w}_k + \alpha_l\boldsymbol{v}_l + \beta_l\boldsymbol{w}_l) \\ + 2\beta(a\otimes\mathbf{I})\boldsymbol{\xi})'\left((A_{kl}+\beta aa')^{-1}\otimes\mathbf{I}\right)(\alpha_k\boldsymbol{v}_k + \beta_k\boldsymbol{w}_k + \alpha_l\boldsymbol{v}_l + \beta_l\boldsymbol{w}_l + 2\beta(a\otimes\boldsymbol{I})\boldsymbol{\xi})\Bigg]. \end{aligned} \tag{60}$$

Taking the limit of this expression when $\beta\to\infty$ requires expanding the argument of the exponent. Again, we can use the Sherman-Morrison formula and replace

$$(A_{kl}+\beta aa')^{-1} = A_{kl}^{-1} - \frac{\beta A_{kl}^{-1}aa'A_{kl}^{-1}}{1+\beta a'A_{kl}^{-1}a}.$$

1. Case $(x_{i_k} x_{j_k} | x_{i_l} x_{j_l})$

After taking the limit, we obtain the following expression for the matrix elements of the δ function with generator functions (18):

$$\begin{aligned} \langle\varphi_k|\delta((a\otimes\mathbf{I})'\boldsymbol{r} - \boldsymbol{\xi})|\varphi_l\rangle = \frac{\pi^{3(n-1)/2}}{|A_{kl}|^{3/2}}\frac{1}{\left(a'A_{kl}^{-1}a\right)^{3/2}}\exp\Bigg[\frac{1}{a'A_{kl}^{-1}a}\Bigg(\frac{1}{4}(\alpha_k v_k + \beta_k w_k + \alpha_l v_l + \beta_l w_l)'\left(\left[a'A_{kl}^{-1}a\right]A_{kl}^{-1} - A_{kl}^{-1}aa'A_{kl}^{-1}\right) \\ \times(\alpha_k v_k + \beta_k w_k + \alpha_l v_l + \beta_l w_l) + (\alpha_k v_k + \beta_k w_k + \alpha_l v_l + \beta_l w_l)'A_{kl}^{-1}a(\boldsymbol{\epsilon}^{x\prime}\boldsymbol{\xi}) - \boldsymbol{\xi}^2\Bigg)\Bigg]. \end{aligned} \tag{61}$$

To find the expectation value with basis functions (17), we still need to take derivatives. After some lengthy but straightforward calculations we find

$$\langle\phi_k|\delta((a\otimes\mathbf{I})'\boldsymbol{r} - \boldsymbol{\xi})|\phi_l\rangle = \frac{\pi^{3(n-1)/2}}{4\,|A_{kl}|^{3/2}\,\omega^{3/2}}\exp\left[-\frac{\boldsymbol{\xi}^2}{\omega}\right]\left(\left[\gamma - \frac{\kappa}{\omega} + \frac{\rho}{\omega^2}\right] + \frac{2}{\omega^2}\left[\kappa - \frac{2\rho}{\omega}\right](\boldsymbol{\epsilon}^{x\prime}\boldsymbol{\xi})^2 + \frac{4\rho}{3\omega^4}(\boldsymbol{\epsilon}^{x\prime}\boldsymbol{\xi})^4\right). \tag{62}$$

Here for various quantities we adopted the notations from Eqs. (34), (36), (45), and (53). However, one must replace $J_{ij} \to aa'$ everywhere in those equations.

Expression (62) provides a general formula for the matrix elements of the pair correlation functions and densities of particles in the center-of-mass frame and in other reference frames. These quantities, in turn, can be used to calculate matrix elements of such operators as powers of the interparticle distances, i.e., $|\boldsymbol{r}_i|^\lambda$ and $|\boldsymbol{r}_{ij}|^\lambda$, where λ is any real number that satisfies $\lambda > -3$. To do this, we consider partial cases of $a = j^i$ and $a = j^j - j^i$, respectively. Here j^i is an n-component vector whose ith component is one, while all others are zeros. One may note that $j^i(j^i)' = J_{ii}$ and $(j^j - j^i)(j^j - j^i)' = J_{ij}$.

The final expression for the matrix elements of $|\boldsymbol{r}_{ij}|^\lambda$ is

$$\langle\phi_k|r_{ij}^\lambda|\phi_l\rangle = \frac{\pi^{(3n-1)/2}}{2\,|A_{kl}|^{3/2}}\,\Gamma\left(\frac{\lambda+3}{2}\right)\omega^{\lambda/2}\left[\left(\gamma - \frac{\kappa}{\omega} + \frac{\rho}{\omega^2}\right) + \frac{\lambda+3}{3\omega}\left(\kappa - \frac{2\rho}{\omega}\right) + \frac{(\lambda+5)(\lambda+3)\rho}{15\omega^2}\right], \tag{63}$$

where $\Gamma(\cdots)$ stands for the Euler gamma function. One may easily check that the above expression reduces to formula (52) when $\lambda = -1$. Similarly, when $\lambda = 2$, it reproduces matrix elements of $\boldsymbol{r}'\boldsymbol{X}\boldsymbol{r}$ in Eq. (41) if we set $\boldsymbol{X} = \boldsymbol{J}_{ij}$. Lastly, the case $\lambda = 0$ yields the overlap integral in Eq. (35).

2. *Case* $(\boldsymbol{x}_{i_k}\boldsymbol{y}_{j_k}|\boldsymbol{y}_{i_l}\boldsymbol{x}_{j_l})$

Here the expressions for the matrix elements of the Dirac δ function and r_{ij}^λ are

$$\begin{aligned}\langle\phi_k|\delta\big((a\otimes\mathbf{I})'\boldsymbol{r} - \boldsymbol{\xi}\big)|\phi_l\rangle = {}& \frac{\pi^{3(n-1)/2}}{4\,|A_{kl}|^{3/2}}\frac{\gamma_1\gamma_2}{(a'A_{kl}^{-1}a)^{3/2}}\exp\left[-\frac{\boldsymbol{\xi}^2}{a'A_{kl}^{-1}a}\right]\left[1 + \frac{1}{a'A_{kl}^{-1}a}\frac{v_k'A_{kl}^{-1}aa'A_{kl}^{-1}v_l}{v_k'A_{kl}^{-1}v_l}\left(2\frac{(\boldsymbol{\epsilon}^{x\prime}\boldsymbol{\xi})^2}{a'A_{kl}^{-1}a} - 1\right)\right]\\ &\times\left[1 + \frac{1}{a'A_{kl}^{-1}a}\frac{w_k'A_{kl}^{-1}aa'A_{kl}^{-1}w_l}{w_k'A_{kl}^{-1}w_l}\left(2\frac{(\boldsymbol{\epsilon}^{y\prime}\boldsymbol{\xi})^2}{a'A_{kl}^{-1}a} - 1\right)\right],\end{aligned} \tag{64}$$

$$\begin{aligned}\langle\phi_k|r_{ij}^\lambda|\phi_l\rangle = {}& \frac{\pi^{(3n-1)/2}}{2\,|A_{kl}|^{3/2}}\,\Gamma\left(\frac{\lambda+3}{2}\right)\omega^{\lambda/2}\gamma_1\gamma_2\left\{\left(1 - \frac{\rho_2}{\omega\gamma_1}\right)\left(1 - \frac{\rho_1}{\omega\gamma_2}\right) + \left(\frac{\lambda+3}{2}\right)\frac{2}{3}\left[\left(1 - \frac{\rho_2}{\omega\gamma_1}\right)\frac{\rho_1}{\omega\gamma_2} + \frac{\rho_2}{\omega\gamma_1}\left(1 - \frac{\rho_1}{\omega\gamma_2}\right)\right]\right.\\ &\left.+\frac{(\lambda+5)(\lambda+3)}{15\omega^2}\frac{\rho_1\rho_2}{\gamma_1\gamma_2}\right\}.\end{aligned} \tag{65}$$

3. *Case* $(\boldsymbol{x}_{i_k}\boldsymbol{x}_{j_k}|\boldsymbol{y}_{i_l}\boldsymbol{y}_{j_l})$

In this case, the expression for the matrix elements of the Dirac δ function is similar to the previous case:

$$\begin{aligned}\langle\phi_k|\delta((a\otimes\mathbf{I})'\boldsymbol{r} - \boldsymbol{\xi})|\phi_l\rangle = {}& \frac{\pi^{3(n-1)/2}}{4\,|A_{kl}|^{3/2}}\frac{\gamma_3\gamma_4}{(a'A_{kl}^{-1}a)^{3/2}}\exp\left[-\frac{\boldsymbol{\xi}^2}{a'A_{kl}^{-1}a}\right]\left[1 + \frac{1}{a'A_{kl}^{-1}a}\frac{v_k'A_{kl}^{-1}aa'A_{kl}^{-1}w_k}{v_k'A_{kl}^{-1}w_k}\left(2\frac{(\boldsymbol{\epsilon}^{x\prime}\boldsymbol{\xi})^2}{a'A_{kl}^{-1}a} - 1\right)\right]\\ &\times\left[1 + \frac{1}{a'A_{kl}^{-1}a}\frac{v_l'A_{kl}^{-1}aa'A_{kl}^{-1}w_l}{v_l'A_{kl}^{-1}w_l}\left(2\frac{(\boldsymbol{\epsilon}^{y\prime}\boldsymbol{\xi})^2}{a'A_{kl}^{-1}a} - 1\right)\right].\end{aligned} \tag{66}$$

The matrix elements of r_{ij}^λ are

$$\begin{aligned}\langle\phi_k|r_{ij}^\lambda|\phi_l\rangle = {}& \frac{\pi^{(3n-1)/2}}{2\,|A_{kl}|^{3/2}}\,\Gamma\left(\frac{\lambda+3}{2}\right)\omega^{\lambda/2}\gamma_3\gamma_4\left\{\left(1 - \frac{\rho_4}{\omega\gamma_3}\right)\left(1 - \frac{\rho_3}{\omega\gamma_4}\right) + \frac{2}{3}\left(\frac{\lambda+3}{2}\right)\left[\left(1 - \frac{\rho_4}{\omega\gamma_3}\right)\frac{\rho_3}{\omega\gamma_4} + \frac{\rho_4}{\omega\gamma_3}\left(1 - \frac{\rho_3}{\omega\gamma_4}\right)\right]\right.\\ &\left.+\frac{(\lambda+5)(\lambda+3)}{15}\frac{\rho_3\rho_4}{\omega^2\gamma_3\gamma_4}\right\}.\end{aligned} \tag{67}$$

VII. ENERGY GRADIENT

The formulas for the matrix elements of the Hamiltonian and overlap derived in the preceding sections should, in principle, be sufficient for performing the energy calculations and generate wave functions. However, in practical calculations aiming to solve the Schrödinger equation with high accuracy, many thousands of ECGs need to be generated. What is even more important, the nonlinear parameters of these thousands of ECGs must be optimized (at least to some extent) for each system and state of interest. This process may be extremely time consuming even on powerful parallel computer systems. The use of the finite difference energy gradient in the process of the energy minimization is very costly and limits the number of basis functions that can be used in the calculations. It has been demonstrated in many applications that employing the analytic energy gradient instead, which can be constructed and computed at a cost that exceeds the cost of a single-point energy evaluation only by a relatively small factor (of the order of 10) that does not grow with the basis size, can greatly expedite the optimization [13–15]. In this section we will explain how the analytic energy gradient can be constructed and evaluate all necessary matrix elements with basis functions (17) that are needed for it.

In the Rayleigh-Ritz variational method the wave function of the system, $\psi(\boldsymbol{r})$, is expanded in terms of some suitable basis functions ϕ_k,

$$\psi(\boldsymbol{r}) = \sum_{k=1}^{\mathcal{N}} \mathtt{c}_k |\phi_k(\{\alpha\}_k)\rangle, \tag{68}$$

where each basis function $\phi_k(\{\alpha\}_k)$ may contain some unique adjustable parameters $\{\alpha_k\}$. Minimizing the energy with respect to linear coefficients $\mathtt{c}_k$ results in a generalized eigenvalue problem

$$\mathtt{Hc} = \varepsilon \mathtt{Sc}, \tag{69}$$

where $\mathtt{H}$ and $\mathtt{S}$ are $\mathcal{N} \times \mathcal{N}$ Hamiltonian and overlap matrices, respectively, and $\mathtt{c}$ is an $\mathcal{N}$-component vector of the linear coefficients. The elements of these matrices are

$$\mathtt{H}_{kl} = \langle \phi_k | \mathcal{H} | \phi_l \rangle, \quad \mathtt{S}_{kl} = \langle \phi_k | \phi_l \rangle. \tag{70}$$

The solutions to the generalized eigenvalue problem, ε (there are $\mathcal{N}$ of them), yield variational upper bounds to the exact energies of ground and excited states.

When it comes to minimizing the energy with respect to parameters α_k, we end up with a much more difficult nonlinear optimization problem.

Let us take the differential of Eq. (69):

$$\mathrm{d}(\mathtt{H} - \varepsilon \mathtt{S})\mathtt{c} = (\mathrm{d}\mathtt{H})\mathtt{c} - (\mathrm{d}\varepsilon)\mathtt{Sc} - \varepsilon(\mathrm{d}\mathtt{S})\mathtt{c} + (\mathtt{H} - \varepsilon\mathtt{S})\mathrm{d}\mathtt{c}.$$

Multiplying this equation by $\mathtt{c}'$ from the left and assuming the wave function is normalized, i.e., $\mathtt{c}'\mathtt{Sc} = 1$, gives

$$\mathrm{d}\varepsilon = \mathtt{c}'(\mathrm{d}\mathtt{H} - \varepsilon \mathrm{d}\mathtt{S})\mathtt{c}. \tag{71}$$

Because only the tth row and tth column of matrices $\mathtt{H}$ and $\mathtt{S}$ depend on parameters $\{\alpha_t\}$ that are unique to basis function $\phi_t(\{\alpha_t\})$, we can write the following expressions for the derivatives:

$$\begin{aligned} \frac{\partial \mathtt{H}_{kl}}{\partial \alpha_t} &= \frac{\partial \mathtt{H}_{kl}}{\partial \alpha_t}(\delta_{kt} + \delta_{lt} - \delta_{kt}\delta_{lt}), \\ \frac{\partial \mathtt{S}_{kl}}{\partial \alpha_t} &= \frac{\partial \mathtt{S}_{kl}}{\partial \alpha_t}(\delta_{kt} + \delta_{lt} - \delta_{kt}\delta_{lt}). \end{aligned}$$

In the above expressions α_t represents any parameter from the set $\{\alpha_t\}$, while indices k and l range from 1 to $\mathcal{N}$. Using these expressions and relation (71), the derivative of the energy eigenvalue ε with respect to parameter α_t can be written as

$$\frac{\partial \varepsilon}{\partial \alpha_t} = 2\mathtt{c}_t' \sum_{l=1}^{\mathcal{N}} \mathtt{c}_l \left(\frac{\partial \mathtt{H}_{tl}}{\partial \alpha_t} - \varepsilon \frac{\partial \mathtt{S}_{tl}}{\partial \alpha_t} \right) - \mathtt{c}_t \mathtt{c}_t' \left(\frac{\partial \mathtt{H}_{tt}}{\partial \alpha_t} - \varepsilon \frac{\partial \mathtt{S}_{tt}}{\partial \alpha_t} \right). \tag{72}$$

To obtain the entire energy gradient (i.e., all of its components), it is necessary to calculate the derivatives with respect to each α_t. This is suboptimal if we compute a derivative for each individual parameter $(L_k)_{11}, (L_k)_{21}, \ldots, (L_k)_{nn}$ separately. Fortunately, we can evaluate all the derivatives with respect to all components of the vech L_k vector in a single step. For that we need the derivatives (more precisely $[n(n+1)/2]$-component gradient vectors) listed below:

$$\frac{\partial \mathtt{H}_{kl}}{\partial(\mathrm{vech}\, L_k)}, \quad \frac{\partial \mathtt{H}_{kl}}{\partial(\mathrm{vech}\, L_l)}, \tag{73}$$

$$\frac{\partial \mathtt{S}_{kl}}{\partial(\mathrm{vech}\, L_k)}, \quad \frac{\partial \mathtt{S}_{kl}}{\partial(\mathrm{vech}\, L_l)}. \tag{74}$$

In the following sections we will present compact vector and matrix expressions for these derivatives for the case of basis functions (17).

A. Overlap gradient

First, let us give the differentials of matrices A_k and A_l that appear in the bra and ket basis functions. They are

$$\begin{aligned} \mathrm{d}A_k &= (\mathrm{d}L_k)L_k' + L_k \mathrm{d}L_k', \\ \mathrm{d}A_l &= (\mathrm{d}L_l)L_l' + L_l(\mathrm{d}L_l'). \end{aligned} \tag{75}$$

Similarly, the differential of matrix A_{kl} is given by

$$\mathrm{d}A_{kl} = (\mathrm{d}L_k)L_k' + L_k \mathrm{d}L_k' + (\mathrm{d}L_l)L_l' + L_l(\mathrm{d}L_l'). \tag{76}$$

1. Case $(x_{i_k}x_{j_k}|x_{i_l}x_{j_l})$

The expression for the differential of the overlap integral is

$$\begin{aligned} \mathrm{d}\mathtt{S}_{kl} = -\mathtt{S}_{kl} \bigg[& \frac{3}{2} \operatorname{tr}\left[A_{kl}^{-1} \mathrm{d}A_{kl}\right] + \frac{1}{\gamma} \Big(\gamma_2 \operatorname{tr}\left[A_{kl}^{-1} v_l v_k' A_{kl}^{-1} \mathrm{d}A_{kl}\right] \\ & + \gamma_1 \operatorname{tr}\left[A_{kl}^{-1} w_l w_k' A_{kl}^{-1} \mathrm{d}A_{kl}\right] + \gamma_4 \operatorname{tr}\left[A_{kl}^{-1} w_k v_k' A_{kl}^{-1} \mathrm{d}A_{kl}\right] \\ & + \gamma_3 \operatorname{tr}\left[A_{kl}^{-1} w_l v_l' A_{kl}^{-1} \mathrm{d}A_{kl}\right] + \gamma_6 \operatorname{tr}\left[A_{kl}^{-1} w_l v_k' A_{kl}^{-1} \mathrm{d}A_{kl}\right] \\ & + \gamma_5 \operatorname{tr}\left[A_{kl}^{-1} v_l w_k' A_{kl}^{-1} \mathrm{d}A_{kl}\right] \Big) \bigg]. \end{aligned} \tag{77}$$

For convenience, let us define several matrices

$$\begin{aligned} K_{kl}^{(1)} &= A_{kl}^{-1} w_l w_k' A_{kl}^{-1}, \\ K_{kl}^{(2)} &= A_{kl}^{-1} v_l v_k' A_{kl}^{-1}, \\ K_{kl}^{(3)} &= A_{kl}^{-1} w_l v_l' A_{kl}^{-1}, \\ K_{kl}^{(4)} &= A_{kl}^{-1} w_k v_k' A_{kl}^{-1}, \\ K_{kl}^{(5)} &= A_{kl}^{-1} v_l w_k' A_{kl}^{-1}, \\ K_{kl}^{(6)} &= A_{kl}^{-1} w_l v_k' A_{kl}^{-1}, \\ K_{kl} &= \left(\gamma_1 K_{kl}^{(1)} + \gamma_2 K_{kl}^{(2)} + \gamma_3 K_{kl}^{(3)} + \gamma_4 K_{kl}^{(4)} \right. \\ &\quad \left. + \gamma_5 K_{kl}^{(5)} + \gamma_6 K_{kl}^{(6)}\right), \\ F_{kl} &= \tfrac{3}{2} A_{kl}^{-1} + \frac{K_{kl}}{\gamma}. \end{aligned} \tag{78}$$

Using these definitions, the differential of the overlap integral can be written as

$$\mathrm{d}\mathtt{S}_{kl} = -\mathtt{S}_{kl} \operatorname{tr}\left[L_k'(F_{kl} + F_{kl}')\mathrm{d}L_k + L_l'(F_{kl} + F_{kl}')\mathrm{d}L_l\right]. \tag{79}$$

When evaluating derivatives with respect to the elements of matrices, it is convenient to use vec and vech operations that transform matrices to long vectors, as defined in Eq. (1). Note that if X and Y are arbitrary complex square matrices, and L is the lower triangular matrix, then the following relations hold:

$$\begin{aligned}\mathrm{tr}[X'Y] &= (\mathrm{vec}\,X)'\,\mathrm{vec}\,Y,\\ (\mathrm{vec}\,X)'\,\mathrm{vec}\,L &= (\mathrm{vech}\,X)'\,\mathrm{vech}\,L.\end{aligned} \tag{80}$$

Using these relations, expression (77) becomes

$$\begin{aligned}\mathrm{d}\mathtt{S}_{kl} = -\mathtt{S}_{kl}\big[&\,\mathrm{vech}((F_{kl}+F'_{kl})L_k)'\mathrm{d}\,\mathrm{vech}\,L_k\\ &+\mathrm{vech}\,\big((F_{kl}+F'_{kl})L_l\big)'\mathrm{d}\,\mathrm{vech}\,L_l\big].\end{aligned} \tag{81}$$

Consequently, the sought derivatives for the overlap integral are

$$\frac{\partial \mathtt{S}_{kl}}{\partial(\mathrm{vech}\,L_k)} = -\mathtt{S}_{kl}\,\mathrm{vech}((F_{kl}+F'_{kl})L_k), \tag{82}$$

$$\frac{\partial \mathtt{S}_{kl}}{\partial(\mathrm{vech}\,L_l)} = -\mathtt{S}_{kl}\,\mathrm{vech}((F_{kl}+F'_{kl})L_l). \tag{83}$$

2. Case $(x_{i_k}y_{j_k}|y_{i_l}x_{j_l})$

In this case γ and K_{kl} reduce to

$$\gamma = \gamma_1\gamma_2, \quad K_{kl} = \left(\gamma_1 K_{kl}^{(1)} + \gamma_2 K_{kl}^{(2)}\right), \tag{84}$$

while the expressions for the derivatives of $\mathtt{S}_{kl}$ remain the same.

3. Case $(x_{i_k}x_{j_k}|y_{i_l}y_{j_l})$

In this case we have

$$\gamma = \gamma_3\gamma_4, \quad K_{kl} = \left(\gamma_3 K_{kl}^{(3)} + \gamma_4 K_{kl}^{(4)}\right). \tag{85}$$

The expressions for the derivatives of $\mathtt{S}_{kl}$ also remain unchanged.

B. Kinetic energy gradient

1. Case $(x_{i_k}x_{j_k}|x_{i_l}x_{j_l})$

The expression for the kinetic energy matrix elements in this case is

$$\mathtt{T}_{kl} = \mathtt{S}_{kl}\left(6\tau + \frac{4}{\gamma}\zeta\right). \tag{86}$$

Let us list the differentials for γ_i's, τ, and η_i's that are defined in Eqs. (34) and (45):

$$\begin{aligned}\mathrm{d}\gamma_1 &= -\,\mathrm{tr}\left[K_{kl}^{(2)}\mathrm{d}A_{kl}\right],\\ \mathrm{d}\gamma_2 &= -\,\mathrm{tr}\left[K_{kl}^{(1)}\mathrm{d}A_{kl}\right],\\ \mathrm{d}\gamma_3 &= -\,\mathrm{tr}\left[K_{kl}^{(4)}\mathrm{d}A_{kl}\right],\\ \mathrm{d}\gamma_4 &= -\,\mathrm{tr}\left[K_{kl}^{(3)}\mathrm{d}A_{kl}\right],\\ \mathrm{d}\gamma_5 &= -\,\mathrm{tr}\left[K_{kl}^{(6)}\mathrm{d}A_{kl}\right],\\ \mathrm{d}\gamma_6 &= -\,\mathrm{tr}\left[K_{kl}^{(5)}\mathrm{d}A_{kl}\right],\end{aligned} \tag{87}$$

$$\begin{aligned}\mathrm{d}\tau &= \mathrm{tr}\left[A_{kl}^{-1}A_lMA_l'A_{kl}^{-1}\mathrm{d}A_k' + A_{kl}^{-1}A_kMA_k'A_{kl}^{-1}\mathrm{d}A_l\right],\\ \mathrm{d}\eta_i &= \mathrm{tr}\left[\left(K_{kl}^{(i)}A_l'MA_lA_{kl}^{-1} - A_{kl}^{-1}A_l'MA_kK_{kl}^{(i)}\right)\mathrm{d}A_k'\right.\\ &\quad\left.+\left(A_{kl}^{-1}A_k'MA_kK_{kl}^{(i)} - K_{kl}^{(i)}A_l'MA_kA_{kl}^{-1}\right)\mathrm{d}A_l\right],\\ &\quad i = 1, 2, 5, 6,\\ \mathrm{d}\eta_3 &= \mathrm{tr}\left[\left(A_{kl}^{-1}A_l'MA_kK_{kl}^{(3)} + K_{kl}^{(3)}A_k'MA_lA_{kl}^{-1}\right)\mathrm{d}A_k'\right.\\ &\quad\left.-\left(A_{kl}^{-1}A_k'MA_kK_{kl}^{(3)} + K_{kl}^{(3)}A_k'MA_kA_{kl}^{-1}\right)\mathrm{d}A_l'\right],\\ \mathrm{d}\eta_4 &= \mathrm{tr}\left[\left(A_{kl}^{-1}A_k'MA_lK_{kl}^{(4)} + K_{kl}^{(4)}A_l'MA_kA_{kl}^{-1}\right)\mathrm{d}A_l\right.\\ &\quad\left.-\left(A_{kl}^{-1}A_l'MA_lK_{kl}^{(4)} + K_{kl}^{(4)}A_l'MA_lA_{kl}^{-1}\right)\mathrm{d}A_k'\right].\end{aligned} \tag{88}$$

Using the above differentials, $\mathrm{d}\mathtt{T}_{kl}$ can be written as

$$\mathrm{d}\mathtt{T}_{kl} = \frac{\mathtt{T}_{kl}}{\mathtt{S}_{kl}}\mathrm{d}\mathtt{S}_{kl} + \mathtt{S}_{kl}\left(6\mathrm{d}\tau + \frac{4}{\gamma}\mathrm{d}\zeta - \frac{4\zeta}{\gamma^2}\mathrm{d}\gamma\right). \tag{89}$$

Now let us define two more matrices,

$$U_{kl} = 6A_{kl}^{-1}A_lMA_l'A_{kl}^{-1} + \frac{4}{\gamma}X_{kl} + 4\frac{\zeta}{\gamma^2}K_{kl}, \tag{90}$$

$$W_{kl} = 6A_{kl}^{-1}A_kMA_k'A_{kl}^{-1} + \frac{4}{\gamma}Y_{kl} + 4\frac{\zeta}{\gamma^2}K_{kl}, \tag{91}$$

where

$$\begin{aligned}X_{kl} = &\sum_{\substack{i=1\\ i\neq 3,4}}^{6}\gamma_i\left(K_{kl}^{(i)}A_l'MA_lA_{kl}^{-1} - A_{kl}^{-1}A_l'MA_kK_{kl}^{(i)}\right)\\ &+G_{kl} + \gamma_4\left(A_{kl}^{-1}A_l'MA_lK_{kl}^{(4)} + K_{kl}^{(4)}A_l'MA_lA_{kl}^{-1}\right)\\ &-\gamma_3\left(A_{kl}^{-1}A_l'MA_kK_{kl}^{(3)} + K_{kl}^{(3)}A_k'MA_lA_{kl}^{-1}\right),\end{aligned} \tag{92}$$

$$\begin{aligned}Y_{kl} = &\sum_{\substack{i=1\\ i\neq 3,4}}^{6}\gamma_i\left(A_{kl}^{-1}A_k'MA_kK_{kl}^{(i)} - K_{kl}^{(i)}A_l'MA_kA_{kl}^{-1}\right)\\ &+G_{kl} + \gamma_3\left(A_{kl}^{-1}A_k'MA_kK_{kl}^{(3)} + K_{kl}^{(3)}A_k'MA_kA_{kl}^{-1}\right)\\ &-\gamma_4\left(A_{kl}^{-1}A_k'MA_lK_{kl}^{(4)} + K_{kl}^{(4)}A_l'MA_kA_{kl}^{-1}\right),\end{aligned} \tag{93}$$

and

$$\begin{aligned}G_{kl} = &\left(\eta_3K_{kl}^{(4)} + \eta_4K_{kl}^{(3)} - \eta_1K_{kl}^{(2)}\right.\\ &\left.-\eta_2K_{kl}^{(1)} - \eta_5K_{kl}^{(6)} - \eta_6K_{kl}^{(5)}\right).\end{aligned} \tag{94}$$

With these the differential of the kinetic energy can be cast in the following form:

$$\begin{aligned}\mathrm{d}\mathtt{T}_{kl} = &\frac{\mathtt{T}_{kl}}{\mathtt{S}_{kl}}\mathrm{d}\mathtt{S}_{kl} + \mathtt{S}_{kl}\,\mathrm{tr}[\mathrm{vech}((U_{kl}+U_{kl}')L_k)'\mathrm{d}\,\mathrm{vech}\,L_k\\ &+\mathrm{vech}((W_{kl}+W_{kl}')L_l)'\mathrm{d}\,\mathrm{vech}\,L_l].\end{aligned} \tag{95}$$

The derivatives of the kinetic energy matrix elements are then given by

$$\frac{\partial\mathtt{T}_{kl}}{\partial(\mathrm{vech}\,L_k)} = \frac{\mathtt{T}_{kl}}{\mathtt{S}_{kl}}\frac{\partial\mathtt{S}_{kl}}{\partial(\mathrm{vech}\,L_k)} + \mathtt{S}_{kl}\,\mathrm{vech}((U_{kl}+U_{kl}')L_k), \tag{96}$$

$$\frac{\partial\mathtt{T}_{kl}}{\partial(\mathrm{vech}\,L_l)} = \frac{\mathtt{T}_{kl}}{\mathtt{S}_{kl}}\frac{\partial\mathtt{S}_{kl}}{\partial(\mathrm{vech}\,L_l)} + \mathtt{S}_{kl}\,\mathrm{vech}((W_{kl}+W_{kl}')L_l). \tag{97}$$

2. *Case* $(x_{i_k}y_{j_k}|y_{i_l}x_{j_l})$

All of the final formulas for derivatives of the kinetic energy matrix elements remain the same. However, matrices X_{kl}, Y_{kl}, and G_{kl} acquire a different form:

$$X_{kl} = G_{kl} + \sum_{i=1}^{2} \gamma_i \big(K_{kl}^{(i)} A_l' M A_l A_{kl}^{-1} - A_{kl}^{-1} A_l' M A_k K_{kl}^{(i)}\big), \quad (98)$$

$$Y_{kl} = G_{kl} + \sum_{i=1}^{2} \gamma_i \big(A_{kl}^{-1} A_k' M A_k K_{kl}^{(i)} - K_{kl}^{(i)} A_l' M A_k A_{kl}^{-1}\big), \quad (99)$$

$$G_{kl} = -\eta_1 K_{kl}^{(2)} - \eta_2 K_{kl}^{(1)}, \quad \zeta = \gamma_1\eta_1 + \gamma_2\eta_2. \quad (100)$$

3. *Case* $(x_{i_k}x_{j_k}|y_{i_l}y_{j_l})$

In this case the expressions for the derivatives of $\mathtt{T}_{kl}$ also remain unchanged, while X_{kl}, Y_{kl}, and G_{kl} are

$$\begin{aligned} X_{kl} = {} & G_{kl} + \gamma_4\big(A_{kl}^{-1} A_l' M A_l K_{kl}^{(4)} + K_{kl}^{(4)} A_l' M A_l A_{kl}^{-1}\big) \\ & - \gamma_3\big(A_{kl}^{-1} A_l' M A_k K_{kl}^{(3)} + K_{kl}^{(3)} A_k' M A_l A_{kl}^{-1}\big), \end{aligned} \quad (101)$$

$$\begin{aligned} Y_{kl} = {} & G_{kl} + \gamma_3\big(A_{kl}^{-1} A_k' M A_k K_{kl}^{(3)} + K_{kl}^{(3)} A_k' M A_k A_{kl}^{-1}\big) \\ & - \gamma_4\big(A_{kl}^{-1} A_k' M A_l K_{kl}^{(4)} + K_{kl}^{(4)} A_l M A_k' A_{kl}^{-1}\big), \end{aligned} \quad (102)$$

$$G_{kl} = \eta_3 K_{kl}^{(4)} + \eta_4 K_{kl}^{(3)}, \quad \zeta = -\gamma_3\eta_3 - \gamma_4\eta_4. \quad (103)$$

C. Potential energy gradient

1. *Case* $(x_{i_k}x_{j_k}|x_{i_l}x_{j_l})$

The differential of $\mathtt{R}_{kl}^{ij}$ is

$$\mathrm{d}\mathtt{R}_{kl}^{ij} = \frac{\mathtt{R}_{kl}^{ij}}{\mathtt{S}_{kl}}\mathrm{d}\mathtt{S}_{kl} + \frac{2}{\sqrt{\pi}}\frac{\mathtt{S}_{kl}}{\omega^{3/2}}\left(\frac{1}{2}\left[\frac{\kappa}{\omega\gamma} - \frac{\rho}{\omega^2\gamma} - 1\right]\mathrm{d}\omega + \left[\frac{\kappa}{3\gamma^2} - \frac{\rho}{5\omega\gamma^2}\right]\mathrm{d}\gamma - \frac{\mathrm{d}\kappa}{3\gamma} + \frac{\mathrm{d}\rho}{5\omega\gamma}\right). \quad (104)$$

The differentials of $\mathrm{d}\rho_q$ $(q = 1, \dots, 6)$, ω, ρ, and κ are

$$\mathrm{d}\rho_q = -\operatorname{tr}\big[\big(A_{kl}^{-1} J_{ij} K_{kl}^{(q)} + K_{kl}^{(q)} J_{ij} A_{kl}^{-1}\big)(\mathrm{d}A_k' + \mathrm{d}A_l)\big], \quad q = 1, \dots, 6, \quad (105)$$

$$\mathrm{d}\omega = -\operatorname{tr}\big[A_{kl}^{-1} J_{ij} A_{kl}^{-1} \mathrm{d}A_{kl}\big], \quad (106)$$

$$\begin{aligned} \mathrm{d}\rho = {} & -\Big(\rho_1 \operatorname{tr}\big[\big(A_{kl}^{-1} J_{ij} K_{kl}^{(2)} + K_{kl}^{(2)} J_{ij} A_{kl}^{-1}\big)\mathrm{d}A_{kl}\big] + \rho_2 \operatorname{tr}\big[\big(A_{kl}^{-1} J_{ij} K_{kl}^{(1)} + K_{kl}^{(1)} J_{ij} A_{kl}^{-1}\big)\mathrm{d}A_{kl}\big] + \rho_3 \operatorname{tr}\big[\big(A_{kl}^{-1} J_{ij} K_{kl}^{(4)} + K_{kl}^{(4)} J_{ij} A_{kl}^{-1}\big)\mathrm{d}A_{kl}\big] \\ & + \rho_4 \operatorname{tr}\big[\big(A_{kl}^{-1} J_{ij} K_{kl}^{(3)} + K_{kl}^{(3)} J_{ij} A_{kl}^{-1}\big)\mathrm{d}A_{kl}\big] + \rho_5 \operatorname{tr}\big[\big(A_{kl}^{-1} J_{ij} K_{kl}^{(6)} + K_{kl}^{(6)} J_{ij} A_{kl}^{-1}\big)\mathrm{d}A_{kl}\big] + \rho_6 \operatorname{tr}\big[\big(A_{kl}^{-1} J_{ij} K_{kl}^{(5)} + K_{kl}^{(5)} J_{ij} A_{kl}^{-1}\big)\mathrm{d}A_{kl}\big]\Big), \end{aligned} \quad (107)$$

$$\begin{aligned} \mathrm{d}\kappa = {} & -\Big(\rho_1 \operatorname{tr}\big[K_{kl}^{(2)}\mathrm{d}A_{kl}\big] + \rho_2 \operatorname{tr}\big[K_{kl}^{(1)}\mathrm{d}A_{kl}\big] + \rho_3 \operatorname{tr}\big[K_{kl}^{(4)}\mathrm{d}A_{kl}\big] + \rho_4 \operatorname{tr}\big[K_{kl}^{(3)}\mathrm{d}A_{kl}\big] + \rho_5 \operatorname{tr}\big[K_{kl}^{(6)}\mathrm{d}A_{kl}\big] + \rho_6 \operatorname{tr}\big[K_{kl}^{(5)}\mathrm{d}A_{kl}\big] \\ & + \sum_{q=1}^{6} \gamma_q \operatorname{tr}\big[\big(A_{kl}^{-1} J_{ij} K_{kl}^{(q)} + K_{kl}^{(q)} J_{ij} A_{kl}^{-1}\big)\mathrm{d}A_{kl}\big]\Big). \end{aligned} \quad (108)$$

Here we will also define the following matrix:

$$\begin{aligned} Q_{kl} = {} & \frac{2}{\sqrt{\pi}\,\omega^{3/2}}\left(\frac{1}{2}\left[1 + \frac{\rho}{\omega^2\gamma} - \frac{\kappa}{\omega\gamma}\right]A_{kl}^{-1} J_{ij} A_{kl}^{-1} + \left[\frac{\rho}{5\omega\gamma^2} - \frac{\kappa}{3\gamma^2}\right]K_{kl} + \frac{1}{3\gamma}\left[\rho_1 K_{kl}^{(2)} + \rho_2 K_{kl}^{(1)} + \rho_3 K_{kl}^{(4)} + \rho_4 K_{kl}^{(3)} + \rho_5 K_{kl}^{(6)}\right.\right. \\ & \left. + \rho_6 K_{kl}^{(5)} + \sum_{q=1}^{6} \gamma_q \big(A_{kl}^{-1} J_{ij} K_{kl}^{(q)} + K_{kl}^{(q)} J_{ij} A_{kl}^{-1}\big)\right] - \frac{1}{5\omega\gamma}\big[\rho_1\big(A_{kl}^{-1} J_{ij} K_{kl}^{(2)} + K_{kl}^{(2)} J_{ij} A_{kl}^{-1}\big) + \rho_2\big(A_{kl}^{-1} J_{ij} K_{kl}^{(1)} + K_{kl}^{(1)} J_{ij} A_{kl}^{-1}\big) \\ & + \rho_3\big(A_{kl}^{-1} J_{ij} K_{kl}^{(4)} + K_{kl}^{(4)} J_{ij} A_{kl}^{-1}\big) + \rho_4\big(A_{kl}^{-1} J_{ij} K_{kl}^{(3)} + K_{kl}^{(3)} J_{ij} A_{kl}^{-1}\big) + \rho_5\big(A_{kl}^{-1} J_{ij} K_{kl}^{(6)} + K_{kl}^{(6)} J_{ij} A_{kl}^{-1}\big) \\ & \left. + \rho_6\big(A_{kl}^{-1} J_{ij} K_{kl}^{(6)} + K_{kl}^{(6)} J_{ij} A_{kl}^{-1}\big)\big]\right). \end{aligned} \quad (109)$$

Using the above definitions the differential and derivatives of $\mathtt{R}_{kl}^{ij}$ can be written as

$$\mathrm{d}\mathtt{R}_{kl}^{ij} = \frac{\mathtt{R}_{kl}^{ij}}{\mathtt{S}_{kl}}\mathrm{d}\mathtt{S}_{kl} + \mathtt{S}_{kl} \operatorname{tr}\big[Q_{kl}(\mathrm{d}A_k' + \mathrm{d}A_l)\big], \quad (110)$$

$$\frac{\partial \mathtt{R}_{kl}^{ij}}{\partial(\operatorname{vech} L_k)} = \frac{\mathtt{R}_{kl}^{ij}}{\mathtt{S}_{kl}}\frac{\partial \mathtt{S}_{kl}}{\partial(\operatorname{vech} L_k)} + \mathtt{S}_{kl} \operatorname{vech}((Q_{kl} + Q_{kl}')L_k), \quad (111)$$

$$\frac{\partial \mathtt{R}_{kl}^{ij}}{\partial(\operatorname{vech} L_l)} = \frac{\mathtt{R}_{kl}^{ij}}{\mathtt{S}_{kl}}\frac{\partial \mathtt{S}_{kl}}{\partial(\operatorname{vech} L_l)} + \mathtt{S}_{kl} \operatorname{vech}((Q_{kl} + Q_{kl}')L_l). \quad (112)$$

TABLE I. CPU time (in seconds) spent on evaluating 10^8 sets of matrix elements of the Hamiltonian + overlap (Gradient = no) and Hamiltonian + overlap + gradient (Gradient = yes) using two different schemes: the new one implemented in this work and the old one that is based on derivations in Refs. [19,31]. All calculations were performed on a single-core of Intel Xeon W-2145 CPU using GNU FORTRAN compiler version 9.5.

		$L = 1$			$L = 2$		
System	Gradient	New	Old	Ratio	New	Old	Ratio
Li	yes	96.06	5202	54.15	97.02	8184	84.35
	no	11.39	511	44.86	16.25	820	50.46
Be	yes	305.20	15062	49.35	310.20	25781	83.11
	no	24.50	945	38.57	30.00	1462	48.73
B	yes	620.62	43610	70.27	628.75	65984	104.94
	no	41.36	1820	44.00	60.33	2524	41.83
C	yes	1062.60	120297	113.21	1081.70	154922	143.22
	no	71.94	3023	42.02	82.05	3733	45.50

2. Case $(x_{i_k}y_{j_k}|y_{i_l}x_{j_l})$

In this case, we have

$$\begin{aligned}\kappa &= \gamma_1\rho_1 + \gamma_2\rho_2, \quad \rho = \rho_1\rho_2,\\ \gamma &= \gamma_1\gamma_2, \quad K_{kl} = \big(\gamma_1 K_{kl}^{(1)} + \gamma_2 K_{kl}^{(2)}\big),\end{aligned} \tag{113}$$

$$\begin{aligned}Q_{kl} = \frac{2}{\sqrt{\pi}\omega^{3/2}}\Bigg(&\frac{1}{2}\left[1 + \frac{\rho}{\omega^2\gamma} - \frac{\kappa}{\omega\gamma}\right]A_{kl}^{-1}J_{ij}A_{kl}^{-1} + \left[\frac{\rho}{5\omega\gamma^2} - \frac{\kappa}{3\gamma^2}\right]K_{kl} + \frac{1}{3\gamma}\Bigg[\rho_1 K_{kl}^{(2)} + \rho_2 K_{kl}^{(1)} + \sum_{q=1}^{2}\gamma_q\big(A_{kl}^{-1}J_{ij}K_{kl}^{(q)}\\ &+ K_{kl}^{(q)}J_{ij}A_{kl}^{-1}\big)\Bigg] - \frac{1}{5\omega\gamma}\big[\rho_1\big(A_{kl}^{-1}J_{ij}K_{kl}^{(2)} + K_{kl}^{(2)}J_{ij}A_{kl}^{-1}\big) + \rho_2\big(A_{kl}^{-1}J_{ij}K_{kl}^{(1)} + K_{kl}^{(1)}J_{ij}A_{kl}^{-1}\big)\big]\Bigg),\end{aligned} \tag{114}$$

while expressions for the differential and derivatives of R_{kl}^{ij} are the same.

3. Case $(x_{i_k}x_{j_k}|y_{i_l}y_{j_l})$

Similarly, here the expressions for the differential and derivatives of R_{kl}^{ij} are the same. The only thing that changes is the form of expressions for κ, γ, K_{kl}, and Q_{kl}:

$$\begin{aligned}\kappa &= [\gamma_3\rho_3 + \gamma_4\rho_4], \quad \rho = \rho_3\rho_4,\\ \gamma &= \gamma_3\gamma_4, \quad K_{kl} = \big(\gamma_3 K_{kl}^{(3)} + \gamma_4 K_{kl}^{(4)}\big),\end{aligned} \tag{115}$$

$$\begin{aligned}Q_{kl} = \frac{2}{\sqrt{\pi}\omega^{3/2}}\Bigg(&\frac{1}{2}\left[1 + \frac{\rho}{\omega^2\gamma} - \frac{\kappa}{\omega\gamma}\right]A_{kl}^{-1}J_{ij}A_{kl}^{-1} + \left[\frac{\rho}{5\omega\gamma^2} - \frac{\kappa}{3\gamma^2}\right]K_{kl} + \frac{1}{3\gamma}\Bigg[\rho_3 K_{kl}^{(4)} + \rho_4 K_{kl}^{(3)} + \sum_{q=3}^{4}\gamma_q\big(A_{kl}^{-1}J_{ij}K_{kl}^{(q)}\\ &+ K_{kl}^{(q)}J_{ij}A_{kl}^{-1}\big)\Bigg] - \frac{1}{5\omega\gamma}\big[\rho_3\big(A_{kl}^{-1}J_{ij}K_{kl}^{(4)} + K_{kl}^{(4)}J_{ij}A_{kl}^{-1}\big) + \rho_4\big(A_{kl}^{-1}J_{ij}K_{kl}^{(3)} + K_{kl}^{(3)}J_{ij}A_{kl}^{-1}\big)\big]\Bigg).\end{aligned} \tag{116}$$

VIII. NUMERICAL TESTS

The expressions derived in the previous sections were implemented into our FORTRAN 90 code for variational calculations of quantum few-body systems. We used the Message Passing Interface (MPI) library to facilitate parallel calculations on computer systems with distributed and shared memory architectures. The FORTRAN code for all individual matrix elements as well as for all auxiliary functions and subroutines was validated by comparing the corresponding results with the numerical output generated by the *Mathematica* symbolic algebra package. We also used up to five-point finite difference formulas to check the accuracy of our analytic gradient of the energy with respect to the Gaussian nonlinear parameters.

The scheme for calculating matrix elements adopted in this work has a significant advantage in terms of efficiency as compared to previous implementations [19,31]. This is demonstrated in Table I, which shows the CPU timings for evaluation of two sets of matrix elements (Hamiltonian + overlap and Hamiltonian + overlap + gradient) using two schemes: the new one that represents basis functions as $[\boldsymbol{v}_k'\boldsymbol{r}][\boldsymbol{w}_k'\boldsymbol{r}]\exp[-\boldsymbol{r}'A_k\boldsymbol{r}]$ and the old one where basis functions are represented as $\boldsymbol{r}'W_k\boldsymbol{r}\exp[-\boldsymbol{r}'A_k\boldsymbol{r}]$ [19,31]. Here, $\boldsymbol{W}$ is a sparse $3n \times 3n$ symmetric matrix, which defines the prefactors of ECGs in Eqs. (13)–(16). We measured the time of evaluating matrix elements for several systems: Li, Be, B, and C atoms. In the table, $L = 1$ denotes the case of basis function

$$(x_{i_k}y_{j_k} - y_{i_k}x_{j_k})\exp[-\boldsymbol{r}'A_k\boldsymbol{r}],$$

while $L = 2$ corresponds to

$$(x_{i_k}x_{j_k} + y_{i_k}y_{j_k} - 2z_{i_k}z_{j_k})\exp[-\boldsymbol{r}'A_k\boldsymbol{r}] \quad \text{or}$$
$$\left(x_{i_k}^2 + y_{i_k}^2 - 2z_{i_k}^2\right)\exp[-\boldsymbol{r}'A_k\boldsymbol{r}].$$

In Table I one can see that the ratio of times spent on evaluating matrix elements based on two different schemes (old versus new) is very large. The use of the new scheme results in roughly a two-order-of-magnitude speedup. Moreover, the ratio increases with the size of the system, i.e., with the number of electrons. For D states of carbon atom ($L = 2$) the ratio of old to new reaches the value of 143 when the matrix elements of the gradient are included. When matrix elements of the gradient are not evaluated, it is roughly 45 times faster without gradients. A higher ratio in the case when the gradient is included can be explained by the fact that the expressions for matrix elements of the analytic gradient derived within the scheme where the basis functions are represented as $\boldsymbol{r}'W_k\boldsymbol{r}\exp[-\boldsymbol{r}'A_k\boldsymbol{r}]$ [31] are somewhat more complicated (from a computational standpoint) than those derived in this work. It should be noted that in practical calculations with ECGs, the optimization of the nonlinear parameters of the Gaussians is very essential and usually represents the most time-consuming part of the calculations. Therefore, doing the optimization of the variational wave functions essentially becomes impractical without using the analytic energy gradient. Thus the timings for the case without gradient seen in Table I serve more of an illustrative purpose.

For timings shown in Table I, each set of matrix elements did not include any summation over permutations of identical particles. Therefore, in order to assess how the computational cost of computing matrix elements grows with the system size one would need to make ratios of the corresponding timings each multiplied by the corresponding factorial of the number of identical particles in the system (or products of factorials, if there is more than one set of identical particles in the system). For example, when we go from a three-electron lithium ($3! = 6$) to a six-electron carbon ($6! = 720$), the cost increases by a factor of approximately 1300 in the case of the Hamiltonian + overlap + gradient matrix elements for both $L = 1$ and $L = 2$ states.

Next, let us show the results of several test calculations on specific states of Li, B, and C atoms that can be treated using the new algorithms derived in this work. We include the finite nuclear mass effects in our computations. However, for the purpose of comparing our numbers with the benchmark data available in the literature, we also report our nonrelativistic energies for the infinite nuclear mass "isotopes." Since the atomic wave function undergoes a relatively minor change when the nuclear mass is changed, there is no need to reoptimize the ECG basis set for a different nuclear mass. The energy shift caused by the nuclear mass change is captured very well by simply adjusting the linear variational parameters (expansion coefficients) of the basis functions when the modified Hamiltonian for a specific isotope is diagonalized in a given ECG basis.

TABLE II. Convergence of the total nonrelativistic energies (in a.u.) of the lithium atom in the lowest ^2D state with the number of ECG basis functions.

	Basis size	$1s^23d$ (^2D)
$^\infty$Li	500	−7.335 523 143 44
	1000	−7.335 523 501 56
	1500	−7.335 523 535 12
	2000	−7.335 523 540 22
	2500	−7.335 523 541 79
	3000	−7.335 523 542 61
^{7}Li	3000	−7.334 927 305 61
	Hylleraas-type basis [33]	
$^\infty$Li	32760	−7.335 523 543 524 685

We would like to emphasize that the test calculations presented below were performed only for the purpose of validating the new algorithm for computing matrix elements of the overlap, kinetic energy, potential energy, and their gradients. In future works we are planning to focus on specific applications that target bound states of atoms, ions, and systems with exotic particles such as positrons. More details and comprehensive numerical data will be presented then. Apart from the new scheme of deriving the matrix elements, the algorithm implemented in this work includes an independent optimization routine for integer indices labeling the particles.

The first test system for validating the new algorithm is the Li atom in its lowest doublet D state. For the nuclear mass of ^{7}Li we adopted the value $m_0 = 12\,786.392\,282$ a.u., which was determined based on the corresponding atomic mass value from Ref. [32]. The results of our test calculations in which we generated basis sets of up to 3000 ECGs are shown in Table II. In this table one can see that with just 3000 ECGs we recover nine (in fact, almost ten) decimal figures in the total nonrelativistic energy of this state. This is confirmed by comparing with the most accurate value reported to date by Wang *et al.* [33] that was obtained in a large-scale variational calculation using Hylleraas-type basis functions. Such level of convergence and very good agreement with the benchmark data for a small, simple, yet nontrivial state of a three-electron system provides solid evidence that the algorithm presented in this work has been implemented correctly. The basis set generation used in this work for all test systems was based on the following strategy that was adopted from our previous works on ECGs and proved its robustness. It begins with choosing randomly the nonlinear parameters of one or a few Gaussians, i.e., the elements of matrix L_k, and then optimizing them with an approach that utilizes the analytic energy gradient. Then new basis functions are added one by one, first by randomly selecting their nonlinear parameters and then the best candidate that lowers the energy the most is optimized further using the analytic energy gradient. At

TABLE III. Convergence of the total nonrelativistic energies (in a.u.) of the boron atom in the lowest ${}^4P$ and ${}^2D$ states with the number of basis functions and comparison with the most accurate literature values.

	Basis size	$1s^22s2p^2\,({}^4P)$	$1s^22s2p^2\,({}^2D)$
${}^\infty$B	3000	$-24.522\,039\,882$	$-24.435\,953\,545$
	4000	$-24.522\,040\,952$	$-24.435\,968\,363$
	5000	$-24.522\,041\,368$	$-24.435\,974\,673$
	6000	$-24.522\,041\,568$	$-24.435\,977\,820$
	7000	$-24.522\,041\,671$	$-24.435\,979\,472$
	8000	$-24.522\,041\,732$	$-24.435\,980\,588$
	9000	$-24.522\,041\,775$	$-24.435\,981\,350$
^{11}B	9000	$-24.520\,827\,254$	$-24.434\,765\,417$
		ECG lobe functions [34]	
${}^\infty$B	4672	$-24.522\,041\,430$	
	8231		$-24.435\,981\,009$

the stage of random selection the nonlinear parameters are sampled from the distribution that is based on the nonlinear parameters of the ECGs that are already in the basis. Typically, a few hundred random basis function candidates are tested before the best one is selected. After a certain number of new functions (typically 5–50) have been added to the basis, we perform a cycle in which the parameters of all basis functions are tuned further, one function at a time. The reason why only one (or few, at most) function is being optimized at a time is because it allows considerable computational savings as in such case there is no need to recompute all matrix elements and solve the generalized eigenvalue problem (which yields linear variational parameters) from scratch. When the optimization cycle is completed, the procedure is repeated; i.e., new functions are added and then a new optimization cycle occurs. This way the basis is grown in a controlled way until its size reaches a large predefined value. In this process of gradual growing and optimizing the ECG basis we also check for possible linear dependencies among basis functions to maintain numerical stability of the calculations. If at some point it is determined that a new function candidate or a modification of already included basis function causes severe linear dependencies (based on a predefined threshold), such additions or modifications are discarded.

Another test calculation that we performed was of the boron atom. We considered two states for this system: the lowest ${}^4P$ state and the lowest ${}^2D$ state. Our data for the total nonrelativistic energies are shown in Table III and we compare them to the most recent and accurate values reported by Strasburger [34]. The nuclear mass of ^{11}B that we adopted is $m_0 = 20\,063.736\,943$ a.u., which was computed using the known value of the atomic mass [32]. Based on the convergence patterns and comparison with the data in Ref. [34] we can conclude that the nonrelativistic energies for both states are converged to about eight decimal figures. Moreover, our values corresponding to the largest generated basis sets of 9000 ECGs provide a marginal improvement over the variational upper bounds reported by Strasburger.

The last test case that we present concerns the six-electron carbon atom. For this system we considered two triplet states: the ground state with the dominant electronic configuration $1s^22s^22p^2\,({}^3P)$ and the first excited state of the same multiplicity and rotational symmetry corresponding to the configuration $1s^22s^22p3p\,({}^3P)$. For the nuclear mass of ^{12}C we used the value $m_0 = 21\,868.663\,851$ a.u. Using up to 8000 ECGs for each state we were able to improve the ground state energy of the carbon atom obtained by Strasburger in a benchmarked variational calculation in Ref. [35] that employed Gaussian lobe basis functions. With this many ECGs the ground state energy is converged to seven decimal figures. It is remarkable that the Gaussian lobe functions, which are variants of ECGs with shifted centers that are spatially projected according to the point group representations of the considered states, provide a rather fast convergence rate to the infinite basis set limit. Apparently, part of the reason for this is due to the fact that the spatial symmetry projector in the case of the ${}^3P$ state of carbon consists of eight terms; i.e., eight ECGs with shifted centers are combined in order to form a single basis function (and this does not take into account the permutational symmetry of the electrons). Yet, ECG lobe functions are not eigenfunctions of the square of the total orbital angular momentum operator. Therefore, they cannot be easily used for calculations of excited states such as the Rydberg series of states with the dominant configurations $1s^22s^22pnp\,({}^3P)$ $(n = 3, 4, \ldots)$. The ECG basis functions adopted in this work do not have this deficiency and can be used for that purpose (i.e., for calculations of excited states with S^e, P^e, and D^e symmetry).

This is demonstrated by our data for the first excited $1s^22s^22p3p\,({}^3P)$ state in Table IV, where the level of convergence reached for the total nonrelativistic energy is very similar (just a tad worse, as expected due to the increased complexity of the wave function) to what we have for the ground state. Let us also note that in spite of a rather moderate size of the basis we generated in this numerical test

TABLE IV. Convergence of the total nonrelativistic energies (in a.u.) of the carbon atom in the ground and excited ${}^3P$ states with the number of basis functions and comparison with the most accurate literature values. Note that the energy uncertainty from Ref. [36] is of purely statistical nature and does not account for the error due to the wave function *Ansatz* employed in that work.

	Basis size	$1s^22s^22p^2\,({}^3P)$	$1s^22s^22p3p\,({}^3P)$
${}^\infty$C	1000	$-37.844\,310\,1$	$-37.518\,749\,7$
	2000	$-37.844\,736\,9$	$-37.519\,414\,4$
	3000	$-37.844\,822\,2$	$-37.519\,552\,9$
	4000	$-37.844\,851\,8$	$-37.519\,600\,6$
	5000	$-37.844\,869\,6$	$-37.519\,630\,8$
	6000	$-37.844\,880\,2$	$-37.519\,648\,7$
	7000	$-37.844\,887\,0$	$-37.519\,660\,0$
	8000	$-37.844\,891\,3$	
^{12}C	7000		$-37.517\,946\,1$
	8000	$-37.843\,179\,3$	
		ECG lobe functions [35]	
${}^\infty$C	5896	$-37.844\,889\,402$	
		Variational Monte Carlo [36]	
${}^\infty$C			$-37.496\,97(5)$

calculation, to the best of our knowledge, this is by far most accurate nonrelativistic energy ever reported for this state.

IX. CONCLUSION

All-particle explicitly correlated Gaussians with prefactors in the form of bipolar harmonics is a powerful tool that can be used in benchmark variational calculations of atomic S^e, P^e, and D^e states, where the dominant configuration contains two *p* electrons or a single *d* electron. The algorithm developed in this work provides a fast and efficient way of computing the Hamiltonian and overlap matrix elements as well as the analytic energy gradient with respect to the nonlinear parameters of the Gaussians. All necessary analytical expressions are presented in compact matrix form for easier implementation and theoretical analysis. Such basis functions can be extended to perform bound state calculations for other types of systems, such as those containing more than one heavy particle or nucleus. They are also expected to be very effective in calculations of excited Rydberg states of atoms with some moderately large values of the principal quantum number of the Rydberg electron. Future possible developments with these basis functions will include the evaluation of the expectation values of effective operators representing leading relativistic and quantum electrodynamics corrections, interstate transition energies, and oscillator strengths.

ACKNOWLEDGMENTS

This work has been supported by Nazarbayev University FDCRG Grant No. 021220FD3651. We acknowledge the use of the Shabyt high performance computer cluster at Nazarbayev University, where most of the numerical calculations reported in this work were carried out.

[1] E. A. Hylleraas, Neue Berechnung der Energie des Heliums im Grundzustande, sowie des tiefsten Terms von Ortho-Helium, Z. Phys. **54**, 347 (1929).

[2] C. Schwartz, Further computations of the He atom ground state, arXiv:math-ph/0605018.

[3] C. Schwartz, Experiment and theory in computations of the He atom ground state, Int. J. Mod. Phys. E **15**, 877 (2006).

[4] H. Nakashima and H. Nakatsuji, Solving the Schrödinger equation for helium atom and its isoelectronic ions with the free iterative complement interaction (ICI) method, J. Chem. Phys. **127**, 224104 (2007).

[5] V. I. Korobov, Coulomb three-body bound-state problem: Variational calculations of nonrelativistic energies, Phys. Rev. A **61**, 064503 (2000).

[6] S. F. Boys, The integral formulae for the variational solution of the molecular many-electron wave equations in terms of Gaussian functions with direct electronic correlation, Proc. R. Soc. London A **258**, 402 (1960).

[7] K. Singer, The use of Gaussian (exponential quadratic) wave functions in molecular problems. I. General formulae for the evaluation of integrals, Proc. R. Soc. London A **258**, 412 (1960).

[8] V. I. Kukulin and V. M. Krasnopol'sky, A stochastic variational method for few-body systems, J. Phys. G **3**, 795 (1977).

[9] K. Varga and Y. Suzuki, Precise solution of few-body problems with the stochastic variational method on a correlated Gaussian basis, Phys. Rev. C **52**, 2885 (1995).

[10] Y. Suzuki and K. Varga, *Stochastic Variational Approach to Quantum-Mechanical Few-Body Problems*, Lecture Notes in Physics Monographs Vol. 54 (Springer, Berlin, 1998).

[11] J. Rychlewski, *Explicitly Correlated Wave Functions in Chemistry and Physics: Theory and Applications*, Progress in Theoretical Chemistry and Physics Vol. 13 (Springer, Dordrecht, 2003).

[12] J. Mitroy, S. Bubin, W. Horiuchi, Y. Suzuki, L. Adamowicz, W. Cencek, K. Szalewicz, J. Komasa, D. Blume, and K. Varga, Theory and application of explicitly correlated Gaussians, Rev. Mod. Phys. **85**, 693 (2013).

[13] S. Bubin, M. Pavanello, W.-C. Tung, K. L. Sharkey, and L. Adamowicz, Born-Oppenheimer and non-Born-Oppenheimer, atomic and molecular calculations with explicitly correlated Gaussians, Chem. Rev. **113**, 36 (2013).

[14] S. Bubin and L. Adamowicz, Matrix elements of *N*-particle explicitly correlated Gaussian basis functions with complex exponential parameters, J. Chem. Phys. **124**, 224317 (2006).

[15] S. Bubin and L. Adamowicz, Energy and energy gradient matrix elements with *N*-particle explicitly correlated complex Gaussian basis functions with $L = 1$, J. Chem. Phys. **128**, 114107 (2008).

[16] K. L. Sharkey, M. Pavanello, S. Bubin, and L. Adamowicz, Algorithm for quantum-mechanical finite-nuclear-mass variational calculations of atoms with two *p* electrons using all-electron explicitly correlated Gaussian basis functions, Phys. Rev. A **80**, 062510 (2009).

[17] K. L. Sharkey, N. Kirnosov, and L. Adamowicz, An algorithm for quantum mechanical finite-nuclear-mass variational calculations of atoms with $L = 3$ using all-electron explicitly correlated Gaussian basis functions, J. Chem. Phys. **138**, 104107 (2013).

[18] J. Komasa, Dipole and quadrupole polarizabilities and shielding factors of beryllium from exponentially correlated Gaussian functions, Phys. Rev. A **65**, 012506 (2001).

[19] K. L. Sharkey, S. Bubin, and L. Adamowicz, Analytical energy gradient in variational calculations of the two lowest ^3P states of the carbon atom with explicitly correlated Gaussian basis functions, J. Chem. Phys. **132**, 184106 (2010).

[20] K. L. Sharkey, S. Bubin, and L. Adamowicz, Refinement of the experimental energy levels of higher ^2D Rydberg states of the lithium atom with very accurate quantum mechanical calculations, J. Chem. Phys. **134**, 194114 (2011).

[21] K. L. Sharkey, S. Bubin, and L. Adamowicz, Singlet-triplet energy splitting between ^1D and ^3D $(1s^2 2snd)$ $n = 3$, 4, 5, and 6, Rydberg states of the beryllium atom (^{9}Be) calculated with all-electron explicitly correlated Gaussian functions, Chem. Phys. Lett. **616-617**, 254 (2014).

[22] M. Stanke, E. Palikot, K. L. Sharkey, and L. Adamowicz, Benchmark calculations of the ^1D Rydberg spectrum of beryllium, Chem. Phys. Lett. **779**, 138823 (2021).

[23] M. Stanke, E. Palikot, K. L. Sharkey, and L. Adamowicz, Benchmark calculations of the ^2D Rydberg spectrum of lithium, Mol. Phys. **119**, e1925765 (2021).

[24] M. Puchalski, J. Komasa, and K. Pachucki, Fine and hyperfine splitting of the low-lying states of ^{9}Be, Phys. Rev. A **104**, 022824 (2021).

[25] T. Joyce and K. Varga, Matrix elements of explicitly correlated Gaussian basis functions with arbitrary angular momentum, J. Chem. Phys. **144**, 184106 (2016).

[26] T. Kato, On the eigenfunctions of many-particle systems in quantum mechanics, Commun. Pure Appl. Math. **10**, 151 (1957).

[27] J. R. Magnus and H. Neudecker, *Matrix-Differential Calculus with Applications in Statistics and Econometrics* (Wiley, Chichester, UK, 1988).

[28] M. Hamermesh, *Group Theory and Its Application to Physical Problems* (Addison-Wesley, Reading, MA, 1962).

[29] R. Pauncz, *Spin Eigenfunctions* (Plenum, New York, 1979).

[30] R. Pauncz, *The Symmetric Group in Quantum Chemistry* (CRC Press, Boca Raton, FL, 1995).

[31] K. L. Sharkey, S. Bubin, and L. Adamowicz, An algorithm for calculating atomic D states with explicitly correlated Gaussian functions, J. Chem. Phys. **134**, 044120 (2011).

[32] M. Wang, W. Huang, F. Kondev, G. Audi, and S. Naimi, The AME 2020 atomic mass evaluation (II). Tables, graphs and references, Chin. Phys. C **45**, 030003 (2021).

[33] L. M. Wang, Z.-C. Yan, H. X. Qiao, and G. W. F. Drake, Variational energies and the Fermi contact term for the low-lying states of lithium: Basis-set completeness, Phys. Rev. A **85**, 052513 (2012).

[34] K. Strasburger, Energy difference between the lowest doublet and quartet states of the boron atom, Phys. Rev. A **102**, 052806 (2020).

[35] K. Strasburger, Explicitly correlated wave functions of the ground state and the lowest quintuplet state of the carbon atom, Phys. Rev. A **99**, 052512 (2019).

[36] P. Maldonado, A. Sarsa, E. Buendía, and F. J. Gálvez, Jastrow correlated and quantum Monte Carlo calculations for the low-lying states of the carbon atom, J. Chem. Phys. **134**, 134102 (2011).