THE JOURNAL OF CHEMICAL PHYSICS **128**, 114107 (2008)

Energy and energy gradient matrix elements with *N*-particle explicitly correlated complex Gaussian basis functions with *L*=1

Sergiy Bubin[1,a)] and Ludwik Adamowicz[1,2]
[1] *Department of Chemistry, University of Arizona, Tucson, Arizona 85721, USA*
[2] *Department of Physics, University of Arizona, Tucson, Arizona 85721, USA*

(Received 31 January 2008; accepted 14 February 2008; published online 20 March 2008)

In this work we consider explicitly correlated complex Gaussian basis functions for expanding the wave function of an N-particle system with the $L=1$ total orbital angular momentum. We derive analytical expressions for various matrix elements with these basis functions including the overlap, kinetic energy, and potential energy (Coulomb interaction) matrix elements, as well as matrix elements of other quantities. The derivatives of the overlap, kinetic, and potential energy integrals with respect to the Gaussian exponential parameters are also derived and used to calculate the energy gradient. All the derivations are performed using the formalism of the matrix differential calculus that facilitates a way of expressing the integrals in an elegant matrix form, which is convenient for the theoretical analysis and the computer implementation. The new method is tested in calculations of two systems: the lowest P state of the beryllium atom and the bound P state of the positronium molecule (with the negative parity). Both calculations yielded new, lowest-to-date, variational upper bounds, while the number of basis functions used was significantly smaller than in previous studies. It was possible to accomplish this due to the use of the analytic energy gradient in the minimization of the variational energy. © *2008 American Institute of Physics*.
[DOI: 10.1063/1.2894866]

I. INTRODUCTION

Variational calculations on small atoms, molecules, and other bound quantum systems employing correlated basis functions explicitly dependent on the interparticle distances have proven to be a powerful tool for obtaining highly accurate solutions of the Schrödinger equation. Perhaps the most promising basis functions in such calculations are explicitly correlated Gaussians (ECGs). Functions of this type possess a very attractive feature—all the necessary matrix elements with them can be evaluated analytically for an arbitrary number of particles. In the case of other types of basis functions it often becomes difficult, if not impossible, to extend the number of particles beyond three or four. This unique property of the Gaussians has stimulated an interest in using them in very accurate atomic and molecular calculations.

In one of our previous works[1] we considered spherically symmetric ECGs with complex exponential parameters. As we showed, these types of functions can be used to calculate bound states with zero total orbital angular momentum of both atoms and molecules using an approach that does not assume the Born–Oppenheimer approximation regarding the separability of the nucleus-electron motion. The use of the complex exponential parameters in the Gaussians gives these basis functions the necessary flexibility in expanding the wave function that simultaneously depends on the coordinates of the electrons and the nuclei (nucleus in the atomic case). In a subsequent work[2] we showed that the complex Gaussians can be also employed in calculating the leading molecular relativistic corrections. In the variational optimization of the wave function in the applications presented in Refs. 1 and 2 we used the analytic gradient of the energy with respect to complex Gaussian exponential parameters. The formulas for calculating derivatives of the integrals involved in the energy expression and used to calculate the gradient were also presented in those works.

In this work we extend the approach of very accurate calculations of atomic and molecular systems with complex Gaussians to states with the $L=1$ total orbital angular momentum. Although atomic calculations with ECGs for states with nonzero total angular momentum have been done before by other authors (see, for example Refs. 3 and 4), neither of those works involved complex Gaussians nor the analytical gradient of the energy with respect to the Gaussian exponential parameters.

The presentation of this work includes the following. First we introduce the Hamiltonian, the notation used in the derivations, and the basis functions for representing the $L=1$ states. Next we explain the action of the permutation operators applied to the basis functions to achieve the correct permutational symmetry of the wave function. Then we present the derivation of the Hamiltonian and overlap matrix elements, as well as matrix elements for calculating some other quantities that are commonly computed in the calculations of small atoms and molecules. Next, we derive the matrix elements needed for computing the analytic gradient of the energy with respect to the Gaussian exponential parameters. In the last part of the work we show some test numerical examples.

[a)] Electronic mail: bubin@email.arizona.edu.

II. HAMILTONIAN

We consider a nonrelativistic Coulombic system consisting of N particles. Let $\mathbf{R}_i$ be the position vector of the ith particle in the laboratory Cartesian coordinate frame, M_i be its mass, and Q_i be its charge. The nonrelativistic Hamiltonian of the system is

$$\hat{H}_{\mathrm{lab}} = -\sum_{i=1}^{N} \frac{1}{2M_i}\nabla^2_{\mathbf{R}_i} + \sum_{i=1}^{N}\sum_{j>i}^{N} \frac{Q_iQ_j}{R_{ij}}. \tag{1}$$

Here $\nabla_{\mathbf{R}_i}$ is the gradient with respect to $\mathbf{R}_i$ and $R_{ij}=|\mathbf{R}_j-\mathbf{R}_i|$ is the distance between the ith and jth particle.

As we are interested in the bound states of the system, in the first step we need to separate out the motion of the center of mass from the laboratory frame Hamiltonian (1). In the approach we use for this separation, we place one particle in the origin of the new, internal, Cartesian coordinate system. This particle is called the reference particle. Then we refer the other particles to the reference particle using the relative coordinates $\mathbf{r}_i=\mathbf{R}_{i+1}-\mathbf{R}_1$. These coordinates, together with the three coordinates describing the position of the center of mass, $\mathbf{r}_0$, are our new coordinates. With the total mass of the system denoted as $M_{\mathrm{tot}}=\Sigma_{i=1}^{N}M_i$, the coordinate transform looks as follows:

$$\begin{aligned}
\mathbf{r}_0 &= \frac{M_1}{M_{\mathrm{tot}}}\mathbf{R}_1 + \frac{M_2}{M_{\mathrm{tot}}}\mathbf{R}_2 + \cdots + \frac{M_N}{M_{\mathrm{tot}}}\mathbf{R}_N,\\
\mathbf{r}_1 &= -\mathbf{R}_1 + \mathbf{R}_2,\\
\mathbf{r}_2 &= -\mathbf{R}_1 + \mathbf{R}_3,\\
&\vdots\\
\mathbf{r}_n &= -\mathbf{R}_1 + \mathbf{R}_N,
\end{aligned} \tag{2}$$

while the inverse coordinate transformation is

$$\begin{aligned}
\mathbf{R}_1 &= \mathbf{r}_0 - \frac{M_2}{M_{\mathrm{tot}}}\mathbf{r}_1 - \frac{M_3}{M_{\mathrm{tot}}}\mathbf{r}_2 - \cdots - \frac{M_N}{M_{\mathrm{tot}}}\mathbf{r}_n,\\
\mathbf{R}_2 &= \mathbf{r}_0 + \left(1-\frac{M_2}{M_{\mathrm{tot}}}\right)\mathbf{r}_1 - \frac{M_3}{M_{\mathrm{tot}}}\mathbf{r}_2 - \cdots - \frac{M_N}{M_{\mathrm{tot}}}\mathbf{r}_n,\\
\mathbf{R}_3 &= \mathbf{r}_0 - \frac{M_2}{M_{\mathrm{tot}}}\mathbf{r}_1 + \left(1-\frac{M_3}{M_{\mathrm{tot}}}\right)\mathbf{r}_2 - \cdots - \frac{M_N}{M_{\mathrm{tot}}}\mathbf{r}_n,\\
&\vdots\\
\mathbf{R}_N &= \mathbf{r}_0 - \frac{M_2}{M_{\mathrm{tot}}}\mathbf{r}_1 - \frac{M_3}{M_{\mathrm{tot}}}\mathbf{r}_2 - \cdots + \left(1-\frac{M_N}{M_{\mathrm{tot}}}\right)\mathbf{r}_n.
\end{aligned} \tag{3}$$

Upon the transformation (2), the laboratory frame Hamiltonian (1) separates into a Hamiltonian describing the motion of the center of mass of the system and the following "internal" Hamiltonian:

$$\hat{H} = -\frac{1}{2}\left(\sum_{i=1}^{n}\frac{1}{m_i}\nabla^2_{\mathbf{r}_i} + \sum_{i\neq j}^{n}\frac{1}{m_0}\nabla'_{\mathbf{r}_i}\nabla_{\mathbf{r}_j}\right) + \sum_{i=1}^{n}\frac{q_0q_i}{r_i} + \sum_{i<j}^{n}\frac{q_iq_j}{r_{ij}}, \tag{4}$$

where $n=N-1$, the prime symbol denotes the matrix/vector transposition, $r_{ij}=|\mathbf{r}_j-\mathbf{r}_i|$, $m_i=M_{i+1}$, $q_i=Q_{i+1}$, and $\mu_i=m_0m_i/(m_0+m_i)$. The Hamiltonian (4) describes the motion of n pseudoparticles with masses m_i and charges q_i in the central field of the reference particle. The motions of the pseudoparticles are coupled through the mass polarization term $\Sigma_{i\neq j}^{n}(1/m_0)\nabla'_{\mathbf{r}_i}\nabla_{\mathbf{r}_j}$ and through the Coulombic interactions dependent on the distances between the pseudoparticles and the origin of the internal coordinate system, r_i, and on the relative distances between the pseudoparticles, r_{ij}.

For convenience, the Hamiltonian (4) can be rewritten in the matrix form. To do that we combine the coordinates of the pseudoparticle positions and the corresponding gradients into two $3n$-component column vectors,

$$\mathbf{r} = \begin{pmatrix}\mathbf{r}_1\\ \mathbf{r}_2\\ \vdots\\ \mathbf{r}_n\end{pmatrix}, \quad \nabla_{\mathbf{r}} = \begin{pmatrix}\nabla_{\mathbf{r}_1}\\ \nabla_{\mathbf{r}_2}\\ \vdots\\ \nabla_{\mathbf{r}_n}\end{pmatrix}. \tag{5}$$

With that we have

$$\hat{H} = -\nabla'_{\mathbf{r}}\mathbf{M}\nabla_{\mathbf{r}} + \sum_{i=1}^{n}\frac{q_0q_i}{r_i} + \sum_{i<j}^{n}\frac{q_iq_j}{r_{ij}}. \tag{6}$$

Here $\mathbf{M}=M\otimes I_3$ is the Kronecker product of the $n\times n$ matrix M and the 3×3 identity matrix I_3. The diagonal elements of matrix M are $1/(2m_1), 1/(2m_2), \ldots, 1/(2m_n)$, while all off-diagonal elements are equal to $1/m_0$.

III. THE NOTATION

Throughout this work we express the formulas in the matrix form that has the advantage in terms of simplicity and convenience for the theoretical analysis and the computer implementation. The derivations in the matrix form are based on the formalism of the matrix differential calculus. A detailed introduction to this subject can be found, for example, in Ref. 5 The notation we use is similar to that used in our previous works,[1,6–8] although due to the peculiarities of the complex Gaussian basis functions used in this work, the notation had to be modified somewhat.

The problem of N particles, or, as described by the internal Hamiltonian (4), the problem of n pseudoparticles moving in the three-dimensional (3D) space involves dealing with two kinds of vectors and matrices. Vectors of the first kind have n components (n^2 for matrices) and each of them represents a quantity associated with a certain particle. The second kind of vectors and matrices represent quantities in the 3D space. These two kinds of vectors and matrices need to be clearly distinguished in our notation. Moreover, we often have to deal with direct products of the n-dimensional particle space and the 3D space, which result in the appearance of Kronecker products of $n\times n$ and 3×3 matrices. In addition, the derivations also involve matrices whose size is

$K\times K$, where K is the length of the the expansion of the wave function in terms of the basis functions. For clarity we adopt the following notation rules:

- α, β, ξ, etc.—lower-case Greek letters are used for scalars.
- a, b, x, etc.—lower-case Latin letters are used for column vectors in the particle space. These vectors have n components.
- $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$, $\boldsymbol{\xi}$, etc.—bold font in the lower-case Greek letters is used to denote vectors in the 3D space. These have three components.
- $\mathbf{a}$, $\mathbf{b}$, $\mathbf{x}$, etc.—bold font in the lower-case Latin letters is used to denote vectors in the $n\times 3$ space. These have $3n$ components. For example, $\mathbf{r}'=(r_{1x},r_{1y},r_{1z},r_{2x},\ldots,r_{nz})$.
- Λ, Ω, Ξ, etc.—upper-case Greek letters are used for matrices in the 3D space. These are 3×3 matrices.
- A, B, X, etc.—upper-case Latin letters are used for matrices in the particle space. These are $n\times n$ matrices.
- $\mathbf{A}$, $\mathbf{B}$, $\mathbf{X}$, etc.—bold font in upper-case Latin letters is used to denote matrices in $3n$-dimensional space. These are $3n\times 3n$ matrices.
- $A\otimes\Lambda$, $x\otimes\xi$, etc.—$\otimes$ stands for the Kronecker product of two matrices (this also applies to vectors, as they are considered to be rectangular matrices).
- A', $\mathbf{A}'$, a', α' etc.—the prime stands for matrix or vector transpose.
- A^*, $\mathbf{A}^*$, a^*, α^* etc.—the star stands for the complex conjugate. If applied to a matrix (vector) then there is no transposition involved; i.e., the result is a matrix (vector) whose elements are complex conjugates of the elements of the original matrix (vector).
- $A^\dagger$, $\mathbf{A}^\dagger$, $a^\dagger$, $\alpha^\dagger$ etc.—the dagger stands for matrix (vector) transpose followed by the complex conjugation.
- A^{-1}, $\mathbf{A}^{-1}$, Λ^{-1}—stands for the inverse of the matrix.
- $\operatorname{diag} A$, $\operatorname{diag} \mathbf{A}$, $\operatorname{diag} \Lambda$—the action of the diag operation results in all off-diagonal elements of a matrix set to zero, while the diagonal elements remain unchanged.
- $|A|$, $|\mathbf{A}|$, $|\Lambda|$—vertical bars stand for the determinant of the matrix. However, if the object in between the vertical bars is a vector or a scalar then the bars denote the absolute value of the vector (scalar).
- $\operatorname{tr} A$, $\operatorname{tr} \mathbf{A}$,$\operatorname{tr} \Lambda$—tr stands for the trace of the matrix.
- I_3, I_n—letter I is only used for identity matrices, so that I_3 is the 3×3 identity matrix, I_n is the $n\times n$ identity matrix, etc.
- A, B, X, etc.—upper-case Latin letters written with the sans-serif font are used for matrices in the space of the basis functions. These are $K\times K$ matrices, where K is the length of the the variational expansion (basis size). For example, S_{kl} denotes the overlap matrix element between the kth and lth basis functions.
- a, b, x, etc.—lower-case Latin letters written with the sans-serif font are used for vectors in the space of the basis functions. These vectors have K components.
- The common rules of the matrix multiplication are assumed. For example, the AB product of two matrices is an $n\times n$ matrix, the $a'b$ product of two vectors is a scalar, the ab' product of two vectors is an $n\times n$ matrix, etc.

The above notation convention applies to all formulas presented in this work. There are, however, a few exceptions, where we kept the traditionally used symbols for some quantities. These few exceptions are obvious and should not cause any confusion to the reader.

In the section of this work that deals with the energy derivatives with respect to the exponential parameters of the basis functions (the energy gradient) we use two operators, vec and vech. They both transform a matrix into a vector. The first operator, vec, stacks the columns of a matrix one underneath the other. Thus, it transforms a $n\times n$ matrix into an n^2-component vector. For example, if X is a 2×2 matrix with elements X_{ij}, then $\operatorname{vec} X$ is the following four-component vector:

$$\operatorname{vec} X=\begin{pmatrix} X_{11}\\ X_{12}\\ X_{21}\\ X_{22}\end{pmatrix}. \tag{7}$$

The second operator, vech (its abbreviature comes from "vector half"), also stacks the columns of a matrix, one underneath the other, but for each column only the diagonal element of the matrix present in that column and all the elements positioned below this diagonal element are used in the stacking. Hence, vech transforms an $n\times n$ matrix into a $n(n+1)/2$-component vector. For example, if X is a 3×3 matrix with elements X_{ij}, then

$$\operatorname{vech} X=\begin{pmatrix} X_{11}\\ X_{12}\\ X_{13}\\ X_{22}\\ X_{23}\\ X_{33}\end{pmatrix}. \tag{8}$$

The vech operator is used for symmetric matrices; in this case $\operatorname{vech} X$ contains only independent elements of X.

Finally, let us list a few useful properties of the vec and vech operators. If X and Y are arbitrary complex square matrices, and L is a complex lower triangular matrix then

$$(\operatorname{vec} X)'\operatorname{vec} Y=\operatorname{tr}[X'Y], \tag{9}$$

$$(\mathrm{vec}\,X)'\mathrm{vec}\,L = (\mathrm{vech}\,X)'\mathrm{vech}\,L. \tag{10}$$

In the case when X is an arbitrary complex matrix and Y is a symmetric complex matrix we have the following relationship:

$$(\mathrm{vec}\,X)'\mathrm{vec}\,Y = (\mathrm{vech}(X + X' - \mathrm{diag}\,X))'\mathrm{vech}\,Y. \tag{11}$$

IV. BASIS FUNCTIONS

The general form of the basis functions for describing the $L=1$ states used in this work is

$$\phi_k = z_{m_k}\exp[-\mathbf{r}'\mathbf{C}_k\mathbf{r}] \equiv z_{m_k}\exp[-\mathbf{r}'((A_k + iB_k)\otimes I_3)\mathbf{r}]. \tag{12}$$

Here m_k is an integer that depends on k with values from 1 to n. A_k and B_k are $n\times n$ symmetric matrices that represent the real and imaginary parts of C_k, respectively. The subscript k in each of these matrices reflects the fact that they are unique for each basis function. In general, neither A_k nor B_k has to be symmetric, but one can always rearrange their elements in such a way that they become symmetric without changing the quadratic form $\mathbf{r}'\mathbf{C}_k\mathbf{r}$. Since it is more convenient to deal with symmetric matrices, we assume in this work that A_k and B_k are symmetric.

Any basis function used in a bound state calculation must be square integrable. This effectively imposes a restriction on the elements of A_k (but not B_k). Matrix A_k must be positive definite. To assure positive definiteness, we represent A_k in the form of the Cholesky decomposition, $A_k = L_kL_k'$, where L_k is a triangular matrix (the so-called Cholesky factor). As any positive definite matrix can be represented in the Cholesky-factored form, such a representation automatically makes A_k positive definite. There is a significant practical advantage of using the Cholesky-factored form for A_k. That rests in the fact that there is no need to impose any restrictions on the elements of L_k. Thus, in the variational optimization of the basis functions, the elements of L_k can be varied with no restrictions in the range $[-\infty,+\infty]$. Thus, in the variational optimization of the L_k matrices we can use very fast and efficient algorithms designed for the unconstrained optimization. This would not be the case if the variational parameters were the elements of matrix A_k.

In the derivations we will often use the following alternative representation of the basis functions (12):

$$\begin{aligned}\phi_k &= \frac{\partial}{\partial\alpha_k}\exp[-\mathbf{r}'\mathbf{C}_k\mathbf{r} + \alpha_k z_{m_k}]\Big|_{\alpha_k=0}\\ &= \frac{\partial}{\partial\alpha_k}\exp[-\mathbf{r}'\mathbf{C}_k\mathbf{r} + \alpha_k(\mathbf{v}^k)'\mathbf{r}]\Big|_{\alpha_k=0},\end{aligned} \tag{13}$$

where α_k is a parameter and $\mathbf{v}^k$ is a vector whose all components are zeros, except the $3m_k$ component, which is set to one. For example, in the case when $n=2$, and $m_k=1$, the six-component vector $\mathbf{v}^k$ is

$$\mathbf{v}^k = \begin{pmatrix}0\\0\\1\\0\\0\\0\end{pmatrix}.$$

Furthermore, it is convenient to define the following "generator" function for ϕ_k in Eq. (13):

$$\varphi_k = \exp[-\mathbf{r}'\mathbf{C}_k\mathbf{r} + \alpha_k(\mathbf{v}^k)'\mathbf{r}]. \tag{14}$$

V. PERMUTATIONAL SYMMETRY

The trial wave function, and consequently the basis functions used in the wave function expansion, should possess a certain symmetry with respect to the permutations of identical particles involved in the system. This symmetry depends on the nature of the particles (i.e., whether or not they are bosons or fermions). When the Hamiltonian of the system is spin independent, it is possible to completely eliminate the spin from the consideration and to only deal with the spatial parts of the basis functions. The spin elimination often leads to nontrivial permutational properties of the spatial parts. These properties and their implementation can be handled by applying certain symmetry projectors, which are linear combinations of permutational operators. For a specific state of the system an appropriate symmetry operator can be defined and its application to the basis functions [such as Eq. (12)] generates functions that possess the right symmetry for that state. A symmetry operator is a sum of elemental operators, $\hat{P}_\alpha$, which are products of some transpositions (i.e., pair permutations). We do not intend to discuss here the ways of building the symmetry projectors. Such methods have been developed and are well known (see, for example, Ref. 9). Here we only wish to discuss the action of the permutational operators on the basis functions (12) and (14), as well as some details related to the implementation of the permutational symmetry.

A permutation of the real particles (i.e., not the pseudoparticles) involved in a permutational operator $\hat{P}$ can be represented as a linear transformation of the laboratory frame coordinates $\mathbf{R}$ of the particles. Since the relation between the laboratory coordinates $\mathbf{R}$ and the internal coordinates $\mathbf{r}$ is linear, the transformation of the internal coordinates under the permutation of the particles is also linear. Hence, it can be described by a permutation matrix, $\mathbf{P} = P\otimes I_3$. The application of $\hat{P}$ to the basis functions (12) gives

$$\begin{aligned}\hat{P}\phi_l &= \hat{P}\frac{\partial}{\partial\alpha_l}\exp[-\mathbf{r}'\mathbf{C}_l\mathbf{r} + \alpha_l(\mathbf{v}^l)'\mathbf{r}]\Big|_{\alpha_l=0}\\ &= \frac{\partial}{\partial\alpha_l}\exp[-(\mathbf{P}\mathbf{r})'\mathbf{C}_l(\mathbf{P}\mathbf{r}) + \alpha_l(\mathbf{v}^l)'(\mathbf{P}\mathbf{r})]\Big|_{\alpha_l=0}\\ &= \frac{\partial}{\partial\alpha_l}\exp[-\mathbf{r}'(\mathbf{P}'\mathbf{C}_l\mathbf{P})\mathbf{r} + \alpha_l(\mathbf{P}'\mathbf{v}^l)'\mathbf{r}]\Big|_{\alpha_l=0}.\end{aligned} \tag{15}$$

As one can see from Eq. (15), the symmetry transformation of ϕ_l is equivalent to a similarity transformation of the matrix of the nonlinear parameters for that basis function ($\mathbf{A}_l \rightarrow \mathbf{P}'\mathbf{A}_l\mathbf{P}$, $\mathbf{B}_l \rightarrow \mathbf{P}'\mathbf{B}_l\mathbf{P}$) and a linear transformation of the corresponding vectors, $\mathbf{v}^l \rightarrow \mathbf{P}'\mathbf{v}^l$. Based on this, a procedure for implementing the symmetry in calculating matrix elements with symmetry-projected basis functions can be developed. When evaluating these elements, one simply needs to compute a sum of elemental matrix elements, each of which being determined with kth and lth basis functions whose matrices of the nonlinear parameters, $\mathbf{A}_k$ and $\mathbf{A}_l$, and the corresponding vectors $\mathbf{v}^l$ were transformed according to Eq. (15).

Furthermore, the operators whose matrix elements are needed in the variational calculation (such as the Hamiltonian, etc.) usually commute with all the permutation operators involved in the symmetry projector. Due to this, the evaluation of the matrix elements of such an operator can be further simplified because the number of terms to calculate can be usually reduced from g^2 to g (where g is the number of terms in the symmetry projector, i.e., the elemental permutational operators). Moreover, due to the hermicity of the operator, one can always restrict the calculation to only those terms that result from applying the permutational operators to the *ket* because

$$\langle \hat{P}_\beta \phi_k | \hat{O} | \hat{P}_\gamma \phi_l \rangle = \langle \phi_k | \hat{O} | \hat{P}_\beta^\dagger \hat{P}_\gamma \phi_l \rangle = \langle \phi_k | \hat{O} | \hat{P}_\delta \phi_l \rangle. \tag{16}$$

It is often important (particularly when computing the derivatives of the Hamiltonian and overlap matrix elements with respect to nonlinear parameters) to keep track of whether the matrices and the vectors of the nonlinear parameters ($\mathbf{A}_l$, $\mathbf{B}_l$, and $\mathbf{v}^l$) in a particular component of the matrix element were transformed or not transformed with a permutation symmetry operator. For this purpose we use the tilde symbol to denote the transformed objects. For example, when a basis function in the *ket* is transformed by a permutational operator we write

$$|\tilde{\phi}_l\rangle \equiv \hat{P}|\phi_l\rangle. \tag{17}$$

The matrices and the vector of the nonlinear parameters of the basis function $\tilde{\phi}_l$ are

$$\tilde{\mathbf{A}}_l \equiv \mathbf{P}'\mathbf{A}_l\mathbf{P}, \quad \tilde{\mathbf{B}}_l \equiv \mathbf{P}'\mathbf{B}_l\mathbf{P}, \quad \tilde{\mathbf{L}}_l \equiv \mathbf{P}'\mathbf{L}_l\mathbf{P}, \quad \tilde{\mathbf{v}}^l \equiv \mathbf{P}'\mathbf{v}^l. \tag{18}$$

Since we often need to add matrices of the *bra* basis function (the kth basis function) and to the corresponding matrices of the *ket* basis function (the lth basis function), let us also define the following:

$$\mathbf{A}_{kl} \equiv \mathbf{A}_k + \mathbf{A}_l, \quad \tilde{\mathbf{A}}_{kl} \equiv \mathbf{A}_k + \tilde{\mathbf{A}}_l, \tag{19}$$

$$\mathbf{B}_{kl} \equiv -\mathbf{B}_k + \mathbf{B}_l, \quad \tilde{\mathbf{B}}_{kl} \equiv -\mathbf{B}_k + \tilde{\mathbf{B}}_l, \tag{20}$$

$$\tilde{\mathbf{C}}_{kl} \equiv \mathbf{C}_k^* + \tilde{\mathbf{C}}_l = \tilde{\mathbf{A}}_{kl} + i\tilde{\mathbf{B}}_{kl}. \tag{21}$$

VI. MATRIX ELEMENTS

Two short general comments regarding the evaluation of matrix elements need to be made. First, the matrix elements of some operator $\hat{O}$ with the basis functions (12) can, in most cases, be evaluated as derivatives of the corresponding matrix elements calculated with the generator functions (14), i.e.,

$$\langle \phi_k | \hat{O} | \tilde{\phi}_l \rangle = \left. \frac{\partial}{\partial \alpha_k} \frac{\partial}{\partial \alpha_l} \langle \varphi_k | \hat{O} | \tilde{\varphi}_l \rangle \right|_{\alpha_k, \alpha_l = 0}. \tag{22}$$

This simple relation is frequently used in this work. Second, when evaluating matrix elements with Gaussian basis functions one often encounters the following general p-dimensional integral:

$$\int_{-\infty}^{+\infty} \exp[-x'Cx + y'x]\mathrm{d}x = \frac{\pi^{p/2}}{|C|^{1/2}} \exp\left[\frac{1}{4} y' C^{-1} y\right], \tag{23}$$

where x is a p-component vector of variables, y is a p-component vector of constant values, and C is a symmetric $p \times p$ matrix with a positive definite real part.

A. Overlap integral

We first evaluate the overlap integral between two generator functions (14). This can be easily done by directly applying formula (23),

$$\begin{aligned} \langle \varphi_k | \tilde{\varphi}_l \rangle &= \langle \varphi_k | \hat{P} | \varphi_l \rangle \\ &= \int \exp[-\mathbf{r}'\mathbf{C}_k^*\mathbf{r} + \alpha_k \mathbf{v}^{k\prime}\mathbf{r}]\hat{P} \\ &\quad \times \exp[-\mathbf{r}'\mathbf{C}_l\mathbf{r} + \alpha_k \mathbf{v}^{l\prime}\mathbf{r}]\mathrm{d}\mathbf{r} \\ &= \int \exp[-\mathbf{r}'\tilde{\mathbf{C}}_{kl}\mathbf{r} + (\alpha_k\mathbf{v}^k + \alpha_l\tilde{\mathbf{v}}^l)'\mathbf{r}]\mathrm{d}\mathbf{r} \\ &= \frac{\pi^{3n/2}}{|\tilde{\mathbf{C}}_{kl}|^{1/2}} \exp\left[\frac{1}{4}(\alpha_k\mathbf{v}^k + \alpha_l\tilde{\mathbf{v}}^l)'\tilde{\mathbf{C}}_{kl}^{-1}(\alpha_k\mathbf{v}^k + \alpha_l\tilde{\mathbf{v}}^l)\right]. \end{aligned} \tag{24}$$

It is convenient to introduce an n-component vector v^k that has only one nonzero element in the m_kth position. This nonzero element is equal to 1. v^k is related to $\mathbf{v}^k$ by

$$\mathbf{v}^k = v^k \otimes \epsilon^z, \tag{25}$$

where

$$\epsilon^z = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}. \tag{26}$$

With this we can reduce the size of the matrices and the vectors in Eq. (24) from $3n$ to n,

$$\langle \varphi_k | \tilde{\varphi}_l \rangle = \frac{\pi^{3n/2}}{|\tilde{C}_{kl}|^{3/2}} \exp\left[\frac{1}{4} (\alpha_k v^k + \alpha_l \tilde{v}^l)' \tilde{C}_{kl}^{-1} (\alpha_k v^k + \alpha_l \tilde{v}^l) \right]. \tag{27}$$

In the above equation we used the fact that $\tilde{\mathbf{C}}_{kl}^{-1} \equiv (\tilde{C}_{kl} \otimes I_3)^{-1} = \tilde{C}_{kl}^{-1} \otimes I_3$ and $|\tilde{\mathbf{C}}_{kl}| \equiv |\tilde{C}_{kl} \otimes I_3| = |\tilde{C}_{kl}|^3$. The overlap integral between two basis functions, ϕ_k and ϕ_l, can now be obtained by a simple differentiation,

$$\langle \phi_k | \tilde{\phi}_l \rangle = \left. \frac{\partial}{\partial \alpha_k} \frac{\partial}{\partial \alpha_l} \langle \varphi_k | \tilde{\varphi}_l \rangle \right|_{\alpha_k, \alpha_l = 0} = \frac{\pi^{3n/2}}{2} \frac{v^{k\prime} \tilde{C}_{kl}^{-1} \tilde{v}^l}{|\tilde{C}_{kl}|^{3/2}}. \tag{28}$$

Since in most situations matrix elements with normalized basis functions are preferable we will present our final results with such functions. The normalized overlap matrix element is

$$\begin{aligned} \mathsf{S}_{kl} &\equiv \frac{\langle \phi_k | \tilde{\phi}_l \rangle}{(\langle \phi_k | \phi_k \rangle \langle \phi_l | \phi_l \rangle)^{1/2}} \\ &= \frac{(|C_{kk}|^{3/2} |C_{ll}|^{3/2})^{1/2}}{|\tilde{C}_{kl}|^{3/2}} \frac{v^{k\prime} \tilde{C}_{kl}^{-1} \tilde{v}^l}{(v^{k\prime} C_{kk}^{-1} v^k v^{l\prime} C_{ll}^{-1} v^l)^{1/2}} \\ &= 2^{3n/2} \left(\frac{\big| |L_k| \big| \big| |L_l| \big|}{|\tilde{C}_{kl}|} \right)^{3/2} \frac{v^{k\prime} \tilde{C}_{kl}^{-1} \tilde{v}^l}{(v^{k\prime} C_{kk}^{-1} v^k v^{l\prime} C_{ll}^{-1} v^l)^{1/2}}. \end{aligned} \tag{29}$$

B. Kinetic energy integral

Before we proceed to evaluating the kinetic energy matrix element with the basis functions (12), we need to evaluate the following auxiliary integral: $\langle \phi_k | \mathbf{r}' \mathbf{W} \mathbf{r} | \tilde{\phi}_l \rangle$, where $\mathbf{W} = W \otimes I_3$ is a complex symmetric matrix. Notice that the requirement of the symmetrical character of matrix $\mathbf{W}$ is not necessary. We use it here to simplify the derivation. By applying the transformation $\mathbf{W} \leftarrow (\mathbf{W} + \mathbf{W}')/2$ matrix $\mathbf{W}$ can always be made symmetric without changing the quadratic form $\mathbf{r}' \mathbf{W} \mathbf{r}$. The above integral can be calculated as follows:

$$\begin{aligned} \langle \phi_k | \mathbf{r}' \mathbf{W} \mathbf{r} | \tilde{\phi}_l \rangle &= \left. - \frac{\partial}{\partial \beta} \langle \phi_k | \exp[-\beta \mathbf{r}' \mathbf{W} \mathbf{r}] | \tilde{\phi}_l \rangle \right|_{\beta = 0} \\ &= \left. - \frac{\partial}{\partial \beta} \frac{\pi^{3n/2}}{2} \frac{v^{k\prime} (\tilde{C}_{kl} + \beta W)^{-1} \tilde{v}^l}{|\tilde{C}_{kl} + \beta W|^{3/2}} \right|_{\beta = 0}. \end{aligned} \tag{30}$$

Here we need to recall two relations from the matrix differential calculus. If X is an arbitrary matrix then the differential of its determinant and its inverse are given by

$$\mathrm{d}|X| = |X| \mathrm{tr}[X^{-1} \mathrm{d}X], \tag{31}$$

$$\mathrm{d}(X^{-1}) = - X^{-1} (\mathrm{d}X) X^{-1}. \tag{32}$$

By applying Eqs. (31) and (32) to (30) we get

$$\begin{aligned} \langle \phi_k | \mathbf{r}' \mathbf{W} \mathbf{r} | \tilde{\phi}_l \rangle &= \frac{\pi^{3n/2}}{2} \frac{1}{|\tilde{C}_{kl}|^{3/2}} \left(\frac{3}{2} \mathrm{tr}[\tilde{C}_{kl}^{-1} W] v^{k\prime} \tilde{C}_{kl}^{-1} \tilde{v}^l \right. \\ &\quad \left. + v^{k\prime} \tilde{C}_{kl}^{-1} W \tilde{C}_{kl}^{-1} \tilde{v}^l \right) \\ &= \langle \phi_k | \tilde{\phi}_l \rangle \left(\frac{3}{2} \mathrm{tr}[\tilde{C}_{kl}^{-1} W] + \frac{v^{k\prime} \tilde{C}_{kl}^{-1} W \tilde{C}_{kl}^{-1} \tilde{v}^l}{v^{k\prime} \tilde{C}_{kl}^{-1} \tilde{v}^l} \right). \end{aligned} \tag{33}$$

Now, when the $\nabla_{\mathbf{r}}$ operator acts on function (14) we get

$$\nabla_{\mathbf{r}} \varphi_k = (-2 \mathbf{C}_k \mathbf{r} + \alpha_k \mathbf{v}^k) \varphi_k. \tag{34}$$

Using this relation, the matrix element of the kinetic energy operator with functions (14) can be written as

$$\begin{aligned} \langle \varphi_k | - \nabla_{\mathbf{r}}' \mathbf{M} \nabla_{\mathbf{r}} | \tilde{\varphi}_l \rangle &= \langle (-2 \mathbf{C}_k \mathbf{r} + \alpha_k \mathbf{v}^k) \varphi_k | \mathbf{M} | (-2 \tilde{\mathbf{C}}_l \mathbf{r} + \alpha_l \tilde{\mathbf{v}}^l) \tilde{\varphi}_l \rangle \\ &= 4 \langle \varphi_k | \mathbf{r}' \mathbf{C}_k^\dagger \mathbf{M} \tilde{\mathbf{C}}_l \mathbf{r} | \tilde{\varphi}_l \rangle \\ &\quad - 2 \alpha_l \langle \varphi_k | \mathbf{r}' \mathbf{C}_k^\dagger \mathbf{M} \tilde{\mathbf{v}}^l | \tilde{\varphi}_l \rangle \\ &\quad - 2 \alpha_k \langle \varphi_k | \mathbf{v}^{k\prime} \mathbf{M} \tilde{\mathbf{C}}_l \mathbf{r} | \tilde{\varphi}_l \rangle \\ &\quad + \alpha_k \alpha_l \mathbf{v}^{k\prime} \mathbf{M} \tilde{\mathbf{v}}^l \langle \varphi_k | \tilde{\varphi}_l \rangle. \end{aligned} \tag{35}$$

To get the kinetic energy matrix element with basis functions (12) from the matrix element with the generator functions (14) we differentiate the last expression with respect to the parameters α_k and α_l and set both parameters to zero at the end. The first term on the right-hand side of Eq. (35) for functions (12) is equal to $4 \langle \phi_k | \mathbf{r}' \mathbf{C}_k^\dagger \mathbf{M} \tilde{\mathbf{C}}_l \mathbf{r} | \tilde{\phi}_l \rangle$ or $2 \langle \phi_k | \mathbf{r}' (\mathbf{C}_k^\dagger \mathbf{M} \tilde{\mathbf{C}}_l + \tilde{\mathbf{C}}_l \mathbf{M} \mathbf{C}_k^\dagger) \mathbf{r} | \tilde{\phi}_l \rangle$. The latter expression contains a symmetric matrix in parentheses and can be evaluated using formula (33). The second and third terms in Eq. (35) are $-2 \langle \varphi_k | (\mathbf{v}^{k\prime} \mathbf{r}) (\mathbf{r}' \mathbf{C}_k^\dagger \mathbf{M} \tilde{\mathbf{v}}^l) | \tilde{\varphi}_l \rangle$ and $-2 \langle \varphi_k | (\mathbf{v}^{k\prime} \mathbf{M} \tilde{\mathbf{C}}_l \mathbf{r}) \times (\tilde{\mathbf{v}}^{l\prime} \mathbf{r}) | \tilde{\varphi}_l \rangle$ (in both α_k and α_l are set to zero). These are very similar to the overlap integral $\langle \phi_k | \tilde{\phi}_l \rangle$. The difference is only the fact that in the second term we need to replace vector $\mathbf{v}^l$ with $\mathbf{C}_k^\dagger \mathbf{M} \tilde{\mathbf{v}}^l$, and in the third term vector $(\mathbf{v}^k)'$ with $\mathbf{v}^{k\prime} \mathbf{M} \mathbf{C}_l$. The fourth term on the right hand side of Eq. (35) is $\mathbf{v}^{k\prime} \mathbf{M} \tilde{\mathbf{v}}^l \langle \varphi_k | \tilde{\varphi}_l \rangle$ after setting $\alpha_k = \alpha_l = 0$. Now, putting all the terms together we get

$$\begin{aligned} \langle \phi_k | - \nabla_{\mathbf{r}}' \mathbf{M} \nabla_{\mathbf{r}} | \tilde{\phi}_l \rangle &= 4 \langle \phi_k | \tilde{\phi}_l \rangle \left(\frac{3}{4} \mathrm{tr}[\tilde{C}_{kl}^{-1} (C_k^\dagger M \tilde{C}_l + \tilde{C}_l M C_k^\dagger)] \right. \\ &\quad \left. + \frac{1}{2} \frac{v^{k\prime} \tilde{C}_{kl}^{-1} (C_k^\dagger M \tilde{C}_l + \tilde{C}_l M C_k^\dagger) \tilde{C}_{kl}^{-1} \tilde{v}^l}{v^{k\prime} \tilde{C}_{kl}^{-1} \tilde{v}^l} \right) \\ &\quad - \pi^{3n/2} \frac{v^{k\prime} \tilde{C}_{kl}^{-1} C_k^\dagger M \tilde{v}^l}{|\tilde{C}_{kl}|^{3/2}} \\ &\quad - \pi^{3n/2} \frac{v^{k\prime} M \tilde{C}_l \tilde{C}_{kl}^{-1} \tilde{v}^l}{|\tilde{C}_{kl}|^{3/2}} + \pi^{3n/2} \frac{v^{k\prime} M \tilde{v}^l}{|\tilde{C}_{kl}|^{3/2}}. \end{aligned} \tag{36}$$

After some rearrangement we obtain the following final formula for the matrix element of the kinetic energy operator with the normalized basis functions

$$\mathsf{T}_{kl}=\mathsf{S}_{kl}\left(6\,\mathrm{tr}[\tilde{C}_{kl}^{-1}C_lM\tilde{C}_k^{\dagger}]+4\frac{\upsilon^{k\prime}\tilde{C}_{kl}^{-1}\tilde{C}_lMC_k^{\dagger}\tilde{C}_{kl}^{-1}\tilde{\upsilon}^l}{\upsilon^{k\prime}\tilde{C}_{kl}^{-1}\tilde{\upsilon}^l}\right).\tag{37}$$

C. Potential energy integral

Let us begin with defining the matrix $\mathbf{J}_{ij}=J_{ij}\otimes I_3$, which will be useful in the further derivations. The matrix is

$$J_{ij}=\begin{cases}E_{ii} & \text{if } i=j\\ E_{ii}+E_{jj}-E_{ij}-E_{ji} & \text{if } i\neq j,\end{cases}\tag{38}$$

where E_{ij} is a matrix with 1 in the i,jth position and 0's elsewhere. Using $\mathbf{J}_{ij}$ we can now write the squares of the interparticle distances as the following quadratic forms:

$$r_i^2=\mathbf{r}'\mathbf{J}_{ii}\mathbf{r},$$

$$r_{ij}^2=\mathbf{r}'\mathbf{J}_{ij}\mathbf{r}.$$

In order to evaluate the matrix elements of the potential energy operator, we need to determine the following Coulomb integrals $\langle\phi_k|1/r_i|\tilde{\phi}_l\rangle$ and $\langle\phi_k|1/r_{ij}|\tilde{\phi}_l\rangle$. To do this we use the following transformation:

$$\left\langle\phi_k\left|\frac{1}{r_{ij}}\right|\tilde{\phi}_l\right\rangle=\frac{2}{\sqrt{\pi}}\int_0^{\infty}\langle\phi_k|\exp[-\beta^2\mathbf{r}'\mathbf{J}_{ij}\mathbf{r}]|\tilde{\phi}_l\rangle\mathrm{d}\beta$$

$$=\pi^{(3n-1)/2}\int_0^{\infty}\frac{\upsilon^{k\prime}(\tilde{C}_{kl}+\beta^2J_{ij})^{-1}\tilde{\upsilon}^l}{|\tilde{C}_{kl}+\beta^2J_{ij}|^{3/2}}\mathrm{d}\beta.\tag{39}$$

Notice that the vector-matrix-vector product in the above expression can be written as the following trace:

$$\upsilon^{k\prime}(\tilde{C}_{kl}+\beta^2J_{ij})^{-1}\tilde{\upsilon}^l=\mathrm{tr}[(\tilde{C}_{kl}+\beta^2J_{ij})^{-1}\tilde{\upsilon}^l\upsilon^{k\prime}].\tag{40}$$

Moreover, it is easy to show that the following equality holds true:

$$\frac{\mathrm{tr}[(\tilde{C}_{kl}+\beta^2J_{ij})^{-1}\tilde{\upsilon}^l\upsilon^{k\prime}]}{|\tilde{C}_{kl}+\beta^2J_{ij}|^{3/2}}=-\frac{2}{3}\frac{\partial}{\partial\gamma}\frac{1}{|\tilde{C}_{kl}+\gamma\tilde{\upsilon}^l\upsilon^{k\prime}+\beta^2J_{ij}|^{3/2}}\Bigg|_{\gamma=0}.\tag{41}$$

The determinant on the right-hand side of Eq. (41) can be expressed as

$$|\tilde{C}_{kl}+\gamma\tilde{\upsilon}^l\upsilon^{k\prime}+\beta^2J_{ij}|=|\tilde{C}_{kl}+\gamma\tilde{\upsilon}^l\upsilon^{k\prime}||I+\beta^2(\tilde{C}_{kl}+\gamma\tilde{\upsilon}^l\upsilon^{k\prime})^{-1}J_{ij}|$$

$$=|\tilde{C}_{kl}+\gamma\tilde{\upsilon}^l\upsilon^{k\prime}|(1+\beta^2\mathrm{tr}[(\tilde{C}_{kl}+\gamma\tilde{\upsilon}^l\upsilon^{k\prime})^{-1}J_{ij}]),\tag{42}$$

which is a consequence of the fact that matrix $(\tilde{C}_{kl}+\gamma\tilde{\upsilon}^l\upsilon^{k\prime})^{-1}J_{ij}$ has rank one and its only nonzero eigenvalue is equal to its trace. After all the above-mentioned manipulations, the expression for the Coulomb integral becomes

$$\left\langle\phi_k\left|\frac{1}{r_{ij}}\right|\tilde{\phi}_l\right\rangle=-\frac{2\pi^{(3n-1)/2}}{3}\frac{\partial}{\partial\gamma}\frac{1}{|\tilde{C}_{kl}+\gamma\tilde{\upsilon}^l\upsilon^{k\prime}|^{3/2}}\times\int_0^{\infty}\frac{\mathrm{d}\beta}{(1+\beta^2\mathrm{tr}[(\tilde{C}_{kl}+\gamma\tilde{\upsilon}^l\upsilon^{k\prime})^{-1}J_{ij}])^{3/2}}\Bigg|_{\gamma=0}$$

$$=\frac{\pi^{(3n-1)/2}}{|\tilde{C}_{kl}|^{3/2}}\left(\frac{\mathrm{tr}[\tilde{C}_{kl}^{-1}\tilde{\upsilon}^l\upsilon^{k\prime}]}{(\mathrm{tr}[\tilde{C}_{kl}^{-1}J_{ij}])^{1/2}}-\frac{1}{3}\frac{\mathrm{tr}[\tilde{C}_{kl}^{-1}\tilde{\upsilon}^l\upsilon^{k\prime}\tilde{C}_{kl}^{-1}J_{ij}]}{(\mathrm{tr}[\tilde{C}_{kl}^{-1}J_{ij}])^{3/2}}\right).\tag{43}$$

Using the normalized basis functions this matrix element becomes (we denote it as R_{kl}^{ij}):

$$\mathsf{R}_{kl}^{ij}=\frac{2}{\sqrt{\pi}}\frac{\mathsf{S}_{kl}}{(\mathrm{tr}[\tilde{C}_{kl}^{-1}J_{ij}])^{1/2}}\times\left(1-\frac{1}{3\,\mathrm{tr}[\tilde{C}_{kl}^{-1}J_{ij}]}\frac{\upsilon^{k\prime}\tilde{C}_{kl}^{-1}J_{ij}\tilde{C}_{kl}^{-1}\tilde{\upsilon}^l}{\upsilon^{k\prime}\tilde{C}_{kl}^{-1}\tilde{\upsilon}^l}\right).\tag{44}$$

The matrix element R_{kl}^{i} corresponding to the Coulomb term $1/r_i$ can be obtained by simply replacing $J_{ij}\rightarrow J_{ii}$ in the expression (44).

The R_{kl}^{i}'s and R_{kl}^{ij}'s can now be used to determine the complete potential energy matrix element as

$$\mathsf{V}_{kl}=\sum_{i=1}^{n}q_0q_i\mathsf{R}_{kl}^{i}+\sum_{i<j}^{n}q_iq_j\mathsf{R}_{kl}^{ij}.\tag{45}$$

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

Computing many important properties of a quantum mechanical system often involves the evaluation of matrix elements with the 3D Dirac delta function. In this section we consider a few such quantities. In the three-dimensional delta function the argument of the function is a vector in the 3D space. If, for example, $\boldsymbol{\alpha}$ is this argument, then by definition $\delta(\boldsymbol{\alpha})\equiv\delta(\alpha_x)\delta(\alpha_y)\delta(\alpha_z)$.

We first derive a general expression for the matrix element with the Dirac delta function in the following form:

$$\delta(a_1\mathbf{r}_1+a_2\mathbf{r}_2+\cdots+a_n\mathbf{r}_n-\boldsymbol{\xi})=\delta((a\otimes I_3)'\mathbf{r}-\boldsymbol{\xi}),\tag{46}$$

where a is a real n-component vector and ξ is a real 3D parameter. Notice that matrix $a\otimes I_3$ on the right-hand side of Eq. (46) is a $n\times 3$ rectangular matrix.

We use the following representation of the delta function:

$$\delta((a\otimes I_3)'\mathbf{r}-\boldsymbol{\xi})=\lim_{\beta\rightarrow\infty}\left(\frac{\beta}{\pi}\right)^{3/2}\times\exp[-\beta((a\otimes I_3)'\mathbf{r}-\boldsymbol{\xi})^2].\tag{47}$$

The exponent in the above expression can be rewritten using the mixed-product property. This property states that if X, Y, Z, and W are matrices of such size that they can form products XZ and YW, then $(X\otimes Y)(Z\otimes W)=XZ\otimes YW$ and

$$\exp[-\beta\mathbf{r}'(aa'\otimes I_3)\mathbf{r}+2\beta((a\otimes I_3)\boldsymbol{\xi})'\mathbf{r}-\beta\boldsymbol{\xi}^2]. \tag{48}$$

The matrix element of the delta function with functions (14) is then

$$\begin{aligned}\langle\varphi_k|\delta((a\otimes I_3)'\mathbf{r}-\boldsymbol{\xi})|\tilde{\varphi}_l\rangle\\ =\lim_{\beta\to\infty}\left(\frac{\beta}{\pi}\right)^{3/2}\exp[-\beta\boldsymbol{\xi}^2]\langle\varphi_k|\exp[-\beta\mathbf{r}'(aa'\otimes I_3)\mathbf{r}\\ +2\beta((a\otimes I_3)\boldsymbol{\xi})'\mathbf{r}]|\tilde{\varphi}_l\rangle.\end{aligned} \tag{49}$$

The integral on the right-hand side of Eq. (49) is similar to the overlap integral [Eq. (24)] and can be easily evaluated. Thus we get

$$\begin{aligned}\langle\varphi_k|\delta((a\otimes I_3)'\mathbf{r}-\boldsymbol{\xi})|\tilde{\varphi}_l\rangle\\ =\lim_{\beta\to\infty}\left(\frac{\beta}{\pi}\right)^{3/2}\exp[-\beta\boldsymbol{\xi}^2]\frac{\pi^{3n/2}}{|\tilde{C}_{kl}+\beta aa'|^{3/2}}\\ \times\exp\bigg[\frac{1}{4}(\alpha_k\mathbf{v}^k+\alpha_l\tilde{\mathbf{v}}^l+2\beta(a\otimes I_3)\boldsymbol{\xi})'\\ \times((\tilde{C}_{kl}+\beta aa')^{-1}\otimes I_3)(\alpha_k\mathbf{v}^k+\alpha_l\tilde{\mathbf{v}}^l+2\beta(a\otimes I_3)\boldsymbol{\xi})\bigg].\end{aligned} \tag{50}$$

Here, similarly to the transformations performed in Eq. (42), we can write the determinant in the last expression as

$$\begin{aligned}|\tilde{C}_{kl}+\beta aa'|&=|\tilde{C}_{kl}|(1+\beta\,\mathrm{tr}[\tilde{C}_{kl}^{-1}aa'])\\ &=|\tilde{C}_{kl}|(1+\beta a'\tilde{C}_{kl}^{-1}a).\end{aligned} \tag{51}$$

Also, we can evaluate the inverse matrix in the second exponent by using the Sherman–Morrison formula, which states that for an arbitrary invertible matrix X and vectors y and z, such that $1+z'X^{-1}y\neq 0$, the following applies:

$$(X+yz')^{-1}=X^{-1}-\frac{X^{-1}yz'X^{-1}}{1+z'X^{-1}y}. \tag{52}$$

Finally, we just need to take the limit of the entire expression for $\beta\to\infty$. This, though a bit lengthy, is quite straightforward. At the end, after some simplifications, we arrive with the following expression for the matrix element:

$$\begin{aligned}\langle\varphi_k|\delta((a\otimes I_3)'\mathbf{r}-\boldsymbol{\xi})|\tilde{\varphi}_l\rangle\\ =\frac{\pi^{3(n-1)/2}}{|\tilde{C}_{kl}|^{3/2}(a'\tilde{C}_{kl}^{-1}a)^{3/2}}\\ \times\exp\bigg[\frac{1}{a'\tilde{C}_{kl}^{-1}a}\bigg(\frac{1}{4}(\alpha_k v^k+\alpha_l\tilde{v}^l)'([a'\tilde{C}_{kl}^{-1}a]\tilde{C}_{kl}^{-1}\\ -\tilde{C}_{kl}^{-1}aa'\tilde{C}_{kl}^{-1})(\alpha_k v^k+\alpha_l\tilde{v}^l)\\ +(\alpha_k v^k+\alpha_l\tilde{v}^l)'\tilde{C}_{kl}^{-1}a(\epsilon^{z\prime}\boldsymbol{\xi})-\boldsymbol{\xi}^2\bigg)\bigg].\end{aligned} \tag{53}$$

In order to compute the matrix element with functions (12), we now need to differentiate the above matrix element with respect to α_k and α_l and set both parameters to zero. This yields the following:

$$\begin{aligned}\langle\phi_k|\delta((a\otimes I_3)'\mathbf{r}-\boldsymbol{\xi})|\tilde{\phi}_l\rangle\\ =\frac{\langle\phi_k|\phi_l\rangle}{\pi^{3/2}(a'\tilde{C}_{kl}^{-1}a)^{3/2}}\\ \times\Bigg\{1+\frac{1}{a'\tilde{C}_{kl}^{-1}a}\frac{v^{k\prime}\tilde{C}_{kl}^{-1}aa'\tilde{C}_{kl}^{-1}\tilde{v}^l}{v^{k\prime}\tilde{C}_{kl}^{-1}\tilde{v}^l}\\ \times\left(2\frac{(\epsilon^{z\prime}\boldsymbol{\xi})^2}{a'\tilde{C}_{kl}^{-1}a}-1\right)\Bigg\}\exp\left[-\frac{\boldsymbol{\xi}^2}{a'\tilde{C}_{kl}^{-1}a}\right].\end{aligned} \tag{54}$$

The formula (54) represents a general case and, with an appropriate choice of vector a, it can be used to calculate matrix elements of operators representing the pair correlation functions, the densities of pseudoparticles, etc.

The definition of the pair correlation function is

$$g^i(\boldsymbol{\xi})=\langle\psi|\delta(\mathbf{r}_i-\boldsymbol{\xi})|\psi\rangle,\quad i=1,\ldots,n, \tag{55}$$

$$g^{ij}(\boldsymbol{\xi})=\langle\psi|\delta(\mathbf{r}_{ij}-\boldsymbol{\xi})|\psi\rangle,\quad i,j=1,\ldots,n,\quad i\neq j, \tag{56}$$

where ψ is the wave function of the system. The pair correlation function represents the probability density of particles 1 and $i+1$ [formula (55)] or particles $i+1$ and $j+1$ [formula (56)] to have relative separation $\boldsymbol{\xi}$. Notice that the lower-case Latin symbol g does not indicate here that the corresponding quantity has n components but is used following the traditional notation to denote the correlation function. The matrix elements $g_{kl}^{i}(\boldsymbol{\xi})$ and $g_{kl}^{ij}(\boldsymbol{\xi})$ are obtained by setting a in Eq. (54) to be equal to $a=j^i$ and $a=j^j-j^i$, respectively, where j^i is an n-component vector whose ith component is one, while all others are zeros. It is worth noting that vectors j^i and j^j-j^i have the following direct relation to matrices J_{ii} and J_{ij}:

$$j^i(j^i)'=J_{ii},\quad (j^j-j^i)(j^j-j^i)'=J_{ij}. \tag{57}$$

Using this, we can write $g_{kl}^{ij}(\xi)$ as

$$\begin{aligned}g_{kl}^{ij}(\boldsymbol{\xi})=\frac{\mathsf{S}_{kl}}{\pi^{3/2}\,\mathrm{tr}[\tilde{C}_{kl}^{-1}J_{ij}]^{3/2}}\\ \times\Bigg\{1+\frac{1}{\mathrm{tr}[\tilde{C}_{kl}^{-1}J_{ij}]}\frac{v^{k\prime}\tilde{C}_{kl}^{-1}J_{ij}\tilde{C}_{kl}^{-1}\tilde{v}^l}{v^{k\prime}\tilde{C}_{kl}^{-1}\tilde{v}^l}\\ \times\left(2\frac{(\epsilon^{z\prime}\boldsymbol{\xi})^2}{\mathrm{tr}[\tilde{C}_{kl}^{-1}J_{ij}]}-1\right)\Bigg\}\exp\left[-\frac{\boldsymbol{\xi}^2}{\mathrm{tr}[\tilde{C}_{kl}^{-1}J_{ij}]}\right].\end{aligned} \tag{58}$$

The definition of the density of particle i in the center-of-mass coordinate frame is the following:

$$\rho^i(\boldsymbol{\xi})=\langle\psi|\delta(\mathbf{R}_i-\mathbf{r}_0-\boldsymbol{\xi})|\psi\rangle,\quad i=1,\ldots,N. \tag{59}$$

This quantity gives the density distribution of the ith particle with respect to the center of mass of the system. $\mathbf{R}_i-\mathbf{r}_0$ can be expressed through the internal coordinates using formulas (3). In this case we get the delta function (46) with a being a vector with the following components:

$$a_j=\delta_{j,i-1}(1-\delta_{i,1})-m_j/M_{\text{tot}}, \tag{60}$$

where $\delta_{i,j}$ is the Kronecker symbol.

When $L=1$, the pair correlation functions and the particle densities in the center-of-mass frame, as well as any other quantity whose matrix elements are related to $\langle\psi|\delta((a\otimes I_3)'\mathbf{r}-\boldsymbol{\xi})|\psi\rangle$, have the axial symmetry. Notice that if the mass of particle 1 is infinite then $\rho^i(\boldsymbol{\xi})=g^i(\boldsymbol{\xi})$, because in this case the center of mass is located on this particle. Similarly, when the mass of particle k is infinite we get $\rho^i(\boldsymbol{\xi})=g^{i-1,k-1}(\boldsymbol{\xi})$.

Equation (54) is important because it can be used to evaluate the matrix elements of any function of $\mathbf{r}_i$ or of a linear combination of $\mathbf{r}_i$'s. Suppose this function is $\chi((a\otimes I_3)'\mathbf{r})$. The matrix elements of χ can be obtained as

$$\langle\phi_k|\chi((a\otimes I_3)'\mathbf{r})|\tilde{\phi}_l\rangle=\int\langle\phi_k|\delta((a\otimes I_3)'\mathbf{r}-\boldsymbol{\xi})\times|\tilde{\phi}_l\rangle\chi(\boldsymbol{\xi})\mathrm{d}\boldsymbol{\xi},\tag{61}$$

which reduces the problem to a 3D integral. In the case when χ depends only on the absolute value of its argument, Eq. (61) simplifies further to a one-dimensional integral that in most cases can be easily evaluated analytically.

E. Interparticle distances

The expectation values of various powers of the interparticle distances are often the quantities of interest in atomic and molecular calculations. The matrix elements of $|\mathbf{r}_i|^\nu$ and $|\mathbf{r}_{ij}|^\nu$, where ν is a real number ($\nu>-3$), can be evaluated using relation (61) with $a=j^i$ and $a=j^j-j^i$ as

$$\langle\phi_k|r_{ij}^\nu|\tilde{\phi}_l\rangle=\int\langle\phi_k|\delta(\mathbf{r}_{ij}-\boldsymbol{\xi})|\tilde{\phi}_l\rangle\xi^\nu\mathrm{d}\boldsymbol{\xi}.\tag{62}$$

After the integration we obtain the following expression:

$$\langle\phi_k|r_{ij}^\nu|\tilde{\phi}_l\rangle=\langle\phi_k|\tilde{\phi}_l\rangle\frac{2}{\sqrt{\pi}}\Gamma\left(\frac{\nu+3}{2}\right)\mathrm{tr}[\tilde{C}_{kl}^{-1}J_{ij}]^{\nu/2}\times\left(1+\frac{\nu}{3}\frac{1}{\mathrm{tr}[\tilde{C}_{kl}^{-1}J_{ij}]}\frac{\upsilon^{k\prime}\tilde{C}_{kl}^{-1}J_{ij}\tilde{C}_{kl}^{-1}\tilde{\upsilon}^l}{\upsilon^{k\prime}\tilde{C}_{kl}^{-1}\tilde{\upsilon}^l}\right),\tag{63}$$

where $\Gamma(\ldots)$ is the Euler gamma function. Notice that, when $\nu=-1$, this formula reproduces Eq. (44). Also, for $\nu=2$, the formula gives the same result as in Eq. (33) with $W=J_{ij}$.

VII. ENERGY GRADIENT

Variational calculations with ECGs can provide very accurate energies and wave functions. However, such calculations usually consume large quantities of computer time. The main reason for this is that the energy depends strongly on the choice of the exponential parameters of the Gaussians. In order to achieve the high accuracy of the results, one needs to perform a computationally very costly optimization of these parameters. It should be said that finding an absolute minimum of the energy with respect to the Gaussian parameters for a long expansion of the wave function in terms of Gaussians (hundreds or thousands of basis functions) is almost a hopeless task. The number of parameters is simply too large and the hypersurface of the energy function is too complicated to find its absolute minimum. For long expansions of the wave function, a single point energy calculation usually takes a considerable amount of time (one has to evaluate all necessary Hamiltonian matrix elements and solve a general symmetric/Hermitian eigenvalue problem) and, as the variational optimization involves multiple energy evaluations, the calculation is never carried to a complete convergence. However, it is never necessary to find the global minimum of the energy. It is usually enough to find a sufficiently low point. Instead of spending large amounts of computational time searching for the global minimum for a particular basis set with a smaller size, it is better to add more basis functions to the set and lower the energy this way. It does not mean, however, that the optimization of the Gaussian exponential parameters can be avoided. Quite the contrary, it still has to be performed quite thoroughly. In general, finding the right balance between the amount of effort spent on the optimization and on increasing the basis size (which also requires large amounts of computer time) is not an easy task and involves guesses, trials, and, most of all, experience.

One of the approaches that aims to simplify the choice of the Gaussian parameters is the so-called stochastic variational method (SVM).[3,10] The idea of the method is simple: one increases the basis by one (or several) function(s) at a time, choosing from a certain number of randomly generated candidates. The chosen function(s) are those that lower the energy the most. Upon including a new basis function to the basis, one can also optimize its parameters using conventional optimization algorithms. The procedure is repeated until the desired basis size or the desired accuracy is reached. The generation of the random candidates may be based on a distribution of the exponential parameters of the basis functions already included in the set. After a certain number of new functions are added to the basis set, one may also apply a procedure which we call the "cyclic optimization." This procedure optimizes the entire basis set by tuning the parameters of only one function at a time, one after another. This cycle may be performed once or repeated several times. Optimizing the parameters of only one function at a time, as well as the random selection of the parameters of only one function at a time, has a certain advantage. In this case, when the parameters of the optimized function change during the optimization, one has to recalculate only one row and one column of the Hamiltonian and overlap matrices. Whether one uses the SVM and optimizes just one function at a time or one optimizes all basis functions simultaneously, the energy has to be evaluated a large number of times. The optimization can be accelerated by utilizing the energy gradient with respect to the exponential parameters of the basis functions. Unfortunately, the calculation of the finite-difference gradient is very costly. To reduce the cost of the optimization, in our previous works[1,6–8] we used the analytic gradient of the energy with respect to the Gaussian exponential parameters. Even though the implementation of the analytic gradient requires additional work, it dramatically speeds up the calculation. The use of the analytic gradient is especially important when a simultaneous optimization of all basis functions is performed. Such optimization is necessary for quantum systems where basis functions are very strongly "coupled" and the strategy based on the optimization of one function at a time is inefficient [the coupling means that when the parameters of one basis function change, the parameters of some other function(s) have to also change in order to lower the energy].

In this section we first explain the procedure for the construction of the analytic gradient and then we derive all the necessary matrix elements for the gradient calculation.

A. Construction of the energy gradient

In the Rayleigh–Ritz variational scheme the wave function of the system, ψ, after the elimination of the spin variables (which can always be done when the interparticle interactions do not depend on the spin), is approximated as a linear combination of K basis functions ϕ_k,

$$\psi(\mathbf{r}) = \sum_{k=1}^{K} \mathsf{c}_k \hat{Y} \phi_k(\mathbf{r}). \tag{64}$$

Here $\hat{Y}$ is some permutational symmetry projector (a linear combination of permutational operators, $\hat{P}_\gamma$) and c_k are the linear variational parameters.

The minimization of the energy functional with respect to the parameters c_k leads to the secular equation

$$(\mathsf{H} - \varepsilon \mathsf{S})\mathsf{c} = 0, \tag{65}$$

where H and S are $K \times K$ Hermitian matrices of the Hamiltonian and overlap integrals, with the elements $\mathsf{H}_{kl} = \langle \phi_k | \hat{H} \hat{Y}^\dagger \hat{Y} | \phi_l \rangle$ and $\mathsf{S}_{kl} = \langle \phi_k | \hat{Y}^\dagger \hat{Y} | \phi_l \rangle$, respectively. c is a K-component vector of the linear parameters c_k.

The solutions of Eq. (65) give upper bounds ε to the exact ground and excited state energies of the system. The corresponding sets of the linear parameters c define the wave functions. There exist K solutions of Eq. (65), but, for the sake of avoiding using an additional index, we will assume that we are interested in a particular solution.

By taking the differential of Eq. (65),

$$\mathrm{d}(\mathsf{H} - \varepsilon \mathsf{S})\mathsf{c} = (\mathrm{d}\mathsf{H})\mathsf{c} - (\mathrm{d}\varepsilon)\mathsf{S}\mathsf{c} - \varepsilon(\mathrm{d}\mathsf{S})\mathsf{c} + (\mathsf{H} - \varepsilon \mathsf{S})\mathrm{d}\mathsf{c} \tag{66}$$

and multiplying by $\mathsf{c}^\dagger$ from the left we obtain

$$\mathrm{d}\varepsilon = \mathsf{c}^\dagger(\mathrm{d}\mathsf{H} - \varepsilon \mathrm{d}\mathsf{S})\mathsf{c}. \tag{67}$$

In the above equation we assumed that the wave function is normalized so that $\mathsf{c}^\dagger \mathsf{S} \mathsf{c} = 1$. The relation (67) is essentially the same as the well known Hellmann–Feynman theorem.

Now let us assume that α_t is some nonlinear parameter, on which the basis function ϕ_t depends. It is obvious that only the elements in the tth row and tth column of the matrices H and S depend on α_t. Thus, the derivative of an arbitrary element of H (as well as S) can be written as

$$\frac{\partial \mathsf{H}_{kl}}{\partial \alpha_t} = \frac{\partial \mathsf{H}_{kl}}{\partial \alpha_t}(\delta_{kt} + \delta_{lt} - \delta_{kt}\delta_{lt}), \quad k,l,t = 1, \ldots, K. \tag{68}$$

From Eqs. (67) and (68) it is easy to find that the derivative of the total energy, ε, with respect to the parameter α_t, is

$$\begin{aligned}
\frac{\partial \varepsilon}{\partial \alpha_t} &= \mathsf{c}_t^* \sum_{l=1}^{K} \mathsf{c}_l \left(\frac{\partial \mathsf{H}_{tl}}{\partial \alpha_t} - \varepsilon \frac{\partial \mathsf{S}_{tl}}{\partial \alpha_t} \right) + \mathsf{c}_t \sum_{l=1}^{K} \mathsf{c}_l^* \left(\frac{\partial \mathsf{H}_{lt}}{\partial \alpha_t} - \varepsilon \frac{\partial \mathsf{S}_{lt}}{\partial \alpha_t} \right) \\
&\quad - \mathsf{c}_t \mathsf{c}_t^* \left(\frac{\partial \mathsf{H}_{tt}}{\partial \alpha_t} - \varepsilon \frac{\partial \mathsf{S}_{tt}}{\partial \alpha_t} \right) \\
&= 2\,\mathrm{Re}\left[\mathsf{c}_t^* \sum_{l=1}^{K} \mathsf{c}_l \left(\frac{\partial \mathsf{H}_{tl}}{\partial \alpha_t} - \varepsilon \frac{\partial \mathsf{S}_{tl}}{\partial \alpha_t} \right) \right] \\
&\quad - \mathsf{c}_t \mathsf{c}_t^* \left(\frac{\partial \mathsf{H}_{tt}}{\partial \alpha_t} - \varepsilon \frac{\partial \mathsf{S}_{tt}}{\partial \alpha_t} \right).
\end{aligned} \tag{69}$$

By calculating such a derivative for each α_k ($k = 1, \ldots, K$) we can get the entire energy gradient. In practice, it is advantageous to evaluate the derivatives of ε with respect to the entire $\mathrm{vech}\, L_k$ and $\mathrm{vech}\, B_k$ vectors rather than doing this separately for individual parameters $(L_k)_{11}, (L_k)_{21}, \ldots, (L_k)_{nn}$ and $(B_k)_{11}, (B_k)_{21}, \ldots, (B_k)_{nn}$. This is because the calculation of the derivatives with respect to the elements of matrices L_k and B_k involves many identical operations and repeating them separately for each element is much less efficient than calculating all of them in a single step.

As it can be seen from Eq. (69), the calculation of the gradient of ε with respect to $\mathrm{vech}\, L_k$ and $\mathrm{vech}\, B_k$ involves the following derivatives of the H and S matrix elements:

$$\begin{gathered}
\frac{\partial \mathsf{H}_{kl}}{\partial(\mathrm{vech}\, L_k)}, \quad \frac{\partial \mathsf{H}_{kl}}{\partial(\mathrm{vech}\, L_l)}, \quad \frac{\partial \mathsf{H}_{kl}}{\partial(\mathrm{vech}\, B_k)}, \quad \frac{\partial \mathsf{H}_{kl}}{\partial(\mathrm{vech}\, B_l)}, \\
\frac{\partial \mathsf{S}_{kl}}{\partial(\mathrm{vech}\, L_k)}, \quad \frac{\partial \mathsf{S}_{kl}}{\partial(\mathrm{vech}\, L_l)}, \quad \frac{\partial \mathsf{S}_{kl}}{\partial(\mathrm{vech}\, B_k)}, \quad \frac{\partial \mathsf{S}_{kl}}{\partial(\mathrm{vech}\, B_l)}.
\end{gathered} \tag{70}$$

Next, we derive the expressions for these derivatives.

B. Overlap derivative

Let us first write out some useful auxiliary relations. The differentials of $C_k^\dagger$ and $\tilde{C}_l$ are

$$\mathrm{d}C_k^\dagger = (\mathrm{d}L_k)L_k' + L_k \mathrm{d}L_k' - i\mathrm{d}B_k, \tag{71}$$

$$\mathrm{d}\tilde{C}_l = (\mathrm{d}\tilde{L}_l)\tilde{L}_l' + \tilde{L}_l \mathrm{d}\tilde{L}_l' + i\mathrm{d}\tilde{B}_l. \tag{72}$$

The differential of $\tilde{C}_{kl}$ is a sum of the two above differentials,

$$\begin{aligned}
\mathrm{d}\tilde{C}_{kl} &= (\mathrm{d}L_k)L_k' + L_k \mathrm{d}L_k' - i\mathrm{d}B_k + P'(\mathrm{d}L_l)L_l' P \\
&\quad + P' L_l (\mathrm{d}L_l') P + iP'(\mathrm{d}B_l)P.
\end{aligned} \tag{73}$$

Hence, the differential of the unnormalized overlap integral is

$$\begin{aligned}\mathrm{d}\langle\phi_k|\tilde{\phi}_l\rangle &= \mathrm{d}\frac{\pi^{3n/2}}{2}\frac{\mathrm{tr}[\tilde{v}^l v^{k\prime} C_{kl}^{-1}]}{|\tilde{C}_{kl}|^{3/2}}\\ &= -\langle\phi_k|\tilde{\phi}_l\rangle\Bigg(\frac{3}{2}\mathrm{tr}[\tilde{C}_{kl}^{-1}\mathrm{d}\tilde{C}_{kl}]\\ &\quad+\frac{1}{\mathrm{tr}[\tilde{v}^l v^{k\prime}\tilde{C}_{kl}^{-1}]}\mathrm{tr}[\tilde{K}_{kl}\mathrm{d}\tilde{C}_{kl}]\Bigg),\end{aligned} \tag{74}$$

where $\tilde{K}_{kl}$ is the following matrix:

$$\tilde{K}_{kl}=\tilde{C}_{kl}^{-1}\tilde{v}^l v^{k\prime}\tilde{C}_{kl}^{-1}. \tag{75}$$

To simplify the further derivation we define the following matrices:

$$\tilde{F}_{kl}=\frac{3}{2}\tilde{C}_{kl}^{-1}+\frac{1}{v^{k\prime}\tilde{C}_{kl}^{-1}\tilde{v}^l}\tilde{K}_{kl} \tag{76}$$

and

$$\tilde{G}_{kl}=P\tilde{F}_{kl}P'. \tag{77}$$

With that we have

$$\begin{aligned}\mathrm{d}\langle\phi_k|\tilde{\phi}_l\rangle &= -\langle\phi_k|\tilde{\phi}_l\rangle\mathrm{tr}[\tilde{F}_{kl}\mathrm{d}\tilde{C}_{kl}]\\ &= -\langle\phi_k|\tilde{\phi}_l\rangle\mathrm{tr}[L_k'(\tilde{F}_{kl}+\tilde{F}_{kl}')\mathrm{d}L_k - i\tilde{F}_{kl}\mathrm{d}B_k\\ &\quad+L_l'(\tilde{G}_{kl}+\tilde{G}_{kl}')\mathrm{d}L_l+i\tilde{G}_{kl}\mathrm{d}B_l].\end{aligned} \tag{78}$$

Using the properties (9)–(11) we can now rewrite the last expression as

$$\begin{aligned}\mathrm{d}\langle\phi_k|\tilde{\phi}_l\rangle &= -\langle\phi_k|\tilde{\phi}_l\rangle[\mathrm{vech}((\tilde{F}_{kl}+\tilde{F}_{kl}')L_k)'\mathrm{d}\,\mathrm{vech}\,L_k\\ &\quad- i\,\mathrm{vech}(\tilde{F}_{kl}+\tilde{F}_{kl}'-\mathrm{diag}\,\tilde{F}_{kl})'\mathrm{d}\,\mathrm{vech}\,B_k\\ &\quad+\mathrm{vech}((\tilde{G}_{kl}+\tilde{G}_{kl}')L_l)'\mathrm{d}\,\mathrm{vech}\,L_l+i\,\mathrm{vech}(\tilde{G}_{kl}\\ &\quad+\tilde{G}_{kl}'-\mathrm{diag}\,\tilde{G}_{kl})'\mathrm{d}\,\mathrm{vech}\,B_l].\end{aligned} \tag{79}$$

It can now be seen that

$$\frac{\partial\langle\phi_k|\tilde{\phi}_l\rangle}{\partial(\,\mathrm{vech}\,L_k)}=-\langle\phi_k|\tilde{\phi}_l\rangle\mathrm{vech}((\tilde{F}_{kl}+\tilde{F}_{kl}')L_k), \tag{80}$$

$$\frac{\partial\langle\phi_k|\tilde{\phi}_l\rangle}{\partial(\mathrm{vech}\,B_k)}=i\langle\phi_k|\tilde{\phi}_l\rangle\mathrm{vech}(\tilde{F}_{kl}+\tilde{F}_{kl}'-\mathrm{diag}\,\tilde{F}_{kl}), \tag{81}$$

$$\frac{\partial\langle\phi_k|\tilde{\phi}_l\rangle}{\partial(\mathrm{vech}\,L_l)}=-\langle\phi_k|\tilde{\phi}_l\rangle\mathrm{vech}((\tilde{G}_{kl}+\tilde{G}_{kl}')L_l), \tag{82}$$

and

$$\frac{\partial\langle\phi_k|\tilde{\phi}_l\rangle}{\partial(\mathrm{vech}\,B_l)}=-i\langle\phi_k|\tilde{\phi}_l\rangle\mathrm{vech}(\tilde{G}_{kl}+\tilde{G}_{kl}'-\mathrm{diag}\,\tilde{G}_{kl}). \tag{83}$$

It should be noted that in the case where $k=l$ and $\hat{P}=1$ we get

$$\frac{\partial\langle\phi_k|\phi_k\rangle}{\partial(\mathrm{vech}\,L_k)}=-4\langle\phi_k|\phi_k\rangle\mathrm{vech}(F_{kk}L_k) \tag{84}$$

and

$$\frac{\partial\langle\phi_k|\phi_k\rangle}{\partial(\mathrm{vech}\,B_k)}=0, \tag{85}$$

with

$$F_{kk}=\frac{3}{2}C_{kk}^{-1}+\frac{1}{v^{k\prime}C_{kk}^{-1}v^k}C_{kk}^{-1}v^k v^{k\prime}C_{kk}^{-1}. \tag{86}$$

Now we can determine the derivative of the normalized overlap integral

$$\begin{aligned}\frac{\partial \mathsf{S}_{kl}}{\partial(\mathrm{vech}\,L_k)} &= \frac{\partial}{\partial(\mathrm{vech}\,L_k)}\frac{\langle\phi_k|\tilde{\phi}_l\rangle}{(\langle\phi_k|\phi_k\rangle\langle\phi_l|\phi_l\rangle)^{1/2}}\\ &= \frac{1}{(\langle\phi_k|\phi_k\rangle\langle\phi_l|\phi_l\rangle)^{1/2}}\frac{\partial\langle\phi_k|\tilde{\phi}_l\rangle}{\partial(\mathrm{vech}\,L_k)}\\ &\quad-\frac{1}{2}\frac{\langle\phi_k|\tilde{\phi}_l\rangle}{(\langle\phi_k|\phi_k\rangle)^{3/2}(\langle\phi_l|\phi_l\rangle)^{1/2}}\frac{\partial\langle\phi_k|\phi_k\rangle}{\partial(\mathrm{vech}\,L_k)}\\ &= \mathsf{S}_{kl}\,\mathrm{vech}((2F_{kk}-\tilde{F}_{kl}-\tilde{F}_{kl}')L_k),\end{aligned} \tag{87}$$

and similarly,

$$\frac{\partial \mathsf{S}_{kl}}{\partial(\mathrm{vech}\,B_k)}=i\mathsf{S}_{kl}\,\mathrm{vech}(\tilde{F}_{kl}+\tilde{F}_{kl}'-\mathrm{diag}\,\tilde{F}_{kl}), \tag{88}$$

$$\frac{\partial \mathsf{S}_{kl}}{\partial(\mathrm{vech}\,L_l)}=\mathsf{S}_{kl}\,\mathrm{vech}((2F_{ll}-\tilde{G}_{kl}-\tilde{G}_{kl}')L_l), \tag{89}$$

and

$$\frac{\partial \mathsf{S}_{kl}}{\partial(\mathrm{vech}\,B_l)}=-i\mathsf{S}_{kl}\,\mathrm{vech}(\tilde{G}_{kl}+\tilde{G}_{kl}'-\mathrm{diag}\,\tilde{G}_{kl}). \tag{90}$$

C. Kinetic energy derivatives

To simplify the expression we are to derive, let us introduce the following abbreviations:

$$\tau_1=\mathrm{tr}[\tilde{C}_{kl}^{-1}C_l M\tilde{C}_k^{\dagger}], \tag{91}$$

$$\tau_2=\mathrm{tr}[\tilde{v}^l v^{k\prime}\tilde{C}_{kl}^{-1}\tilde{C}_l M C_k^{\dagger}\tilde{C}_{kl}^{-1}], \tag{92}$$

$$\tau_3=\mathrm{tr}[\tilde{v}^l v^{k\prime}\tilde{C}_{kl}^{-1}]. \tag{93}$$

With these abbreviations, the matrix element of the kinetic energy operator (37) is

$$\mathsf{T}_{kl}=\mathsf{S}_{kl}\left(6\tau_1+4\frac{\tau_2}{\tau_3}\right), \tag{94}$$

and the differential of the kinetic energy matrix element is

$$\mathrm{d}\mathsf{T}_{kl}=\frac{\mathsf{T}_{kl}}{\mathsf{S}_{kl}}\mathrm{d}\mathsf{S}_{kl}+\mathsf{S}_{kl}\left(6\mathrm{d}\tau_1+\frac{4}{\tau_3}\mathrm{d}\tau_2-4\frac{\tau_2}{\tau_3^2}\mathrm{d}\tau_3\right). \tag{95}$$

The expressions for $\mathrm{d}\tau_1$, $\mathrm{d}\tau_2$, and $\mathrm{d}\tau_3$ in terms of $\mathrm{d}C_k^{\dagger}$ and $\mathrm{d}\tilde{C}_l$ can be easily obtained and they are

$$\mathrm{d}\tau_1 = \mathrm{tr}[\tilde{C}_{kl}^{-1}\tilde{C}_l M \tilde{C}_l \tilde{C}_{kl}^{-1}\mathrm{d}C_k^{\dagger} + \tilde{C}_{kl}^{-1} C_k^{\dagger} M C_k^{\dagger} \tilde{C}_{kl}^{-1}\mathrm{d}\tilde{C}_l], \tag{96}$$

$$\begin{aligned}\mathrm{d}\tau_2 = \mathrm{tr}[&(\tilde{K}_{kl}\tilde{C}_l M \tilde{C}_l \tilde{C}_{kl}^{-1} - \tilde{C}_{kl}^{-1}\tilde{C}_l M C_k^{\dagger}\tilde{K}_{kl})\mathrm{d}C_k^{\dagger} \\ &+ (\tilde{C}_{kl}^{-1} C_k^{\dagger} M C_k^{\dagger}\tilde{K}_{kl} - \tilde{K}_{kl}\tilde{C}_l M C_k^{\dagger}\tilde{C}_{kl}^{-1})\mathrm{d}\tilde{C}_l],\end{aligned} \tag{97}$$

$$\mathrm{d}\tau_3 = -\,\mathrm{tr}[\tilde{K}_{kl}\mathrm{d}C_k^{\dagger} + \tilde{K}_{kl}\mathrm{d}\tilde{C}_l]. \tag{98}$$

Plugging Eqs. (96)–(98) into (95) and denoting

$$\begin{aligned}\tilde{U}_{kl} = {}& 6\tilde{C}_{kl}^{-1}\tilde{C}_l M \tilde{C}_l \tilde{C}_{kl}^{-1} + \frac{4}{\tau_3}(\tilde{K}_{kl}\tilde{C}_l M \tilde{C}_l \tilde{C}_{kl}^{-1} \\ &- \tilde{C}_{kl}^{-1}\tilde{C}_l M C_k^{\dagger}\tilde{K}_{kl}) + 4\frac{\tau_2}{\tau_3^2}\tilde{K}_{kl},\end{aligned} \tag{99}$$

and

$$\begin{aligned}\tilde{W}_{kl} = P\Bigg(& 6\tilde{C}_{kl}^{-1} C_k^{\dagger} M C_k^{\dagger}\tilde{C}_{kl}^{-1} + \frac{4}{\tau_3}(\tilde{C}_{kl}^{-1} C_k^{\dagger} M C_k^{\dagger}\tilde{K}_{kl} \\ &- \tilde{K}_{kl}\tilde{C}_l M C_k^{\dagger}\tilde{C}_{kl}^{-1}) + 4\frac{\tau_2}{\tau_3^2}\tilde{K}_{kl}\Bigg)P',\end{aligned} \tag{100}$$

we get

$$\begin{aligned}\mathrm{d}\mathsf{T}_{kl} &= \frac{\mathsf{T}_{kl}}{\mathsf{S}_{kl}}\mathrm{d}\mathsf{S}_{kl} + \mathsf{S}_{kl}\,\mathrm{tr}[\tilde{U}_{kl}\mathrm{d}C_k^{\dagger} + \tilde{W}_{kl}\mathrm{d}\tilde{C}_l] \\ &= \frac{\mathsf{T}_{kl}}{\mathsf{S}_{kl}}\mathrm{d}\mathsf{S}_{kl} + \mathsf{S}_{kl}\,\mathrm{tr}[\mathrm{vech}((\tilde{U}_{kl} + \tilde{U}_{kl}')L_k)'\mathrm{d}\,\mathrm{vech}\,L_k \\ &\quad - i\,\mathrm{vech}(\tilde{U}_{kl} + \tilde{U}_{kl}' - \mathrm{diag}\,\tilde{U}_{kl})'\mathrm{d}\,\mathrm{vech}\,B_k \\ &\quad + \mathrm{vech}((\tilde{W}_{kl} + \tilde{W}_{kl}')L_l)'\mathrm{d}\,\mathrm{vech}\,L_l + i\,\mathrm{vech}(\tilde{W}_{kl} \\ &\quad + \tilde{W}_{kl}' - \mathrm{diag}\,\tilde{W}_{kl})'\mathrm{d}\,\mathrm{vech}\,B_l].\end{aligned} \tag{101}$$

Hence, the derivatives of the normalized kinetic energy matrix element are

$$\frac{\partial\mathsf{T}_{kl}}{\partial(\mathrm{vech}\,L_k)} = \frac{\mathsf{T}_{kl}}{\mathsf{S}_{kl}}\frac{\partial\mathsf{S}_{kl}}{\partial(\mathrm{vech}\,L_k)} + \mathsf{S}_{kl}\,\mathrm{vech}((\tilde{U}_{kl} + \tilde{U}_{kl}')L_k), \tag{102}$$

$$\begin{aligned}\frac{\partial\mathsf{T}_{kl}}{\partial(\mathrm{vech}\,B_k)} = {}& \frac{\mathsf{T}_{kl}}{\mathsf{S}_{kl}}\frac{\partial\mathsf{S}_{kl}}{\partial(\mathrm{vech}\,B_k)} - i\mathsf{S}_{kl}\,\mathrm{vech}(\tilde{U}_{kl} + \tilde{U}_{kl}' \\ &- \mathrm{diag}\,\tilde{U}_{kl}),\end{aligned} \tag{103}$$

$$\frac{\partial\mathsf{T}_{kl}}{\partial(\mathrm{vech}\,L_l)} = \frac{\mathsf{T}_{kl}}{\mathsf{S}_{kl}}\frac{\partial\mathsf{S}_{kl}}{\partial(\mathrm{vech}\,L_l)} + \mathsf{S}_{kl}\,\mathrm{vech}((\tilde{W}_{kl} + \tilde{W}_{kl}')L_l), \tag{104}$$

$$\begin{aligned}\frac{\partial\mathsf{T}_{kl}}{\partial(\mathrm{vech}\,B_l)} = {}& \frac{\mathsf{T}_{kl}}{\mathsf{S}_{kl}}\frac{\partial\mathsf{S}_{kl}}{\partial(\mathrm{vech}\,B_l)} + i\mathsf{S}_{kl}\,\mathrm{vech}(\tilde{W}_{kl} + \tilde{W}_{kl}' \\ &- \mathrm{diag}\,\tilde{W}_{kl}).\end{aligned} \tag{105}$$

D. Potential energy derivatives

Here, again, to simplify the notation we introduce the following abbreviations:

$$\eta_1 = \mathrm{tr}[J_{ij}\tilde{C}_{kl}^{-1}], \tag{106}$$

$$\eta_2 = \mathrm{tr}[\tilde{v}^l v^{k\prime}\tilde{C}_{kl}^{-1} J_{ij}\tilde{C}_{kl}^{-1}], \tag{107}$$

and

$$\eta_3 = \mathrm{tr}[\tilde{v}^l v^{k\prime}\tilde{C}_{kl}^{-1}]. \tag{108}$$

The differential of the Coulomb integrals,

$$\mathsf{R}_{kl}^{ij} = \frac{2}{\sqrt{\pi}}\mathsf{S}_{kl}\frac{1}{\eta_1^{1/2}}\left(1 - \frac{1}{3}\frac{\eta_2}{\eta_1\eta_3}\right),$$

which make up the potential energy, is

$$\begin{aligned}\mathrm{d}\mathsf{R}_{kl}^{ij} = {}& \frac{\mathsf{R}_{kl}^{ij}}{\mathsf{S}_{kl}}\mathrm{d}\mathsf{S}_{kl} + \frac{2}{\sqrt{\pi}}\mathsf{S}_{kl}\frac{1}{\eta_1^{3/2}}\Bigg(\frac{1}{2}\left(\frac{\eta_2}{\eta_1\eta_3} - 1\right)\mathrm{d}\eta_1 \\ &- \frac{1}{3}\frac{1}{\eta_3}\mathrm{d}\eta_2 + \frac{1}{3}\frac{\eta_2}{\eta_3^2}\mathrm{d}\eta_3\Bigg).\end{aligned} \tag{109}$$

Here

$$\mathrm{d}\eta_1 = -\,\mathrm{tr}[\tilde{C}_{kl}^{-1} J_{ij}\tilde{C}_{kl}^{-1}(\mathrm{d}C_k^{\dagger} + \mathrm{d}\tilde{C}_l)], \tag{110}$$

$$\begin{aligned}\mathrm{d}\eta_2 = -\,\mathrm{tr}[&(\tilde{C}_{kl}^{-1} J_{ij}\tilde{C}_{kl}^{-1}\tilde{v}^l v^{k\prime}\tilde{C}_{kl}^{-1} + \tilde{C}_{kl}^{-1}\tilde{v}^l v^{k\prime}\tilde{C}_{kl}^{-1} J_{ij}\tilde{C}_{kl}^{-1})(\mathrm{d}C_k^{\dagger} \\ &+ \mathrm{d}\tilde{C}_l)],\end{aligned} \tag{111}$$

$$\mathrm{d}\eta_3 = -\,\mathrm{tr}[\tilde{C}_{kl}^{-1}\tilde{v}^l v^{k\prime}\tilde{C}_{kl}^{-1}(\mathrm{d}C_k^{\dagger} + \mathrm{d}\tilde{C}_l)]. \tag{112}$$

If we denote

$$\begin{aligned}\tilde{Q}_{kl} = {}& \frac{2}{\sqrt{\pi}}\frac{1}{\eta_1^{3/2}}\Bigg(\frac{1}{2}\left(1 - \frac{\eta_2}{\eta_1\eta_3}\right)\tilde{C}_{kl}^{-1} J_{ij}\tilde{C}_{kl}^{-1} \\ &+ \frac{1}{3\eta_3}(\tilde{C}_{kl}^{-1} J_{ij}\tilde{C}_{kl}^{-1}\tilde{v}^l v^{k\prime}\tilde{C}_{kl}^{-1} + \tilde{C}_{kl}^{-1}\tilde{v}^l v^{k\prime}\tilde{C}_{kl}^{-1} J_{ij}\tilde{C}_{kl}^{-1}) \\ &- \frac{\eta_2}{3\eta_3^2}\tilde{C}_{kl}^{-1}\tilde{v}^l v^{k\prime}\tilde{C}_{kl}^{-1}\Bigg)\end{aligned} \tag{113}$$

and

$$\tilde{D}_{kl} = P\tilde{Q}_{kl}P', \tag{114}$$

then

$$\begin{aligned}\mathrm{d}\mathsf{R}_{kl}^{ij} = {}& \frac{\mathsf{R}_{kl}^{ij}}{\mathsf{S}_{kl}}\mathrm{d}\mathsf{S}_{kl} + \mathsf{S}_{kl}\,\mathrm{tr}[\tilde{Q}_{kl}(\mathrm{d}C_k^{\dagger} + \mathrm{d}\tilde{C}_l)] = \frac{\mathsf{R}_{kl}^{ij}}{\mathsf{S}_{kl}}\mathrm{d}\mathsf{S}_{kl} \\ &+ \mathsf{S}_{kl}(\mathrm{vech}((\tilde{Q}_{kl} + \tilde{Q}_{kl}')L_k)'\mathrm{d}\,\mathrm{vech}\,L_k - i\,\mathrm{vech}(\tilde{Q}_{kl} \\ &+ \tilde{Q}_{kl}' - \mathrm{diag}\,\tilde{Q}_{kl})'\mathrm{d}\,\mathrm{vech}\,B_k + \mathrm{vech}((\tilde{D}_{kl} \\ &+ \tilde{D}_{kl}')L_l)'\mathrm{d}\,\mathrm{vech}\,L_l + i\,\mathrm{vech}(\tilde{D}_{kl} + \tilde{D}_{kl}' \\ &- \mathrm{diag}\,\tilde{D}_{kl})'\mathrm{d}\,\mathrm{vech}\,B_l).\end{aligned} \tag{115}$$

From this expression we can get the derivatives of the potential energy matrix element,

$$\frac{\partial \mathsf{R}_{kl}^{ij}}{\partial(\mathrm{vech}\, L_k)} = \frac{\mathsf{R}_{kl}^{ij}}{\mathsf{S}_{kl}} \frac{\partial \mathsf{S}_{kl}}{\partial(\mathrm{vech}\, L_k)} + \mathsf{S}_{kl}\, \mathrm{vech}((\tilde{Q}_{kl} + \tilde{Q}_{kl}')L_k), \tag{116}$$

$$\frac{\partial \mathsf{R}_{kl}^{ij}}{\partial(\mathrm{vech}\, B_k)} = \frac{\mathsf{R}_{kl}^{ij}}{\mathsf{S}_{kl}} \frac{\partial \mathsf{S}_{kl}}{\partial(\mathrm{vech}\, B_k)} - i\mathsf{S}_{kl}\, \mathrm{vech}(\tilde{Q}_{kl} + \tilde{Q}_{kl}' - \mathrm{diag}\, \tilde{Q}_{kl}), \tag{117}$$

$$\frac{\partial \mathsf{R}_{kl}^{ij}}{\partial(\mathrm{vech}\, L_l)} = \frac{\mathsf{R}_{kl}^{ij}}{\mathsf{S}_{kl}} \frac{\partial \mathsf{S}_{kl}}{\partial(\mathrm{vech}\, L_l)} + \mathsf{S}_{kl}\, \mathrm{vech}((\tilde{D}_{kl} + \tilde{D}_{kl}')L_l), \tag{118}$$

$$\frac{\partial \mathsf{R}_{kl}^{ij}}{\partial(\mathrm{vech}\, B_l)} = \frac{\mathsf{R}_{kl}^{ij}}{\mathsf{S}_{kl}} \frac{\partial \mathsf{S}_{kl}}{\partial(\mathrm{vech}\, B_l)} + i\mathsf{S}_{kl}\, \mathrm{vech}(\tilde{D}_{kl} + \tilde{D}_{kl}' - \mathrm{diag}\, \tilde{D}_{kl}). \tag{119}$$

Since the potential energy matrix element is a sum of the Coulomb integrals, its derivative is a sum of the derivatives of R_{kl}^{ij} and R_{kl}^{i}.

At the end of this section we would like to make a comment concerning numerical implementation of the formulas for the gradient of the kinetic and potential energies. These contain many matrix products in the form $XY\tilde{v}^l v^{k\prime} Z$ and $XJ_{ij}X$, where X, Y, and Z are some arbitrary matrices. Such products can be evaluated very efficiently if we take into account the structure of the matrices $\tilde{v}^l v^{k\prime}$ and J_{ij}. As commonly applied, the $XY\tilde{v}^l v^{k\prime} Z$ product is evaluated as $(X(Y\tilde{v}^l))(v^{k\prime} Z)$, which only requires matrix-vector and vector-vector multiplications. The elements of the $XJ_{ij}X$ matrix are given by the following simple relation:

$$(XJ_{ii}X)_{pq} = (X)_{pi}(X)_{iq},$$

$$(XJ_{ij}X)_{pq} = (X)_{pi}(X)_{iq} + (X)_{pj}(X)_{jq} - (X)_{pi}(X)_{jq} - (X)_{pj}(X)_{iq}, \quad i \neq j. \tag{120}$$

VIII. NUMERICAL TESTS

In order to test the formulas derived in this work and to illustrate the performance of the method, we have performed calculations of two systems. The first system is the positronium molecule (Ps_2) in the $L=1$ state with the negative parity. The second system is the beryllium atom in its lowest P state. We chose these two systems because they provide nontrivial cases for very accurate variational calculations. Also, very accurate reference energies for these two systems have been published in the literature. It was our goal to test our new method for systems with at least four particles, as such systems are likely to be the main target of the calculations with ECGs we intend to perform in the future. For smaller systems (three to four particles) it is possible to use other types of explicitly correlated basis functions that often perform better than Gaussians.

Even though the positronium molecule consists of only four particles, it is an important model because it represents a case of a fully nonadiabatic system. While the correction to the total energy due to the finite nuclear mass for an adiabatic system (e.g., a molecule) is usually quite small, this correction is much more significant for a nonadiabatic system such as the positronium molecule. Thus, if there were an error in the calculation of the kinetic energy and its gradient, it might be difficult to detect and correct when the test is performed for an adiabatic system. In the positronium molecule the masses of all particles are the same. To the best of our knowledge, this is the only nonadiabatic system consisting of more than three particles whose P-state energy is known with very high accuracy. There are many other four-particle systems (for instance, the Li atom) whose energies of $L=1$ states have been determined very accurately, but all of them are adiabatic or nearly adiabatic.

One more reason for using the Be atom and the Ps_2 molecule in our tests is the fact that prior accurate calculations on these systems were carried out using ECGs. However, the optimizations of the exponential parameters in those calculations were done without the analytic gradient. Thus, the comparison of our results with those obtained in the prior calculations gives us the opportunity to illustrate the advantage of using the analytic gradient in the calculations.

Since the use of complex exponential parameters in the Gaussians in the calculations of the Be atom and the Ps_2 molecule is not strictly required (the use of complex parameters is unavoidable only in non-Born-Oppenheimer calculations on "true" molecules), at this stage of the development we decided to limit the implementation of the algorithms presented in this work to Gaussians with real exponential parameters. This is because the calculations with real parameters are substantially faster as all the required arithmetic operations only involve the real arithmetic, which is significantly less time consuming than the complex arithmetic. In addition, the number of actual parameters that need to be optimized in a calculation with real Gaussians is twice smaller than in the case of complex Gaussians.

A. Positronium molecule (*L*=1, negative parity)

The stability of the positronium molecule in its ground state was predicted by Hylleraas and Ore back in 1947.[13] While numerous high accuracy calculations on Ps_2 have been performed since then, no calculation on an excited state was reported until 1998. In the works of Varga and co-workers[11,12] it was shown that at least one excited state of the positronium molecule should exist. This state has $L=1$, $S=0$, and the negative parity. The reason for its existence is that its decay into the ground states of two separate Ps atoms is forbidden by symmetry. The nonrelativistic dissociation threshold for the Ps_2 state with $L=1$ and negative parity is $\mathrm{Ps}(L=0)+\mathrm{Ps}(L=1)=-0.3125$ a.u.

After elimination of the spin variables, the properly symmetrized basis functions for the calculations of the Ps_2 molecule in the negative-parity state with $L=1$ can be chosen as

$$(1-\hat{P}_{13}\hat{P}_{24})(1+\hat{P}_{12})(1+\hat{P}_{34})\phi_k, \tag{121}$$

where ϕ_k is a nonsymmetrized basis function, which in our case is a Gaussian (12) with B_k set to zero. In Eq. (121) we assumed that particles 1 and 2 are positrons and particles 3

TABLE I. Convergence of the total energy (in a.u.) of the positronium molecule in the state with $L=1$ and negative parity.

Basis size	This work	Varga *et al.*[a]
100	−0.334 400 893	−0.334 399 869
200	−0.334 407 545	−0.334 405 047
300	−0.334 408 147	
400	−0.334 408 266 3	−0.334 407 971
500	−0.334 408 295 5	
800		−0.334 408 177
1200		−0.334 408 234
1600		−0.334 408 265 8

[a] We only show the lowest values from works in Refs. 11 and 12.

TABLE II. Convergence of the total energy (in a.u.) of the beryllium atom in the 2 ^1P state.

Basis size	This work	Komasa *et al.*[a]
100	−14.471 732 323	
150	−14.472 759 879	−14.472 212 762
200	−14.473 101 103	
300	−14.473 326 695	−14.473 207 488
400	−14.473 393 969	
600	−14.473 431 855	−14.473 390 922
800	−14.473 442 537	
1200		−14.473 442 016

[a] Reference 4.

and 4 are electrons. $\hat{P}_{ij}$'s, as before, represent permutation operators of particles i and j.

The Ps$_2$ calculations have been carried out on a parallel computer system. In the calculations we extended the basis set gradually from 0 up to 500 basis functions using the strategy similar to the SVM described above. Both the Gaussian parameters and the indexes m_k of the exponent premultiplies z_{m_k} were subject to the stochastic selection. When optimizing the exponential parameters of a randomly selected basis function, we used the analytic gradient. When the basis size reached 100, 200, 300, 400, and 500 we also performed a gradient-based optimization of the exponential parameters of all basis functions simultaneously. The results of the calculations are shown in Table I. For comparison, in the same table we also present the total energies taken from two other most recent works.[11,12] As one can notice, for the small basis set of 100 functions our energy is only marginally better than that of Varga *et al.*[11] This suggests that the 100 term wave function in that work[11] was very thoroughly optimized using the cyclic optimization described in the book by Suzuki and Varga.[3] However, with the increase of the number of basis functions, the difference between the energies reported in Refs. 11 and 12 and our values becomes quite noticeable. In fact, our 300 term energy is significantly lower then the 400 term value from work of Varga *et al.*[11] and is not much higher than the 800 term energy from the work of Usukura *et al.*[12] With only 400 basis functions in the calculation we were able to get a lower energy than the best result reported by Usukura *et al.*[12] and obtained with 1600 basis functions. Such a large difference in the basis set size (by a factor of 4) between our calculations and the calculations by Usukura *et al.* clearly illustrates the importance of the optimization of the exponential nonlinear parameters in the variational calculation and the benefits of using the analytic gradient in this optimization.

The largest basis set we used in the calculations of Ps$_2$ was 500 Gaussians. With this basis we obtained a new variational lower bound of −0.324 408 265 8 a.u. for the $L=1$ state of this system which we estimate to be converged to the relative accuracy of 5×10^{-8}.

B. 2 ^1P state of Be atom

The second system in our tests, the beryllium atom, has four electrons and one heavy particle—the Be nucleus. In the present work we decided to only consider the $^{\infty}$Be atom, where the nucleus mass is set to infinity, because such a system was studied before with very high accuracy by Komasa and Rychlewski using up to 1200 ECGs.[4] But we should note that we could have easily set the Be nucleus mass to a finite value as our formalism allows us to do that. From the purely technical point of view the actual value of the mass of the nucleus does not have any significant effect on the calculations and the amount of computational work remains essentially the same regardless whether the mass is finite or infinite.

The results of our calculations on the $^{\infty}$Be atom versus the corresponding values taken from the work of Komasa and Rychlewski[4] are presented in Table II. As one can see from the table, for larger basis sizes, the ratio of the number of basis functions that was necessary to reach a similar accuracy in the case of the cyclic optimization (performed in the work of Komasa and Rychlewski[4]) and in the case of the simultaneous optimization of all basis functions using the analytic gradient (performed in our calculations) is approximately 1.5, with our number of basis functions being lower. With the basis set of 800 Gaussians, we already managed to improve the best energy obtained by Komasa and Rychlewski[4] with 1200 basis functions. Our new variational upper bound for the 2 ^1P state of Be is −14.473 442 537 a.u. while their energy was −14.473 442 016 a.u. Though this is only a marginal improvement it is significant that it has been obtained with much fewer functions (800 versus 1200). One notices that in the case of the Be atom the difference in the number of basis functions needed to reach a certain accuracy is not as striking as it was in the case of the Ps$_2$ calculations. It indicates that Komasa and Rychlewski[4] put a lot of effort in the cyclic optimization of their basis sets.

ACKNOWLEDGMENTS

This work has been supported by the National Science Foundation. We would also like to thank University of Arizona Center of Computing and Information Technology for using their supercomputer resources.

[1] S. Bubin and L. Adamowicz, J. Chem. Phys. **124**, 224317 (2006).
[2] S. Bubin, M. Stanke, D. Kędziera, and L. Adamowicz, Phys. Rev. A **76**, 022512 (2007).
[3] Y. Suzuki and K. Varga, *Stochastic Variational Approach to Quantum-Mechanical Few-Body Problems*, Lecture Notes in Physics (Springer,

Berlin, 1998).
[4] J. Komasa and J. Rychlewski, Chem. Phys. Lett. **342**, 185 (2001).
[5] J. R. Magnus and H. Neudecker, *Matrix-Differential Calculus with Applications in Statistics and Econometrics* (Wiley, Chichester, 1988).
[6] D. B. Kinghorn, Int. J. Quantum Chem. **57**, 141 (1996).
[7] D. B. Kinghorn and L. Adamowicz, J. Chem. Phys. **110**, 7166 (1999).
[8] S. Bubin, M. Cafiero, and L. Adamowicz, Adv. Chem. Phys. **131**, 377 (2005).
[9] M. Hamermesh, *Group Theory and Its Application to Physical Problems* (Dover, New York, 1962).
[10] V. I. Kukulin and V. M. Krasnopol'sky, J. Phys. G **3**, 795 (1977).
[11] K. Varga, J. Usukura, and Y. Suzuki, Phys. Rev. Lett. **80**, 1876 (1998).
[12] J. Usukura, K. Varga, and Y. Suzuki, Phys. Rev. A **58**, 1918 (1998).
[13] E. A. Hylleraas and A. Ore, Phys. Rev. **71**, 493 (1947).