

Matrix elements of N -particle explicitly correlated Gaussian basis functions with complex exponential parameters

Sergiy Bubin^{a)} and Ludwik Adamowicz*Department of Physics, University of Arizona, Tucson, Arizona 85721 and Department of Chemistry, University of Arizona, Tucson, Arizona 85721*

(Received 29 September 2005; accepted 19 April 2006; published online 13 June 2006)

In this work we present analytical expressions for Hamiltonian matrix elements with spherically symmetric, explicitly correlated Gaussian basis functions with complex exponential parameters for an arbitrary number of particles. The expressions are derived using the formalism of matrix differential calculus. In addition, we present expressions for the energy gradient that includes derivatives of the Hamiltonian integrals with respect to the exponential parameters. The gradient is used in the variational optimization of the parameters. All the expressions are presented in the matrix form suitable for both numerical implementation and theoretical analysis. The energy and gradient formulas have been programmed and used to calculate ground and excited states of the He atom using an approach that does not involve the Born-Oppenheimer approximation.

© 2006 American Institute of Physics. [DOI: [10.1063/1.2204605](https://doi.org/10.1063/1.2204605)]

I. INTRODUCTION

Explicitly correlated Gaussian functions have become quite popular in variational calculations of various few-particle quantum mechanical systems in the last several decades. The use of these functions represents a relatively simple yet very powerful tool for obtaining highly accurate solutions of the Schrödinger equation. The advantage of using Gaussians in comparison with other basis functions is that all basic matrix elements with Gaussians can be evaluated analytically for an arbitrary number of particles. At the same time, for most other explicitly correlated basis functions such a procedure often leads to rather complicated integrals that are either very hard or impossible to evaluate and implement numerically once one goes beyond systems containing three to four particles. However, simple explicitly correlated Gaussians also have certain shortcomings. One of them is the difficulty to accurately describe the electron-nucleus cusp.

The major research thrust in our group for more than a decade has been to use the explicitly correlated Gaussians in very accurate variational calculations of atomic and molecular systems without assuming the Born-Oppenheimer approximation regarding the separability of the electronic and nuclear motions. When the nuclei and electrons are treated on the same footing the Hamiltonian of the system is isotropic (i.e., fully symmetric with respect to rotations of the internal coordinate system) and the wave functions are eigenfunctions of the group of three-dimensional (3D) rotations. In representing such wave functions one needs to use basis functions that have the symmetry consistent with the symmetry of the Hamiltonian and, at the same time, are capable of describing the highly correlated motions of the electrons and the nuclei. For example, when one considers an atomic system in excited states, the basis functions need to describe

the nodes of the electronic component of the wave function whose number increases with the level of the excitations.

In molecular calculations the situation becomes even more complicated than in atoms. Non-Born-Oppenheimer (BO) calculations of, for example, pure vibrational states require that the basis functions accurately represent the nodes in the wave function that occur along the relative nucleus-nucleus coordinates. Since nuclei are highly localized around their equilibrium positions and they are much heavier than the electrons, the vibrational component of the total wave function has sharp peaks. In a diatomic system described in an internal coordinate frame with one of the nuclei at the coordinate origin (this is the frame we use) those peaks are shifted away from the origin. However, simple one-center Gaussians have their peaks at the origin, where the vibrational component of the wave function has to have an extremely low amplitude. We have proposed different ways to overcome this difficulty. The first is to use shifted Gaussians.¹⁻³ Even though this approach is applicable in some cases, the problem here is that the shifted Gaussians are not eigenfunctions of the total angular momentum operator. Another way is to modify Gaussians by including pre-multipliers in the form of powers of the internuclear distances.^{4,5} This works quite well in the case of diatomic systems.⁶⁻¹² However, the generalization of this approach to triatomic systems encounters serious problems in obtaining compact analytical expressions for matrix elements and not less serious problems in practical implementation.¹³

In this work we consider yet another way of describing the radial nodes in the ground and excited state wave functions of atomic and molecular systems calculated without the use of the Born-Oppenheimer approximation. The approach is based on the use of explicitly correlated Gaussian functions with complex exponential parameters. Such functions are much more flexible than real Gaussians and, put together in certain combinations (to simulate periodic functions such

^{a)}Electronic mail: bubin@email.arizona.edu

as sin or cos), they should be capable of providing a good description of the radial oscillations in the wave function in atomic and molecular excited states. Another application of complex Gaussians may be calculations of quantities whose matrix elements are difficult to evaluate with Gaussians containing premultipliers in the form of powers of internuclear distances. For example, the nucleus-electron correlation function, unlike the nucleus-nucleus correlation function,¹⁴ cannot be easily computed with the basis functions with the premultipliers, but it should be much easier to calculate using the complex Gaussians.

Despite the simplicity of the idea of using complex parameters in Gaussians, it has received almost no attention in the literature. The use of complex parameters has been done in calculations of three- and four-particle systems with the Slater-type basis functions.¹⁵⁻¹⁹ As far as complex Gaussians are concerned, except for one simple calculation of a few four-particle systems,²⁰ there has not been much more done in this area, according to our knowledge. Our work aims to fill this gap.

In this work we first derive the analytical expressions for the Hamiltonian matrix elements needed in N -particle variational calculations with complex Gaussians. Next we derive the expressions for calculating the analytical gradient of the energy with respect to the exponential parameters involved in the Gaussian functions. The gradient is needed to accelerate the variational optimization of the Gaussian exponential parameters. Lastly, we show a numerical example concerning calculations of the ground and excited states of the He atom.

II. HAMILTONIAN

Let us consider a system of N nonrelativistic particles with Coulomb interaction. If \mathbf{R}_i , M_i , and Q_i are the position in the laboratory frame, the mass, and the charge of i th particle, respectively, then the Hamiltonian of such a system has the following form:

$$\hat{H}_{\text{tot}} = - \sum_{i=1}^N \frac{1}{2M_i} \nabla_{\mathbf{R}_i}^2 + \sum_{i=1}^N \sum_{j>i}^N \frac{Q_i Q_j}{R_{ij}}, \quad (1)$$

where $\nabla_{\mathbf{R}_i}$ is the gradient with respect to \mathbf{R}_i , and $R_{ij} = |\mathbf{R}_j - \mathbf{R}_i|$ are interparticle distances. To separate out the translational motion of the center of mass from Hamiltonian (1) we place a particle, let us say particle 1, usually with the heaviest mass, at the origin of the internal coordinate system and we refer the other particles to that center particle using the position vectors $\mathbf{r}_i = \mathbf{R}_{i+1} - \mathbf{R}_1$. The direct and inverse coordinate transforms are given by the following expressions:

$$\begin{aligned} \mathbf{r}_0 &= \frac{M_1}{m_0} \mathbf{R}_1 + \frac{M_2}{m_0} \mathbf{R}_2 + \cdots + \frac{M_N}{m_0} \mathbf{R}_N, \\ \mathbf{r}_1 &= -\mathbf{R}_1 + \mathbf{R}_2, \\ \mathbf{r}_2 &= -\mathbf{R}_1 + \mathbf{R}_3, \\ &\vdots \end{aligned} \quad (2)$$

$$\mathbf{r}_n = -\mathbf{R}_1 + \mathbf{R}_N,$$

$$\begin{aligned} \mathbf{R}_1 &= \mathbf{r}_0 - \frac{M_2}{m_0} \mathbf{r}_1 - \frac{M_3}{m_0} \mathbf{r}_2 - \cdots - \frac{M_N}{m_0} \mathbf{r}_n, \\ \mathbf{R}_2 &= \mathbf{r}_0 + \left(1 - \frac{M_2}{m_0}\right) \mathbf{r}_1 - \frac{M_3}{m_0} \mathbf{r}_2 - \cdots - \frac{M_N}{m_0} \mathbf{r}_n, \\ \mathbf{R}_3 &= \mathbf{r}_0 - \frac{M_2}{m_0} \mathbf{r}_1 + \left(1 - \frac{M_3}{m_0}\right) \mathbf{r}_2 - \cdots - \frac{M_N}{m_0} \mathbf{r}_n, \\ &\vdots \\ \mathbf{R}_N &= \mathbf{r}_0 - \frac{M_2}{m_0} \mathbf{r}_1 - \frac{M_3}{m_0} \mathbf{r}_2 - \cdots + \left(1 - \frac{M_N}{m_0}\right) \mathbf{r}_n. \end{aligned} \quad (3)$$

Here, \mathbf{r}_0 is the position of the center of mass, and $m_0 = \sum_{i=1}^N M_i$ is the total mass of the system. By using internal coordinates we reduce an N -particle problem to an n -pseudoparticle one with $n \equiv N-1$. The resulting internal Hamiltonian is given by

$$\hat{H} = -\frac{1}{2} \left(\sum_{i=1}^n \frac{1}{m_i} \nabla_{\mathbf{r}_i}^2 + \sum_{i \neq j}^n \frac{1}{M_1} \nabla_{\mathbf{r}_i} \nabla_{\mathbf{r}_j} \right) + \sum_{i=1}^n \frac{q_0 q_i}{r_i} + \sum_{i < j}^n \frac{q_i q_j}{r_{ij}}, \quad (4)$$

where the prime stands for matrix-vector transpose, and $r_{ij} = |\mathbf{r}_j - \mathbf{r}_i|$. The masses and charges of the pseudoparticles are defined as $m_i = M_1 M_{i+1} / (M_1 + M_{i+1})$ and $q_i = Q_{i+1}$, respectively. For more details on separating the center of mass motion see Refs. 5, 21, and 22.

We can combine all pseudoparticle positions and the gradients with respect to these positions in n -component vectors

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

The internal Hamiltonian is then conveniently written in the matrix form

$$\hat{H} = -\nabla_{\mathbf{r}}' M \nabla_{\mathbf{r}} + \sum_{i=1}^n \frac{q_0 q_i}{r_i} + \sum_{i < j}^n \frac{q_i q_j}{r_{ij}}, \quad (6)$$

where M is an $n \times n$ matrix with $1/2m_1, 1/2m_2, \dots, 1/2m_n$ on the diagonal and $1/M_1$ for all off-diagonal elements. Strictly speaking, in place of M in expression (6) we should put $M \otimes I_3$, where \otimes denotes the Kronecker product and I_3 is a 3×3 identity matrix. However, for the sake of simplicity, we will skip writing it, as it will be explained in the next section.

III. NOTATIONS

Before we proceed to matrix elements, we need to provide certain definitions and introduce appropriate notations. The formalism of matrix differential calculus will be extensively used throughout the paper. It has been proven to be a

very convenient tool for derivations of matrix elements with different kinds of Gaussian basis functions.^{5,22,23} But, unfortunately, it is rarely applied in physics or chemistry. To avoid possible confusion we will briefly review some basic concepts. For a detailed introduction to this subject we can recommend Ref. 24.

In our problem we deal with two kinds of vectors. First are common three-dimensional vectors representing components of different quantities. The other ones are n -component vectors in particle space, where each component corresponds to a quantity or an operator associated with a certain particle. Since each of those n components is often a three-dimensional vector itself, effectively we have $3n$ components. In mathematical terms it is a direct product of two spaces. As a result of this fact, we need to use the Kronecker products throughout. For example, a vector-matrix-vector product may look as follows:

$$\mathbf{r}'(A \otimes I_3)\mathbf{r} = (x_1, y_1, z_1, x_2, \dots, z_n) \times \begin{pmatrix} A_{11} & 0 & 0 & A_{12} & \cdots & 0 \\ 0 & A_{11} & 0 & 0 & \cdots & 0 \\ 0 & 0 & A_{11} & 0 & \cdots & A_{1n} \\ A_{21} & 0 & 0 & A_{22} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & A_{n1} & 0 & \cdots & A_{nn} \end{pmatrix} \times \begin{pmatrix} x_1 \\ y_1 \\ z_1 \\ x_2 \\ \vdots \\ z_n \end{pmatrix}. \quad (7)$$

Here, A is an $n \times n$ matrix, and A_{ij} are its elements. Writing routinely the Kronecker product is quite tedious and at the same time obvious. To simplify the notations we will use the following matrix-vector multiplication convention instead:

$$\mathbf{r}'A\mathbf{r} = (\mathbf{r}'_1, \mathbf{r}'_2, \dots, \mathbf{r}'_n) \begin{pmatrix} A_{11}\mathbf{r}_1 + A_{12}\mathbf{r}_2 + \cdots + A_{1n}\mathbf{r}_n \\ A_{21}\mathbf{r}_1 + A_{22}\mathbf{r}_2 + \cdots + A_{2n}\mathbf{r}_n \\ \vdots \\ A_{n1}\mathbf{r}_1 + A_{n2}\mathbf{r}_2 + \cdots + A_{nn}\mathbf{r}_n \end{pmatrix} = A_{11}\mathbf{r}'_1\mathbf{r}_1 + A_{12}\mathbf{r}'_1\mathbf{r}_2 + \cdots + A_{nn}\mathbf{r}'_n\mathbf{r}_n, \quad (8)$$

where $\mathbf{r}'_i\mathbf{r}_j = x_i x_j + y_i y_j + z_i z_j$. Thus, we first carry out all matrix-vector multiplications in n -dimensional particle space (or in any other space we deal with) treating three-dimensional vectors \mathbf{r}_i as numbers and then, at the end, we replace each product $\mathbf{r}'_i\mathbf{r}_j$ with a dot product. In other words, everywhere a product of two quantities denoting three-dimensional vectors appears it should be understood as a dot product, unless a special notice is made.

To keep all the derivations clear we will be using the following notations wherever it is appropriate:

- α, β, ξ , etc.—lowercase Greek letters for scalars.
- a, b, x , etc.—lowercase Latin letters for column vectors in particle space. These vectors have n components.
- A, B, X , etc.—uppercase Latin letters for matrices in particle space. These are $n \times n$ matrices.
- $\boldsymbol{\alpha}, \boldsymbol{\beta}, \mathbf{a}_i, \mathbf{b}_j$, etc.—bold font for three-dimensional vectors.
- X', x' , etc.—the prime stands for matrix (vector) transpose.
- X^*, x^*, ξ^* , etc.—the star stands for the complex conjugate. If applied to a matrix (vector) then there is no transposition assumed, i.e., the result is a matrix (vector) whose elements are complex conjugated elements of the initial matrix (vector).
- X^\dagger, x^\dagger , etc.—the dagger stands for matrix (vector) transpose followed by complex conjugation.
- X^{-1} —stands for the inverse matrix.
- $\text{diag } X$ —stands for a matrix whose diagonal elements are equal to the corresponding diagonal elements of X and all off-diagonal elements are zeros.
- $|X|$ —vertical bars stand for the determinant of a matrix. However, if the object in between vertical bars is a three-dimensional vector or a scalar then the absolute value of the vector (scalar) is assumed.
- $\text{tr } X$ —stands for the trace of matrix X .
- $X \otimes Y$ —stands for the Kronecker product of matrices X and Y .

For some quantities the above mentioned notation scheme will not be valid, but there should not be any confusion in those cases.

We also want to introduce two operators, vec and vech . They both transform a matrix into a vector. The vec operator stacks the columns of a matrix, one underneath the other. If X is an $n \times n$ matrix with elements X_{ij} , then $\text{vec } X$ is the following n^2 -component vector:

$$\text{vec } X = \begin{pmatrix} X_{11} \\ X_{12} \\ \vdots \\ X_{1n} \\ X_{21} \\ \vdots \\ X_{nn} \end{pmatrix}. \quad (9)$$

Unlike vec , the vech operator takes only the part of a column, that is beneath the diagonal, including the diagonal element. Hence, it results in a $n(n+1)/2$ -component vector. For example, if X is a 3×3 matrix, then

$$\text{vech } X = \begin{pmatrix} X_{11} \\ X_{12} \\ X_{13} \\ X_{22} \\ X_{23} \\ X_{33} \end{pmatrix}. \quad (10)$$

In the case when matrix X is symmetric, $\text{vech } X$ contains only its independent elements.

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

$$(\text{vec } X)' \text{vec } Y = \text{tr}[X'Y], \quad (11)$$

$$(\text{vec } X)' \text{vec } L = (\text{vech } X)' \text{vech } L. \quad (12)$$

For the case when both X and Y are complex symmetric or antisymmetric we have the following relationship:

$$(\text{vec } X)' \text{vec } Y = (\text{vech}(2X - \text{diag } X))' \text{vech } Y. \quad (13)$$

IV. COMPLEX GAUSSIAN BASIS FUNCTIONS

The basis functions we consider in this paper are explicitly correlated Gaussian functions with complex parameters. The general form of such functions is given by

$$\phi_k = \exp[-\mathbf{r}'(A_k + iB_k)\mathbf{r}] \equiv \exp[-\mathbf{r}'C_k\mathbf{r}]. \quad (14)$$

Here A_k and B_k are $n \times n$ real matrices. Sometimes it is convenient to use one complex matrix instead, which we will denote C_k . The lower index k reflects the fact that these matrices are unique for each basis function.

Although neither of the matrices has to be symmetric, one can always rearrange their elements, without changing the quadratic form itself, in such a way that they will be symmetric. This reduces the number of independent parameters and throughout the paper we will be assuming that all A_k 's, B_k 's, and C_k 's are symmetric. Quantum mechanics requires the total wave function to be finite and square integrable. This imposes a restriction on A_k . It must be positive definite. Notice, however, that the elements of B_k may take any values. To assure the positive definiteness of A_k it is very convenient to represent it as a Cholesky decomposition, $A_k = L_k L_k'$, with L_k being a lower triangular matrix. The advantage of such a representation is that there are no restrictions on the elements of L_k . Hence, if the numerical optimization of the elements is to be performed, we can safely use very efficient algorithms for a nonconstrained optimization.

V. PERMUTATIONAL SYMMETRY

Basis functions used in variational calculations of quantum systems containing identical particles should (in many cases must) possess certain permutational symmetry. This is usually done by applying some linear combination of permutational operators to primitive basis functions that, in general, do not have any specific symmetry with respect to permutations of particles. Each permutational operator may be represented as a product of transpositions (i.e., pair permu-

tations). Below we would like to mention how these operators or transpositions act on primitive complex Gaussian basis functions (14). The permutations of real particles induce transformations on internal coordinates, \mathbf{r}_i . These transformations are linear, as the connection between the laboratory coordinates, \mathbf{R}_i , and the internal coordinates is linear. In the case of complex Gaussian basis functions, the action of a permutational operator \hat{P} can be conveniently represented by the corresponding permutational matrix, P :

$$\begin{aligned} \hat{P} \exp[-\mathbf{r}'C_l\mathbf{r}] &= \exp[-(\mathbf{P}\mathbf{r})'C_l(\mathbf{P}\mathbf{r})] \\ &= \exp[-\mathbf{r}'(P'C_lP)\mathbf{r}]. \end{aligned} \quad (15)$$

As one can see, with the action of the permutational operator the matrix of nonlinear parameters, C_l , undergoes a similarity transformation. It should be noted that matrix P is not in general an elementary transformation matrix. The explicit form of P depends on the operator \hat{P} . In the case of a transposition it has a very simple view. If the transposition involves particle 1 (i.e., the reference particle) and j , then the corresponding matrix P for internal coordinate transformation is an $n \times n$ identity matrix with all elements in the j th column replaced by -1 . For transpositions that involve particles i and j (and neither one is particle 1) the matrix is the identity matrix with interchanged columns $(i-1)$ and $(j-1)$, as well as rows $(i-1)$ and $(j-1)$.

The important consequence of the aforementioned is the fact that the calculation of all matrix elements with complex Gaussians affected by permutational operators is not different from the case when the Gaussians are not affected. One just needs to transform the matrices of nonlinear variational parameters, C_l , accordingly.

As a rule, the operators whose matrix elements are required in variational calculations of a quantum mechanical system commute with all the permutations from the permutational symmetry group of that system. Due to this fact the matrix elements of an operator \hat{O} can be written in the following way:

$$\langle \phi_k | \hat{P}_\alpha \hat{O} \hat{P}_\beta | \phi_l \rangle = \langle \phi_k | \hat{O} \hat{P}_\gamma | \phi_l \rangle, \quad (16)$$

where \hat{P}_α , \hat{P}_β , and \hat{P}_γ are some permutations. Thus, in our derivations we restrict ourselves to the case when only the ket is affected by a permutation operator.

VI. MATRIX ELEMENTS

To begin let us first give the value of a well-known Gaussian integral, which will be used in the derivations below:

$$\int_{-\infty}^{+\infty} \exp[-x'Cx + y'y] dx = \frac{\pi^{n/2}}{|C|^{1/2}} \exp\left[\frac{1}{4}y'C^{-1}y\right]. \quad (17)$$

The integration here is over n variables, and x is an n -component vector of these variables. y is a constant vector, $n \times n$ matrix C is assumed to be symmetric, and its real part is positive definite. Also, here and everywhere below, by the square root one should understand its principal value (i.e., that root whose real part is greater than zero).

The following notations will help keep expressions more compact:

$$|\tilde{\phi}_l\rangle = \hat{P}|\phi_l\rangle, \quad (18)$$

$$\tilde{L}_k = P'L_k, \quad \tilde{A}_k = P'A_kP, \quad \tilde{B}_k = P'B_kP, \quad (19)$$

$$A_{kl} = A_k + A_l, \quad \tilde{A}_{kl} = A_k + \tilde{A}_l, \quad (20)$$

$$B_{kl} = -B_k + B_l, \quad \tilde{B}_{kl} = -B_k + \tilde{B}_l, \quad (21)$$

and

$$\tilde{C}_{kl} = C_k^* + \tilde{C}_l = \tilde{A}_{kl} + i\tilde{B}_{kl}. \quad (22)$$

Here P represents the permutation matrix corresponding to some permutation operator \hat{P} .

A. Overlap

The overlap integral follows directly from formula (17). We just need to keep in mind that the integration is done over $3n$ variables:

$$\begin{aligned} \langle \phi_k | \tilde{\phi}_l \rangle &\equiv \langle \phi_k | \hat{P} | \phi_l \rangle \\ &= \int_{-\infty}^{+\infty} \exp[-\mathbf{r}' C_k^* \mathbf{r}] \hat{P} \exp[-\mathbf{r}' C_l \mathbf{r}] d\mathbf{r} \\ &= \int_{-\infty}^{+\infty} \exp[-\mathbf{r}' C_k^* \mathbf{r}] \exp[-\mathbf{r}' \tilde{C}_l \mathbf{r}] d\mathbf{r} = \frac{\pi^{3n/2}}{|\tilde{C}_{kl}|^{3/2}}. \end{aligned}$$

For the overlap matrix elements with normalized basis functions, which are superior in practical calculations, we have

$$\begin{aligned} S_{kl} &\equiv \frac{\langle \phi_k | \tilde{\phi}_l \rangle}{(\langle \phi_k | \phi_k \rangle \langle \phi_l | \phi_l \rangle)^{1/2}} = \frac{(|C_{kk}|^{3/2} |C_{ll}|^{3/2})^{1/2}}{|\tilde{C}_{kl}|^{3/2}} \\ &= 2^{3n/2} \left(\frac{\|L_k\| \|L_l\|}{|\tilde{C}_{kl}|} \right)^{3/2}. \quad (23) \end{aligned}$$

B. Kinetic energy

As one can easily verify, the action of the $\nabla_{\mathbf{r}}$ operator on the basis functions is as follows:

$$\nabla_{\mathbf{r}} \phi_k = -2 \phi_k C_k \mathbf{r}. \quad (24)$$

Using this result we can obtain the following expression for the kinetic energy operator matrix elements:

$$\begin{aligned} \langle \phi_k | -\nabla_{\mathbf{r}}' M \nabla_{\mathbf{r}} | \tilde{\phi}_l \rangle &= \langle \nabla_{\mathbf{r}} \phi_k | M | \nabla_{\mathbf{r}} \tilde{\phi}_l \rangle \\ &= 4 \langle C_k \mathbf{r} \phi_k | M | \tilde{C}_l \mathbf{r} \tilde{\phi}_l \rangle \\ &= 4 \langle \phi_k | \mathbf{r}' C_k^\dagger M \tilde{C}_l \mathbf{r} | \tilde{\phi}_l \rangle \\ &= 4 \langle \phi_k | \text{tr}[\mathbf{r} \mathbf{r}' C_k^\dagger M \tilde{C}_l] | \tilde{\phi}_l \rangle \\ &= 4 \langle \phi_k | \text{vec}(\mathbf{r} \mathbf{r}')' | \tilde{\phi}_l \rangle \text{vec}(C_k^\dagger M \tilde{C}_l). \quad (25) \end{aligned}$$

The first factor here is

$$\begin{aligned} \langle \phi_k | \text{vec}(\mathbf{r} \mathbf{r}')' | \tilde{\phi}_l \rangle &= \int_{-\infty}^{\infty} \phi_k^* \text{vec}(\mathbf{r} \mathbf{r}')' \tilde{\phi}_l d\mathbf{r} = \int_{-\infty}^{\infty} -\frac{\partial \phi_k^* \tilde{\phi}_l}{\partial \text{vec}(\tilde{C}_{kl})'} d\mathbf{r} = -\frac{\partial}{\partial \text{vec}(\tilde{C}_{kl})'} \int_{-\infty}^{\infty} \phi_k^* \tilde{\phi}_l d\mathbf{r} = -\frac{\partial}{\partial \text{vec}(\tilde{C}_{kl})'} \frac{\pi^{3n/2}}{|\tilde{C}_{kl}|^{3/2}} \\ &= \frac{3}{2} \frac{\pi^{3n/2}}{|\tilde{C}_{kl}|^{5/2}} \frac{\partial |\tilde{C}_{kl}|}{\partial \text{vec}(\tilde{C}_{kl})'} = \frac{3}{2} \frac{\pi^{3n/2}}{|\tilde{C}_{kl}|^{5/2}} |\tilde{C}_{kl}| \text{vec}(\tilde{C}_{kl}^{-1})' = \frac{3}{2} \langle \phi_k | \tilde{\phi}_l \rangle \text{vec}(\tilde{C}_{kl}^{-1})'. \quad (26) \end{aligned}$$

Above we used the facts that for an arbitrary matrix X ,

$$d|X| = |X| \text{tr}[X^{-1} dX] = |X| \text{vec}((X^{-1})')' \text{vec}(dX), \quad (27)$$

$$\frac{\partial |X|}{\partial \text{vec}(X)'} = |X| \text{vec}((X^{-1})')', \quad (28)$$

and that matrix \tilde{C}_{kl}^{-1} is symmetric.

Thus,

$$\begin{aligned} \langle \phi_k | -\nabla_{\mathbf{r}}' M \nabla_{\mathbf{r}} | \tilde{\phi}_l \rangle &= 6 \langle \phi_k | \tilde{\phi}_l \rangle \text{vec}(\tilde{C}_{kl}^{-1})' \text{vec}(C_k^\dagger M \tilde{C}_l) \\ &= 6 \langle \phi_k | \tilde{\phi}_l \rangle \text{tr}[\tilde{C}_{kl}^{-1} C_k^\dagger M \tilde{C}_l] \\ &= 6 \langle \phi_k | \tilde{\phi}_l \rangle \text{tr}[M \tilde{C}_l \tilde{C}_{kl}^{-1} C_k^\dagger]. \quad (29) \end{aligned}$$

In the case of normalized basis functions the kinetic energy matrix element becomes

$$T_{kl} = 6 S_{kl} \text{tr}[M \tilde{C}_l \tilde{C}_{kl}^{-1} C_k^\dagger]. \quad (30)$$

C. Pair correlation function and potential energy

In the beginning let us obtain an auxiliary formula, which will be very useful in further derivations. The formula is for the matrix element of the three-dimensional Dirac delta function

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

where a is a real n -component vector, and $\boldsymbol{\xi}$ is some real three-dimensional parameter.

Using the following representation of the delta function,

$$\delta(a'\mathbf{r} - \boldsymbol{\xi}) = \lim_{\beta \rightarrow \infty} \left(\frac{\beta}{\pi} \right)^{3/2} \exp[-\beta(a'\mathbf{r} - \boldsymbol{\xi})^2], \quad (31)$$

and formula (17) we have

$$\begin{aligned} \langle \phi_k | \delta(a'\mathbf{r} - \boldsymbol{\xi}) | \tilde{\phi}_l \rangle &= \lim_{\beta \rightarrow \infty} \left(\frac{\beta}{\pi} \right)^{3/2} \langle \phi_k | \exp[-\beta \mathbf{r}'(aa')\mathbf{r} \\ &\quad + 2\beta a'\mathbf{r}\boldsymbol{\xi} - \beta \boldsymbol{\xi}^2] | \tilde{\phi}_l \rangle \\ &= \lim_{\beta \rightarrow \infty} \left(\frac{\beta}{\pi} \right)^{3/2} \frac{\pi^{3n/2}}{|\tilde{C}_{kl} + \beta aa'|^{3/2}} \\ &\quad \times \langle \phi_k | \exp[-\beta^2 a'(\tilde{C}_{kl} + \beta aa')^{-1} a \boldsymbol{\xi}^2 \\ &\quad - \beta \boldsymbol{\xi}^2] | \tilde{\phi}_l \rangle. \end{aligned} \quad (32)$$

Since aa' is a rank one matrix we can write the determinant in the last formula as

$$\begin{aligned} |\tilde{C}_{kl} + \beta aa'| &= |\tilde{C}_{kl}| |I + \beta \tilde{C}_{kl}^{-1} aa'| \\ &= |\tilde{C}_{kl}| (1 + \beta \text{tr}[\tilde{C}_{kl}^{-1} aa']), \end{aligned} \quad (33)$$

where I is $n \times n$ identity matrix. Indeed, matrix $\tilde{C}_{kl}^{-1} aa'$ has rank one and its only nonzero eigenvalue is $\text{tr}[\tilde{C}_{kl}^{-1} aa']$, since the trace of a matrix is the sum of its eigenvalues. Adding the identity matrix just increases all the eigenvalues by 1. Hence, the eigenvalues of $I + \beta \tilde{C}_{kl}^{-1} aa'$ are $1 + \beta \text{tr}[\tilde{C}_{kl}^{-1} aa']$ and 1. Since the determinant is the product of all eigenvalues, we have what is given in (33).

As the limit of the preexponential part in (32) is a finite number, the limit of the expression in the exponent must be equal to $-\gamma \boldsymbol{\xi}^2$, with γ being a finite number. Otherwise the entire expression (32) would have had its limit equal to either zero or infinity, which is not the case. Hence,

$$\begin{aligned} \langle \phi_k | \delta(a'\mathbf{r} - \boldsymbol{\xi}) | \tilde{\phi}_l \rangle &= \frac{\pi^{3(n-1)/2}}{|\tilde{C}_{kl}|^{3/2} \text{tr}[\tilde{C}_{kl}^{-1} aa']^{3/2}} \exp[-\gamma \boldsymbol{\xi}^2] \\ &= \langle \phi_k | \tilde{\phi}_l \rangle \frac{1}{\pi^{3/2} \text{tr}[\tilde{C}_{kl}^{-1} aa']^{3/2}} \\ &\quad \times \exp[-\gamma \boldsymbol{\xi}^2]. \end{aligned} \quad (34)$$

Making use of the normalization condition,

$$\int_{-\infty}^{\infty} \langle \phi_k | \delta(a'\mathbf{r} - \boldsymbol{\xi}) | \tilde{\phi}_l \rangle d\boldsymbol{\xi} = \langle \phi_k | \tilde{\phi}_l \rangle, \quad (35)$$

we can find that $\gamma = \text{tr}[\tilde{C}_{kl}^{-1} aa']$. Thus,

$$\begin{aligned} \langle \phi_k | \delta(a'\mathbf{r} - \boldsymbol{\xi}) | \tilde{\phi}_l \rangle &= \langle \phi_k | \tilde{\phi}_l \rangle \frac{1}{\pi^{3/2} \text{tr}[\tilde{C}_{kl}^{-1} aa']^{3/2}} \\ &\quad \times \exp\left[-\frac{\boldsymbol{\xi}^2}{\text{tr}[\tilde{C}_{kl}^{-1} aa']}\right]. \end{aligned} \quad (36)$$

Let us now derive the expression for the matrix elements of the pair correlation function,

$$g^{ij}(\boldsymbol{\xi}) \equiv \langle \psi | \delta(\mathbf{r}_{ij} - \boldsymbol{\xi}) | \psi \rangle, \quad (37)$$

$$g^i(\boldsymbol{\xi}) \equiv \langle \psi | \delta(\mathbf{r}_i - \boldsymbol{\xi}) | \psi \rangle, \quad (38)$$

which in case (37) is the probability density of pseudoparticles i and j , or, equivalently, particles $i+1$ and $j+1$ to have separation $\boldsymbol{\xi}$. Expression (38) is used when the pair correlation function of particles 1 and $i+1$ is needed. Here, ψ is the wave function of the system under consideration. Apart from being a useful quantity itself, the correlation functions (37) and (38) will serve as an auxiliary tool to obtain the expression for the potential energy and any other quantity that depends on \mathbf{r}_i or \mathbf{r}_{ij} . To help preserve all the derivations in the matrix form, we will introduce vector j^i whose only nonzero component is the i th one, which is equal to 1. Also, we will introduce a rank one matrix J_{ij} . Its definition is as follows:

$$J_{ij} = \begin{cases} E_{ii} & \text{if } i = j \\ E_{ii} + E_{jj} - E_{ij} - E_{ji} & \text{if } i \neq j, \end{cases} \quad (39)$$

where E_{ij} is a matrix with 1 in the i, j th position and 0's elsewhere.

Noticing that $\mathbf{r}_{ij} = (j^j - j^i)'\mathbf{r}$, $\mathbf{r}_i = (j^i)'\mathbf{r}$, and also $(j^j - j^i) \times (j^j - j^i)' = J_{ij}$ for $i \neq j$, $j^i(j^i)' = J_{ii}$ we can easily conclude that

$$\begin{aligned} \langle \phi_k | \delta(\mathbf{r}_{ij} - \boldsymbol{\xi}) | \tilde{\phi}_l \rangle &= \langle \phi_k | \tilde{\phi}_l \rangle \frac{1}{\pi^{3/2}} \frac{1}{\text{tr}[\tilde{C}_{kl}^{-1} J_{ij}]^{3/2}} \\ &\quad \times \exp\left[-\frac{\boldsymbol{\xi}^2}{\text{tr}[\tilde{C}_{kl}^{-1} J_{ij}]}\right], \end{aligned} \quad (40)$$

and

$$\begin{aligned} \langle \phi_k | \delta(\mathbf{r}_i - \boldsymbol{\xi}) | \tilde{\phi}_l \rangle &= \langle \phi_k | \tilde{\phi}_l \rangle \frac{1}{\pi^{3/2}} \frac{1}{\text{tr}[\tilde{C}_{kl}^{-1} J_{ii}]^{3/2}} \\ &\quad \times \exp\left[-\frac{\boldsymbol{\xi}^2}{\text{tr}[\tilde{C}_{kl}^{-1} J_{ii}]}\right]. \end{aligned} \quad (41)$$

The above expressions for the matrix elements of the correlation function allow one to evaluate the matrix elements of an arbitrary function $\chi(\mathbf{r}_{ij})$:

$$\begin{aligned} \langle \phi_k | \chi(\mathbf{r}_{ij}) | \tilde{\phi}_l \rangle &= \int_{-\infty}^{\infty} \chi(\mathbf{r}_{ij}) \langle \phi_k | \delta(\mathbf{r}_{ij} - \boldsymbol{\xi}) | \tilde{\phi}_l \rangle d\boldsymbol{\xi} \\ &= \langle \phi_k | \tilde{\phi}_l \rangle \frac{1}{\pi^{3/2} \text{tr}[\tilde{C}_{kl}^{-1} J_{ij}]^{3/2}} \\ &\quad \times \int_{-\infty}^{\infty} \chi(\boldsymbol{\xi}) \exp\left[-\frac{\boldsymbol{\xi}^2}{\text{tr}[\tilde{C}_{kl}^{-1} J_{ij}]}\right] d\boldsymbol{\xi}. \end{aligned} \quad (42)$$

If $\chi(\mathbf{r}_{ij})$ depends only on the absolute value of the inter-pseudoparticle distance, this formula becomes

$$\langle \phi_k | \chi(r_{ij}) | \tilde{\phi}_l \rangle = \langle \phi_k | \tilde{\phi}_l \rangle \frac{4}{\sqrt{\pi} \text{tr}[\tilde{C}_{kl}^{-1} J_{ij}]} \frac{1}{\text{tr}[\tilde{C}_{kl}^{-1} J_{ij}]^{3/2}} \times \int_0^\infty \xi^2 \chi(\xi) \exp\left[-\frac{\xi^2}{\text{tr}[\tilde{C}_{kl}^{-1} J_{ij}]}\right] d\xi. \quad (43)$$

In particular, for $\chi(\xi) = \xi^\nu$, with ν being a real number, $\nu > -3$, using the value of the integral

$$\int_0^\infty \lambda^\kappa e^{-\alpha\lambda^2} d\lambda = \frac{1}{2} \frac{\Gamma((\kappa+1)/2)}{\alpha^{(\kappa+1)/2}}, \quad (44)$$

Re $\kappa > -1$, Re $\alpha > 0$,

where $\Gamma(\kappa)$ is the Euler gamma function, we obtain

$$\langle \phi_k | r_{ij}^\nu | \tilde{\phi}_l \rangle = \langle \phi_k | \tilde{\phi}_l \rangle \frac{2}{\sqrt{\pi}} \Gamma\left(\frac{\nu+3}{2}\right) \text{tr}[\tilde{C}_{kl}^{-1} J_{ij}]^{\nu/2}. \quad (45)$$

If $\nu = -1$, as is the case for the Coulomb potential, then

$$\langle \phi_k | \frac{1}{r_{ij}} | \tilde{\phi}_l \rangle = \langle \phi_k | \tilde{\phi}_l \rangle \frac{2}{\sqrt{\pi} \text{tr}[\tilde{C}_{kl}^{-1} J_{ij}]^{1/2}}, \quad (46)$$

and the corresponding expression with normalized basis functions is

$$R_{kl}^{ij} = \frac{2}{\sqrt{\pi}} S_{kl} \frac{1}{\text{tr}[\tilde{C}_{kl}^{-1} J_{ij}]^{1/2}}. \quad (47)$$

With this we can now write the expression for the potential energy matrix element, V_{kl} , as follows:

$$V_{kl} = \sum_{i=1}^n q_0 q_i R_{kl}^{ii} + \sum_{i<j}^n q_i q_j R_{kl}^{ij}, \quad (48)$$

or

$$V_{kl} = \frac{2S_{kl}}{\sqrt{\pi}} \left(\sum_{i=1}^n \frac{q_0 q_i}{\text{tr}[\tilde{C}_{kl}^{-1} J_{ii}]^{1/2}} + \sum_{i<j}^n \frac{q_i q_j}{\text{tr}[\tilde{C}_{kl}^{-1} J_{ij}]^{1/2}} \right). \quad (49)$$

For practical implementation it is useful to note that the trace of a product of an arbitrary matrix X and J_{ij} has a very simple form in terms of the elements of X :

$$\text{tr}[XJ_{ij}] = \begin{cases} (X)_{ii} & \text{if } i=j \\ (X)_{ii} + (X)_{jj} - (X)_{ij} - (X)_{ji} & \text{if } i \neq j. \end{cases} \quad (50)$$

D. Particle density

The particle densities are tools for analyzing quantum mechanical systems of few particles since the wave function for such systems cannot be visualized due to its multidimensional nature. Below we present the matrix elements of this quantity evaluated with complex Gaussians.

If a system is described by wave function ψ , then the density of particle i in the reference frame where the center of mass of the system, \mathbf{r}_0 , is at rest is given by the formula

$$\rho^i(\boldsymbol{\xi}) \equiv \langle \psi | \delta(\mathbf{R}_i - \mathbf{r}_0 - \boldsymbol{\xi}) | \psi \rangle. \quad (51)$$

Here, \mathbf{R}_i , expressed through internal coordinates, can be taken from (3). The calculation of the density using complex Gaussians requires the following matrix elements:

$$\langle \phi_k | \delta((b^i)' \mathbf{r} - \boldsymbol{\xi}) | \tilde{\phi}_l \rangle, \quad (52)$$

where b^i is an n -component vector whose elements in case when $i > 1$ are

$$(b^i)_j = -M_{j+1}/m_0, \quad j = 1, \dots, n, \quad j \neq i-1,$$

$$(b^i)_{i-1} = 1 - M_i/m_0.$$

If $i=1$, then the elements of b^i are

$$(b^1)_j = -M_{j+1}/m_0, \quad j = 1, \dots, n.$$

In order to evaluate (52) we make use of the previously derived relation (36):

$$\langle \phi_k | \delta((b^i)' \mathbf{r} - \boldsymbol{\xi}) | \tilde{\phi}_l \rangle = \langle \phi_k | \tilde{\phi}_l \rangle \frac{1}{\pi^{3/2}} \frac{1}{\text{tr}[\tilde{C}_{kl}^{-1} b^i (b^i)']^{3/2}} \times \exp\left[-\frac{\xi^2}{\text{tr}[\tilde{C}_{kl}^{-1} b^i (b^i)']}\right]. \quad (53)$$

In the case of the normalized density matrix elements the latter expression becomes

$$\rho_{kl}^i(\boldsymbol{\xi}) = \frac{1}{\pi^{3/2}} S_{kl} \frac{1}{\text{tr}[\tilde{C}_{kl}^{-1} b^i (b^i)']^{3/2}} \exp\left[-\frac{\xi^2}{\text{tr}[\tilde{C}_{kl}^{-1} b^i (b^i)']}\right]. \quad (54)$$

VII. ENERGY GRADIENT

The formulas we derived in the previous section are sufficient for carrying out energy calculations and finding the wave functions. In principle, it is possible to write a computer code that performs such a task using only these formulas. However, if we aim for high precision in calculations, then many basis functions have to be included and the nonlinear parameters contained in these functions (elements of matrices A_k and B_k) have to be optimized, as the accuracy of variational expansions in terms of correlated Gaussians strongly depends on the values of these parameters. The total number of nonlinear parameters may reach many thousands and their simultaneous optimization represents a very time consuming computational task. The use of the finite difference gradient of the energy as a function of nonlinear parameters is very costly and puts serious restrictions on the maximal number of basis functions. The utilization of the analytic gradient helps to relax this restriction and can speed up the calculation process by several orders of magnitude. Thus, the use of the analytic gradient is very important in many practical situations. In this section we shortly describe the procedure of constructing the analytic gradient of the energy functional and derive necessary matrix elements.

A. Constructing the energy gradient

In the Rayleigh-Ritz variational method the wave function of the system under consideration, $\psi(\mathbf{r})$, is approximated as a linear combination of \mathcal{K} basis functions φ_k ,

$$\psi(\mathbf{r}) = \sum_{k=1}^{\mathcal{K}} c_k \varphi_k(\mathbf{r}). \quad (55)$$

The upper bounds, ϵ , to the exact energies are then obtained from the solution of the secular equation

$$(H - \epsilon S)c = 0, \quad (56)$$

where H and S are $\mathcal{K} \times \mathcal{K}$ Hermitian matrices of the Hamiltonian and the overlap, respectively, whose elements are $H = \langle \varphi_k | \hat{H} | \varphi_l \rangle$ and $S = \langle \varphi_k | \varphi_l \rangle$, c is a \mathcal{K} -component complex vector with elements c_k . We will be assuming that ϵ and c correspond to a certain state (quantum number). This will keep the notations simple as we will not need to use any indices for ϵ and c to denote the quantum number(s).

Taking the differential of (56),

$$d[(H - \epsilon S)c] = (dH)c - (d\epsilon)Sc - \epsilon(dS)c + (H - \epsilon S)dc, \quad (57)$$

and multiplying by c^\dagger from the left we obtain

$$d\epsilon = c^\dagger(dH - \epsilon dS)c. \quad (58)$$

In the above expression we assumed that c is normalized in such a way that $c^\dagger S c = 1$. It is worth noting that relation (58) is essentially the same as the Hellmann-Feynman theorem. The differentiation in (58) may be performed with respect to any arbitrary variable, regardless of its nature. It does not

matter whether the quantity is a physical parameter, for instance, a charge, or a variable that characterizes the trial wave function, such as a nonlinear variational parameter(s).

We can rewrite (58) as

$$\begin{aligned} d\epsilon &= \text{tr}[c c^\dagger (dH - \epsilon dS)] \\ &= \text{vec}((c c^\dagger)')' \text{vec}(dH - \epsilon dS) \\ &= \text{vec}(c c^\dagger)^\dagger [d(\text{vec } H) - \epsilon d(\text{vec } S)]. \end{aligned} \quad (59)$$

Now, suppose that a is some vector containing nonlinear variational parameters. For generality, we will not specify how many nonlinear parameters there are (i.e., how many components vector a has) and how basis functions depend on them. The gradient of the energy with respect to these parameters is

$$\nabla_a \epsilon \equiv \frac{\partial \epsilon}{\partial a'} = \text{vec}(c c^\dagger)^\dagger \left(\frac{\partial \text{vec } H}{\partial a'} - \epsilon \frac{\partial \text{vec } S}{\partial a'} \right). \quad (60)$$

In practical applications, elements H_{kl} and S_{kl} often depend only on nonlinear parameters of k th and l th basis functions. Thus, the matrix

$$\frac{\partial \text{vec } H}{\partial a'} - \epsilon \frac{\partial \text{vec } S}{\partial a'}$$

is sparse and this fact should be taken into account when expression (60) is programmed. This situation takes place when we deal with complex Gaussians. An efficient scheme of computing the entire gradient can be utilized in the following manner. First, we calculate $\nabla_{\text{vech } L_1} \epsilon$, $\nabla_{\text{vech } B_1} \epsilon$, $\nabla_{\text{vech } L_2} \epsilon, \dots, \nabla_{\text{vech } B_K} \epsilon$, and then combine them into the total gradient. In turn, the evaluation of $\nabla_{\text{vech } L_k} \epsilon$, as well as $\nabla_{\text{vech } B_k} \epsilon$, can be done using the following expression:

$$\begin{aligned} \nabla_{\text{vech } L_k} \epsilon &\equiv \frac{\partial \epsilon}{\partial (\text{vech } L_k)'} = c_k \sum_{l=1}^{\mathcal{K}} c_l^* \left[\frac{\partial H_{lk}}{\partial (\text{vech } L_k)'} - \epsilon \frac{\partial S_{lk}}{\partial (\text{vech } L_k)'} \right] + c_k^* \sum_{l=1}^{\mathcal{K}} c_l \left[\frac{\partial H_{kl}}{\partial (\text{vech } L_k)'} - \epsilon \frac{\partial S_{kl}}{\partial (\text{vech } L_k)'} \right] \\ &\quad - c_k c_k^* \left[\frac{\partial H_{kk}}{\partial (\text{vech } L_k)'} - \epsilon \frac{\partial S_{kk}}{\partial (\text{vech } L_k)'} \right] = 2 \text{Re} \left(c_k \sum_{l=1}^{\mathcal{K}} c_l^* \left[\frac{\partial H_{lk}}{\partial (\text{vech } L_k)'} - \epsilon \frac{\partial S_{lk}}{\partial (\text{vech } L_k)'} \right] \right) \\ &\quad - |c_k|^2 \left[\frac{\partial H_{kk}}{\partial (\text{vech } L_k)'} - \epsilon \frac{\partial S_{kk}}{\partial (\text{vech } L_k)'} \right]. \end{aligned} \quad (61)$$

Looking at (60) and (61) one can see that the calculation of the energy gradient requires the knowledge of the following derivatives:

$$\frac{\partial H_{kl}}{\partial (\text{vech } L_k)'}, \quad \frac{\partial H_{kl}}{\partial (\text{vech } L_l)'}, \quad \frac{\partial H_{kl}}{\partial (\text{vec } B_k)'}, \quad \text{and} \quad \frac{\partial H_{kl}}{\partial (\text{vec } B_l)'}, \quad (62)$$

as well as the corresponding derivatives of S_{kl} . The next part of our article will be devoted to obtaining the formulas for these derivatives.

B. Overlap derivatives

Let us begin with the differential for the non-normalized overlap:

$$\begin{aligned} d\langle\phi_k|\tilde{\phi}_l\rangle &= d\frac{\pi^{3n/2}}{|\tilde{C}_{kl}|^{3/2}} = -\frac{3}{2}\frac{\pi^{3n/2}}{|\tilde{C}_{kl}|^{3/2}}\text{tr}[\tilde{C}_{kl}^{-1}d\tilde{C}_{kl}] = -\frac{3}{2}\langle\phi_k|\tilde{\phi}_l\rangle\text{tr}[\tilde{C}_{kl}^{-1}d(A_k + \tilde{A}_l - iB_k + i\tilde{B}_l)] = -\frac{3}{2}\langle\phi_k|\tilde{\phi}_l\rangle\text{tr}[\tilde{C}_{kl}^{-1}((dL_k)L_k' + L_kdL_k' \\ &+ (d\tilde{L}_l)\tilde{L}_l' + \tilde{L}_ld\tilde{L}_l' - idB_k + id\tilde{B}_l)] = -\frac{3}{2}\langle\phi_k|\tilde{\phi}_l\rangle\text{tr}[2L_k'\tilde{C}_{kl}^{-1}dL_k + 2\tilde{L}_l'\tilde{C}_{kl}^{-1}d\tilde{L}_l - i\tilde{C}_{kl}^{-1}dB_k + i\tilde{C}_{kl}^{-1}d\tilde{B}_l] \\ &= -\frac{3}{2}\langle\phi_k|\tilde{\phi}_l\rangle\text{tr}[2L_k'\tilde{C}_{kl}^{-1}dL_k + 2L_l'P\tilde{C}_{kl}^{-1}P'dL_l - i\tilde{C}_{kl}^{-1}dB_k + iP\tilde{C}_{kl}^{-1}P'dB_l]. \end{aligned} \quad (63)$$

To simplify the expression it is convenient to denote

$$\tilde{\tilde{C}}_{kl} = P\tilde{C}_{kl}P', \quad (64)$$

so that

$$P\tilde{C}_{kl}^{-1}P' = (P\tilde{C}_{kl}P')^{-1} = \tilde{\tilde{C}}_{kl}^{-1}. \quad (65)$$

Using relations (11)–(13) we can transform (63) in the following manner:

$$\begin{aligned} d\langle\phi_k|\tilde{\phi}_l\rangle &= -\frac{3}{2}\langle\phi_k|\tilde{\phi}_l\rangle[2\text{vec}(\tilde{C}_{kl}^{-1}L_k)'\text{vec}(dL_k) + 2\text{vec}(\tilde{\tilde{C}}_{kl}^{-1}L_l)'\text{vec}(dL_l) - i\text{vec}(\tilde{C}_{kl}^{-1})'\text{vec}(dB_k) + i\text{vec}(\tilde{\tilde{C}}_{kl}^{-1})'\text{vec}(dB_l)] \\ &= -\frac{3}{2}\langle\phi_k|\tilde{\phi}_l\rangle[2\text{vech}(\tilde{C}_{kl}^{-1}L_k)'\text{vech}(dL_k) + 2\text{vech}(\tilde{\tilde{C}}_{kl}^{-1}L_l)'\text{vech}(dL_l) - i\text{vech}(2\tilde{C}_{kl}^{-1} - \text{diag}\tilde{C}_{kl}^{-1})'\text{vech}B_k \\ &+ i\text{vech}(2\tilde{\tilde{C}}_{kl}^{-1} - \text{diag}\tilde{\tilde{C}}_{kl}^{-1})'\text{vech}B_l]. \end{aligned} \quad (66)$$

It follows from the last formula that

$$\frac{\partial\langle\phi_k|\tilde{\phi}_l\rangle}{\partial(\text{vech}L_k)'} = -3\langle\phi_k|\tilde{\phi}_l\rangle\text{vech}(\tilde{C}_{kl}^{-1}L_k)',$$

$$\frac{\partial\langle\phi_k|\tilde{\phi}_l\rangle}{\partial(\text{vech}L_l)'} = -3\langle\phi_k|\tilde{\phi}_l\rangle\text{vech}(\tilde{\tilde{C}}_{kl}^{-1}L_l)', \quad (67)$$

$$\frac{\partial\langle\phi_k|\tilde{\phi}_l\rangle}{\partial(\text{vech}B_k)'} = \frac{3}{2}i\langle\phi_k|\tilde{\phi}_l\rangle\text{vech}(2\tilde{C}_{kl}^{-1} - \text{diag}\tilde{C}_{kl}^{-1})',$$

$$\frac{\partial\langle\phi_k|\tilde{\phi}_l\rangle}{\partial(\text{vech}B_l)'} = -\frac{3}{2}i\langle\phi_k|\tilde{\phi}_l\rangle\text{vech}(2\tilde{\tilde{C}}_{kl}^{-1} - \text{diag}\tilde{\tilde{C}}_{kl}^{-1})'.$$

For the case when $k=l$ and $\hat{P}=1$ we obtain

$$\begin{aligned} \frac{\partial\langle\phi_k|\phi_k\rangle}{\partial(\text{vech}L_k)'} &= -6\langle\phi_k|\phi_k\rangle\text{vech}(C_{kk}^{-1}L_k)' \\ &= -3\langle\phi_k|\phi_k\rangle\text{vech}(A_k^{-1}L_k)' \\ &= -3\langle\phi_k|\phi_k\rangle\text{vech}((L_k^{-1}))', \end{aligned} \quad (68)$$

and

$$\frac{\partial\langle\phi_k|\phi_k\rangle}{\partial(\text{vech}B_k)'} = 0. \quad (69)$$

Knowing the derivatives of the non-normalized overlap matrix elements it is now straightforward to obtain the corresponding expressions for the normalized overlap, S_{kl} :

$$\begin{aligned} \frac{\partial S_{kl}}{\partial(\text{vech}L_k)'} &= \frac{\partial}{\partial(\text{vech}L_k)'} \frac{\langle\phi_k|\tilde{\phi}_l\rangle}{\sqrt{\langle\phi_k|\phi_k\rangle\langle\phi_l|\phi_l\rangle}} \\ &= \frac{1}{\sqrt{\langle\phi_k|\phi_k\rangle\langle\phi_l|\phi_l\rangle}} \frac{\partial\langle\phi_k|\tilde{\phi}_l\rangle}{\partial(\text{vech}L_k)'} \\ &\quad - \frac{1}{2} \frac{\langle\phi_k|\tilde{\phi}_l\rangle}{\sqrt{\langle\phi_k|\phi_k\rangle\langle\phi_l|\phi_l\rangle}} \frac{1}{\langle\phi_k|\phi_k\rangle} \frac{\partial\langle\phi_k|\phi_k\rangle}{\partial(\text{vech}L_k)'} \\ &= S_{kl} \left[-3\text{vech}(\tilde{C}_{kl}^{-1}L_k)' \right. \\ &\quad \left. - \frac{1}{2}(-3)\text{vech}((L_k^{-1}))' \right] \\ &= \frac{3}{2}S_{kl}\text{vech}((L_k^{-1})' - 2\tilde{C}_{kl}^{-1}L_k)'. \end{aligned} \quad (70)$$

Other derivatives of the normalized overlap are

$$\frac{\partial S_{kl}}{\partial(\text{vech}L_l)'} = \frac{3}{2}S_{kl}\text{vech}((L_l^{-1})' - 2\tilde{\tilde{C}}_{kl}^{-1}L_l)',$$

$$\begin{aligned} \frac{\partial S_{kl}}{\partial(\text{vech } B_k)'} &= \frac{1}{\sqrt{\langle \phi_k | \phi_k \rangle \langle \phi_l | \phi_l \rangle}} \frac{\partial \langle \phi_k | \tilde{\phi}_l \rangle}{\partial(\text{vech } B_k)'} \\ &= \frac{3}{2} i S_{kl} \text{vech}(2\tilde{C}_{kl}^{-1} - \text{diag } \tilde{C}_{kl}^{-1})', \end{aligned} \quad (71)$$

$$\frac{\partial S_{kl}}{\partial(\text{vech } B_l)'} = -\frac{3}{2} i S_{kl} \text{vech}(2\tilde{C}_{kl}^{-1} - \text{diag } \tilde{C}_{kl}^{-1})'.$$

C. Kinetic energy derivatives

Taking the differential of the expression for the normalized kinetic energy matrix element (30) gives the following:

$$\begin{aligned} dT_{kl} &= d(6S_{kl} \text{tr}[M\tilde{C}_l\tilde{C}_{kl}^{-1}C_k^\dagger]) \\ &= 6 \text{tr}[M\tilde{C}_l\tilde{C}_{kl}^{-1}C_k^\dagger]dS_{kl} \\ &\quad + 6S_{kl} \text{tr}[Md(\tilde{C}_l\tilde{C}_{kl}^{-1}C_k^\dagger)]. \end{aligned} \quad (72)$$

Using the fact that for an arbitrary square matrix X ,

$$d(X^{-1}) = -X^{-1}(dX)X^{-1}, \quad (73)$$

the differential in the last term of (72) can be expanded as

$$\begin{aligned} d(\tilde{C}_l\tilde{C}_{kl}^{-1}C_k^\dagger) &= d(\tilde{C}_l)\tilde{C}_{kl}^{-1}C_k^\dagger + \tilde{C}_l(d\tilde{C}_{kl}^{-1})C_k^\dagger + \tilde{C}_l\tilde{C}_{kl}^{-1}dC_k^\dagger \\ &= d(\tilde{C}_l)\tilde{C}_{kl}^{-1}C_k^\dagger - \tilde{C}_l\tilde{C}_{kl}^{-1}(d\tilde{C}_{kl})\tilde{C}_{kl}^{-1}C_k^\dagger + \tilde{C}_l\tilde{C}_{kl}^{-1}dC_k^\dagger \\ &= d(\tilde{C}_l)\tilde{C}_{kl}^{-1}C_k^\dagger - \tilde{C}_l\tilde{C}_{kl}^{-1}(dC_k^\dagger)\tilde{C}_{kl}^{-1}C_k^\dagger \\ &\quad - \tilde{C}_l\tilde{C}_{kl}^{-1}(d\tilde{C}_l)\tilde{C}_{kl}^{-1}C_k^\dagger + \tilde{C}_l\tilde{C}_{kl}^{-1}dC_k^\dagger. \end{aligned} \quad (74)$$

Expression (72) then becomes

$$\begin{aligned} dT_{kl} &= \frac{T_{kl}}{S_{kl}}dS_{kl} + 6S_{kl} \text{tr}[MC_k^\dagger\tilde{C}_{kl}^{-1}d\tilde{C}_l - \tilde{C}_{kl}^{-1}C_k^\dagger M\tilde{C}_l\tilde{C}_{kl}^{-1}d\tilde{C}_l + M\tilde{C}_l\tilde{C}_{kl}^{-1}dC_k^\dagger - \tilde{C}_{kl}^{-1}C_k^\dagger M\tilde{C}_l\tilde{C}_{kl}^{-1}dC_k^\dagger] \\ &= \frac{T_{kl}}{S_{kl}}dS_{kl} + 6S_{kl} \text{tr}[\tilde{C}_{kl}^{-1}C_k^\dagger Md\tilde{C}_l - \tilde{C}_{kl}^{-1}C_k^\dagger M\tilde{C}_l\tilde{C}_{kl}^{-1}d\tilde{C}_l + \tilde{C}_{kl}^{-1}\tilde{C}_l MdC_k^\dagger - \tilde{C}_{kl}^{-1}\tilde{C}_l MC_k^\dagger\tilde{C}_{kl}^{-1}dC_k^\dagger] \\ &= \frac{T_{kl}}{S_{kl}}dS_{kl} + 6S_{kl} \text{tr}[\tilde{C}_{kl}^{-1}C_k^\dagger M(1 - \tilde{C}_l\tilde{C}_{kl}^{-1})d\tilde{C}_l + \tilde{C}_{kl}^{-1}\tilde{C}_l M(1 - C_k^\dagger\tilde{C}_{kl}^{-1})dC_k^\dagger] \\ &= \frac{T_{kl}}{S_{kl}}dS_{kl} + 6S_{kl} \text{tr}[\tilde{C}_{kl}^{-1}C_k^\dagger MC_k^\dagger\tilde{C}_{kl}^{-1}d\tilde{C}_l + \tilde{C}_{kl}^{-1}\tilde{C}_l M\tilde{C}_l\tilde{C}_{kl}^{-1}dC_k^\dagger]. \end{aligned} \quad (75)$$

Next, we plug

$$d\tilde{C}_l = (d\tilde{L}_l)\tilde{L}_l' + \tilde{L}_l d\tilde{L}_l' + id\tilde{B}_l, \quad (76)$$

$$dC_k^\dagger = (dL_k)L_k' + L_k dL_k' - idB_k,$$

into (75) and obtain

$$\begin{aligned} dT_{kl} &= \frac{T_{kl}}{S_{kl}}dS_{kl} + 6S_{kl} \text{tr}[2L_k'\tilde{C}_{kl}^{-1}\tilde{C}_l M\tilde{C}_l\tilde{C}_{kl}^{-1}dL_k \\ &\quad + 2\tilde{L}_l'\tilde{C}_{kl}^{-1}C_k^\dagger MC_k^\dagger\tilde{C}_{kl}^{-1}d\tilde{L}_l - i\tilde{C}_{kl}^{-1}\tilde{C}_l M\tilde{C}_l\tilde{C}_{kl}^{-1}dB_k \\ &\quad + i\tilde{C}_{kl}^{-1}C_k^\dagger MC_k^\dagger\tilde{C}_{kl}^{-1}d\tilde{B}_l]. \end{aligned} \quad (77)$$

In order to make the formulas more compact let us introduce the following notations:

$$F = \tilde{C}_{kl}^{-1}\tilde{C}_l M\tilde{C}_l\tilde{C}_{kl}^{-1}, \quad (78)$$

$$G = P\tilde{C}_{kl}^{-1}C_k^\dagger MC_k^\dagger\tilde{C}_{kl}^{-1}P'. \quad (79)$$

With this we have

$$\begin{aligned} dT_{kl} &= \frac{T_{kl}}{S_{kl}}dS_{kl} + 6S_{kl} \text{tr}[2L_k'FdL_k + 2L_l'GdL_l - idB_k \\ &\quad + iGdB_l], \end{aligned} \quad (80)$$

which then can be transformed as follows:

$$\begin{aligned} dT_{kl} &= \frac{T_{kl}}{S_{kl}}dS_{kl} + 6S_{kl}[2 \text{vec}(FL_k)' \text{vec}(dL_k) + 2 \text{vec}(GL_l)' \text{vec}(dL_l) - i \text{vec}(F)' \text{vec}(dB_k) + i \text{vec}(G)' \text{vec}(dB_l)] \\ &= \frac{T_{kl}}{S_{kl}}dS_{kl} + 6S_{kl}[2 \text{vech}(FL_k)' d \text{vech } L_k + 2 \text{vech}(GL_l)' d \text{vech } L_l - i \text{vech}(2F - \text{diag}(F))' d \text{vech } B_k \\ &\quad + i \text{vech}(2G - \text{diag}(G))' d \text{vech } B_l]. \end{aligned} \quad (81)$$

Thus, we can see that the derivatives of the normalized kinetic energy matrix element are

$$\begin{aligned}\frac{\partial T_{kl}}{\partial(\text{vech } L_k)'} &= \frac{T_{kl}}{S_{kl}} \frac{\partial S_{kl}}{\partial(\text{vech } L_k)'} + 12S_{kl} \text{vech}(FL_k)', \\ \frac{\partial T_{kl}}{\partial(\text{vech } L_l)'} &= \frac{T_{kl}}{S_{kl}} \frac{\partial S_{kl}}{\partial(\text{vech } L_l)'} + 12S_{kl} \text{vech}(GL_l)', \\ \frac{\partial T_{kl}}{\partial(\text{vech } B_k)'} &= \frac{T_{kl}}{S_{kl}} \frac{\partial S_{kl}}{\partial(\text{vech } B_k)'} \\ &\quad - 6iS_{kl} \text{vech}(2F - \text{diag}(F))', \\ \frac{\partial T_{kl}}{\partial(\text{vech } B_l)'} &= \frac{T_{kl}}{S_{kl}} \frac{\partial S_{kl}}{\partial(\text{vech } B_l)'} \\ &\quad + 6iS_{kl} \text{vech}(2G - \text{diag}(G))'. \end{aligned} \quad (82)$$

D. Potential energy derivatives

To write out the expression for the derivatives of the potential energy matrix element (48) with respect to nonlinear variational parameters we first need to evaluate the corresponding expressions for R_{kl}^{ij} . We start with taking the differential of (47):

$$\begin{aligned}dR_{kl}^{ij} &= \frac{R_{kl}^{ij}}{S_{kl}} dS_{kl} + \frac{2}{\sqrt{\pi}} S_{kl} d \text{tr}[\tilde{C}_{kl}^{-1} J_{ij}]^{-1/2} \\ &= \frac{R_{kl}^{ij}}{S_{kl}} dS_{kl} - \frac{1}{\sqrt{\pi}} S_{kl} \text{tr}[\tilde{C}_{kl}^{-1} J_{ij}]^{-3/2} \text{tr}[(d\tilde{C}_{kl}^{-1}) J_{ij}] \\ &= \frac{R_{kl}^{ij}}{S_{kl}} dS_{kl} + \frac{1}{\sqrt{\pi}} S_{kl} \text{tr}[\tilde{C}_{kl}^{-1} J_{ij}]^{-3/2} \text{tr}[\tilde{C}_{kl}^{-1} (d\tilde{C}_{kl}) \tilde{C}_{kl}^{-1} J_{ij}]. \end{aligned} \quad (83)$$

After inserting (76) we transform it as follows:

$$\begin{aligned}dR_{kl}^{ij} &= \frac{R_{kl}^{ij}}{S_{kl}} dS_{kl} + \frac{1}{\sqrt{\pi}} S_{kl} \text{tr}[\tilde{C}_{kl}^{-1} J_{ij}]^{-3/2} \text{tr}[\tilde{C}_{kl}^{-1} J_{ij} \tilde{C}_{kl}^{-1} ((dL_k)L_k' + L_k dL_k' + (d\tilde{L}_l)\tilde{L}_l' + \tilde{L}_l d\tilde{L}_l' - idB_k + id\tilde{B}_l)] \\ &= \frac{R_{kl}^{ij}}{S_{kl}} dS_{kl} + \frac{1}{\sqrt{\pi}} S_{kl} \text{tr}[\tilde{C}_{kl}^{-1} J_{ij}]^{-3/2} \text{tr}[2L_k' \tilde{C}_{kl}^{-1} J_{ij} \tilde{C}_{kl}^{-1} dL_k + 2L_l' P \tilde{C}_{kl}^{-1} J_{ij} \tilde{C}_{kl}^{-1} P' dL_l - i\tilde{C}_{kl}^{-1} J_{ij} \tilde{C}_{kl}^{-1} dB_k + iP \tilde{C}_{kl}^{-1} J_{ij} \tilde{C}_{kl}^{-1} P' dB_l]. \end{aligned} \quad (84)$$

If we denote

$$\tilde{Q} = \tilde{C}_{kl}^{-1} J_{ij} \tilde{C}_{kl}^{-1}, \quad (85)$$

$$\tilde{\tilde{Q}} = P \tilde{Q} P', \quad (86)$$

then

$$\begin{aligned}dR_{kl}^{ij} &= \frac{R_{kl}^{ij}}{S_{kl}} dS_{kl} + \frac{1}{\sqrt{\pi}} S_{kl} \text{tr}[\tilde{C}_{kl}^{-1} J_{ij}]^{-3/2} [2 \text{vec}(\tilde{Q}L_k)' \text{vec}(dL_k) + 2 \text{vec}(\tilde{Q}L_l)' \text{vec}(dL_l) - i \text{vec}(\tilde{Q})' \text{vec}(dB_k) + i \text{vec}(\tilde{\tilde{Q}})' \text{vec}(dB_l)] \\ &= \frac{R_{kl}^{ij}}{S_{kl}} dS_{kl} + \frac{1}{\sqrt{\pi}} S_{kl} \text{tr}[\tilde{C}_{kl}^{-1} J_{ij}]^{-3/2} [2 \text{vech}(\tilde{Q}L_k)' d \text{vech } L_k + 2 \text{vech}(\tilde{Q}L_l)' d \text{vech } L_l - i \text{vech}(2\tilde{Q} - \text{diag}(\tilde{Q}))' \text{vech}(dB_k) \\ &\quad + i \text{vech}(2\tilde{\tilde{Q}} - \text{diag}(\tilde{\tilde{Q}}))' \text{vech}(dB_l)]. \end{aligned} \quad (87)$$

Hence, the derivatives of R_{kl}^{ij} ($k \neq l$) are

$$\begin{aligned} \frac{\partial R_{kl}^{ij}}{\partial(\text{vech } L_k)'} &= \frac{R_{kl}^{ij}}{S_{kl}} \frac{\partial S_{kl}}{\partial(\text{vech } L_k)'} \\ &+ \frac{2}{\sqrt{\pi}} S_{kl} \text{tr}[\tilde{C}_{kl}^{-1} J_{ij}]^{-3/2} \text{vech}(\tilde{Q}L_k)', \\ \frac{\partial R_{kl}^{ij}}{\partial(\text{vech } L_l)'} &= \frac{R_{kl}^{ij}}{S_{kl}} \frac{\partial S_{kl}}{\partial(\text{vech } L_l)'} \\ &+ \frac{2}{\sqrt{\pi}} S_{kl} \text{tr}[\tilde{C}_{kl}^{-1} J_{ij}]^{-3/2} \text{vech}(\tilde{Q}L_l)', \\ \frac{\partial R_{kl}^{ij}}{\partial(\text{vech } B_k)'} &= \frac{R_{kl}^{ij}}{S_{kl}} \frac{\partial S_{kl}}{\partial(\text{vech } B_k)'} - \frac{i}{\sqrt{\pi}} S_{kl} \text{tr}[\tilde{C}_{kl}^{-1} J_{ij}]^{-3/2} \\ &\times \text{vech}(2\tilde{Q} - \text{diag}(\tilde{Q}))', \\ \frac{\partial R_{kl}^{ij}}{\partial(\text{vech } B_l)'} &= \frac{R_{kl}^{ij}}{S_{kl}} \frac{\partial S_{kl}}{\partial(\text{vech } B_l)'} + \frac{i}{\sqrt{\pi}} S_{kl} \text{tr}[\tilde{C}_{kl}^{-1} J_{ij}]^{-3/2} \\ &\times \text{vech}(2\tilde{Q} - \text{diag}(\tilde{Q}))'. \end{aligned} \quad (88)$$

The derivatives of the potential energy matrix element (48) are nothing but linear combinations of (88):

$$\begin{aligned} \frac{\partial V_{kl}}{\partial(\text{vech } L_k)'} &= \sum_{i=1}^n q_0 q_i \frac{\partial R_{kl}^{ii}}{\partial(\text{vech } L_k)'} + \sum_{i<j}^n q_i q_j \frac{\partial R_{kl}^{ij}}{\partial(\text{vech } L_k)'} \\ &= \frac{V_{kl}}{S_{kl}} \frac{\partial S_{kl}}{\partial(\text{vech } L_k)'} + \frac{2}{\sqrt{\pi}} S_{kl} \\ &\times \left(\sum_{i=1}^n q_0 q_i \text{tr}[\tilde{C}_{kl}^{-1} J_{ii}]^{-3/2} \text{vech}(\tilde{Q}L_k)' \right. \\ &\left. + \sum_{i<j}^n q_i q_j \text{tr}[\tilde{C}_{kl}^{-1} J_{ij}]^{-3/2} \text{vech}(\tilde{Q}L_k)' \right), \end{aligned} \quad (89)$$

$$\begin{aligned} \frac{\partial V_{kl}}{\partial(\text{vech } L_l)'} &= \frac{V_{kl}}{S_{kl}} \frac{\partial S_{kl}}{\partial(\text{vech } L_l)'} \\ &+ \frac{2}{\sqrt{\pi}} S_{kl} \left(\sum_{i=1}^n q_0 q_i \text{tr}[\tilde{C}_{kl}^{-1} J_{ii}]^{-3/2} \text{vech}(\tilde{Q}L_l)' \right. \\ &\left. + \sum_{i<j}^n q_i q_j \text{tr}[\tilde{C}_{kl}^{-1} J_{ij}]^{-3/2} \text{vech}(\tilde{Q}L_l)' \right), \end{aligned} \quad (90)$$

$$\begin{aligned} \frac{\partial V_{kl}}{\partial(\text{vech } B_k)'} &= \frac{V_{kl}}{S_{kl}} \frac{\partial S_{kl}}{\partial(\text{vech } B_k)'} - \frac{i}{\sqrt{\pi}} S_{kl} \\ &\times \left(\sum_{i=1}^n q_0 q_i \text{tr}[\tilde{C}_{kl}^{-1} J_{ii}]^{-3/2} \text{vech}(2\tilde{Q} - \text{diag}(\tilde{Q}))' \right. \\ &\left. + \sum_{i<j}^n q_i q_j \text{tr}[\tilde{C}_{kl}^{-1} J_{ij}]^{-3/2} \text{vech}(2\tilde{Q} \right. \\ &\left. - \text{diag}(2\tilde{Q}))' \right), \end{aligned} \quad (91)$$

$$\begin{aligned} \frac{\partial V_{kl}}{\partial(\text{vech } B_l)'} &= \frac{V_{kl}}{S_{kl}} \frac{\partial S_{kl}}{\partial(\text{vech } B_l)'} + \frac{i}{\sqrt{\pi}} S_{kl} \\ &\times \left(\sum_{i=1}^n q_0 q_i \text{tr}[\tilde{C}_{kl}^{-1} J_{ii}]^{-3/2} \text{vech}(2\tilde{Q} - \text{diag}(\tilde{Q}))' \right. \\ &\left. + \sum_{i<j}^n q_i q_j \text{tr}[\tilde{C}_{kl}^{-1} J_{ij}]^{-3/2} \text{vech}(2\tilde{Q} \right. \\ &\left. - \text{diag}(2\tilde{Q}))' \right). \end{aligned} \quad (92)$$

One final comment we would like to make regarding the evaluation of the potential energy derivatives is that there are many products $\tilde{C}_{kl}^{-1} J_{ij} \tilde{C}_{kl}^{-1}$ appear in the expressions above. For an arbitrary matrix X the elements of these products can be easily computed using the following relations:

TABLE I. The ground and excited state energies of the He atom (in a.u.) corresponding to the $1s^1 n s^1$ Rydberg states obtained in non-Born-Oppenheimer calculations with complex Gaussians basis sets of different lengths. The energies in the last column were calculated with an infinite nuclear mass. The symbol n denotes the principle quantum number in the $1s^1 n s^1$ state.

n	$K=50$	$K=100$	$K=200$	$K=500$	$K=500, \text{ inf. mass}$
1	-2.903 297 70	-2.903 301 81	-2.903 304 17	-2.903 304 56	-2.903 724 38
2	-2.145 668 44	-2.145 676 15	-2.145 677 76	-2.145 678 58	-2.145 974 04
3	-2.060 970 59	-2.060 985 77	-2.060 988 16	-2.060 989 04	-2.061 271 95
4	-2.033 289 16	-2.033 304 99	-2.033 306 88	-2.033 307 78	-2.033 586 67
4	-2.020 872 26	-2.020 895 92	-2.020 898 31	-2.020 899 65	-2.021 176 78
6	-2.014 215 31	-2.014 280 41	-2.014 285 40	-2.014 286 73	-2.014 562 92
7	-2.010 201 05	-2.010 344 05	-2.010 348 52	-2.010 349 94	-2.010 625 57
8	-2.007 577 31	-2.007 812 61	-2.007 816 66	-2.007 818 05	-2.008 093 33
9	-2.005 357 24	-2.006 086 48	-2.006 092 26	-2.006 094 08	-2.006 369 12
10	-2.004 073 99	-2.004 854 33	-2.004 863 13	-2.004 867 42	-2.005 142 27

TABLE II. The energy difference between the ground and excited Rydberg energies of the He atom (in a.u.), $E_{\text{fin}} - E_{\text{inf}}$, calculated with finite and infinite nuclear mass and 500 basis functions, and its ratio to the total energy of the corresponding state with infinite nuclear mass.

n	$E_{\text{fin}} - E_{\text{inf}}$	$(E_{\text{fin}} - E_{\text{inf}})/E_{\text{inf}}$
1	4.1982×10^{-4}	1.4458×10^{-4}
2	2.9546×10^{-4}	1.3768×10^{-4}
3	2.8291×10^{-4}	1.3725×10^{-4}
4	2.7890×10^{-4}	1.3715×10^{-4}
4	2.7713×10^{-4}	1.3711×10^{-4}
6	2.7619×10^{-4}	1.3709×10^{-4}
7	2.7563×10^{-4}	1.3709×10^{-4}
8	2.7527×10^{-4}	1.3708×10^{-4}
9	2.7503×10^{-4}	1.3708×10^{-4}
10	2.7486×10^{-4}	1.3708×10^{-4}

$$(XJ_{ij}X)_{pq} = (X)_{pi}(X)_{iq}, \quad (93)$$

$$(XJ_{ij}X)_{pq} = (X)_{pi}(X)_{iq} + (X)_{pj}(X)_{jq} - (X)_{pi}(X)_{jq} - (X)_{pj}(X)_{iq},$$

$$i \neq j.$$

VIII. NUMERICAL TESTS

To ensure that the analytic expressions for the derivatives of the Hamiltonian and the overlap matrix elements with respect to nonlinear variational parameters are correct, we first tested them for consistency with the corresponding finite difference derivatives using MATHEMATICA computer algebra package.²⁵ After that we implemented the formulas for both matrix elements and their derivatives numerically and ran several simple non-Born-Oppenheimer calculations to illustrate that they work. Since the main goal of our tests was the verification of the analytic expressions and the capability of the complex Gaussians to recover the nodal structure of the wave function, we did not consider systems that contain more than four particles and did not run calculations with large basis sets as this would have been a rather time

TABLE III. The difference of the ground and excited Rydberg state energies of the He atom obtained with 500 basis functions and the total non-Born-Oppenheimer ground state energy of the He⁺ ion ($E^{\text{He}^+} = -1.999\,725\,851$ a.u.) vs the non-BO energies of the hydrogen like atom (E^{H}/n^2) with a unit charge and the nuclear mass equal to the sum of the He nucleus and one electron mass. All data are in a.u.

n	$E_n - E^{\text{He}^+}$	E^{H}/n^2
1	0.903 999	0.499 931
2	0.146 248	0.124 983
3	0.061 546	0.055 548
4	0.033 861	0.031 246
4	0.021 451	0.019 997
6	0.014 837	0.013 887
7	0.010 900	0.010 203
8	0.008 367	0.007 811
9	0.006 643	0.006 172
10	0.005 416	0.004 999

consuming task. The work on writing an efficient parallel computer code, suitable for large scale calculations with complex Gaussians, is currently underway and the results of such calculations will be published separately, as they become available.

As a test example for the method developed in this work we have used the He atom. The calculations have concerned ten Rydberg singlet states of this system corresponding to the $1s^1ns^1$ ($n=1, \dots, 10$) electronic configuration.

A separate set of calculations was performed for each state. In the calculations we simultaneously optimized all nonlinear Gaussian variational parameters utilizing an approach involving the analytic gradient. For each state we used the basis sets consisting of 50, 100, 200, and 500 functions. The energy results obtained in the calculations are shown in Table I. As can be seen, the complex Gaussian method performs very well for He. The energies are well converged with the number of basis functions, although the convergence is uneven and the ground state energy is converged by at least one order of magnitude better than the energy for the last $1s^110s^1$ state. This can be expected considering larger number of radial nodes in the wave function of the latter state. In the table we also show the energies obtained with the infinite mass of the He nucleus (the Born-Oppenheimer energies).

The difference between the non-BO energies and BO energies provides an account of the nonadiabatic effects. In Table II we show those differences. We also show the ratio of the contribution from the nonadiabatic effects and the total non-BO energy. As one notices, while the absolute nonadiabatic contribution varies particularly for the lower states, the relative contribution is remarkably constant. Finally, in Table III we show the differences of the non-BO energies of the $1s^1ns^1$ states of the He atom and the non-BO energy of the He⁺ ion and the non-BO energies of the ns^1 states of the H atom. As one can see, as the level of the excitation increases, the He/He⁺ energy difference approaches the hydrogen Rydberg state energies. This, of course, can be expected, because in higher $1s^1ns^1$ excited states the ns electron more and more resembles an electron moving in the field of a localized +1 charge.

IX. SUMMARY

In this work we presented formulas for matrix elements and their derivatives involving explicitly correlated Gaussian functions with complex parameters. All the expressions were obtained using matrix differential calculus and are given in a convenient matrix form, which is suitable for both numerical implementation and theoretical analysis. The results of this work will serve as a foundation for future implementation of complex Gaussian basis functions in variational calculations of different quantum mechanical systems, in particular, small molecules where real Gaussian basis functions do not reveal sufficient effectiveness. The use of such functions seems to be appealing due to their flexibility and the ability of evaluating various matrix elements for an arbitrary number of particles. The derivatives of the Hamiltonian matrix elements with respect to nonlinear variational parameters obtained in

this work are helpful in significantly accelerating the process of optimization of the nonlinear parameters. The latter was evident in the non-BO calculations that have been performed for $1s^1ns^1$ Rydberg states of the He atom. The results for ten of those states presented in this work demonstrate that the complex Gaussians are effective in describing multiple radial nodes in the atomic wave function.

- ¹M. Cafiero and L. Adamowicz, Chem. Phys. Lett. **387**, 136 (2004).
- ²M. Cafiero and L. Adamowicz, Phys. Rev. Lett. **88**, 033002 (2002).
- ³M. Cafiero and L. Adamowicz, Phys. Rev. Lett. **89**, 073001 (2002).
- ⁴D. B. Kinghorn and L. Adamowicz, J. Chem. Phys. **106**, 8760 (1997).
- ⁵D. B. Kinghorn and L. Adamowicz, J. Chem. Phys. **110**, 7166 (1999).
- ⁶D. B. Kinghorn and L. Adamowicz, Phys. Rev. Lett. **83**, 2541 (1999).
- ⁷S. Bubin and L. Adamowicz, J. Chem. Phys. **118**, 3079 (2003).
- ⁸S. Bubin and L. Adamowicz, J. Chem. Phys. **120**, 6051 (2004).
- ⁹S. Bubin and L. Adamowicz, J. Chem. Phys. **121**, 6249 (2004).
- ¹⁰S. Bubin, E. Bednarz, and L. Adamowicz, J. Chem. Phys. **122**, 041102 (2005).

- ¹¹M. Pavanello, S. Bubin, M. Molski, and L. Adamowicz, J. Chem. Phys. **123**, 104306 (2005).
- ¹²S. Bubin, L. Adamowicz, and M. Molski, J. Chem. Phys. **123**, 134310 (2005).
- ¹³E. Bednarz, S. Bubin, and L. Adamowicz, Mol. Phys. **103**, 1169 (2005).
- ¹⁴S. Bubin and L. Adamowicz, Chem. Phys. Lett. **403**, 185 (2005).
- ¹⁵T. K. Rebane and O. N. Yusupov, Opt. Spectrosc. **71**, 148 (1991).
- ¹⁶T. K. Rebane, Opt. Spectrosc. **93**, 908 (2002).
- ¹⁷V. S. Zotev and T. K. Rebane, Phys. Rev. A **65**, 062501 (2002).
- ¹⁸A. M. Frolov and V. H. Smith, Jr., J. Phys. B **28**, L449 (1995).
- ¹⁹A. M. Frolov, Phys. Rev. A **67**, 064501 (2003).
- ²⁰T. K. Rebane and P. V. Vitushinskii, Opt. Spectrosc. **92**, 21 (2002).
- ²¹M. Cafiero, S. Bubin, and L. Adamowicz, Phys. Chem. Chem. Phys. **5**, 1491 (2003).
- ²²S. Bubin, M. Cafiero, and L. Adamowicz, Adv. Chem. Phys. **131**, 377 (2005).
- ²³D. B. Kinghorn, Int. J. Quantum Chem. **57**, 141 (1996).
- ²⁴J. R. Magnus and H. Neudecker, *Matrix Differential Calculus with Applications in Statistics and Econometrics* (Wiley, Chichester, 1988).
- ²⁵MATHEMATICA, Version 5.0, Wolfram Research, Inc., Champaign, IL, 2003.