NON-BORN–OPPENHEIMER VARIATIONAL CALCULATIONS OF ATOMS AND MOLECULES WITH EXPLICITLY CORRELATED GAUSSIAN BASIS FUNCTIONS

SERGIY BUBIN

Department of Physics and Department of Chemistry
University of Arizona, Tucson, AZ

MAURICIO CAFIERO

Department of Chemistry, Rhodes College, Memphis, TN

LUDWIK ADAMOWICZ

Department of Chemistry and Department of Physics
University of Arizona, Tucson, AZ

CONTENTS

Advances in Chemical Physics, Volume 131, edited by Stuart A. Rice

I. INTRODUCTION

Soon after the Schrödinger equation was introduced in 1926, several works appeared dealing with the fundamental problem of the nuclear motion in molecules. Very soon after, the relativistic equations were introduced for one- and two-electron systems. The experiments on the Lamb shift stimulated

derivation of the expressions for fine energy corrections related to mass–velocity effects, radiative effects, and so on. Following this development, questions were raised whether an *ab initio* approach, in which only the values of fundamental constants are taken from experiments, is capable of reproducing experimental results with the precision which matches that of the experimental techniques. At first, the questions concerned the accuracy of the calculations, but soon they were extended to testing the model of the molecular electron structure provided by the nonrelativistic and relativistic quantum mechanics.

In order to answer these questions, accurate experimental and theoretical results were needed for representative molecular systems. Theoreticians, for obvious reasons, have favored very simple systems, such as the hydrogen molecular ion (H_2^+) for their calculations. However, with only one electron, this system did not provide a proper test case for the molecular quantum mechanical methods due to the absence of the electron correlation. Therefore, the two-electron hydrogen molecule has served as the system on which the fundamental laws of quantum mechanics have been first tested.

In an attempt to make the quantum-mechanical calculations on molecular systems practical and to provide a more intuitive interpretation of the computed results, it has long been a quest in the electronic structure theory of molecules to establish a solid base for separating the motion of light electrons from the motion of heavier nuclei. It is believed that the original work of Born and Oppenheimer [1] initiated the discussion by the analysis of the diatomic case. Further works of Cobes and Seiler [2], who managed, with the use of singular perturbation theory, to resolve the problem of the diverging series, which appeared in the Born–Oppenheimer (BO) expansion, and particularly of Klein et al. [3], who extended the formalism to polyatomic systems, have brought the consideration of the topic to a level of commonly accepted theory.

Apart from the further refinements of the BO approach, there has been a continuing interest in theoretically describing molecular systems with a method that treats the motions of both nuclei and electrons equivalently. This type of methodology has to entirely depart from the PES concept. It is particularly interesting how this type of approach describes the conventional notions of the molecular and electronic structures. In particular, the concept of chemical bonding, which at the BO level is an electronic phenomenon, has to be described in an approach departing from the BO approximation, as an effect derived from collective dynamical behavior of both electrons and nuclei.

Another motivation for considering molecular systems without assuming the BO approximation stems from the realization that in order to reach ''spectroscopic'' accuracy in quantum-mechanical calculations (i.e., error less than 1 microhartree), one needs to account for the coupling between motions of electrons and nuclei and, in some cases, also for the relativistic effects. Modern experimental techniques, such as gas-phase ion-beam spectroscopy, reach

accuracy on the order of $0.001\,\mathrm{cm}^{-1}$ (5 nanohartrees) [4]. In order for the molecular quantum mechanics to continue providing assistance in resolving and assigning experimental spectra and in studies of reaction dynamics, work has to continue on the development of more refined theoretical methodology, which accounts for nonadiabatic interactions. With such methodology, fundamental concepts of the molecular quantum mechanics can be explored and the basic theoretical framework of the high-resolution molecular spectroscopy can be tested. Recent advances in high-performance computing, especially in the area of massively parallel systems, has given momentum to proceed with the development of quantum-mechanical methods that depart from the BO approximation and describe the motions of the nuclei and electrons with a single wave function. In the context of the non-BO calculations, it is particularly interesting to study highly rho-vibrationally and electronically excited molecules and clusters, where more significant coupling between the two motions can occur. Essentially, whenever the spacings between electronic excitation levels become comparable to the spacings between the vibrational or rotational levels, nonadiabatic effects are likely to be found. Studies of these effects are relevant to astrophysical phenomena, molecule dynamics, and molecular behavior at high temperatures.

If one assumes the BO approximation and considers potential energy surfaces of a molecule, one can usually identify areas where there is a high density of electronic states and areas where the electronic states are more separated from each other. A strong nonadiabatic coupling can be expected to mainly occur in the areas with the high electronic state density. This simple realization has given rise to a considerable body of recent theoretical research on nonadiabatic phenomena, done in relation to conical intersections of molecular PESs [5–14], most notably by Yarkony and his group, among others. The reason for having large or even infinite nonadiabatic coupling terms in the conical areas is that fast-moving electrons may create exceptionally large forces causing the nuclei to strongly accelerate. The terms responsible for this accelerated motion cannot be ignored even in the first approximation. The consequences of the conical PES crossings to the dynamics of molecular reactions have also been considered by a number of groups (see, for example, the work of Hammes-Schiffer [15–26]). In those works, however, the non-BO effects are only considered at the conical intersection of two electronic PESs since the focus there has been more on determining the probability of the process splitting and following two different electronic PESs and less on very accurate global representation of coupled electronic–nuclear states, as it has been in our work.

The nonadiabatic coupling terms can quickly become large or even infinite (or singular) when two successive adiabatic states become degenarate. Such singular nonadiabatic coupling may not only lead to the breakdown of the

Born–Oppenheimer perturbation theory but may also make the application of the perturbation theory inadequate.

Though infrequent, fully non-Born–Oppenheimer high-accuracy calculations on atomic and molecular systems have been increasing in number in recent years. However, besides our group, there are only a few groups doing such calculations, particularly for molecular systems. One should mention the recent works of Frolov and Smith [27–29] and of Frolov [30–34] concerning some exotic systems involving muons and positrons, as well as one electron H_2 isotopomers. Our works, reviewed here, have concerned both atomic and molecular systems. Although for molecules most of our non-BO calculations concerned ground and excited states of diatomic systems, we have also recently extended the non-BO approach beyond molecules with two nuclei. The system with the largest number of particles we have calculated so far had two nuclei and five electrons (the LiH^- anion).

There are several elements in non-BO calculations that distinguish them from the conventional BO calculations. The first one concerns the Hamiltonian. If one neglects the relativistic effects and places the considered system in a laboratory Cartesian coordinate frame, the Hamiltonian has a simple form of a sum of one-particle kinetic energy operators for all particles involved in the system plus a sum representing all pair Coulombic interactions between the particles. It is convenient to separate the center-of-mass kinetic energy from the Hamiltonian. This is usually done by a coordinate transformation that involves choosing a new coordinate system whose first three coordinates are the Cartesian coordinates of the center of mass in the laboratory coordinate system and the remaining $3N - 3$ coordinates are internal coordinates. There are a number of ways to select the internal coordinates. In some works, these have been the Jacobi coordinates, the spherical coordinates, or the coordinates defined with respect with the center of mass of the system. In our approach we have used a cartesian coordinate system with the coordinate origin placed at one of the particles (usually the heaviest one). This will be described later in this chapter.

An important difference between the BO and non-BO internal Hamiltonians is that the former describes only the motion of electrons in the stationary field of nuclei positioned in fixed points in space (represented by point charges) while the latter describes the coupled motion of both nuclei and electrons. In the conventional molecular BO calculations, one typically uses atom-centered basis functions (in most calculations one-electron atomic orbitals) to expand the electronic wave function. The fermionic nature of the electrons dictates that such a function has to be antisymmetric with respect to the permutation of the labels of the electrons. In some high-precision BO calculations the wave function is expanded in terms of basis functions that explicitly depend on the interelectronic distances (so-called explicitly correlated functions). Such

functions usually very effectively describe the electron correlation effects, which need to be included in any high-level BO calculation. An alternative to the explicitly correlated functions is to use expansions in terms of Slater determinants constructed using one electron functions (molecular orbitals). Such an approach is called the configuration interaction (CI) method, and it usually converges much slower than the approach using explicitly correlated basis functions.

Non-BO calculations usually need to be performed to very high precision since only then the non-BO effects, which are usually very small, can be adequately determined. This requires that not only the electron–electron correlation effects are described very accurately, but also the correlation effects due to the nucleus–nucleus interaction and due to the nucleus–electron interaction are accurately represented in the wave function. Since the electrons are light particles, their individual wave functions can strongly overlap, and the probability of finding two electrons (with opposite spins) simultaneously in the same point in space is much higher than for two nuclei, which are much heavier and avoid each other to a much higher degree. One can say that the nuclear correlation effects are much stronger than the electronic ones. Also, the correlation effects associated with the coupled motion of electrons and nuclei are significant because the electrons, particularly the core electrons, follow the nuclei very closely. In order to describe the three types of correlation effects simultaneously with high precision in the non-BO wave function, one needs to use basis functions, which not only depend on the interelectron distances, but also explicitly depend on electron–nucleus and nucleus–nucleus distances. The functions of this type that we have used in our non-BO calculations will be shown later in this chapter. We should mention that the basis set selection is the central point in the non-BO calculation.

After the separation of the kinetic energy operator due to the center-of-mass motion from the Hamiltonian, the Hamiltonian describes the internal motions of electrons and nuclei in the system. These in the BO approximation can be separated into the vibrational and rotational motions of the nuclear frame of the molecule and the electronic motion that only parametrically depends on the instantenous positions of the nuclei. When the BO approximation is removed, the electronic and nuclear motions become coupled and the only good quantum numbers, which can be used to quantize the stationary states of the system, are the principle quantum number, the quantum number quantizing the square of the total (nuclear and electronic) squared angular momentum, and the quantum number quantizing the projection of the total angular momentum vector on a selected direction (usually the z axis). The separation of different rotational states is an important feature that can considerably simplify the calculations.

If in the non-BO calculation one chooses a basis set of eigenfunctions of the operator representing the square of the total angular momentum and the

operator representing the projection of the angular momentum on the selected axis, one can separate the calculation for different rotational states and perform them independently of each other. However, if one uses basis functions that are not rotational eigenfunctions, then the manifold of states that one gets includes all types of internal excitations (i.e., rotational, vibrational, and electronic). Since the rotational state spacing is usually much smaller in comparison with the spacings between the vibrational and electronic levels, the different vibrational and electronic levels (or, as we should more correctly call them, the vibro-electronic levels, because the vibrational and electronic motions are coupled in the non-BO calculation) are separated by a large number of rotational levels. This creates a problem if the calculations are done with the use of the variational method, because in order to determine, say, the vibrational spectrum of the molecule corresponding to the zero angular momentum, one needs to ''fish'' them out from the very high density spectrum of all internal states. There are two ways to overcome this obvious difficulty in the calculation. One, as mentioned, is to use in the calculation the basis functions of rotational eigenstates. The second is based on including in the Hamiltonian an operator that artificially shifts up the energies of states with the rotational quantum numbers different from the quantum number of the states targeted by the calculation. The first approach is what we have used in the calculations shown in this chapter. The second approach is currently being developed for calculating rotationally excited states using explicitly correlated basis functions that incorporate centers displaced away from the origin of the coordinate systems (due to this displacement, these functions are not rotational eigenfunctions).

We start this chapter by showing the coordinate transformation that allows us to separate the operator representing the kinetic energy of the center-of-mass motion from the total Hamiltonian expressed in terms of the Cartesian laboratory coordinates. Next we discuss the symmetry of the internal Hamiltonian and the spatial and permutational symmetry of the wave functions. In the following section we describe the algorithms involved in calculating Hamiltonian matrix elements and their derivatives with respect to the nonlinear parameters involved in the basis functions. We start the discussion of the numerical results by showing some atomic calculations. Next, we present some of our recent calculations on diatomic systems and we discuss their accuracy and the nonadiabatic effects they describe. In the following section we consider the interaction of a molecular system described without assuming the Born–Oppenheimer approximation with an external stationary electric field. In that section we also review some of our calculations concerning electrical properties of some small diatomic molecules. The field-dependent calculations have been done using basis functions whose centers are allowed to ''float'' away from the origin of the coordinate system to describe the polarization of the molecule along the direction of the field. In the following section we describe the use of

the "floating" basis functions in field-independent non-BO calculations. The chapter is concluded with (a) a description of directions for our future works on the non-BO approach and (b) outstanding problems that, in our view, need to be addressed in further advancing the development and the implementation of the non-BO molecular quantum mechanics.

II. HAMILTONIAN, SEPARATION OF THE CENTER OF MASS, INTERACTION WITH ELECTRIC FIELD

A. Nonrelativistic Hamiltonian

A system of $n + 1$ particles of masses M_i and charges Q_i may be described at any point in time by the $n + 1$ vectors, R_i, describing the positions of the particles:

$$R_i = \begin{pmatrix} x_i \\ y_i \\ z_i \end{pmatrix} \tag{1}$$

and the $n + 1$ vectors, P_i, describing the momenta of the particles:

$$P_i = \begin{pmatrix} P_{x,i} \\ P_{y,i} \\ P_{z,i} \end{pmatrix} \tag{2}$$

For convenience we collect the R_i vectors together:

$$R = \begin{pmatrix} R_1 \\ R_2 \\ \vdots \\ R_{n+1} \end{pmatrix} \tag{3}$$

and similarly collect the momenta together:

$$P = \begin{pmatrix} P_1 \\ P_2 \\ \vdots \\ P_{n+1} \end{pmatrix} \tag{4}$$

The kinetic energy of this system is given by

$$T = \sum_{i=1}^{n+1} \frac{P_i^2}{2M_i} \tag{5}$$

If we assume only Coulombic interactions between the particles, the potential energy is given by

$$V = \sum_{i=1}^{n+1} \sum_{j>i}^{n+1} \frac{Q_i Q_j}{r_{ij}} \tag{6}$$

where r_{ij} is the magnitude of the distance vector between particles i and j.

The total Hamiltonian for this system is thus

$$H = \sum_{i=1}^{n+1} \frac{P_i^2}{2M_i} + \sum_{i=1}^{n+1} \sum_{j>i}^{n+1} \frac{1}{4\pi\epsilon_o} \frac{Q_i Q_j}{r_{ij}} \tag{7}$$

We may transform this to the quantum-mechanical Hamiltonian operator by substitution of the configuration space operators

$$\hat{x} \rightarrow x \tag{8}$$

$$\hat{P}_x \rightarrow \frac{1}{i} \frac{\partial}{\partial x} \tag{9}$$

and we get

$$H = \sum_{i=1}^{n+1} \frac{-1}{2M_i} \nabla_i^2 + \sum_{i=1}^{n+1} \sum_{j>i}^{n+1} \frac{Q_i Q_j}{r_{ij}} \tag{10}$$

B. The Dipole Approximation

The interaction of matter particles and light calls for treating the light quantum-mechanically. It is often sufficient, though, to treat the matter quantum mechanically and the light classically via the semiclassical approach. The light, then, is treated as perpendicular oscillating electric and magnetic fields. The effect of the magnetic portion on matter is usually less than the effect of the electric portion, and so for the work presented here we may neglect the magnetic portion. If we further assume a static field, then we may express the interaction as the scalar product of the total dipole moment of the system and the field:

$$E = \mu \cdot \varepsilon \tag{11}$$

For a derivation of the above Hamiltonian, please see, for example, the book by Schatz and Ratner [71]. We may also express the total quantum-mechanical

operator as

$$\mu \cdot \varepsilon = \sum_{i=1}^{n+1} \varepsilon \cdot R_i Q_i$$

where each term in the sum is the interaction of each particle with the electric field.

The Hamiltonian for the system in an electric field is thus

$$H = \sum_{i=1}^{n+1} \frac{P_i^2}{2M_i} + \sum_{i=1}^{n+1} \sum_{j>i}^{n+1} \frac{Q_i Q_j}{r_{ij}} - \sum_{i=1}^{n+1} \varepsilon \cdot R_i Q_i \tag{12}$$

C. Transformation to Center-of-Mass Coordinates

The number of internal degrees of freedom for any system may be reduced by a transformation to center-of-mass coordinates. For example, the system of $n+1$ particles with $3(n+1)$ degrees of freedom is reduced n pseudoparticles with $3n$ degrees of freedom, with the 3 leftover degrees of freedom describing the motion of the center of mass.

For the R vector described above, it may be shown that the appropriate transformation is

$$T = \begin{pmatrix} \frac{M_1}{M_T} & \frac{M_2}{M_T} & \frac{M_3}{M_T} & \frac{M_4}{M_T} & \cdots & \frac{M_{n+1}}{M_T} \\ -1 & 1 & 0 & 0 & 0 & \cdots \\ -1 & 0 & 1 & 0 & 0 & \cdots \\ -1 & 0 & 0 & 1 & 0 & \cdots \\ -1 & 0 & 0 & 0 & 1 & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \end{pmatrix} \otimes I_3 \tag{13}$$

in the sense that T transforms R, the set of all $3(n+1)$ coordinates, into r_0, the set of coordinates describing the position of the center of mass, and r, the set of $3n$ coordinates describing the positions of the n pseudoparticles:

$$TR = \begin{pmatrix} r_0 \\ r_1 \\ \vdots \\ r_n \end{pmatrix} \tag{14}$$

or

$$TR = \begin{pmatrix} r_0 \\ r \end{pmatrix} \tag{15}$$

where r is defined as

$$r = \begin{pmatrix} r_1 \\ r_2 \\ \vdots \\ r_n \end{pmatrix} \tag{16}$$

In the above, M_T is the sum of all $n+1$ masses in the original system.

The momenta are transformed by the inverse transformation:

$$T^{-1}P = \begin{pmatrix} p_0 \\ p_1 \\ \vdots \\ p_n \end{pmatrix} \tag{17}$$

or

$$T^{-1}P = \begin{pmatrix} p_0 \\ p \end{pmatrix} \tag{18}$$

where p_0 describes the momentum of the center of mass and p, the vector of momenta of the pseudoparticles, is defined as

$$p = \begin{pmatrix} p_1 \\ p_2 \\ \vdots \\ p_n \end{pmatrix} \tag{19}$$

The inverse transformation T^{-1} is given by

$$\begin{pmatrix} 1 & -\frac{M_2}{M_T} & -\frac{M_3}{M_T} & -\frac{M_4}{M_T} & \cdots & -\frac{M_{n+1}}{M_T} \\ 1 & \frac{M_T-M_2}{M_T} & -\frac{M_3}{M_T} & -\frac{M_4}{M_T} & -\frac{M_5}{M_T} & \cdots \\ 1 & -\frac{M_2}{M_T} & \frac{M_T-M_3}{M_T} & -\frac{M_4}{M_T} & -\frac{M_5}{M_T} & \cdots \\ 1 & -\frac{M_2}{M_T} & -\frac{M_3}{M_T} & \frac{M_T-M_4}{M_T} & -\frac{M_5}{M_T} & \cdots \\ 1 & -\frac{M_2}{M_T} & -\frac{M_3}{M_T} & -\frac{M_4}{M_T} & \frac{M_T-M_5}{M_T} & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \end{pmatrix} \otimes I_3 \tag{20}$$

The charges map directly $Q_i \rightarrow q_{i-1}$, with the change on the particle at the center of mass mapping to a central potential.

These transformations result in the internal Hamiltonian (in atomic units):

$$\hat{H} = -\frac{1}{2}\left(\sum_{i}^{n}\frac{1}{m_i}\nabla_i^2 + \sum_{i\neq j}^{n}\frac{1}{M_0}\nabla_i'\nabla_j\right) + \sum_{i=1}^{n}\frac{q_0 q_i}{r_i} + \sum_{i<j}^{n}\frac{q_i q_j}{r_{ij}} - \sum_{i=1}^{n}\varepsilon\cdot r_i q_i \tag{21}$$

which may be written as

$$\hat{H} = -\nabla_r'\bar{M}\nabla_r + \sum_{i=1}^{n}\frac{q_0 q_i}{r_i} + \sum_{i<j}^{n}\frac{q_i q_j}{r_{ij}} - \sum_{i=1}^{n}\varepsilon\cdot r_i q_i$$

The mass matrix M enters the Hamiltonian for convenience of expression and is an $n \times n$ matrix with $\frac{1}{2m_i}$ on the diagonal elements and $\frac{1}{2M_0}$ on all of the off-diagonal elements; the $\bar{M}$ notation for any matrix will mean a Kronecker product with the 3×3 identity matrix, $\bar{M} = M \otimes I_3$.

III. PERMUTATIONAL SYMMETRY

Determination of a wave function for a system that obeys the correct permutational symmetry may be ensured by projection onto the irreducible representations of the symmetry groups to which the systems in question belong. For each subset of identical particles i, we can implement the desired permutational symmetry into the basis functions by projection onto the irreducible representation of the permutation group, $\mathscr{S}_{n_i}$, for total spin S_i using the appropriate projection operator $\hat{Y}_i$. The total projection operator would then be a product:

$$\hat{Y} = \prod_i \hat{Y}_i \tag{22}$$

For fermions, the projection operators are simply Young operators, derived from the appropriate Young tableau, as will be shown below.

A. Projection onto the Irreducible Representations of the *n*th-Order Symmetric Group

The energy of a quantum system is invariant to permutations of identical particles in the system. Thus, the Hamiltonian for a system with n identical particles can be said to commute with the elements of the nth-order symmetric group:

$$[\hat{H}, S_n] = 0 \tag{23}$$

This requires that the eigenfunctions of the Hamiltonian are simultaneously eigenfunctions of both the Hamiltonian and the symmetric group. This may be accomplished by taking the basis functions used in the calculations, which may be called primitive basis functions, and projecting them onto the appropriate irreducible representation of the symmetric group. After this treatment, we may call the basis functions symmetry-projected basis functions.

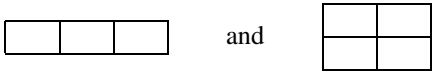
The projection operator takes the form of a sum of all of the possible permutations of the identical particles, $\hat{P}_i$, each multiplied by an appropriate constant, a_i:

$$\hat{Y} = \sum_{i=1}^{n!} a_i \hat{P}_i \tag{24}$$

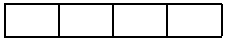
The format we will use to indicate the permutations is $(abc\ldots)$. For example, the permutation of particles 1 and 2 is (12). The permutation of particles 1, 3, and 4 is (134). The material below concerns how to obtain the coefficients, a_i.

The Pauli antisymmetry principle tells us that the wave function (including spin degrees of freedom), and thus the basis functions, for a system of identical particles must transform like the totally antisymmetric irreducible representation in the case of fermions, or spin $\frac{1}{2}k$ (for odd k) particles, and like the totally symmetric irreducible representation in the case of bosons, or spin k particles (where k may take on only integer values).

Projection operators for irreducible representations of the symmetric group are obtained easily from their corresponding Young tableaux. A Young tableau is created from a Young frame. A Young frame is a series of connected boxes such as



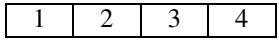
The shape of the Young frame corresponding to the desired irreducible representation of the symmetric group is obtained from the physics of the system. For example, for the totally antisymmetric representation of a group S_n we use a frame that is completely horizontal with n boxes. For example, for four particles we have



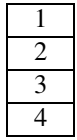
We then have to determine the dimension of the representation. This is done by filling in the boxes in the Young frame with the integers $1\ldots n$ according to the following rules:

- Numbers must increase to the right.
- Numbers must increase down.

Obviously for the above frame, we can only fill in the numbers 1, 2, 3, and 4 according to the rules in one way:



Thus this corresponds to a one-dimensional representation. The totally symmetric representation is created in a similar fashion, but with all of the boxes arranged vertically:



This is also a one-dimensional representation.

The above operators apply only to primitive basis functions that have the spin degree of freedom included. In the current work we follow the work of Matsen and use a spin-free Hamiltonian and spin-free basis functions. This approach is valid for systems wherein spin–orbit type perturbations are not considered. In this case we must come up with a different way of obtaining the Young tableaux, and thus the correct projection operators.

We explain here how to obtain the needed tableaux for fermions. We begin with the number of identical particles under consideration, n, and their total spin quantum number, s. We then calculate the symmetry quantum number, $p = \frac{n}{2} - s$. We then define a partition, μ, as

$$\mu = [2^p 1^{n-2p}] \tag{25}$$

This partition tells us to build a Young frame with

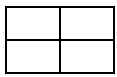
- 2 boxes in the first p rows
- 1 box in the remaining $n - 2p$ rows

We then fill in the numbers according to the above rules.

For example, for a system of four identical fermions (such as the four electrons in LiH) with all spins paired (i.e., $s = 0$), we have

$$\mu = [2^2 1^0] \tag{26}$$

and the Young frame:



Filling this in according to the rules, we have

1	2
3	4

and

1	3
2	4

Since we have two ways of filling in the Young frame (i.e., we have two tableaux), we have a two-dimensional representation.

B. Projection Operators Obtained by the Operator Method

Now that we have established how to create the Young tableaux, we must outline a method of obtaining the desired projection operators from them. Please note that some of the following material has been adapted from Pauncz [73]. A simple way to do this for small n is described as follows. We define an operator A to be the antisymmetrizer for rows of the Young tableaux:

$$A = \prod^{rows} \sum_{i=1}^{n!} \delta \hat{P}_i, \tag{27}$$

where δ is positive for odd permutations and negative for even permutations. We also define an operator S to be the symmetrizer for columns of the Young tableaux:

$$S = \prod^{columns} \sum_{i=1}^{n!} \hat{P}_i. \tag{28}$$

The Young operator, or simply the projection operator, is then the product $\hat{Y} = AS$.

For example, for the four-electron case mentioned above, we take the first Young tableau

1	2
3	4

and generate the following operators:

$$\begin{aligned} A &= \{[E - (12)][E - (34)]\} \\ S &= \{[E + (13)][E + (24)]\} \end{aligned} \tag{29}$$

where we have used E to mean the identity operator. We can then simply build the projection operator using the rules of multiplication of permutations.

C. Projection Operators Obtained by the Representation Method

Another method that may be used to generate the projection operator involves the use a matrix representation of the operator. In particular, we will use the orthogonal representation. First we must assign a Yamanouchi symbol to each tableau we have created. This is done by going through the numbers from 1 to n in each tableau and writing down in which row the number occurs. Thus if we assign names to the above tableaux:

$$T(a) = \begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 & 4 \\ \hline \end{array} \qquad \text{and} \qquad T(b) = \begin{array}{|c|c|}\hline 1 & 3 \\ \hline 2 & 4 \\ \hline \end{array}$$

we have the Yamanouchi symbols: $YS[T(a)] = 1122$ and $YS[T(b)] = 1212$. We may use these Yamanouchi symbols to order the tableaux. We start with the last number in the symbol and see which tableau has the highest number. This will the be first tableau. If two or more Tableaux have the last numbers equal, then we go to the second to last numbers and use these as the criteria. This procedure is done recursively until all of the tableaux are assigned numbers. In the case of the tableaux we are using for examples, we must go the second to last numbers to find the ordering. Thus we have

$$T_1 = \begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 & 4 \\ \hline \end{array} \qquad \text{and} \qquad T_2 = \begin{array}{|c|c|}\hline 1 & 3 \\ \hline 2 & 4 \\ \hline \end{array}$$

Now for each tableaux i, we may define the axial distance between two adjacent numbers p and $p+1$, $d^i_{p,p+1}$ (for $1 \leq p < n$) to be the number of boxes that must be traversed to reach the box containing $p+1$ starting from the box containing p. The sign assigned to the axial distance will be positive if one had to go left and down, and negative if one had to go right and up.

Now we may find the matrix representation, U, of the operators. The dimensions of the matrices will be the same as the dimensions of the irreducible representation used. The matrix representation of the identity operator, $U(\hat{E})$, will of course be the identity matrix. If it is noted that any permutation may be written as a product of transpositions (permutations of order 2), and any transposition may be written as a product of elementary transpositions $(p\ p+1)$ [74], then it is only nessesary to find matrix representations of the elementary transpositions. The diagonal elements of the elementary transposition $(p\ p+1)$ are given by

$$U[(p\ p+1)]_{ii} = \frac{-1}{d^i_{p,p+1}} = D(i,k) \tag{30}$$

Thus one needs to know all of the i Young tableaux, as well as all of the axial distances in order to construct the matrix. The off-diagonal element is given by

$$U[(p\ p+1)]_{ij} = \sqrt{1 - D(i,k)^2} \tag{31}$$

if the tableaux i and j differ only by the transposition of p and $p+1$. Otherwise the off-diagonal elements are zero. The nonelementary transpositions may be obtained from the elementary transpositions with the use of the identity:

$$(i\ j+1) = (j\ j+1)(i\ j)(j\ j+1) \tag{32}$$

For the above tableaux and the associated symmetric group S_4, we only need the matrices $U[(12)], U[(23)]$, and $U[(34)]$. Thus for the two tableaux, we have only six axial distances: $d^1_{(12)} = -1, d^1_{(23)} = 2, d^1_{(34)} = -1, d^2_{(12)} = 1, d^2_{(23)} = -2$, and $d^2_{(34)} = 1$. Using these, one can find all of the representation matrices.

In the operator method described above, one may obtain a projection operator for each tableaux, so in the representation method, the same must be true. The projection operators are thus

$$\hat{Y}_{pp} = \sum_{i=1}^{n!} U[\hat{P}_i]_{pp} \hat{P}_i \tag{33}$$

where the sum is over the permutations, and $1 \leq p \leq$ dimension of the representation.

D. Decomposition of S_n in Cosets of S_{n-1}

The process of obtaining all of the representation matrices can become quite tedious as n increases, since the number of permutations increases as $n!$. This may be simplified by the decomposition of larger groups in cosets of smaller groups.

If we have a group, G of order $|g|$, and a subgroup, H of order $|h|$, where $|h| < |g|$, there is some set of elements g_i such that $g_i \in G$, but $g_i \notin H$. For the elements of H, h_j, it can be shown that $p_k = g_i \times h_j$ is not an element of H. Thus if we chose some g_i and multiply it by all of the elements of H, g_iH, this generates another subgroup, H' of order $|h|$, whose elements h_j' are not in H. If all of the elements of G are used up in H and H'—that is, if $|g| = 2\ |h|$—then we are done. If, for example, $|g|/|h| = 3$, then we can choose some other element g_k, where $g_k \notin H$ and $g_k \notin H'$ and form another subgroup $g_kH = H''$. We may continue this procedure until all of the elements of G are used up.

These subgroups of G generated by multiplication of some element of G by the subgroup H are called cosets of H in G. If the multiplication is carried out on the left, they are called left cosets, and vice versa for right multiplication.

If we wanted to generate the representation matrices of S_4, we would find the three elementary transpositions and the identity and then generate the other 20 matrices. On the other hand, we could find the representations of the group S_2 in S_4. This consists of two matrices, $U[(E)]$ and $U[(12)]$. We could then use $U[(23)]$ and $U[(13)]$ to generate all of the six elements in S_3. We could then use $U[(34)]$, $U[(14)]$, and $U[(24)]$ to generate the rest of S_4. While this may seem at first more time-consuming, it is much more easily automated than the brute force approach.

E. Effects of Permutations on Basis Functions

The permutations discussed above act on the particle coordinates. In a less symbolic, more mathematical footing, we can consider the permutations as transformation matrices, $\hat{P}$, which act on the coordinate vector, R, turning them into the permuted coordinates. For example, if we consider the H_2^+ molecule with the coordinate vector

$$R = \begin{pmatrix} R_{H1} \\ R_{H2} \\ R_{e1} \end{pmatrix} \tag{34}$$

then the transformation matrix permuting the two protons would be

$$\hat{P} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \otimes I_3 \tag{35}$$

so that we have

$$\hat{P}R = \begin{pmatrix} R_{H2} \\ R_{H1} \\ R_{e1} \end{pmatrix} \tag{36}$$

Other permutations are done in a similar manner. When a transformation to the center of mass ($\hat{T}$) is performed, this also affects the permutations: $\bar{P} = \hat{T}^{-1}\hat{P}\hat{T}$. After this transformation, the new permutation acts on r: $\hat{P}R = \bar{P}r$. From now on we will refer to all permutations as $\hat{P}$, and center-of-mass transformation will have to be inferred from the context.

These permutations on coordinates are equivalent to operations on the basis functions. We will use shifted spherical Gaussians for this example (these functions will be discussed in a detailed way below in this chapter):

$$\begin{aligned} Pg &= \exp\{-(PR - s)'\bar{A}(PR - s)\} \\ &= \exp\{-(PR - PP^{-1}s)'\bar{A}(PR - PP^{-1}s)\} \\ &= \exp\{-[P(R - P^{-1}s)]'\bar{A}P(R - P^{-1}s)\} \\ &= \exp\{-(R - P^{-1}s)'P'\bar{A}P(R - P^{-1}s)\} \end{aligned} \tag{37}$$

Here we dropped the hat on the P for convenience. So in general we need not only the permutations, but also their inverses. The inverses are easy to obtain, however.

In the case of transpositions, such as that described above, it is obvious that $P = P^{-1} = P'$. Also, for products of transpositions, P_a and P_b, where $[P_a, P_b] = 0$, we also have $P = P^{-1}$. Furthermore, still for transpositions, $\bar{P} = \bar{P}^{-1}$, since

$$\begin{aligned} PP &= 1 \\ PTT^{-1}P &= 1 \\ T^{-1}PTT^{-1}PT &= T^{-1}T \\ \bar{P}\bar{P} &= 1 \end{aligned} \tag{38}$$

and thus $\bar{P} = \bar{P}^{-1}$.

Going from transpositions to permutations of higher order, we make use of the fact that any permutation of order n may be written as a product of $n - 1$ transpositions. For the permutation of order 3, we have $P = P_aP_b$ where P_a and P_b are transpositions that do not necessarily commute. We find the inverse easily:

$$\begin{aligned} P &= P_aP_b \\ P^{-1} &= (P_aP_b)^{-1} \\ P^{-1} &= P_b^{-1}P_a^{-1} \end{aligned}$$

but since P_a and P_b are transpositions,

$$\begin{aligned} P^{-1} &= P_bP_a \\ P^{-1} &= P_b'P_a' \\ P^{-1} &= P' \end{aligned} \tag{39}$$

The same can be shown to be true for higher-order permutations as well.

IV. ATOMIC NON-BO CALCULATIONS

The symmetry requirements and the need to very effectively describe the correlation effects have been the main motivations that have turned our attention to explicitly correlated Gaussian functions as the choice for the basis set in the atomic and molecular non-BO calculations. These functions have been used previously in Born–Oppenheimer calculations to describe the electron correlation in molecular systems using the perturbation theory approach [35–42]. While in those calculations, Gaussian pair functions (geminals), each dependent only on a single interelectron distance in the exponential factor, $\exp(-\beta r_{ij}^2)$, were used, in the non-BO calculations each basis function needs to depend on distances between all pairs of particles forming the system.

In our non-BO calculations performed so far, we have considered atomic systems with only *s*-electrons and molecular systems with only σ-electrons. The atomic non-BO calculations are much less complicated than the molecular calculations. After separation of the center-of-mass motion from the Hamiltonian and placing the atom nucleus in the center of the coordinate system, the internal Hamiltonian describes the motion of light pseudoelectrons in the central field on a positive charge (the charge of the nucleus) located in the origin of the internal coordinate system. Thus the basis functions in this case have to be able to accurately describe only the electronic correlation effect and the spherically symmetric distribution of the electrons around the central positive charge.

In our atomic calculations, the *s*-type explicitly correlated Gaussian functions have the following form:

$$\phi_k = \exp[-\mathbf{r}'(A_k \otimes I_3)\mathbf{r}] \tag{40}$$

The above function is a one-center correlated Gaussian with exponential coefficients forming the symmetric matrix A_k. ϕ_k are rotationally invariant functions as required by the symmetry of the problem—that is, invariant with respect to any orthogonal transformation. To show the invariance, let U be any 3×3 orthogonal matrix (any proper or improper rotation in 3-space) that is applied to rotate the $\mathbf{r}$ vector in the 3–D space. Prove the invariance:

$$\begin{aligned}((I_n \otimes U)\mathbf{r})'(A_k \otimes I_3)(I_n \otimes U)\mathbf{r} &= \mathbf{r}'(I_n \otimes U')(A_k \otimes I_3)(I_n \otimes U)\mathbf{r} &(41)\\ &= \mathbf{r}'(A_k \otimes U'U)\mathbf{r} &(42)\\ &= \mathbf{r}'(A_k \otimes I_3)\mathbf{r} &(43)\end{aligned}$$

The n-particle one-center correlated Gaussians, ϕ_k, can also be expressed in the more conventional form used in the electronic structure calculations as

$$\begin{aligned}\phi_k = \exp\big[&-\alpha_{1k}r_1^2 - \alpha_{2k}r_2^2 - \cdots - \alpha_{nk}r_n^2\\ &-\beta_{12,k}r_{12}^2 - \beta_{13,k}r_{13}^2 - \cdots - \beta_{nn-1,k}r_{nn-1}^2\big]\end{aligned} \tag{44}$$

In this form, the n-particle correlated Gaussian is a product of n orbital Gaussians centered at the origin of the coordinate system and $n(n-1)/2$ Gaussian pair functions (geminals).

To describe bound stationary states of the system, the ϕ_k's have to be square-normalizable functions. The square-integrability of these functions may be achieved using the following general form of an n-particle correlated Gaussian with the negative exponential of a positive definite quadratic form in $3n$ variables:

$$\phi_k = \exp\left[-\mathbf{r}'\left(L_k L_k' \otimes I_3\right)\mathbf{r}\right] \tag{45}$$

Here $\mathbf{r}$ is a $3n \times 1$ vector of Cartesian coordinates for the n pseudoelectrons, and L_k is an $n \times n$ lower triangular matrix of rank n whose elements may vary in the range $[-\infty, \infty]$.

As mentioned, most calculations we have done so far have concerned molecular systems. However, prior to development of the non-BO method for the diatomic systems, we performed some very accurate non-BO calculations of the electron affinities of H, D, and T [43]. The difference in the electron affinities of the three systems is a purely nonadiabatic effect resulting from different reduce masses of the pseudoelectron. The pseudoelectrons are the heaviest in the T/T^- system and the lightest in the H/H^- system. The calculated results and their comparison with the experimental results of Lineberger and coworkers [44] are shown in Table I. The calculated results include the relativistic, relativistic recoil, Lamb shift, and finite nuclear size corrections labeled ΔE_{corr} calculated by Drake [45]. The agreement with the experiment for H and D is excellent. The 3.7-cm^{-1} increase of the electron affinity in going from H to D is very well reproduced by the calculations. No experimental EA value is available for T.

TABLE I
Electron Affinity of Hydrogen, Deuterium, and Tritium Atoms Obtained with 300 Explicitly Correlated Gaussian Functions[a]

	Hydrogen	Deuterium	Tritium
$E_H - E_{H^-}$	6083.4058 cm^{-1}	6087.0201 cm^{-1}	6088.2233 cm^{-1}
ΔE_{corr}[b]	0.307505 cm^{-1}	0.307589 cm^{-1}	0.307616 cm^{-1}
EA	6083.0983 cm^{-1}	6086.7126 cm^{-1}	6087.9157 cm^{-1}
Lykke et al. (experiment)[c]	6082.99 ± 0.15 cm^{-1}	6086.2 ± 0.6 cm^{-1}	

[a]The term ΔE_{corr} contains relativistic, relativistic recoil, Lamb shift, and finite nuclear size corrections [43].
[b]Reference 45.
[c]Reference 44.

The H/H$^-$, D/D$^-$, and T/T$^-$ calculations were done using the variational method and 300 Gaussian functions per system. While these many functions ensure adequate convergence of the calculation for small atoms, it is usually far from adequate for even the smallest diatomic molecules. Later in this chapter, we will show calculations for HD^+ and H_2 systems where 2000 and even more basis functions were required.

V. DIATOMIC NON-BO CALCULATIONS

A. Correlated Gaussian Basis Set

The general form of an n-pseudoparticle correlated Gaussian function is given by

$$\phi_k = \exp[-\mathbf{r}' A_k \mathbf{r}] \tag{46}$$

Here, $\mathbf{r}$ is a $n \times 1$ vector of Cartesian coordinates of pseudoparticles, such that

$$\mathbf{r}' \equiv (\mathbf{r}'_1, \mathbf{r}'_2, \ldots, \mathbf{r}'_n) \tag{47}$$

and A_k is a symmetric $n \times n$ matrix of nonlinear variational parameters,

$$A_k \equiv \begin{pmatrix} (A_k)_{1,1} & (A_k)_{1,2} & \cdots & (A_k)_{1,n} \\ (A_k)_{2,1} & (A_k)_{2,2} & \cdots & (A_k)_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ (A_k)_{n,1} & (A_k)_{n,2} & \cdots & (A_k)_{n,n} \end{pmatrix}$$

Effectively, vector $\mathbf{r}$ has $3 \times n \times 1$ components since each $\mathbf{r}_i$ in (47) is itself a three-dimensional vector. Technically speaking, in place of A_k in (46), one should write the Kronecker product $A_k \otimes I_3$ with I_3 being the 3×3 identity matrix. However, to simplify notations and avoid writing routinely this obvious Kronecker product, below in this section we will be using the following convention for matrix–vector multiplications involving such vectors:

$$\begin{aligned} \mathbf{r}' A_k \mathbf{r} &= (\mathbf{r}'_1, \mathbf{r}'_2, \ldots, \mathbf{r}'_n) \begin{pmatrix} (A_k)_{1,1}\mathbf{r}_1 + (A_k)_{1,2}\mathbf{r}_2 + \cdots + (A_k)_{1,n}\mathbf{r}_n \\ (A_k)_{2,1}\mathbf{r}_1 + (A_k)_{2,2}\mathbf{r}_2 + \cdots + (A_k)_{2,n}\mathbf{r}_n \\ \vdots \\ (A_k)_{n,1}\mathbf{r}_1 + (A_k)_{n,2}\mathbf{r}_2 + \cdots + (A_k)_{n,n}\mathbf{r}_n \end{pmatrix} \\ &= (A_k)_{1,1}\mathbf{r}'_1\mathbf{r}_1 + (A_k)_{1,2}\mathbf{r}'_1\mathbf{r}_2 + \cdots + (A_k)_{n,n}\mathbf{r}'_n\mathbf{r}_n \end{aligned}$$

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 treating $\mathbf{r}_i$ 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 considered as a dot product if no special notice is made.

A set of nonlinear parameters A_k, in general case, is unique for each function ϕ_k. To satisfy the requirement of square integrability of the wave function, each matrix A_k must be positively defined. It imposes certain restrictions on the values that the elements of matrix A_k may take. To ensure the positive definiteness and to simplify some calclations, it is very convenient to represent matrix A_k in a Cholesky factored form,

$$A_k = L_k L_k' \tag{48}$$

where the elements of lower triangular matrix L_k may take any real values.

Functions (46) have been succesfully used in numerous quantum-mechanical variational calculations of atomic and exotic systems where there is, at most, one particle (nuclei), which is substantially heavier than other constituents. However, as is well known, simple correlated Gaussian functions centered at the origin cannot provide a satisfactory convergence rate for nearly adiabatic systems, such as molecules, containing two or more heavy particles. In the diatomic case, which we we will mainly be concerned with in this section, one may introduce in basis functions (46) additional factors of powers of the internuclear distance. Such factors shift the peaks of Gaussians to some distance from the origin. This allows us to adequately describe the localization of nuclei around their equilibrium position.

If we label the two heavy particles as one and two, then the distance between these particles is given by $r_1 = |\mathbf{R_1} - \mathbf{R_2}|$ and the basis functions have the following form:

$$\varphi_k = r_1^{m_k} \exp[-\mathbf{r}' A_k \mathbf{r}] \tag{49}$$

$r_1^{m_k}$ can be conveniently written as the square root of a quadratic form in $\mathbf{r}$ using the matrix J_{11}, which is defined as an $n \times n$ matrix with 1 in the 1, 1 position and with 0's elsewhere:

$$r_1^m = [\mathbf{r}' J_{11} \mathbf{r}]^{m/2} \tag{50}$$

Similarly, any preudoparticle coordinate $\mathbf{r}_i$ or interpseudoparticle distance $\mathbf{r}_{ij}$ can be represented as

$$r_{ij} = \left[\mathbf{r}' J_{ij} \mathbf{r}\right]^{1/2} \tag{51}$$

$$J_{ij} = \begin{cases} E_{ii} & \text{if } i = j \\ E_{ii} + E_{jj} - E_{ij} - E_{ji} & \text{if } i \neq j \end{cases} \tag{52}$$

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

B. Implementation of the Permutational Symmetry

Permutations of real particles induce transformations on internal coordinates. Let P be a permutation of real particles; then P transforms basis functions (49) as

$$P\varphi_k = r_1^{m_k} \exp\left[-\mathbf{r}'\left(\tau_P' A_k \tau_P\right)\mathbf{r}\right] \tag{53}$$

Here, the $n \times n$ matrix τ_P is not an elementary permutation matrix. In case when P is a transposition corresponding to the interchange of 1st and jth particle ($P \equiv P_{1j}$), the matrix $\tau_{P_{1j}}$ is the identitity matrix with all elements in $(j-1)$th column replaced with -1. If P is a transposition corresponding to the interchange of ith and jth particle ($i \neq 1$ and $i \neq 1$), then $\tau_{P_{ij}}$ is the identity matrix whose $(i-1),(i-1)$th and $(j-1),(j-1)$th elements are replaced with 0's and $(i-1),(j-1)$th and $(j-1),(i-1)$th elements are 1's. All transformation matrices τ_P for permutations that are not transpositions can be represented as products of τ_P corresponding to a certain sequence of transpositions. The latter is due to the fact that any permutation can be represented as a product of certain transpositions.

Note that the factor $r_1^{m_k}$ is invariant under any permutation on systems of particles for which φ_k is a valid basis.

A symmetry projector, $\mathscr{P}$, for an irreducible representation of the permutational symmetry group of a system is given by

$$\mathscr{P} = \sum_P \chi_P P \mapsto \sum_P \chi_P \tau_P \tag{54}$$

Hence, $\mathscr{P}$ acts on φ_k as

$$\mathscr{P}\varphi_k = \sum_P \chi_P \; r_1^{m_k} \exp\left[-\mathbf{r}'\left(\tau_P' A_k \tau_P\right)\mathbf{r}\right] \tag{55}$$

The coefficients χ_P are from the matrix elements of the irreducible representation for the desired state.

Computational effort for computing matrix elements with symmetry-projected basis functions can be reduced by a factor equal to the order of the group by exploiting commutation of the symmetry projectors with the Hamiltonian and identity operators. In general,

$$\langle \mathscr{P}\varphi_k | H | \mathscr{P}\varphi_l \rangle = \langle \varphi_k | H \big| \mathscr{P}^\dagger \mathscr{P} \varphi_l \rangle \tag{56}$$

Thus, symmetry projection need only be performed on the ket. Typically, projection operators are Hermitian and essentially idempotent: $\mathscr{P}^\dagger \mathscr{P} \propto \mathscr{P}$ in any case, and we will simply write $\mathscr{P}$ for ket projector.

The matrix elements needed in calculations are of the form

$$O_{kl} = \langle \varphi_k | O | \mathscr{P} \varphi_l \rangle = \sum_P \chi_P \left\langle r_1^{m_k} \exp[-\mathbf{r}' A_k \mathbf{r}] \middle| O \middle| r_1^{m_l} \exp\left[-\mathbf{r}'\left(\tau_P' A_l \tau_P\right)\mathbf{r}\right] \right\rangle$$

where O is some operator that commutes with all the permutations from the symmetry group of the system under consideration. We will not write the summation over terms in the symmetry projector in the formulas that follow and simply note that these terms are accounted for by making the following substitution in the integral formulae:

$$A_l \mapsto \tau_P' A_l \tau_P = \tau_P' L_l \left(\tau_P' L_l\right)' \tag{57}$$

No other modification to the integral formulae need be made.

C. Integrals and Energy Derivatives

The evaluation of matrix elements for explicitly correlated Gaussians (46) and (49) can be done in a very elegant and relatively simple way using matrix differential calculus. A systematic description of this very powerful mathematical tool is given in the book by Magnus and Neudecker [105]. The use of matrix differential calculus allows one to obtain compact expressions for matrix elements in the matrix form, which is very suitable for numerical computations [116, 118] and perhaps facilitates a new theoretical insight. The present section is written in the spirit of Refs. 116 and 118, following most of the notation conventions therein. Thus, the reader can look for information about some basic ideas presented in these references if needed.

1. Some Notations

The vec operator transforms a matrix into a vector by stacking the columns of the matrix one underneath the other. Let B be an $m \times n$ matrix and let b_j be its jth column; then vec B is the $mn \times 1$ vector

$$\operatorname{vec} B = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} \tag{58}$$

An operator similar to vec is the vech, ''vector half,'' operator. Let B be a square $n \times n$ matrix . Then vech B is the $n(n+1)/2 \times 1$ vector obtained by stacking the

lower triangular elements of B. For example, if $n = 3$, then

$$\operatorname{vech} B = \begin{bmatrix} B_{11} \\ B_{21} \\ B_{31} \\ B_{22} \\ B_{32} \\ B_{33} \end{bmatrix} \tag{59}$$

For symmetric B, $\operatorname{vech} B$ contains the independent elements of B.

Other notation used: $\operatorname{diag} B$ is the diagonal $n \times n$ matrix consisting of the diagonal elements of the square matrix B. The trace of B is denoted $\operatorname{tr} B$, and the determinant of B is denoted $|B|$. The Kronecker product of two matrices is denoted by symbol $\otimes$. Other notation will be introduced as needed.

2. *Overlap Matrix Elements*

The following well-known integral is used in the derivations below:

$$\int_{-\infty}^{\infty} \exp[-x'Ax + b'x]dx_1 dx_2 \ldots dx_n = \frac{\pi^{n/2}}{|A|^{1/2}} \exp\left[\frac{1}{4} b'A^{-1}b\right] \tag{60}$$

Here, A is a positive definite $n \times n$ matrix, A^{-1} is the inverse of A, b is $n \times 1$ vector, and x is an $n \times 1$ vector variable.

The overlap of simple Gaussians (46) follows directly from (60). If $A_{kl} = A_k + A_l$,

$$\langle \phi_k | \phi_l \rangle = \frac{\pi^{3n/2}}{|A_{kl}|^{3/2}} \tag{61}$$

For further derivations we will need the matrix element of the Dirac delta function, $\delta(\mathbf{r}_{ij} - \boldsymbol{\xi})$. Using the following representation of the delta function,

$$\delta(\mathbf{r}_{ij} - \boldsymbol{\xi}) = \lim_{\alpha \to \infty} \left(\frac{\alpha}{\pi}\right)^{3/2} \exp\left[-\alpha(\mathbf{r}_{ij} - \boldsymbol{\xi})^2\right]$$

we can obtain

$$\begin{aligned} \left\langle \phi_k \middle| \delta(\mathbf{r}_{ij} - \boldsymbol{\xi}) \middle| \phi_l \right\rangle &= \lim_{\alpha \to \infty} \frac{\alpha^{3/2}}{\pi^{3/2}} \left\langle \phi_k \middle| \exp\left[-\alpha r_{ij}^2 + 2\alpha \mathbf{r}_{ij}\boldsymbol{\xi} - \alpha\xi^2\right] \middle| \phi_l \right\rangle \\ &= \lim_{\alpha \to \infty} \frac{\alpha^{3/2}}{\pi^{3/2}} \left\langle \phi_k \middle| \exp\left[-\alpha \mathbf{r}' J_{ij} \mathbf{r} + 2\alpha(j_i'\mathbf{r} - j_j'\mathbf{r})\boldsymbol{\xi} - \alpha\xi^2\right] \middle| \phi_l \right\rangle \\ &= \lim_{\alpha \to \infty} \frac{\alpha^{3/2}}{\pi^{3/2}} \frac{\pi^{3n/2}}{|A_{kl} + \alpha J_{ij}|^{3/2}} \exp\left[\alpha^2 \xi^2 (j_i' - j_j')(A_{kl} + \alpha J_{ij})^{-1}(j_i - j_j) - \alpha\xi^2\right] \end{aligned} \tag{62}$$

In the second line we introduced an $n \times 1$ vector j_i whose ith component is 1 and all others are zeros. Since J_{ij} is a rank 1 matrix, we can rewrite the determinant in the last formula in the following form:

$$|A_{kl} + \alpha J_{ij}| = |A_{kl}||I + \alpha A_{kl}^{-1} J_{ij}| = |A_{kl}|(1 + \alpha \mathrm{tr}[A_{kl}^{-1} J_{ij}])$$

where I is $n \times n$ identity matrix. Because the limit of the preexponential part of (62) is a finite number, the limit of the exponent must be $-\beta\xi^2$, with β being some finite number:

$$\begin{aligned}\langle \phi_k | \delta(\mathbf{r}_{ij} - \boldsymbol{\xi}) | \phi_l \rangle &= \frac{\pi^{3(n-1)/2}}{|A_{kl}|^{3/2}} \frac{1}{\mathrm{tr}[A_{kl}^{-1} J_{ij}]^{3/2}} \exp[-\beta\xi^2] \\ &= \langle \phi_k | \phi_l \rangle \frac{1}{\pi^{3/2}} \frac{1}{\mathrm{tr}[A_{kl}^{-1} J_{ij}]^{3/2}} \exp[-\beta\xi^2] \end{aligned} \tag{63}$$

Using the normalization condition,

$$\int \langle \phi_k | \delta(\mathbf{r}_{ij} - \boldsymbol{\xi}) | \phi_l \rangle d\boldsymbol{\xi} = \langle \phi_k | \phi_l \rangle$$

We can easily see that $\beta = \mathrm{tr}[A_{kl}^{-1} J_{ij}]^{-1}$. Thus,

$$\langle \phi_k | \delta(\mathbf{r}_{ij} - \boldsymbol{\xi}) | \phi_l \rangle = \langle \phi_k | \phi_l \rangle \frac{1}{\pi^{3/2}} \frac{1}{\mathrm{tr}[A_{kl}^{-1} J_{ij}]^{3/2}} \exp\left[- \frac{\xi^2}{\mathrm{tr}[A_{kl}^{-1} J_{ij}]} \right] \tag{64}$$

The last relationship allows one to evaluate the matrix element of an arbitrary function $f(\mathbf{r}_{ij})$, which depends on a single pseudoparticle coordinate or a single interpseudoparticle coordinate,

$$\begin{aligned}\langle \phi_k | f(\mathbf{r}_{ij}) | \phi_l \rangle &= \int f(\mathbf{r}_{ij}) \langle \phi_k | \delta(\mathbf{r}_{ij} - \boldsymbol{\xi}) | \phi_l \rangle d\boldsymbol{\xi} \\ &= \langle \phi_k | \phi_l \rangle \frac{1}{\pi^{3/2}} \int f\left(\frac{\boldsymbol{\xi}}{\mathrm{tr}[A_{kl}^{-1} J_{ij}]^{1/2}} \right) e^{-\xi^2} d\boldsymbol{\xi} \end{aligned} \tag{65}$$

In the most important case, when f depends only on the absolute value of the interpseudoparticle distance, this formula becomes

$$\langle \phi_k | f(r_{ij}) | \phi_l \rangle = \langle \phi_k | \phi_l \rangle \frac{4}{\sqrt{\pi}} \int_0^\infty f\left(\mathrm{tr}[A_{kl}^{-1} J_{ij}]^{1/2} \xi \right) \xi^2 e^{-\xi^2} d\xi \tag{66}$$

Having this general expression we now can obtain the overlap matrix element for basis functions with premultipliers (49):

$$\begin{aligned}\langle\varphi_k|\varphi_l\rangle &= \left\langle\phi_k\middle|r_1^{m_k+m_l}\middle|\phi_l\right\rangle \\ &= \frac{2}{\sqrt{\pi}}\Gamma\left(\frac{m_{kl}+3}{2}\right)\mathrm{tr}[A_{kl}^{-1}J_{11}]^{m_{kl}/2}\langle\phi_k|\phi_l\rangle \\ &= \frac{2}{\sqrt{\pi}}\Gamma\left(\frac{m_{kl}+3}{2}\right)(A_{kl}^{-1})_{11}^{m_{kl}/2}\langle\phi_k|\phi_l\rangle \end{aligned} \tag{67}$$

Here, $\Gamma(x)$ is the Euler gamma function and $m_{kl} = m_k + m_l$.

In real calculations, it is advisable to use normalized basis functions in order to avoid problems with numerical instabilities. Therefore, the overlap matrix elements, S_{kl}, are defined using normalized basis functions. After simplification, we obtain

$$\begin{aligned}S_{kl} &= \frac{\langle\varphi_k|\varphi_l\rangle}{\left(\langle\varphi_k|\varphi_k\rangle\langle\varphi_l|\varphi_l\rangle\right)^{1/2}} \\ &= \gamma_1(m_k,m_l)2^{3n/2}\left[\left(\frac{(A_{kl}^{-1})_{11}}{(A_k^{-1})_{11}}\right)^{m_k}\left(\frac{(A_{kl}^{-1})_{11}}{(A_l^{-1})_{11}}\right)^{m_l}\left(\frac{||L_k||\ ||L_k||}{|A_{kl}|}\right)^3\right]^{1/2}\end{aligned} \tag{68}$$

where $||L_k||$ denotes the absolute value of the determinant of the Cholesky factor of A_k, and

$$\gamma_1(m_k,m_l) = 2^{\frac{m_{kl}}{2}}\frac{\Gamma\left(\frac{m_{kl}+3}{2}\right)}{\left[\Gamma\left(m_k+\frac{3}{2}\right)\Gamma\left(m_l+\frac{3}{2}\right)\right]^{1/2}} \tag{69}$$

is a constant that can be precomputed for a range of m_k and m_l values to speed up matrix element calculations.

3. *Kinetic Energy Matrix Elements*

To evaluate the integral

$$\left\langle\varphi_k\middle|-\nabla'_{\mathbf{r}}M\nabla_{\mathbf{r}}\middle|\varphi_l\right\rangle = \left\langle\nabla'_{\mathbf{r}}r_1^{m_k}\phi_k|M|\nabla_{\mathbf{r}}r_1^{m_l}\phi_l\right\rangle$$

a few preliminary results will be needed. Recalling (50), we first evaluate the gradient of φ_k with respect to vector $\mathbf{r}$:

$$\begin{aligned}\nabla_{\mathbf{r}}\phi_k &= -2\phi_k A_k\mathbf{r} \\ \nabla_{\mathbf{r}}\varphi_k &= \nabla_{\mathbf{r}}[\mathbf{r}'J_{11}\mathbf{r}]^{m_k/2}\phi_k = [\mathbf{r}'J_{11}\mathbf{r}]^{m_k/2}\phi_k\left(\frac{m_k}{r_1^2}J_{11}\mathbf{r} - 2A_k\mathbf{r}\right)\end{aligned}$$

Hence,

$$\begin{aligned}
&\langle\varphi_k|-\nabla'_{\mathbf{r}}M\nabla_{\mathbf{r}}|\varphi_l\rangle\\
&=\left\langle\phi_k\left|[\mathbf{r}'J_{11}\mathbf{r}]^{m_k/2}\left(\frac{m_k}{r_1^2}J_{11}\mathbf{r}-2A_k\mathbf{r}\right)M[\mathbf{r}'J_{11}\mathbf{r}]^{m_l/2}\left(\frac{m_l}{r_1^2}J_{11}\mathbf{r}-2A_l\mathbf{r}\right)\right|\phi_l\right\rangle\\
&=\left\langle\phi_k\left|[\mathbf{r}'J_{11}\mathbf{r}]^{m_{kl}/2}\frac{m_km_l}{r_1^4}\mathbf{r}'J_{11}MJ_{11}\mathbf{r}\right|\phi_l\right\rangle-2\left\langle\phi_k\left|[\mathbf{r}'J_{11}\mathbf{r}]^{m_{kl}/2}\frac{m_k}{r_1^2}\mathbf{r}'J_{11}MA_l\mathbf{r}\right|\phi_l\right\rangle\\
&\quad-2\left\langle\phi_k\left|[\mathbf{r}'J_{11}\mathbf{r}]^{m_{kl}/2}\frac{m_l}{r_1^2}\mathbf{r}'A_kMJ_{11}\mathbf{r}\right|\phi_l\right\rangle+4\left\langle\phi_k\left|[\mathbf{r}'J_{11}\mathbf{r}]^{m_{kl}/2}\mathbf{r}'A_kMA_l\mathbf{r}\right|\phi_l\right\rangle
\end{aligned}$$

Making use of the fact that for an arbitrary matrix B,

$$\mathbf{r}'B\mathbf{r}=\mathrm{tr}[B\mathbf{r}\mathbf{r}']=(\mathrm{vec}\,B')'\mathrm{vec}(\mathbf{r}\mathbf{r}')$$

and replacing $\mathbf{r}'J_{11}MJ_{11}\mathbf{r}$ with $r_1^2M_{11}$, we now have

$$\begin{aligned}
&\langle\varphi_k|-\nabla'_{\mathbf{r}}M\nabla_{\mathbf{r}}|\varphi_l\rangle\\
&=m_km_lM_{11}\left\langle\varphi_k\left|\frac{1}{r_1^2}\right|\varphi_l\right\rangle-2m_k(\mathrm{vec}(A_lMJ_{11}))'\left\langle\varphi_k\left|\frac{\mathrm{vec}(\mathbf{r}\mathbf{r}')}{r_1^2}\right|\varphi_l\right\rangle\\
&\quad-2m_l(\mathrm{vec}(J_{11}MA_k))'\left\langle\varphi_k\left|\frac{\mathrm{vec}(\mathbf{r}\mathbf{r}')}{r_1^2}\right|\varphi_l\right\rangle+4(\mathrm{vec}(A_lMA_k))'\langle\varphi_k|\mathrm{vec}(\mathbf{r}\mathbf{r}')|\varphi_l\rangle\\
&=m_km_lM_{11}\left\langle\varphi_k\left|\frac{1}{r_1^2}\right|\varphi_l\right\rangle+2m_k(\mathrm{vec}(A_lMJ_{11}))'\frac{\partial}{\partial\mathrm{vec}A_{kl}}\left\langle\varphi_k\left|\frac{1}{r_1^2}\right|\varphi_l\right\rangle\\
&\quad+2m_l(\mathrm{vec}(J_{11}MA_k))'\frac{\partial}{\partial\mathrm{vec}A_{kl}}\left\langle\varphi_k\left|\frac{1}{r_1^2}\right|\varphi_l\right\rangle-4(\mathrm{vec}(A_lMA_k))'\frac{\partial}{\partial\mathrm{vec}A_{kl}}\langle\varphi_k|\varphi_l\rangle
\end{aligned}$$

In the last formula we used the relation $\langle\varphi_k|\mathrm{vec}(\mathbf{r}\mathbf{r}')|\varphi_l\rangle=-\frac{\partial}{\partial\mathrm{vec}A_{kl}}\langle\varphi_k|\varphi_l\rangle$. The gradient of the overlap with respect to $(\mathrm{vec}A_{kl})'$ is

$$\begin{aligned}
&\frac{\partial}{\partial(\mathrm{vec}A_{kl})'}\langle\varphi_k|\varphi_l\rangle=2\pi^{(3n-1)/2}\Gamma\left(\frac{m_{kl}+3}{2}\right)\\
&\times\left\{\mathrm{tr}[A_{kl}^{-1}J_{11}]^{m_{kl}/2}\frac{\partial}{\partial(\mathrm{vec}A_{kl})'}\frac{1}{|A_{kl}|^{3/2}}+\frac{1}{|A_{kl}|^{3/2}}\frac{\partial}{\partial(\mathrm{vec}A_{kl})'}\mathrm{tr}[A_{kl}^{-1}J_{11}]^{m_{kl}/2}\right\}
\end{aligned}\tag{70}$$

It is known from matrix differential calculus that for a matrix variable X and a constant matrix C the following is true:

$$\begin{aligned}
d|X|=|X|\mathrm{tr}[X^{-1}dX]=|X|(\mathrm{vec}(X^{-1})')'d\mathrm{vec}(X)\\
\frac{\partial|X|}{\partial(\mathrm{vec}X)'}=|X|(\mathrm{vec}(X^{-1})')'
\end{aligned}$$

and

$$d\mathrm{tr}[CX^{-1}] = (\mathrm{vec}C')'\mathrm{vec}(X^{-1}) = (\mathrm{vec}C')'\left[-(X')^{-1}\otimes X^{-1}\right]d\mathrm{vec}X,$$

$$\frac{\partial \mathrm{tr}[CX^{-1}]}{\partial \mathrm{vec}(X)'} = -(\mathrm{vec}C')'\left[(X')^{-1}\otimes X^{-1}\right]$$

Applying these differentiation rules to the expression (70), one obtains

$$\frac{\partial}{\partial(\mathrm{vec}A_{kl})'}\langle\varphi_k|\varphi_l\rangle = \langle\varphi_k|\varphi_l\rangle\left\{-\frac{3}{2}(\mathrm{vec}(A_{kl}^{-1}))' - \frac{m_{kl}}{2}\frac{1}{\mathrm{tr}[A_{kl}^{-1}J_{11}]}(\mathrm{vec}J_{11})'\left[A_{kl}^{-1}\otimes A_{kl}^{-1}\right]\right\}$$

or

$$\frac{\partial}{\partial \mathrm{vec}A_{kl}}\langle\varphi_k|\varphi_l\rangle = \langle\varphi_k|\varphi_l\rangle\left\{-\frac{3}{2}\mathrm{vec}(A_{kl}^{-1}) - \frac{m_{kl}}{2}\frac{1}{(A_{kl}^{-1})_{11}}\left[A_{kl}^{-1}\otimes A_{kl}^{-1}\right]\mathrm{vec}J_{11}\right\}$$

Hence, the matrix elements appearing in the kinetic energy are

$$\left\langle\varphi_k\left|\frac{1}{r_1^2}\right|\varphi_l\right\rangle = \frac{2}{m_{kl}+1}\frac{1}{(A_{kl}^{-1})_{11}}\langle\varphi_k|\varphi_l\rangle$$

$$\frac{\partial}{\partial \mathrm{vec}A_{kl}}\langle\varphi_k|\varphi_l\rangle = \left\{-\frac{3}{2}\mathrm{vec}(A_{kl}^{-1}) - \frac{m_{kl}}{2}\frac{1}{(A_{kl}^{-1})_{11}}\mathrm{vec}(A_{kl}^{-1}J_{11}A_{kl}^{-1})\right\}\langle\varphi_k|\varphi_l\rangle$$

$$\frac{\partial}{\partial \mathrm{vec}A_{kl}}\left\langle\varphi_k\left|\frac{1}{r_1^2}\right|\varphi_l\right\rangle = \left\{-\frac{3}{2}\mathrm{vec}(A_{kl}^{-1}) - \frac{m_{kl}-2}{2}\frac{1}{(A_{kl}^{-1})_{11}}\mathrm{vec}(A_{kl}^{-1}J_{11}A_{kl}^{-1})\right\} \times\frac{2}{m_{kl}+1}\frac{\langle\varphi_k|\varphi_l\rangle}{(A_{kl}^{-1})_{11}}$$

where we used the fact that $[X'\otimes Y]\mathrm{vec}(Z) = \mathrm{vec}(YZX)$. The kinetic energy matrix element is then

$$\begin{aligned}
&\langle\varphi_k|-\nabla_{\mathbf{r}}'M\nabla_{\mathbf{r}}|\varphi_l\rangle \\
&= m_k m_l M_{11}\frac{2}{m_{kl}+1}\frac{1}{(A_{kl}^{-1})_{11}}\langle\varphi_k|\varphi_l\rangle \\
&\quad - m_k(\mathrm{vec}(A_lMJ_{11}))'\left\{3\mathrm{vec}(A_{kl}^{-1}) + \frac{m_{kl}-2}{(A_{kl}^{-1})_{11}}\mathrm{vec}(A_{kl}^{-1}J_{11}A_{kl}^{-1})\right\}\frac{2}{m_{kl}+1}\frac{\langle\varphi_k|\varphi_l\rangle}{(A_{kl}^{-1})_{11}} \\
&\quad - m_l(\mathrm{vec}(J_{11}MA_k))'\left\{3\mathrm{vec}(A_{kl}^{-1}) + \frac{m_{kl}-2}{(A_{kl}^{-1})_{11}}\mathrm{vec}(A_{kl}^{-1}J_{11}A_{kl}^{-1})\right\}\frac{2}{m_{kl}+1}\frac{\langle\varphi_k|\varphi_l\rangle}{(A_{kl}^{-1})_{11}} \\
&\quad + 2(\mathrm{vec}(A_lMA_k))'\left\{3\mathrm{vec}(A_{kl}^{-1}) + \frac{m_{kl}}{(A_{kl}^{-1})_{11}}\mathrm{vec}(A_{kl}^{-1}J_{11}A_{kl}^{-1})\right\}\langle\varphi_k|\varphi_l\rangle \qquad (71)
\end{aligned}$$

After simplification and rearrangement the final expression for the kinetic energy takes the following form:

$$T_{kl} = 6\mathrm{tr}[A_k M A_l A_{kl}^{-1}] S_{kl} + \frac{2}{(A_{kl}^{-1})_{11}} \left[\frac{m_k m_l M_{11}}{m_{kl}+1} - m_k (A_{kl}^{-1} A_l M A_l A_{kl}^{-1})_{11} - m_l (A_{kl}^{-1} A_k M A_k A_{kl}^{-1})_{11} \right] S_{kl} \tag{72}$$

4. *Potential Energy Matrix Elements*

We will derive the potential energy components by finding the integral $R_{kl}^{ij} = \langle 1/r_{ij} \rangle$. There are two cases, m_{kl} even and m_{kl} odd.

Let us first introduce some simplifying definitions:

$$a = \mathrm{tr}\left[J_{11} A_{kl}^{-1}\right] \tag{73}$$

$$b = \mathrm{tr}\left[J_{ij} A_{kl}^{-1}\right] \tag{74}$$

$$c = \mathrm{tr}\left[J_{11} A_{kl}^{-1} J_{ij} A_{kl}^{-1}\right] \tag{75}$$

which for implementation purposes can be written as

$$a = \left(A_{kl}^{-1}\right)_{11} \tag{76}$$

$$b = \begin{cases} \left(A_{kl}^{-1}\right)_{ii}, & i = j \\ \left(A_{kl}^{-1}\right)_{ii} + \left(A_{kl}^{-1}\right)_{jj} - 2\left(A_{kl}^{-1}\right)_{ij}, & i \neq j \end{cases} \tag{77}$$

and

$$c = \begin{cases} \left(A_{kl}^{-1}\right)_{1i}^2, & i = j \\ \left(\left(A_{kl}^{-1}\right)_{1i} - \left(A_{kl}^{-1}\right)_{j1}\right)^2, & i \neq j \end{cases} \tag{78}$$

Let m_{kl} be even. If $p = m_{kl}/2$ with $p = 0, 1, 2, \ldots$, then

$$\langle \varphi_k | 1/r_{ij} | \varphi_l \rangle = \langle \phi_k | r_1^{2p}/r_{ij} | \phi_l \rangle \tag{79}$$

Using an integral transformation for $1/r_{ij}$,

$$\langle \varphi_k | 1/r_{ij} | \varphi_l \rangle = \frac{2}{\sqrt{\pi}} \int_0^\infty \langle \varphi_k | \exp\left[-v^2 \mathbf{r}' J_{ij} \mathbf{r}\right] | \varphi_l \rangle dv \tag{80}$$

followed by a differential transformation for r_1^{2p} and then integration over $\mathbf{r}$, we have

$$\begin{aligned}
&\langle\varphi_k|1/r_{ij}|\varphi_l\rangle \\
&\quad= \frac{2}{\sqrt{\pi}}(-1)^p \int_0^\infty \langle\phi_k| \frac{\partial^p}{\partial u^p} \exp[-u\mathbf{r}'J_{11}\mathbf{r}] \exp\left[-v^2\mathbf{r}'J_{ij}\mathbf{r}\right]|\phi_l\rangle dv \Big|_{u=0} \\
&\quad= \frac{2}{\sqrt{\pi}}(-1)^p \frac{\partial^p}{\partial u^p} \int_0^\infty \int_{-\infty}^\infty \exp\left[-\mathbf{r}'\left(A_{kl} + uJ_{11} + v^2J_{ij}\right)\mathbf{r}\right] d\mathbf{r}\, dv \Big|_{u=0} \\
&\quad= \frac{2}{\sqrt{\pi}}(-1)^p \frac{\partial^p}{\partial u^p} \int_0^\infty \frac{\pi^{3n/2}}{\left|A_{kl} + uJ_{11} + v^2J_{ij}\right|^{3/2}} dv \Big|_{u=0} \\
&\quad= \frac{2}{\sqrt{\pi}}\langle\phi_k|\phi_l\rangle(-1)^p \frac{\partial^p}{\partial u^p} \int_0^\infty \frac{1}{\left|I_n + uJ_{11}A_{kl}^{-1} + v^2J_{ij}A_{kl}^{-1}\right|^{3/2}} dv \Big|_{u=0}
\end{aligned} \tag{81}$$

Now, since J_{11} and J_{ij} are rank-one matrices, we can write the determinant in the integral above as a polynomial in the traces we defined in Refs. 73–75 and then integrate over v, yielding

$$\begin{aligned}
&\langle\varphi_k|1/r_{ij}|\varphi_l\rangle \\
&\quad= \frac{2}{\sqrt{\pi}}\langle\phi_k|\phi_l\rangle(-1)^p \frac{\partial^p}{\partial u^p} \int_0^\infty \frac{1}{(1 + ua + v^2b + uv^2(ab - c))^{3/2}} dv \Big|_{u=0} \\
&\quad= \frac{2}{\sqrt{\pi}}\langle\phi_k|\phi_l\rangle(-1)^p \frac{\partial^p}{\partial u^p} \frac{1}{(1 + au)(b + u(ab - c))^{1/2}} \Big|_{u=0}
\end{aligned} \tag{82}$$

Then, differentiating p times with respect to u, setting u to zero, and simplifying gives the final result,

$$R_{kl}^{ij} = S_{kl} \frac{\gamma_2(p)}{\sqrt{b}} \sum_{q=0}^{p} \gamma_3(q) \left(1 - \frac{c}{ab}\right)^q \tag{83}$$

where

$$\gamma_2(p) = \frac{\Gamma(p+1)}{\Gamma(p+3/2)} \tag{84}$$

and

$$\gamma_3(q) = \frac{\Gamma(q+1/2)}{\Gamma(q+1)\Gamma(1/2)} \tag{85}$$

The case when m_{kl} is odd involves an additional integral transformation, which unfortunately makes these terms somewhat more complicated. The integral evaluation is similar to that for the even case and proceeds as follows.

Let $(m_{kl}+1)/2 = p$ with $p = 1, 2, 3, \ldots$, then

$$\langle \varphi_k | 1/r_{ij} | \varphi_l \rangle = \langle \phi_k | r_1^{2p-1}/r_{ij} | \phi_l \rangle = \left\langle \phi_k \left| \frac{r_1^{2p}}{r_1 r_{ij}} \right| \phi_l \right\rangle \tag{86}$$

Making transformations as for the even case, but with the addition of an extra integral transformation for $1/r_1$, we obtain

$$\begin{aligned}
&\langle \varphi_k | 1/r_{ij} | \varphi_l \rangle \\
&= \left(\frac{2}{\sqrt{\pi}}\right)^2 (-1)^p \int_0^\infty \int_0^\infty \langle \phi_k | \frac{\partial^p}{\partial u^p} \exp[-u\mathbf{r}' J_{11}\mathbf{r}] \exp\left[-w^2 \mathbf{r}' J_{11}\mathbf{r}\right] \exp\left[-v^2 \mathbf{r}' J_{ij}\mathbf{r}\right] | \phi_l \rangle dw dv \Big|_{u=0} \\
&= \frac{4}{\pi}(-1)^p \frac{\partial^p}{\partial u^p} \int_0^\infty \int_0^\infty \int_{-\infty}^\infty \exp\left[-\mathbf{r}'\left(A_{kl} + uJ_{11} + w^2 J_{11} + v^2 J_{ij}\right)\mathbf{r}\right] d\mathbf{r}\, dw\, dv \Big|_{u=0} \\
&= \frac{4}{\pi}\langle \phi_k | \phi_l \rangle (-1)^p \frac{\partial^p}{\partial u^p} \int_0^\infty \int_0^\infty \frac{1}{\left| I_n + (u+w^2) J_{11} A_{kl}^{-1} + v^2 J_{ij} A_{kl}^{-1} 1 \right|^{3/2}} dw\, dv \Bigg|_{u=0}
\end{aligned} \tag{87}$$

Again, reducing the determinant to a polynomial in traces and integrating gives

$$\begin{aligned}
&\langle \varphi_k | 1/r_{ij} | \varphi_l \rangle \\
&= \frac{4}{\pi}\langle \phi_k | \phi_l \rangle (-1)^p \frac{\partial^p}{\partial u^p} \int_0^\infty \int_0^\infty \frac{1}{\left(1 + (u+w^2)a + v^2 b + uv^2(ab-c)\right)^{3/2}} dw\; dv \Bigg|_{u=0} \\
&= \frac{4}{\pi}\langle \phi_k | \phi_l \rangle (-1)^p \frac{\partial^p}{\partial u^p} \int_0^\infty \frac{1}{\left(1 + a(u+w^2)\right)\left(b + (u+w^2)(ab-c)\right)^{1/2}} dw \Bigg|_{u=0} \\
&= \frac{4}{\pi}\langle \phi_k | \phi_l \rangle (-1)^p \frac{\partial^p}{\partial u^p} \frac{1}{\left(c(1+au)\right)^{1/2}} \arcsin\left[\left(\frac{c}{a(b + u(ab-c))} \right)^{1/2} \right] \Bigg|_{u=0}
\end{aligned} \tag{88}$$

Differentiating and simplifying gives the final result for odd m_k values:

$$R_{kl}^{ij} = \frac{2}{\sqrt{\pi}} S_{kl} \sqrt{a} \left[\gamma_3(p) \frac{\arcsin\left[\sqrt{\frac{c}{ab}}\right]}{\sqrt{c}} + \frac{1}{2\sqrt{ab-c}} \sum_{q=1}^{p} \sum_{t=0}^{q-1} \frac{1}{q} \gamma_3(p-q)\gamma_3(t) \left(1 - \frac{c}{ab}\right)^{q-t} \right] \tag{89}$$

The inclusion of the arcsin and the double summation in this formula unfortunately complicates these odd power terms compared to the even power case. The implementation of odd powers m_k requires significantly more computer time due to the complexity of this formula. Furthermore, we found that variation of near optimal m_k by plus or minus one had negligeable effect on energy convergence. Therefore, in our calculations utilizing gradient formulas for energy optimization, we excluded the odd power case.

5. *Some Other Matrix Elements*

The matrix elements of r_1 and r_1^2, which in the case of a diatomic molecule are the internuclear distance and its square, can easily be obtained using relationship (66):

$$\langle\varphi_k|r_1|\varphi_l\rangle = \langle\phi_k|r_1^{m_{kl}+1}|\phi_l\rangle = \langle\phi_k|\phi_l\rangle \frac{2}{\sqrt{\pi}} \mathrm{tr}\left[J_{11}A_{kl}^{-1}\right]^{(m_{kl}+1)/2} \Gamma\left(\frac{m_{kl}+4}{2}\right)$$
$$= \langle\varphi_k|\varphi_l\rangle \frac{m_{kl}+2}{2} \gamma_2\left(\frac{m_{kl}}{2}\right) \sqrt{\left(A_{kl}^{-1}\right)_{11}} \tag{90}$$

$$\langle\varphi_k|r_1^2|\varphi_l\rangle = \langle\phi_k|r_1^{m_{kl}+2}|\phi_l\rangle = \langle\varphi_k|\varphi_l\rangle \frac{m_{kl}+3}{2} \left(A_{kl}^{-1}\right)_{11} \tag{91}$$

The evaluation of the expectation values of other than r_1 distances and their squares can be done by differentiating expressions (83) and (67), respectively. We will restrict ourselves to the case when m_k's are even, so that $p = m_{kl}/2$ with $p = 0, 1, 2, \ldots$. Let us first consider the simpler case of r_{ij}^2:

$$\langle\varphi_k|r_{ij}^2|\varphi_l\rangle = -\frac{\partial}{\partial u}\langle\varphi_k|\exp\left[-u\,\mathbf{r}'J_{ij}\mathbf{r}\right]|\varphi_l\rangle\Big|_{u=0}$$
$$= -\frac{\partial}{\partial u}\frac{2}{\sqrt{\pi}}\,\Gamma\left(p+\frac{3}{2}\right)\mathrm{tr}[(A_{kl}+uJ_{ij})^{-1}J_{11}]^p \frac{\pi^{3n/2}}{\left|A_{kl}+uJ_{ij}\right|^{3/2}}\Bigg|_{u=0} \tag{92}$$

Here, we will need some simple facts from matrix differential calculus. If X is a matrix variable and β is a parameter that X depends on, then

$$\frac{\partial|X|}{\partial\beta} = |X|\mathrm{tr}\left[X^{-1}\frac{\partial X}{\partial\beta}\right] \tag{93}$$

$$\frac{\partial\mathrm{tr}[X]}{\partial\beta} = \mathrm{tr}\left[\frac{\partial X}{\partial\beta}\right] \tag{94}$$

$$\frac{\partial(X^{-1})}{\partial\beta} = -X^{-1}\frac{\partial X}{\partial\beta}X^{-1} \tag{95}$$

Using these formulae, one can show that

$$\frac{\partial}{\partial u}\mathrm{tr}[(A_{kl}+uJ_{ij})^{-1}J_{11}] = -\mathrm{tr}[(A_{kl}+uJ_{ij})^{-1}J_{ij}(A_{kl}+uJ_{ij})^{-1}J_{11}] \tag{96}$$

and

$$\langle\varphi_k|r_{ij}^2|\varphi_l\rangle = \langle\varphi_k|\varphi_l\rangle\left(p\frac{c}{a}+b\right) \tag{97}$$

In the same manner for the first power of r_{ij}, one obtains

$$\begin{aligned}
&\langle\varphi_k|r_{ij}|\varphi_l\rangle \\
&= -\frac{\partial}{\partial u}\left\langle\phi_k\left|\frac{\exp\left[-u\,\mathbf{r}'J_{ij}\mathbf{r}\right]}{r_{ij}}\right|\phi_l\right\rangle\Bigg|_{u=0} \\
&= -\frac{\partial}{\partial u}\frac{2}{\sqrt{\pi}}\Gamma\left(p+\frac{3}{2}\right)\mathrm{tr}[(A_{kl}+uJ_{ij})^{-1}J_{11}]^p\frac{\pi^{3n/2}}{\left|A_{kl}+uJ_{ij}\right|^{3/2}}\frac{\gamma_2(p)}{\mathrm{tr}[(A_{kl}+uJ_{ij})^{-1}J_{ij}]^{1/2}} \\
&\quad\times\sum_{q=0}^{p}\gamma_3(q)\left(1-\frac{\mathrm{tr}[(A_{kl}+uJ_{ij})^{-1}J_{ij}(A_{kl}+uJ_{ij})^{-1}J_{11}]}{\mathrm{tr}[(A_{kl}+uJ_{ij})^{-1}J_{11}]\mathrm{tr}[(A_{kl}+uJ_{ij})^{-1}J_{ij}]}\right)^q\Bigg|_{u=0} \\
&= \langle\varphi_k|1/r_{ij}|\varphi_l\rangle\left\{-p\frac{c}{a}-\frac{3}{2}b+\frac{1}{2}\frac{c}{b}+\left(\frac{2t}{ab}-\frac{c^2}{a^2b}-\frac{ch}{ab^2}\right)\sum_{q=1}^{p}\gamma_3(q)\,q\left(1-\frac{c}{ab}\right)^{q-1}\right\}
\end{aligned} \tag{98}$$

where

$$\begin{aligned}
h &= \mathrm{tr}[A_{kl}^{-1}J_{ij}A_{kl}^{-1}J_{ij}] \\
t &= \mathrm{tr}[A_{kl}^{-1}J_{ij}A_{kl}^{-1}J_{ij}A_{kl}^{-1}J_{11}]
\end{aligned}$$

The last two traces are equal to

$$h=\begin{cases}\left(A_{kl}^{-1}\right)_{ii}^2, & i=j\\ \left[\left(A_{kl}^{-1}\right)_{ii}+\left(A_{kl}^{-1}\right)_{jj}-2\left(A_{kl}^{-1}\right)_{ij}\right]^2, & i\neq j\end{cases} \tag{99}$$

$$t=\begin{cases}0, & i\neq 1 \text{ and } j\neq 1\\ \left[\left(A_{kl}^{-1}\right)_{11}-\left(A_{kl}^{-1}\right)_{1j}\right]^2\left[\left(A_{kl}^{-1}\right)_{11}+\left(A_{kl}^{-1}\right)_{jj}-2\left(A_{kl}^{-1}\right)_{1j}\right], & i=1 \text{ and } i\neq j\\ \left(A_{kl}^{-1}\right)_{11}^3, & i=j=1\end{cases} \tag{100}$$

In order to perform calculations of the correlation function of particles 1 and 2 (which is the same quantity as the density of the 1st pseudoparticle, $g_1(\boldsymbol{\xi})$) using basis (49), we need the matrix elements of the delta function, $\delta(\mathbf{r}_1 - \boldsymbol{\xi})$. They can be evaluated by replacing $A_{kl} \to A_{kl} + uJ_{11}$ in expression (64) and then differentiating p times with respect to u.

$$\langle \varphi_k | \delta(\mathbf{r}_1 - \boldsymbol{\xi}) | \varphi_l \rangle = (-1)^p \frac{\partial^p}{\partial u^p} \frac{\pi^{3n/2}}{|A_{kl} + uJ_{11}|^{3/2}} \frac{1}{\pi^{3/2}} \frac{1}{\mathrm{tr}[(A_{kl} + uJ_{11})^{-1} J_{11}]^{3/2}} \times \exp\left[- \frac{\xi^2}{\mathrm{tr}[(A_{kl} + uJ_{11})^{-1} J_{11}]} \right] \Bigg|_{u=0}$$

Applying formula (96) and using the fact that $\mathrm{tr}[XJ_{ii}XJ_{ii}] = \mathrm{tr}[XJ_{ii}]^2$ for an arbitrary matrix X reduces the final result to

$$\langle \varphi_k | \delta(\mathbf{r}_1 - \boldsymbol{\xi}) | \varphi_l \rangle = \langle \varphi_k | \varphi_l \rangle \frac{1}{2\pi} \frac{1}{\Gamma\left(p + \frac{3}{2}\right)} \frac{1}{(A_{kl}^{-1})_{11}^{3/2}} \left(\frac{\xi^2}{(A_{kl}^{-1})_{11}} \right)^p \exp\left[- \frac{\xi^2}{(A_{kl}^{-1})_{11}} \right] \tag{101}$$

The matrix elements of $\delta(\mathbf{r}_{ij})$ can be easily obtained by straightforward integration. The procedure is very similar to the evaluation of the overlap integral and yields

$$\langle \varphi_k | \delta(\mathbf{r}_{ij}) | \varphi_l \rangle = \frac{2}{\sqrt{\pi}} \Gamma\left(p + \frac{3}{2} \right) (D_{kl}^{-1})_{11}^p \frac{\pi^{3(n-1)/2}}{|D_{kl}|^{3/2}} = \langle \varphi_k | \varphi_l \rangle \frac{1}{\pi^{3/2}} \left(\frac{|A_{kl}|}{|D_{kl}|} \right)^{3/2} \left(\frac{(D_{kl}^{-1})_{11}}{(A_{kl}^{-1})_{11}} \right)^p \tag{102}$$

where D_{kl} is an $(n-1) \times (n-1)$ matrix formed from A_{kl} by adding the jth row to the ith one, then adding and jth column to the ith column, and then crossing out the jth column and row.

6. *Energy Gradient*

The integral formulas above are sufficient for performing energy calculations and evaluating some expectation values. However, optimization of the many nonlinear parameters contained in the exponent matrices L_k demands excessive computational resources if an approximate numerical energy gradient is used. Several orders of magnitude of computational effort can be saved by utilizing analytic gradients. Additionally, more thorough optimization can be achieved

due to increased accuracy of analytic gradients compared with numerical approximations. We feel that analytic gradient formulas are essential in many practical situations. To this end, the energy gradient formulas are now presented.

We begin with the derivative of the secular equation with respect to energy eigenvalues. For some background on matrix differential calculus, see the Refs. 116 and 117.

The secular equation,

$$(H - \epsilon S)c = 0 \tag{103}$$

defines the energy ϵ as an implicit function of the $N \times N$ matrices H and S, where N is the number of basis functions. H and S are themselves functions of the $Nn(n+1)/2 \times 1$ vector $L = \left[(\text{vech } L_1)', \cdots, (\text{vech } L_N)'\right]$ of nonlinear exponential parameters contained in the matrices L_k; recall that $A_k = L_k L_k'$. The energy gradient with respect to L is then

$$g = \nabla_L \epsilon = \frac{1}{c'Sc}\left(\frac{\partial \text{vech } H}{\partial L'} - \epsilon \frac{\partial \text{vech } S}{\partial L'}\right)' (\text{vech}[2cc' - \text{diag } cc']) \tag{104}$$

The matrix $(\partial \text{vech } H/\partial L' - \epsilon\, \partial \text{vech } S/\partial L') = \partial \text{vech}(H - \epsilon S)/\partial L'$ in the gradient above is sparse and has dimension $N(N+1)/2 \times Nn(n+1)/2$. This sparse matrix together with the eigenvector c can be collapsed to a dense partitioned form with dimension $N \times Nn(n+1)/2$,

$$G = \begin{pmatrix} c_1^2 \frac{\partial (H-\epsilon S)_{11}}{\partial (\text{vech } L_1)'} & 2c_1c_2 \frac{\partial (H-\epsilon S)_{12}}{\partial (\text{vech } L_2)'} & \cdots & 2c_1c_N \frac{\partial (H-\epsilon S)_{1N}}{\partial (\text{vech } L_N)'} \\ 2c_2c_1 \frac{\partial (H-\epsilon S)_{21}}{\partial (\text{vech } L_1)'} & c_2^2 \frac{\partial (H-\epsilon S)_{22}}{\partial (\text{vech } L_2)'} & \cdots & 2c_2c_N \frac{\partial (H-\epsilon S)_{2N}}{\partial (\text{vech } L_N)'} \\ \vdots & \vdots & \ddots & \vdots \\ 2c_Nc_1 \frac{\partial (H-\epsilon S)_{N1}}{\partial (\text{vech } L_1)'} & 2c_Nc_2 \frac{\partial (H-\epsilon S)_{N2}}{\partial (\text{vech } L_2)'} & \cdots & c_N^2 \frac{\partial (H-\epsilon S)_{NN}}{\partial (\text{vech } L_N)'} \end{pmatrix} \tag{105}$$

With G defined as above, the gradient can be computed by summing over the rows of G. That is

$$g_i = \frac{1}{c'Sc} \sum_j G_{ji}$$

The nonzero terms in the matrices $\partial \text{vech } H/\partial a'$ and $\partial \text{vech } S/\partial a'$ are contained in the $1 \times n(n+1)/2$ vectors $\partial H_{kl}/\partial (\text{vech } L_k)'$ and $\partial S_{kl}/\partial (\text{vech } L_k)'$ and also in $\partial H_{kl}/\partial (\text{vech } L_l)'$ and $\partial S_{kl}/\partial (\text{vech } L_l)'$. Matrix derivatives depend on the arrangement of elements in a matrix variable; therefore, since symmetry

projection on kets effectively reorders elements of the exponent matrix L_l, the formulas for derivatives with respect to the exponent matrices in the ket, vech $[L_l]$, are different from those with respect to the exponent matrices in the bra, vech $[L_k]$. We assume that under particle permutations, A_l transforms as

$$\tau'_P A_l \tau_P = \tau'_P L_l L'_l \tau_P = \tau'_P L_l \left(\tau'_P L_l\right)'$$

where τ_P are some transformation matrices. Also, the matrix derivatives for the diagonal blocks of G are complicated somewhat by the symmetry projection on the kets. However, they can be computed using the following relationship, for example:

$$\frac{\partial H_{kk}}{\partial(\text{vech } L_k)'} = \left.\frac{\partial H_{kl}}{\partial(\text{vech } L_k)'}\right|_{l=k} + \left.\frac{\partial H_{kl}}{\partial(\text{vech } L_l)'}\right|_{l=k} \tag{106}$$

Thus, only two sets of formulas for the derivatives need be computed.

Matrix elements are scalar-valued matrix functions of the exponent matrices L_k. Therefore, the appropriate mathematical tool for finding derivatives is the matrix differential calculus [116, 118]. Using this, the derivations are nontrivial but straightforward. We will only present the final results of the derivations. The reader wishing to derive these formulas, or other matrix derivatives, is referred to the Ref. 116 and references therein.

We want to note that only one term in the symmetry projection will be represented. As was the case for the integral formulas, the symmetry terms require the substitution $A_l \mapsto \tau'_P A_l \tau_P = \tau'_P L_l \left(\tau'_P L_l\right)'$ or, more generally, $L_l \mapsto \tau'_P L_l$. This is required for derivatives with respect to both vech $[L_k]$ and vech $[L_l]$. The derivatives with respect to vech $[L_l]$ will require further modification, and this will be noted in the formulas below.

Using the normalized overlap formula, Eq. (68), the derivative with respect to the nonzero terms of the lower triangular matrix L_k is

$$\begin{aligned}\frac{\partial S_{kl}}{\partial(\text{vech } L_k)'} = {} & \frac{3}{2} S_{kl} \text{ vech } \left[\left(L_k^{-1}\right)' - 2A_{kl}^{-1} L_k\right]' \\ & + S_{kl} \frac{m_k}{\left(A_k^{-1}\right)_{11}} \text{vech}\left[A_k^{-1} J_{11} A_k^{-1} L_k\right]' \\ & - S_{kl} \frac{m_{kl}}{\left(A_{kl}^{-1}\right)_{11}} \text{vech}\left[A_{kl}^{-1} J_{11} A_{kl}^{-1} L_k\right]'\end{aligned} \tag{107}$$

For the derivative, with respect to the elements vech L_l, we account for the symmetry terms by making a multiplication by τ_P in addition to the substitutions

described above. Thus,

$$\frac{\partial S_{kl}}{\partial(\text{vech } L_l)'} = \frac{3}{2} S_{kl}\ \text{vech}\left[\tau_P \left(L_l^{-1}\right)' - 2\tau_P A_{kl}^{-1} L_l\right]' + S_{kl} \frac{m_l}{\left(A_l^{-1}\right)_{11}} \text{vech}\left[\tau_P A_l^{-1} J_{11} A_l^{-1} L_l\right]' - S_{kl} \frac{m_{kl}}{\left(A_{kl}^{-1}\right)_{11}} \text{vech}\left[\tau_P A_{kl}^{-1} J_{11} A_{kl}^{-1} L_l\right]' \tag{108}$$

The kinetic energy gradient components are obtained by differentiating Eq. (72) with respect to vech L_k and vech L_l:

$$\begin{aligned}
&\frac{\partial T_{kl}}{\partial(\text{vech } L_k)'} \\
&= \frac{\partial S_{kl}}{\partial(\text{vech } L_k)'} \frac{T_{kl}}{S_{kl}} + 2S_{kl} \Bigg[6\ \text{vech}\left[A_{kl}^{-1} A_l M A_l A_{kl}^{-1} L_k\right]' \\
&\quad + 2\left(A_{kl}^{-1}\right)_{11}^{-2} \left(\frac{m_k m_l M_{11}}{m_{kl}+1} - m_k \left(A_{kl}^{-1} A_l M\, A_l A_{kl}^{-1}\right)_{11} - m_l \left(A_{kl}^{-1} A_k M\, A_k A_{kl}^{-1}\right)_{11} \right) \\
&\qquad \times \text{vech}\left[A_{kl}^{-1} J_{11} A_{kl}^{-1} L_k\right]' \\
&\quad - 2\left(A_{kl}^{-1}\right)_{11}^{-1} m_l\ \text{vech}[A_{kl}^{-1} J_{11} A_{kl}^{-1} A_k M A_l A_{kl}^{-1} L_k + A_{kl}^{-1} A_l M A_k A_{kl}^{-1} J_{11} A_{kl}^{-1} L_k] \\
&\quad + 2\left(A_{kl}^{-1}\right)_{11}^{-1} m_k\ \text{vech}[A_{kl}^{-1} J_{11} A_{kl}^{-1} A_l M A_l A_{kl}^{-1} L_k + A_{kl}^{-1} A_l M A_l A_{kl}^{-1} J_{11} A_{kl}^{-1} L_k] \Bigg]
\end{aligned} \tag{109}$$

and the derivative with respect to vech L_l including the symmetry projector term τ_P is given by

$$\begin{aligned}
&\frac{\partial T_{kl}}{\partial(\text{vech } L_l)'} \\
&= \frac{\partial S_{kl}}{\partial(\text{vech } L_l)'} \frac{T_{kl}}{S_{kl}} + 2S_{kl} \Bigg[6\ \text{vech}\left[\tau_P A_{kl}^{-1} A_k M A_k A_{kl}^{-1} L_l\right]' \\
&\quad + 2\left(A_{kl}^{-1}\right)_{11}^{-2} \left(\frac{m_k m_l M_{11}}{m_{kl}+1} - m_k \left(A_{kl}^{-1} A_l M A_l A_{kl}^{-1}\right)_{11} - m_l \left(A_{kl}^{-1} A_k M A_k A_{kl}^{-1}\right)_{11} \right) \\
&\qquad \times \text{vech}\left[\tau_P A_{kl}^{-1} J_{11} A_{kl}^{-1} L_l\right]' \\
&\quad - 2\left(A_{kl}^{-1}\right)_{11}^{-1} m_k\ \text{vech}[\tau_P A_{kl}^{-1} J_{11} A_{kl}^{-1} A_l M A_k A_{kl}^{-1} L_l + \tau_P A_{kl}^{-1} A_k M A_l A_{kl}^{-1} J_{11} A_{kl}^{-1} L_l] \\
&\quad + 2\left(A_{kl}^{-1}\right)_{11}^{-1} m_l\ \text{vech}[\tau_P A_{kl}^{-1} J_{11} A_{kl}^{-1} A_k M A_k A_{kl}^{-1} L_l + \tau_P A_{kl}^{-1} A_k M A_k A_{kl}^{-1} J_{11} A_{kl}^{-1} L_l] \Bigg]
\end{aligned} \tag{110}$$

The purpose of the gradient formulas is to enhance optimization efforts, and we have found that well-optimized wave functions utilizing a mixture of even and odd powers, m_k, did not produce results any better than using even powers only. Hence, for this reason, along with reasons that were stated at the end of the section describing the potential energy matrix elements, we will not derive the gradient terms for the potential energy matrix elements that utilize odd m_k.

Using the even m_k potential energy integral formula, Eq. (83), the definitions for a, b, and c given in Eqs. (73)–(75), and the definitions for γ_2 and γ_3 given in Eqs. (84 and 85), the gradient terms for the potential, with $m_{kl} = 2p$, are given by

$$\begin{aligned}
&\frac{\partial R_{kl}^{ij}}{\partial(\operatorname{vech} L_k)'} \\
&= \frac{\partial S_{kl}}{\partial(\operatorname{vech} L_k)'}\frac{R_{kl}^{ij}}{S_{kl}} - \frac{\partial b}{\partial(\operatorname{vech} L_k)'}\frac{R_{kl}^{ij}}{2b} + \gamma_2(p) S_{kl} b^{-1/2} \qquad (111)\\
&\times \left[\sum_{q=1}^{p} \gamma_3(q) q \left(1 - \frac{c}{ab}\right)^{q-1} \frac{c}{(ab)^2}\left(\frac{\partial a}{\partial(\operatorname{vech} L_k)'} b + a\frac{\partial b}{\partial(\operatorname{vech} L_k)'} - \frac{ab}{c}\frac{\partial c}{\partial(\operatorname{vech} L_k)'}\right)\right]
\end{aligned}$$

where

$$\frac{\partial a}{\partial(\operatorname{vech} L_k)'} = -2 \operatorname{vech}\left[A_{kl}^{-1} J_{11} A_{kl}^{-1} L_k\right] \qquad (112)$$

$$\frac{\partial b}{\partial(\operatorname{vech} L_k)'} = -2 \operatorname{vech}\left[A_{kl}^{-1} J_{ij} A_{kl}^{-1} L_k\right] \qquad (113)$$

and

$$\begin{aligned}
\frac{\partial c}{\partial(\operatorname{vech} L_k)'} = &-2 \operatorname{vech}\left[A_{kl}^{-1} J_{11} A_{kl}^{-1} J_{ij} A_{kl}^{-1} L_k\right] \\
&- 2 \operatorname{vech}\left[A_{kl}^{-1} J_{ij} A_{kl}^{-1} J_{11} A_{kl}^{-1} L_k\right] \qquad (114)
\end{aligned}$$

The derivative with respect to vech L_l has the same form as the above but with L_k replaced by $\tau_P' L_l$ and each of the expressions vech $\left[A_{kl}^{-1} \cdots\right]$ replaced by vech $\left[\tau_P A_{kl}^{-1} \cdots\right]$.

D. Variational Method and Minimization of the Energy Functional

The energy in variational calculations is obtained by minimizing the Rayleigh quotient. In the case of basis set (49), this quotient has the following form:

$$E(\{c_k\}, \{m_k\}, \{L_k\}) = \min_{\{\{c_k\},\{m_k\},\{L_k\}\}} \frac{c' H(\{m_k\}, \{L_k\}) c}{c' S(\{m_k\}, \{L_k\}) c} \qquad (115)$$

where the minimization takes place with respect to the linear coefficients, $\{c_k\}$ of the wave function expansion in terms of the basis functions, and with respect to the nonlinear exponential parameters, $\{L_k\}$, and preexponential powers, $\{m_k\}$, of the basis functions. H and S are the Hamiltonian and overlap matrices of size $N \times N$ (N is the number of basis functions). The finding of the linear coefficients, $\{c_k\}$, is usually done by solving secular equation (103). However, in ground-state calculations one may minimize the above quotient directly, without diagonalization. In some cases such a procedure, when $\{c_k\}$ and $\{L_k\}$ are considered to be independent, may have certain advantages. But in many situations we do reduce the minimization with respect to $\{c_k\}$ to the generalized symmetric eigenvalue problem. In our calculations we used the inverse iteration method to solve this problem. The idea of the method consists in performing the iterations,

$$(H - \epsilon_{\text{appr}} S) c_{k+1} = S c_k \tag{116}$$

where ϵ_{appr} is an approximate value of the exact solution of the generalized symmetric eigenvalue problem. The starting vector c_0 can be chosen randomly. As long as the exact eigenvalue we need to obtain is closer to ϵ_{appr} than any other eigenvalue, the iterations (116) will converge. Typically, just a few iterations are needed to find ϵ with sufficient accuracy. Each iteration in (116) is performed in a few steps:

$$H - \epsilon_{\text{appr}} S = LDL' \tag{117}$$

$$Lx = S c_k \tag{118}$$

$$Dy = x \tag{119}$$

$$L' c_{k+1} = y \tag{120}$$

Here, L is a lower triangular matrix (not to be confused with L_k, the Cholesky factor of the matrix of nonlinear parameters A_k), and D is a diagonal matrix. The scheme of the solution of the generalized symmetric eigenvalue problem above has proven to be very efficient and accurate in numerous calculations. But the main advantage of this scheme is revealed when one has to routinely solve the secular equation with only one row and one column of matrices H and S changed. In this case, the update of factorization (117) requires only $\propto N^2$ arithmetic operations while, in general, the solution "from scratch" needs $\propto N^3$ operations.

It is well known that the convergence of variational expansions in terms of correlated Gaussians, both the simple ones and those with premultipliers, strongly depends on how one selects the nonlinear parameters in the Gaussian

exponentials. In order to get high-accuracy results in the calculations, one needs to perform optimizations of those parameters at some level. Due to a usually large number of basis functions in non-BO calculations and, consequently, a larger number of the exponential parameters, this task represents a serious computational problem. The two most commonly applied approaches to the parameter optimization are (a) a full optimization, which is very effective when the analytical gradient of the variational energy functional with respect to the parameters is available, and (b) the method based on a stochastic selection of the parameters.

We found that in many practical situations a hybrid method that combines the gradient-driven optimization with the stochastic selection method turns out to be very efficient. In this approach, we first generate a relatively small basis set for each of the studied systems using the full gradient optimization. This generates a good starting point for each system for the next step of the procedure. In this next step we apply the following strategy. We incrementally increase the size of the basis set by including additional basis functions, one by one, with randomly selected values of the nonlinear parameters and values of the pre-exponential powers. After including a function into the basis set, we first optimize the power of its preexponential factor using the finite difference approach and then optimize the nonlinear parameters in its exponent using the analytical gradient approach. After adding several new basis functions using this approach (the number varies depending on the desired degree of optimization), the whole basis is reoptimized by means of the gradient approach applied in sequence to each basis functions, one function at a time. This continues until the number of basis functions reached a certain limit or until the necessary accuracy is achieved. Although this procedure has been proven to be quite efficient in optimizations of large basis sets of correlated Gaussians, it still requires a lot of computational resources, especially for systems with a large number of particles and a large number of particle permutations in the Young symmetry operator. In addition to that, as we found from our experience, a full (simultaneous) optimization of all nonlinear parameters may still be very desirable for highly vibrationally excited states where the Gaussians tend to be very strongly coupled. The way to partially overcome this problem of high computational demands is extensive parallelization of the computer code for use on multinode computational systems. For this purpose we used the message passing interface (MPI), a widely used tool in the world of parallel computations, and were able to achieve a sufficient parallelization level of the code for runs with several processors. This development enabled us to significanly extend our capabilities of optimizing large basis sets.

To illustrate the capabilities of the variational method, we will present later the results and discuss the details of some diatomic non-BO calculations on small molecules, which were carried out by our group.

E. The Ground and Excited States of H_2

Due to a small number of electrons and nuclei, the H_2 molecule often serves as the first target for testing different methods in quantum chemistry. Despite the fact that this system has been very well studied since the early days of quantum mechanics and the pioneering work of Heitler and London [106], very few studies did not invoke Born–Oppenheimer approximation. Moreover, all of the non-BO calculations usually deal with the ground or few low-lying states only. In this part we present highly accurate, nonrelativistic, variational, non-BO calculations of the "vibrational spectrum" of the H_2 molecule [121]. Although we use the traditional term "vibrational spectrum," the states we have calculated can be better characterized as states with zero total angular momentum, which is the sum of the angular momenta of the electrons and the nuclei.

The non-BO wave functions of different excited states have to differ from each other by the number of nodes along the internuclear distance, which in the case of basis (49) is r_1. To accurately describe the nodal structure in all 15 states considered in our calculations, a wide range of powers, m_k, had to be used. While in the calculations of the H_2 ground state [119], the power range was 0–40, in the present calculations it was extended to 0–250 in order to allow pseudoparticle 1 density (i.e., nuclear density) peaks to be more localized and sharp if needed. We should notice that if one aims for highly accurate results for the energy, then the wave function of each of the excited states must be obtained in a separate calculation. Thus, the optimization of nonlinear parameters is done independently for each state considered.

Table II contains total variational energies of the lowest 15 states corresponding to the rotational ground state ($J = 0$) calculated with 3000 basis functions each. We also show the expectation values of the internuclear distance and its square calculated as average values using the optimized wave function of each state ($\langle r_1 \rangle$ and $\langle r_1^2 \rangle$). The energy for the ground state of -1.1640250300 hartree is noticeably lower than the previously reported upper bound [119] of -1.1640250232 hartree. We are certain that for at least a few lowest excited states, the quality of the results is very similar to that for the ground state. However, for the highest states, where the number of nodes in the wave function is much higher, the quality of the calculations decreases, but we believe that it still allows determination of the transition energies, with the accuracy similar to the experimental uncertainty, if not higher.

In Table II we also compare our total variational energies with the energies obtained by Wolniewicz. In his calculations Wolniewicz employed an approach wherein the zeroth order the adiabatic approximation for the wave function was used (i.e., the wave function is a product of the ground-state electronic wave function and a vibrational wave function) and he calculated the nonadiabatic effects as corrections [107, 108]. In general the agreement between our results

TABLE II
Nonadiabatic Variational Energies for 15 States of the H_2 Molecule with Zero Total Angular Momentum (the Ground Rotational States) Obtained with 3000 Basis Functions for Each State and Expectation Values of the Internuclear Distance and the Square of the Internuclear Distance, $\langle r_{H-H} \rangle$ $\langle r^2_{H-H} \rangle$[a]

v	E	$\langle r_{H-H} \rangle$	$\langle r^2_{H-H} \rangle$	E, Wolniewicz [107]
0	−1.1640250300	1.4487380	2.1270459	−1.1640250185
1	−1.1450653676	1.5453495	2.4739967	−1.1450653629
2	−1.1271779152	1.6460579	2.8568172	−1.1271779324
3	−1.1103404429	1.7517082	3.2814143	−1.1103404855
4	−1.0945391187	1.8634245	3.7556995	−1.0945391940
5	−1.0797693217	1.9827332	4.2905417	−1.0797694803
6	−1.0660370737	2.1117587	4.9013207	−1.0660372849
7	−1.0533604890	2.2535349	5.6105163	−1.0533608258
8	−1.0417726950	2.4125952	6.4525567	−1.0417731139
9	−1.0313249454	2.5958940	7.4823876	−1.0313254708
10	−1.0220917849	2.8149490	8.7946796	−1.0220924876
11	−1.0141782601	3.0901798	10.566922	−1.0141791536
12	−1.0077301951	3.4627010	13.181814	−1.0077311979
13	−1.0029493758	4.0342373	17.680148	−1.0029504633
14	−1.0001150482	5.2110181	28.919890	−1.0001159762
—[b]	−0.9994556794			

[a]Also, the nonrelativistic energies of Wolniewicz are presented for comparison. All quantities in atomic units.
[b]Nonrelativistic dissociation threshold.

and the results of Wolniewicz is very good. However, one notices that the agreement is much better for the lower energies than for the higher ones. While for the two lowest states our energies are lower than those obtained by Wolniewicz, the energies for the higher states are progressively higher.

Included in Table III is the comparison of the transition frequencies calculated from the energies obtained in our calculations with the experimental transition frequencies of Dabrowski [125]. To convert theoretical frequencies into wavenumbers, we used the factor of 1 hartree = 219474.63137 cm^{-1}. For all the frequencies our results are either within or very close to the experimental error bracket of 0.1 cm^{-1}. We hope that the advances in high-resolution spectroscopy will inspire remeasurements of the vibrational spectrum of H_2 with the accuracy lower than 0.1 cm^{-1}. With such high-precision results, it would be possible to verify whether the larger differences between the calculated and the experimental frequencies for higher excitation levels, which now appear, are due to the relativistic and radiative effects.

Finally, in Table III we also compare our results for the transition energies with the results obtained by Wolniewicz [107,108]. Wolniewicz also calculated

TABLE III

Comparison of Vibrational Frequencies $E_{v+1} - E_v$ (in cm^{-1}) of H_2 Calculated from Non-Born–Oppenheimer Energies [121] with the Experimental Values of Dabrowski [125] and with the results of Wolniewicz, Obtained Using the Conventional Approach Based on the Potential Energy Curve[a]

v	Dabrowski	Work [121]	Wolniewicz[b] (Diff.)	Wolniewicz[c] (Diff.)
0	4161.14	4161.165 (+0.025)	4161.163 (+0.023)	4161.167 (+0.027)
1	3925.79	3925.842 (+0.052)	3925.837 (+0.047)	3925.836 (+0.046)
2	3695.43	3695.398 (−0.032)	3695.392 (−0.038)	3695.389 (−0.041)
3	3467.95	3467.990 (+0.040)	3467.983 (+0.033)	3467.976 (+0.026)
4	3241.61	3241.596 (−0.014)	3241.577 (−0.033)	3241.564 (−0.046)
5	3013.86	3013.880 (+0.020)	3013.869 (+0.009)	3013.851 (−0.009)
6	2782.13	2782.189 (+0.059)	2782.161 (+0.031)	2782.136 (+0.006)
7	2543.25	2543.227 (−0.023)	2543.209 (−0.041)	2543.175 (−0.075)
8	2292.93	2293.016 (+0.086)	2292.993 (+0.063)	2292.950 (+0.020)
9	2026.38	2026.445 (+0.064)	2026.406 (+0.026)	2026.351 (−0.029)
10	1736.66	1736.818 (+0.158)	1736.776 (+0.116)	1736.707 (+0.047)
11	1415.07	1415.187 (+0.117)	1415.163 (+0.093)	1415.076 (+0.006)
12	1049.16	1049.269 (+0.109)	1049.250 (+0.090)	1049.139 (−0.021)
13	622.02	622.063 (+0.043)	622.098 (+0.078)	621.956 (−0.064)

[a]Differences between the calculated and the experimental results are shown in parentheses.
[b]Private communications (nonrelativistic values).
[c]From Ref. 108 (includes relativistic and radiative corrections).

transition energies corrected for the relativistic effects, and these results are also shown in Table III. Upon comparing the results, one notices that our transition energies are, in general, very similar to Wolniewicz's nonrelativistic results. Both sets of results show higher positive discrepancies in comparison with the experimental values for the higher excitation levels. These discrepancies decrease somewhat when the relativistic effects are included. However, the insufficient accuracy of the experiment, as well as perhaps that of some of the theoretical results, does not allow us to carry out a more detail analysis of the remaining discrepancies.

F. Charge Asymmetry in HD^+ Molecular Ion

The lack of a center of symmetry in HD^+, due to the difference in nuclear masses, creates a particularly interesting situation that requires a theoretical approach that may differ from those used to describe the parent cation, H_2^+, and its symmetric isotopomer, D_2^+. The asymmetry of the HD^+ system has been investigated both experimentally [109,110] and theoretically [111–114]. In recent work, Ben-Itzhak et al. [109] studied the dissociation of the electronic ground state of HD^+ following ionization of HD by fast proton impact and found the $H^+ + D(1s)$ dissociation channel is more likely than the $H(1s) + D^+$ dissociation channel by about 7%. They attributed this asymmetry breakdown to

the finite nuclear mass correction to the Born–Oppenheimer (BO) approximation, which makes the $1s\sigma$ state 3.7 meV lower than the $2p\sigma$ state at the dissociation limit.

Near the dissociation threshold the density of bound states in the HD^+ spectrum increases. If we remain in the ground rotational state manifold (i.e., consider only states with total angular momentum equal to zero), there is a place in the spectrum where the spacing between the consecutive levels becomes comparable to the difference between the binding energies of the H and D atoms (equal to 29.84 cm^{-1}). The D atom is energetically more stable because it has slightly larger reduced mass than H, which makes the electron slightly closer (on average) approach the nucleus, resulting in stronger Coulombic attraction and a lower binding energy. According to the previous calculations based on the adiabatic approximation [113], the H/D energy gap is approximately matched by the energy difference between the $v = 20$ and $v = 21$ levels. In this region the vibrational wave function that corresponds to, say, $v = 20$ level, combined with the ground-state electronic wave function that places the electron at the proton, has energy similar to that of the wave function, with the vibrational component corresponding to $v = 21$ and with the electronic component localizing the electron at d. Since such two-wave functions have the same symmetry, their mixing can occur. This nonadiabatic coupling must be included in the calculation of the dissociation of HD^+ that yields a proton plus a deuterium atom since the electron favors the heavier nucleus.

In nearly all theoretical calculations of H_2^+ and its isotopes' spectra reported in the literature, a body-fixed coordinate system with the origin at the geometric center of the nuclei has been used. For example, in the recent work of Esry and Sadeghpour [113], as well as the works of Moss [111, 112, 114], the starting point was the H_2^+ BO Hamiltonian in prolate spheroidal coordinates (PSC); and electronic wave functions and energies were first obtained as a function of the internuclear distance. Here we demonstrate the capability of the variational method to treat such systems fully nonadiabatically [124]. The very high powers m_k in the preexponential multipliers, which, as in case of H_2 molecule calculations, ranged from 0 to 250, allow one to describe very sharp peaks in the ''vibrational'' part of the wave function.

The effort in the first stage of the calculations has been focused on generating very accurate variational wave functions and energies for the rotationless vibrational states of the HD^+ ion. As mentioned, this system has been studied by many researchers and very accurate, virtually exact nonrelativistic energies have been published in the literature [112]. This includes the energy for the highest vibrational $v = 22$ state, which is only about 0.4309 cm^{-1} below the $D + H^+$ dissociation limit.

The basis set for each vibrational state was generated in a separate calculation. To achieve a similar level of accuracy as obtained in the best

TABLE IV

Comparison of the Total Nonrelativistic Nonadiabatic Variational Energies, E (in a.u.), and Dissociation Energies, D (in cm^{-1}), of HD^+ Vibrational States with Zero Total Angular Momentum, Obtained in Work [124] and the Corresponding Quantities Obtained by Moss [111][a]

v	E, Work [124]	D, Work [124]	E, Moss	D, Moss
0	−0.5978979685	21516.0096	−0.5978979686	21516.0096
1	−0.5891818291	19603.0382	−0.5891818295	19603.0382
2	−0.5809037001	17786.1989	−0.5809037006	17786.1989
3	−0.5730505464	16062.6308	−0.5730505469	16062.6309
4	−0.5656110418	14429.8483	−0.5656110424	14429.8484
5	−0.5585755200	12885.7298	−0.5585755213	12885.7300
6	−0.5519359482	11428.5122	−0.5519359493	11428.5124
7	−0.5456859137	10056.7882	−0.5456859158	10056.7886
8	−0.5398206394	8769.5092	−0.5398206420	8769.5098
9	−0.5343370110	7565.9919	−0.5343370137	7565.9925
10	−0.5292336317	6445.9296	−0.5292336357	6445.9305
11	−0.5245109059	5409.4111	−0.5245109104	5409.4121
12	−0.5201711374	4456.9421	−0.5201711482	4456.9444
13	−0.5162186988	3589.4820	−0.5162187105	3589.4846
14	−0.5126601767	2808.4767	−0.5126601926	2808.4802
15	−0.5095046270	2115.9136	−0.5095046516	2115.9190
16	−0.5067638344	1514.3792	−0.5067638779	1514.3887
17	−0.5044526466	1007.1321	−0.5044526992	1007.1436
18	−0.5025891815	598.1488	−0.5025892341	598.1603
19	−0.5011947323	292.1025	−0.5011947991	292.1172
20	−0.5002924017	94.0638	−0.5002924543	94.0754
21	−0.4999103339	10.2097	−0.4999103614	10.2157
22	−0.4998657692	0.4288	−0.4998657786	0.4309
$D + H^+$	−0.4998638152			

[a]Total energies were calculated using dissociation energies, mass values, and hartree to cm^{-1} conversion factor from Ref. 111.

previous calculations [112], we used 2000 basis functions for each state—except the ground and first excited states, where we limited ourselves to 1000-term expansions because the energies for those states were essentially converged with this number of functions.

In Table IV we compare our variational energies with the best literature values of Moss [112]. As one can see, the values agree very well. The agreement is consistently very good for all the states calculated.

After the wave functions for all 23 $(v = 0, \ldots, 22)$ states were generated, we calculated the expectation values of the internuclear d–p distance, $\langle r_1 \rangle$, the deuteron–electron (d–e) distance, $\langle r_2 \rangle$, and the proton–electron (p–e) distance, $\langle r_{12} \rangle$, for each state. The expectation values of the squares of the distances were also computed.

TABLE V
Expectation Values of the Deuteron–Proton Distance, r_{d-p}, the Deuteron–Electron Distance, r_{d-e}, and the Proton–Electron Distance, r_{p-e}, and Their Squares for the Vibrational Levels of HD^+ in the Rotational Ground State[a]

v	$\langle r_{d-p} \rangle$	$\langle r_{d-e} \rangle$	$\langle r_{p-e} \rangle$	$\langle r_{d-p}^2 \rangle$	$\langle r_{d-e}^2 \rangle$	$\langle r_{p-e}^2 \rangle$
0	2.055	1.688	1.688	4.268	3.534	3.537
1	2.171	1.750	1.750	4.855	3.839	3.843
2	2.292	1.813	1.814	5.492	4.169	4.173
3	2.417	1.880	1.881	6.185	4.526	4.531
4	2.547	1.948	1.950	6.942	4.915	4.921
5	2.683	2.020	2.022	7.771	5.339	5.346
6	2.825	2.095	2.097	8.682	5.804	5.813
7	2.975	2.175	2.177	9.689	6.318	6.329
8	3.135	2.259	2.261	10.81	6.888	6.902
9	3.305	2.348	2.351	12.06	7.527	7.545
10	3.489	2.445	2.448	13.48	8.250	8.272
11	3.689	2.549	2.554	15.09	9.074	9.105
12	3.909	2.664	2.670	16.96	10.03	10.07
13	4.154	2.791	2.799	19.16	11.15	11.21
14	4.432	2.934	2.946	21.79	12.49	12.57
15	4.754	3.099	3.116	25.01	14.13	14.26
16	5.138	3.292	3.319	29.11	16.20	16.41
17	5.611	3.527	3.572	34.55	18.92	19.30
18	6.227	3.821	3.910	42.25	22.66	23.47
19	7.099	4.198	4.421	54.35	28.13	30.38
20	8.550	4.569	5.516	77.74	35.66	46.64
21	12.95	2.306	12.19	176.0	12.94	168.2
22	28.53	1.600	28.46	900.4	4.261	901.8
D atom[b]		1.500			3.002	

[a]All quantities are in atomic units.
[b]In the ground state.

The results are shown in Table V. As can be expected, the average internuclear distance increases with the rising level of excitation. This increase becomes more prominent at the levels near the dissociation threshold. For example, in going from $v = 21$ to $v = 22$ the average internuclear distance increases more than twofold from 12.95 a.u. to 28.53 a.u. In the $v = 22$ state the HD^+ ion is almost dissociated. However, the most striking feature that becomes apparent upon examining the results is a sudden increase of the asymmetry between the deuteron–electron and proton–electron average distances above the $v = 20$ excitation level. In levels up to $v = 20$, there is some asymmetry of the electron distribution with the *p–e* distance being slightly longer than the *d–e* distance. For example, in the $v = 20$ state the *d–e* average distance is 4.569 a.u. and the *p–e* distance is 5.516 a.u. The situation becomes completely different for the $v = 21$ state. Here the *p–e* distance of 12.19 a.u. is almost equal to the

average value of the internuclear distance, but the *d–e* distance becomes much smaller and equals only 2.306 a.u. It is apparent that in this state the electron is essentially localized at the deuteron and the ion becomes highly polarized. An analogous situation also occurs for the $v = 22$ state. Here, again, the *p–e* average distance is very close to the internuclear distance while the *d–e* distance is close to what it is in an isolated D atom.

As has been mentioned above, the inclusion of basis functions (49) with high power values, m_k, is very essential for the calculations of molecular systems. It is especially important for highly vibrationally excited states where there are many highly localized peaks in the nuclear correlation function. To illustrate this point, we calculated this correlation function (it corresponds to the internuclear distance, $r_{d-p} = r_1$), which is the same as the probability density of pseudoparticle 1. The definition of this quantity is as follows:

$$g_1(\boldsymbol{\xi}) = \langle \delta(\mathbf{r}_1 - \boldsymbol{\xi}) \rangle$$

In essence, it is the probability density of the two nuclei to have relative separation $\boldsymbol{\xi}$. Since the orientation of the molecule is not fixed (nuclei are not fixed any more if we deal with an non-BO approach), $g_1(\boldsymbol{\xi})$ is a spherically symmetric function. The plots of $g_1(\boldsymbol{\xi})$ are presented in Figs. 1–4. It should be noted that all the correlation functions shown are normalized in such a way that

$$4\pi \int_0^\infty g_1(\xi)\xi^2 \, d\xi = 1$$

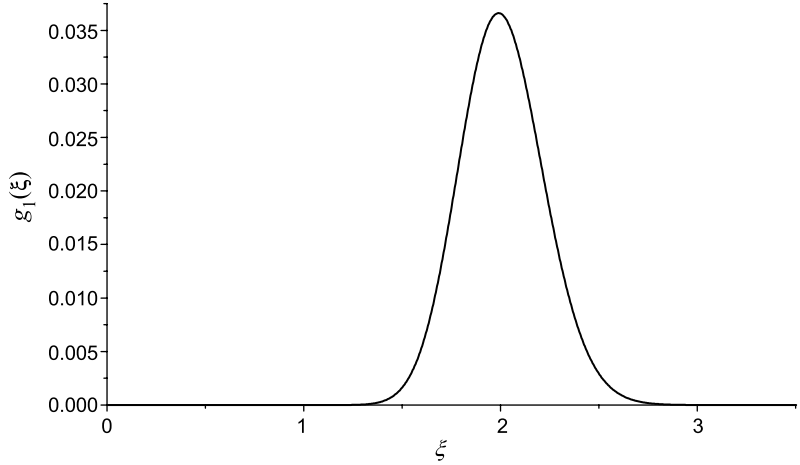


Figure 1. The correlation function $g_1(\xi)$ for $v = 0$ vibrational state of HD^+. All quantities are in a.u.

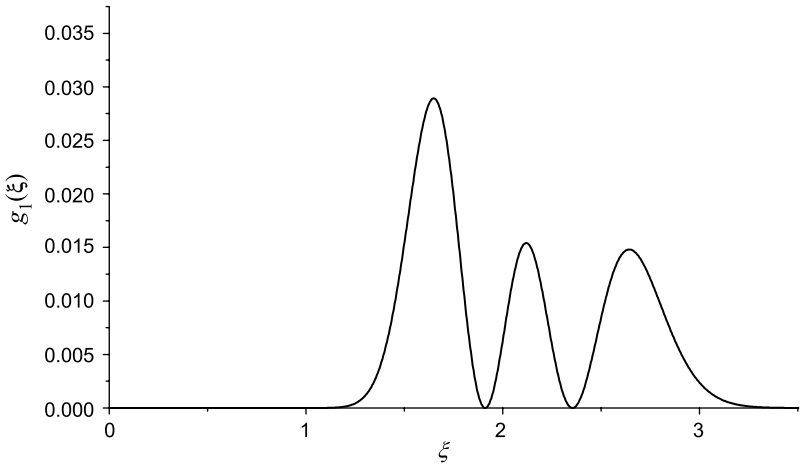


Figure 2. The correlation function $g_1(\xi)$ for $v = 2$ vibrational state of HD^+. All quantities are in a.u.

Upon looking at the graphs, one can clearly see how strongly localized the peaks of $g_1(\xi)$ are—in particular, for the case of the $v = 22$ state. It even surprises us, to some degree, how well a basis set consisting of just 2000 basis functions (49) is capable of fitting the wave function, so that it reproduces seven decimal figures in the total energy.

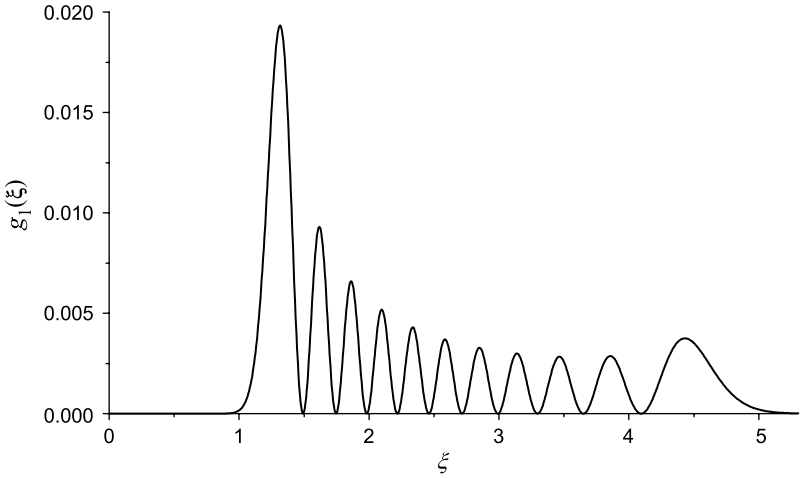


Figure 3. The correlation function $g_1(\xi)$ for $v = 10$ vibrational state of HD^+. All quantities are in a.u.

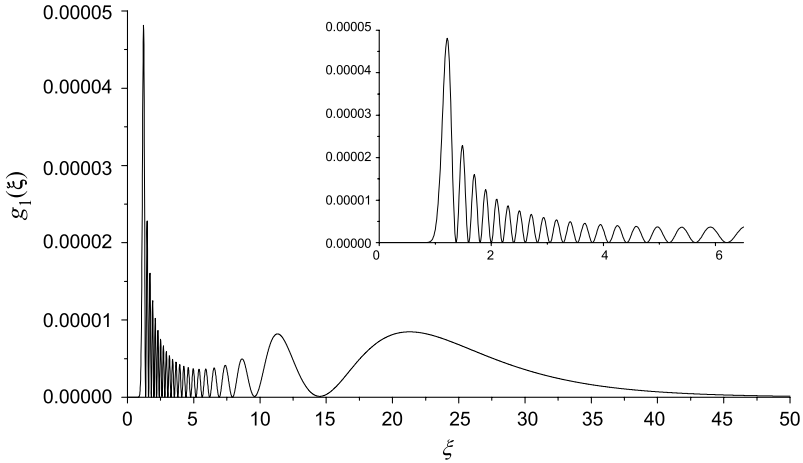


Figure 4. The correlation function $g_1(\xi)$ for $v = 22$ vibrational state of HD^+. All quantities are in a.u.

G. LiH and LiD Electron Affinity

The determination of electron affinities (EAs) is one of the most serious problems in quantum chemistry. While the Hartree–Fock electron affinity can be easily evaluated, most anions turn out to be unbound at this level of theory. Thus, the correlation effects are extremely crucial in evaluating EAs. At this point, lithium hydride and lithium hydride anion make up a very good benchmark system because they are still small enough yet exhibit features of more complicated systems. Four and five electrons, respectively, give rise to higher-order correlation effects that are not possible in H_2.

The theoretical interest in the LiH^- has increased since the electron affinity of LiH and its deuterated counterpart, LiD, were measured with the use of the photoelectron spectroscopy by Bowen and co-workers [126]. The adiabatic electron affinities of ^{7}LiH and ^{7}LiD determined in that experiment were 0.342 ± 0.012 eV for the former and 0.337 ± 0.012 eV for the latter system. The appearance of these data posed a challenge for theory to reproduce those values in rigorous calculations based on the first principles. Since the two systems are small, it has been particularly interesting to see if the experimental EAs can be reproduced in calculations where the BO approximation is not assumed [123].

Since the time that Bowen's and co-workers' article was published, the theory based on the BO approximation, except for one very recent multi-reference configuration interaction (MRCI) calculation by Chang et al. [127], has been unable to produce a value of the LiH adiabatic electron affinity that

falls within the experimental uncertainty bracket. More than that, in many works where authors managed to obtain an EA that was relatively close to the experimentally determined one, the success was often due to fortuitous cancellation of errors in the total energies of LiH and LiH^- rather than the accuracy achieved in calculations. This inability seems somewhat odd, since there is nothing particularly unusual about and fundamentally difficult to describe in the LiH^-/LiH system. Since the LiH molecule is a polar system ($Li^{\delta+}$–$H^{\delta-}$), an excess electron in the process of attachment localizes on the electropositive alkali atom in a nonbonding orbital of the neutral molecule.

We would like to mention that one should not expect the nonadiabatic effects to play a noticeable role in the LiH and LiD electron affinity calculations. However, by applying the non-BO approach, we can directly determine the total energies of the anion and the neutral system in one-step calculations, and we do not need to resort to calculating the electronic potential energy curves for the anion and the neutral first and using them for calculating their nuclear vibration energies in the next step, as is done in the BO approach. Thus, using the non-BO method, we not only make the calculation free of any artifacts that may result from the two-step procedure used in the BO approximation, but we also obtain total and relative energies that, if the basis sets become more complete, approach the true nonrelativistic limits of those quantities, free of any approximations.

In comparison with H_2 and HD^+, LiH and LiH^- represent a significantly more challenging case from the point of view of computational demands, even if we restrict ourselves to the consideration of the ground state only. The total number of particles is six in the case of LiH and seven in the case of LiH^-, which means that effectively we deal with a 15- and an 18-dimensional problem, respectively. Considering that the number of permutations in the Young operator also increases as the number of identical particles increases (this number is 24 LiH and 120 for LiH^-), the amount of computational work required for performing sound calculations rises by a factor of several orders of magnitude. To attempt such calculations at the present time, one has to gain access to large-scale parallel computer systems. However, with advances in computer hardware, this kind of calculation may soon become quite ordinary and even routine.

The results of our calculations for the energy are shown in Table VI. In Table VII we show the values of the LiH and LiD electron affinities calculated as the difference of the energies of the anion and the neutral system for all lengths of the basis set for which the total energies are reported. A question can be raised whether it is appropriate to use the total energies obtained with the same length of the basis set for LiH^- and LiH (or LiD^- and LiD) in the electron affinity calculation. Since LiH^- has one more electron than LiH, it should require more basis functions for LiH^- than for LiH to achieve a similar level of

TABLE VI
Nonadiabatic Variational Ground-State Energies of the LiH^-, LiH, LiD^-, and LiD Molecules Obtained with Different Basis Set Sizes [123][a]

Basis Size	LiH	LiD	LiH^-	LiD^-
1000	−8.06632055	−8.06725433	−8.07778128	−8.07859843
1200	−8.06634454	−8.06728273	−8.07794715	−8.07875507
1400	−8.06636491	−8.06730593	−8.07807991	−8.07888651
1600	−8.06638295	−8.06732805	−8.07822084	−8.07903483
1800	−8.06639498	−8.06734216	−8.07830536	−8.07912238
2000	−8.06640408	−8.06735333	−8.07837566	−8.07919877
2200	−8.06641099	−8.06736128	−8.07842711	−8.07925549
2400	−8.06641554	−8.06736613	−8.07846014	−8.07929081
2600	−8.06642068	−8.06737208	−8.07848753	−8.07931985
2800	−8.06642353	−8.06737514	−8.07851089	−8.07934370
3000	−8.06642581	−8.06737749	−8.07852933	−8.07936248
3200	−8.06642787	−8.06737962	−8.07854406	−8.07937759
3400	−8.06642941	−8.06738123	−8.07855805	−8.07939162
3600	−8.06643070	−8.06738251	−8.07856887	−8.07940445

[a]All energies are in atomic units.

TABLE VII
Convergence of the Electron Affinities of LiH and LiD (in eV) in Terms of the Number of the Basis Functions [123] and the Corresponding Experimental Values of Sarkas et al.

Basis Size	LiH	LiD
1000	0.31186	0.30869
1200	0.31572	0.31218
1400	0.31878	0.31512
1600	0.32213	0.31856
1800	0.32410	0.32056
2000	0.32576	0.32233
2200	0.32698	0.32366
2400	0.32775	0.32449
2600	0.32836	0.32512
2800	0.32891	0.32568
3000	0.32935	0.32613
3200	0.32970	0.32648
3400	0.33004	0.32682
3600	0.33030	0.32713
Experiment [126]	0.342 ± 0.012	0.337 ± 0.012

accuracy of the results. This is indeed showing in the convergence of the results presented in Table VI. Certainly the 3600-term energy for LiH is better converged than the 3600-term energy for LiH^-. It should be mentioned that the LiH and LiD variational energies obtained in work [123] are the lowest non-BO ground-state energies ever obtained for these systems. It is significantly lower than the previous non-BO value of −8.06615576 hartree, which was also obtained in our group [120]. The same is true about the LiH^- and LiD^- energies, though in this case the final values are not as tightly converged as the energies for LiH and LiD. Also, work [123], to date, is the only one where a variational, non-BO calculation of the ground state of a five-electron system such as LiH^- or LiD^- has been attempted.

Since the variational approach has been used in this work and the LiH (LiD) energy at a particular basis set size is better converged than the LiH^- (LiD^-) energy, the calculated electron affinity is always lower than the result would be in the limit of the complete basis set. Thus, by calculating the EA at the basis set with the same size for the anion and for neutral system and by incrementally increasing the size of the basis, we could monitor the EA convergence and be sure that all the EA values we calculated including the final value (this being the 3600-term result) are lower bounds to the true EA.

As shown in Table VII, our best results for LiH and LiD electron affinities obtained with the 3600-term expansions of the wave functions for LiH^-/LiH and LiD^-/LiD are 0.33030 and 0.32713 eV, respectively. Even though, as stated, both values represent lower bounds to the true EAs, they both are within the uncertainty brackets of the experimental results of 0.342 ± 0.012 eV (LiH) and 0.337 ± 0.012 eV (LiD) obtained by Bowen and co-workers [126].

In the calculations, the powers, m_k, in the preexponential factors in the basis functions (49) were allowed to vary in the interval of 0–200. The obtained distributions of m_k's have the mean values of 70.0, 74.1, 67.6, and 67.9, along with the standard deviations of 49.7, 50.4, 49.3, and 49.7, for LiH, LiD, LiH^-, and LiD^-, respectively. A slightly higher mean power for LiD than for LiH can be explained by a more localized vibrational component of the wave function for the former than for the latter system. Lower mean powers for the anions than for the neutral systems result from two opposing effects. First, the bond lengths for the anions are slightly longer than for the neutral counterparts (see the next paragraph), which should require larger powers. Second, due the weakening of the bonds in the anions, the vibrational components of their wave functions become more delocalized, resulting in lowering of the powers. Apparently the second effect dominates over the first one.

Finally, Table VIII shows the expectation values of the internuclear distance and its square for LiH, LiD, LiH^-, and LiD^-. Here we see trends that can be easily understood considering that the attachment of an excess electron weakens slightly the Li–H bond and that increasing the mass of H by switching to D

TABLE VIII
Expectation Values of the Internuclear Distance, $r_1 \equiv r_{LiH}$, and Its Square (in a.u.) for LiH, LiD, LiH^-, and LiD^-, Calculated with the 3600-Term Basis Sets [123]

System	$\langle r_1 \rangle$	$\langle r_1^2 \rangle$
LiH^-	3.214708	10.39391
LiH	3.061047	9.419733
LiD^-	3.199737	10.28362
LiD	3.049131	9.334349

results in making the vibrational component of the total wave function more localized, resulting in a slight contraction of the average length of the bond. This contraction is very similar in going from LiH and LiD and from LiH^- to LiD^-.

H. Molecules Containing Positron: e^+LiH

During the last several years, much advance has been made concerning the study of bound states of the positron with small systems. The ability of various atoms, ions, and molecules to bind a positron is now well established and represent a popular subject of research. However, most of the calculations performed were done on atomic systems (or those containing just one particle significanly heavier than the positron). For example, there has been considerable interest to the positronium hydride, HPs ($p^+e^+e^-e^-$), and its isotopomers (see Refs. 128–130 and references therein). Molecular systems containing a positron have been mainly treated with the use of the Quantum Monte Carlo [131, 132] or variational method [133–135] assuming BO approximation. An attempt of calculating e^+LiH was made in Ref. 136, where the authors used simple Gaussians (without premultipliers). However, the convergence of the total energy turned out to be very slow while the basis size was not large, which resulted in relatively low accuracy of the calculations.

In Ref. 122 we raised the question of whether one can use basis functions (49) in non-BO quantum mechanical calculations of molecular diatomic systems containing positrons and whether those functions are capable of providing a proper representation for the positron–nucleus and positron–electron correlation effects in a diatomic system. e^+LiH was chosen as a target system. Along with e^+LiH, we also performed the calculations of HPs and Li^+ because the total energies of these systems are needed for determining the dissotiation energy.

The convergence of the energy values for HPs, LiH, and e^+LiH in terms of the number of the basis functions is shown in Table IX. In the case of e^+LiH, as

TABLE IX
Total Non-Born–Oppenheimer Energies (in a.u.) of HPs, LiH, and e^+LiH as a Function of Basis Size [122]

Basis Size	HPs	LiH	e^+LiH
800	−0.7888705040	−8.066278419	−8.103075429
1000	−0.7888705983	−8.066320545	−8.103572816
1200	−0.7888706398	−8.066344535	−8.103905788
1400	−0.7888706611	−8.066364905	−8.104113213
1600	−0.7888706790	−8.066382950	−8.104256550
1800	−0.7888706877	−8.066394978	−8.104372009
2000	−0.7888706940	−8.066404077	−8.104478249
2200	−0.7888706984	−8.066410987	−8.104543434
2400	−0.7888707014	−8.066415542	−8.104598552
2600	−0.7888707036	−8.066420678	−8.104645276
2800	−0.7888707057	−8.066423527	−8.104683502
3000	−0.7888707062	−8.066425806	−8.104713922
3200	−0.7888707066	−8.066427866	−8.104739913

well as for LiH case, the powers m_k in (49) were selected from the interval 0–200. Although the positronium hydride wave function can be obtained with very high precision even without using powers of the hydrogen–positron distance in the preexponential factors in the basis functions, we did include some functions with small preexponential powers ranging from 0 to 10 to ensure better numerical stability in the calculations.

From the lowest energy values shown in Table IX, one can determine that the positron detachment energy of e^+LiH, $\mathrm{PDE} = E(e^+\mathrm{LiH}) - E(\mathrm{LiH})$, is 0.038312 hartree. The lowest-energy fragmentation of e^+LiH corresponds to dissociation of the system into HPs + Li^+. To calculate the dissociation energy, $\mathrm{DE} = E(e^+\mathrm{LiH}) - E(i^+) - E(\mathrm{HPs})$, one needs to determine the total energy of the Li^+ ion. Since the non-BO calculation of this quantity is very simple, rather than taking it from the literature, we recalculated it using our method. A 400-term expansion was sufficient to obtain a highly accurate result of −7.279321518 hartree, where, we believe, all the significant figures shown are exact.

Given the values of the HPs and Li^+ energies calculated in this work in addition to that of e^+LiH, our dissociation energy is 0.036548 hartree. This value qualitatively agrees with the value of 0.0382(2) hartree obtained in the Born–Oppenheimer calculations by Mella and et al. [131] using the Quantum Monte Carlo method. It also agrees with the Born–Oppenheimer result of 0.036936 obtained by Strasburger [135] with the use of explicitly correlated Gaussians and the variational method.

The lowest variational energy upper-bound for the Born–Oppenheimer LiH ground-state energy to date is −8.070538 hartree [137]. Assuming that the energies of Li and H with infinitely heavy nuclei are −7.4780603 [138] and −0.5 hartree, respectively, one obtains the infinite-mass dissociation energy of LiH of −0.0924777 hartree. The finite-mass energy of LiH can be estimated by subtracting this number from the sum of the finite-mass energies of Li and H atoms and by adding to the result the zero-point LiH energy. Using the finite-mass energy of Li of −7.4774519 hartree which one can calculate by using the expansion from the article of Yan and Drake [138], the corresponding value for the H atom of −0.4997278 hartree, and the zero-point LiH energy of 0.0031981 hartree (see Ref. 139), we obtain the LiH ground-state energy corrected for the finite nuclear masses equal to −8.066459 hartree. The use of experimentally determined zero-point energy of −0.0031799 [140] shifts this value to −8.066478 hartree. The difference between this value and our non-BO result of −8.066427866 hartree is larger than the estimated sum of their inaccuracies. Although, perhaps, this may partially be attributed to relatively low accuracy of the zero-point energy, it is clear that the nonadiabatic effect of the coupled electron–nuclear motion must play a role in the difference.

In Table X we present expectation values of the internuclear distance and its square for LiH and e^+LiH as well as the electron–positron contact densities for HPs and e^+LiH evaluated with the largest basis set of 3200 basis functions obtained in the calculations. It should be noted that the mean internuclear distance of LiH calculated here is slightly higher than the known value, 3.015 bohr, of the equilibrium nuclear distance—that is, the distance where the potential energy curve reaches its minimum. This is, obviously, an expected result since larger distances contribute more to the mean distance when one averages the internuclear distance over the ''vibrational'' part of the wave function, even if a purely Born–Oppenheimer calculation is carried out, and, hence, the mean distance is always larger than the distance corresponding the peak of the vibrational component of the wave function (minimum of the potential energy curve). Thus the discrepancy becomes more and more

TABLE X
Expectation Values of the Li–H Internuclear Distance, Its Square, and Electron–Positron Contact Densities Evaluated at 3200-Function Basis Size [122][a]

System	$\langle r_{\mathrm{LiH}} \rangle$	$\langle r^2_{\mathrm{LiH}} \rangle$	$\langle \delta(\mathbf{r}_{e^-e^+}) \rangle$
HPs	—	—	2.44855×10^{-2}
LiH	3.06105	9.41977	—
e^+LiH	3.44470	11.9397	7.08879×10^{-3}

[a]All quantities are in atomic units.

noticeable as the nuclear masses decrease. For e^+LiH, our $r_{iH}(\equiv r_1)$ mean distance of 3.445 bohr agrees well with the the equilibrium internuclear separation of 3.348 bohr obtained by Strasburger [135] in the Born–Oppenheimer calculations.

An important characteristic of positronic systems relevant to the experiment is their lifetimes. The expectation value of the electron–positron contact density allows us to evaluate the two-photon annihilation rate for a positronic system using the expression

$$\Gamma_{2\gamma} = n\frac{\pi\alpha^4 c}{a_0}\langle\delta(\mathbf{r}_{e^-e^+})\rangle$$

where α is the fine structure constant, a_0 is the Bohr radius, c is the velocity of light, and n denotes the number of electron–positron pairs in the system (2 and 4 in the case of HPs and e^+LiH, respectively). The two-photon annihilation rates we obtained with 3200 basis functions are $2.4716 \times 10^9\ s^{-1}$ for HPs and $1.4311 \times 10^9\ s^{-1}$ for e^+LiH. This indicates that a positron attached to LiH survives much longer than in the HPs system. The HPs annihilation rate can be compared with the result of Yan and Ho obtained in a finite-mass calculation using Hylleraas coordinates [141] (HPs) and with the explicitly correlated Gaussian calculation [129] performed for $^\infty$HPs (HPs with infinetely heavy proton), both of which yielded the value of $2.4722 \times 10^9\ s^{-1}$. In the case of e^+LiH, we can make a comparison with the Born–Oppenheimer Quantum Monte Carlo result of Mella et al. [132], which yielded $1.49 \times 10^9\ s^{-1}$ (the vibrationally averaged result), and with the Born–Oppenheimer explicitly correlated Gaussian result of Strasburger [135] where the value of $1.375 \times 10^9\ s^{-1}$ was obtained at the e^+LiH equilibrium distance of $R = 3.348$ bohr.

VI. NON-BO CALCULATIONS OF DIATOMIC MOLECULES IN ELECTRIC FIELD WITH SHIFTED GAUSSIANS

When a diatomic molecule enters an electric field, the rotational symmetry of the wave function is broken. The wave function is no longer an eigenfunction of the total angular momentum operator, $\hat{\boldsymbol{J}}$, but, if we assume the electric field to be directed along the molecular z axis, is only an eigenfunction of $\hat{\boldsymbol{J}}_z$. Thus, we cannot use the above spherically symmetric basis functions to expand the wave function. We find that expansion of the electric-field-perturbed wave function in terms of spherical Gaussians with floating centers provides a good description of the electric field effect.

It can be shown that the basis of spherical explicitly correlated Gaussian functions with floating centers (FSECG) form a complete set. These functions

have the form

$$g_k(r) = \prod_{i=1}^{n} \exp(-\alpha_i^k (r_i - R_i^k)^2) \prod_{j>i}^{n} \exp(-\beta_{ij}^k (r_i - r_j)^2) \tag{121}$$

which may equivalently be expressed as

$$g_k(r) = \exp(-(r - s_k)'(A_k \otimes I_3)(r - s_k)) \tag{122}$$

The Kronecker product with the identity ensures rotational invariance (sphericalness); elliptical Gaussians could be obtained by using a full $n \times n$ A matrix. In the former formulation of the basis function, it is difficult to ensure the square integrability of the functions, but this becomes easy in the latter formulation. In this format, all that is required is that the matrix, A_k, be positive definite. This may be achieved by constructing the matrix from a Cholesky decomposition: $A_k = L_k' L_k$. Later in this work we will use the notation $\bar{A} = A \otimes I_3$ to indicate the Kronecker product with the 3×3 identity matrix.

The correct symmetry of the system may be ensured by projection onto the irreducible representations, as described above. Thus, when taking symmetry into account, the final form of the basis function is

$$\hat{Y} g_k(r) = \prod_i \hat{Y}_i \exp\{-(r - s_k)'[(L_k L_k') \otimes I_3](r - s_k)\} \tag{123}$$

and the spin-free spatial wave function has the form

$$\Psi = \sum_{k=1}^{m} c_k \hat{Y} g_k \tag{124}$$

where m is the size of the basis.

A. Integrals and Energy Derivatives

1. *Born–Oppenheimer Integrals over Correlated Gaussians*

Later we will discuss conventional Born–Oppenheimer calculations used in conjunction with the current work. Thus, for completeness, we will cover here the integrals needed for these calculations. These integrals are quite similar to the ones used in the non-Born–Oppenheimer calculaions, as will be shown below. We show first the integrals over the Born–Oppenheimer Hamiltonian:

$$\hat{H} = -\frac{1}{2} \nabla_r' \nabla_r + \sum_{j>i}^{n} \frac{1}{r_{ij}} - \sum_{i=1}^{n} \sum_{t=1}^{N} \frac{Z_t}{r_{it}} + \sum_{u>t}^{N} \frac{Z_t Z_u}{r_{tu}} \tag{125}$$

where explicit separation of the electronic and nuclear coordinates has been performed. In this case, n indicates only the number of electrons, rather than indicating the number of particles, while N indicates the number of nuclei. These molecular integrals have been published before in several formats [48, 52, 61, 75], but the format presented here is most useful for deriving the analytical gradients used in the optimization of the wave function.

Some definitions used in the integral formulas are

$$\begin{aligned} \bar{A}_{kl} &= \bar{A}_k + \bar{A}_l \\ e &= s_k' \bar{A}_k + s_l' \bar{A}_l \\ s &= \bar{A}_{kl}^{-1} e \\ p_{kl} &= \frac{|A_k|^{1/2} |A_l|^{1/2}}{|A_{kl}|} \end{aligned}$$

and

$$\gamma = -s_k' \bar{A}_k s_k - s_l' \bar{A}_l s_l + e' \bar{A}_{kl}^{-1} e$$

Also, we will use the J_{ij} matrix given by (52). We define a $3n$ vector t as containing the three coordinates of the tth nucleus repeated n times:

$$t = \begin{pmatrix} r_t \\ r_t \\ \vdots \\ r_t \end{pmatrix}$$

Repeated use is made of the following formula:

$$\int_{-\infty}^{\infty} \exp(-x'Mx + iz'x)dx = \pi^{n/2} |M|^{-1/2} \exp(-z'M^{-1}z/4) \tag{126}$$

where M is an $n \times n$ matrix, and x and z are vectors of size n.

The overlap integral is the simplest and serves as a basis for the rest of the integrals. The product of two functions is

$$\begin{aligned} g_k g_l &= \exp(-(r - s_k)' \bar{A}_k (r - s_k) - (r - s_l)' \bar{A}_l (r - s_l)) \\ &= \exp(-r' (\bar{A}_k + \bar{A}_l) r + 2(s_k' \bar{A}_k + s_l' \bar{A}_l) r) \times \exp(-s_k' \bar{A}_k s_k - s_l' \bar{A}_l s_l) \end{aligned}$$

Using (38), we get the following expression for the overlap integral:

$$\langle g_k | g_l \rangle = \pi^{3n/2} |A_{kl}|^{-3/2} \exp(-s_k' \bar{A}_k s_k - s_l' \bar{A}_l s_l + e' \bar{A}_{kl}^{-1} e) \tag{127}$$

Using the same method for the integrals $\langle g_k|g_k\rangle$ and $\langle g_l|g_l\rangle$, we find the overlap integral for the normalized basis functions to be

$$S_{kl} = 2^{3n}\left(\frac{|A_k|^{1/2}|A_l|^{1/2}}{|A_{kl}|}\right)^{3/2}\exp(\gamma) \tag{128}$$

The kinetic energy integral follows easily from the overlap integral. We write the kinetic energy operator

$$\hat{T} = -\frac{1}{2}\nabla_r \cdot \nabla_r \tag{129}$$

where the subscript r denotes the gradient with respect to all $3n$ coordinates of the n electrons. The integral is thus

$$\begin{aligned} T_{kl} &= \langle g_k|\hat{T}|g_l\rangle \\ &= -\frac{1}{2}\langle g_k|\nabla_r \cdot \nabla_r|g_l\rangle \\ &= \frac{1}{2}\langle \nabla'_r g_k|\nabla_r g_l\rangle \end{aligned}$$

Looking at (2), it is evident that $\nabla_r g_k = -\nabla_{s_k} g_k$, thus our integral becomes

$$\begin{aligned} T_{kl} &= \frac{1}{2}\langle \nabla'_{s_k} g_k|\nabla_{s_l} g_l\rangle \\ &= \frac{1}{2}\langle g_k|\nabla'_{s_k}\nabla_{s_l}|g_l\rangle \\ &= \frac{1}{2}\mathrm{tr}[\nabla_{s_l}\nabla'_{s_k}\langle g_k|g_l\rangle] \end{aligned}$$

The two derivatives of the overlap integral are easily evaluated directly from (40) and the result is

$$\begin{aligned} \nabla'_{s_k}\langle g_k|g_l\rangle &= S_{kl}(-2\bar{A}_k s_k + 2\bar{A}_k s) \\ \nabla_{s_l}\nabla'_{s_k}\langle g_k|g_l\rangle &= S_{kl}[(2\bar{A}_l s - 2\bar{A}_l s_l)'(2\bar{A}_k s - 2\bar{A}_k s_k) + 2\bar{A}_k\bar{A}_{kl}^{-1}\bar{A}_l] \end{aligned}$$

So the final expression for the kinetic energy integral is

$$T_{kl} = S_{kl}(2(s - s_k)'\bar{A}_k\bar{A}_l(s - s_l) + 3\mathrm{tr}[A_k A_{kl}^{-1}A_l]) \tag{130}$$

The two potential energy integrals, electron repulsion and nuclear attraction, are derived quite similarly. Both involve the Gaussian transformation

$$\frac{1}{r_{ij}} = \frac{2}{\pi^{1/2}} \int_0^\infty \exp(-u^2 r_{ij}^2)\, du \tag{131}$$

and the identity

$$\int_0^\infty (1+\alpha u^2)^{-3/2} \exp\left(-\frac{a u^2}{1+\alpha u^2}\right) du = \frac{\pi^{1/2}}{2a^{1/2}} \operatorname{erf}\left(\left(\frac{a}{\alpha}\right)^{1/2}\right) \tag{132}$$

The electron repulsion integral is the more straightforward of the two potential integrals. It should be recognized that r_{ij}^2 can be written as $r' \bar{J}_{ij} r$. Thus the integral, using (43), becomes

$$\begin{aligned} ER_{kl}^{ij} &= \langle g_k | \frac{1}{r_{ij}} | g_l \rangle \\ &= \frac{2}{\pi^{1/2}} \exp(\gamma) \int_{-\infty}^\infty dr \int_0^\infty du\ \exp(-(r-s)' \bar{A}_{kl} (r-s) - u^2 r' \bar{J}_{ij} r) \end{aligned}$$

which can be further rearranged to give

$$ER_{kl}^{ij} = \frac{2}{\pi^{1/2}} \exp(\gamma) \exp(-s' \bar{A}_{kl} s) \int_{-\infty}^\infty dr \int_0^\infty du\ \exp(-r'(\bar{A}_{kl} + u^2 \bar{J}_{ij}) r + 2 s' \bar{A}_{kl} r)$$

Applying (38) to the integral over r, we get

$$ER_{kl}^{ij} = \frac{2}{\pi^{1/2}} \exp(\gamma) \pi^{3n/2} \int_0^\infty du |\bar{A}_{kl} + u^2 \bar{J}_{ij}|^{-1/2} \exp\left(-\frac{u^2 s' \bar{J}_{ij} s}{1 + u^2 \operatorname{tr}[J_{ij} A_{kl}]}\right)$$

Following the approach given in the work of Kinghorn [116], we further get

$$ER_{kl}^{ij} = \frac{2}{\pi^{1/2}} S_{kl} \int_0^\infty du (1 + u^2 \operatorname{tr}[J_{ij} A_{kl}])^{-3/2} \exp\left(-\frac{u^2 s' \bar{J}_{ij} s}{1 + u^2 \operatorname{tr}[J_{ij} A_{kl}]}\right)$$

By applying (44), we get

$$ER_{kl}^{ij} = S_{kl} \left(\frac{1}{s' \bar{J}_{ij} s}\right)^{1/2} \operatorname{erf}\left(\left(\frac{s' \bar{J}_{ij} s}{\operatorname{tr}[J_{ij} A_{kl}]}\right)^{1/2}\right) \tag{133}$$

In order to solve the nuclear attraction integral, we use the t vector. We recognize that $r_{it}^2 = (r-t)'\bar{J}_{ii}(r-t)$. The integral can thus be written as

$$\begin{aligned} NA_{kl}^{it} &= \langle g_k | \frac{1}{r_{it}} | g_l \rangle \\ &= \frac{2}{\pi^{1/2}} \exp(\gamma) \int_{-\infty}^{\infty} dr \int_0^{\infty} du \, \exp(-(r-s)'\bar{A}_{kl}(r-s) - u^2(r-t)'\bar{J}_{ii}(r-t)) \end{aligned}$$

which, with a little rearrangement, can be put into the form

$$\begin{aligned} NA_{kl}^{it} = {} & \frac{2}{\pi^{1/2}} \exp(\gamma) \exp(-(s-t)'\bar{A}_{kl}(s-t)) \\ & \times \int_{-\infty}^{\infty} dr \int_0^{\infty} du \exp(-(r-t)'(\bar{A}_{kl} + u^2\bar{J}_{ii})(r-t) + 2(s-t)'\bar{A}_{kl}(r-t)) \end{aligned}$$

By changing the integration variable from r to $(r-t)$ and applying (38), we get

$$NA_{kl}^{it} = \frac{2}{\pi^{1/2}} \exp(\gamma) \pi^{3n/2} \int_0^{\infty} du |\bar{A}_{kl} + u^2\bar{J}_{ii}|^{-1/2} \exp\left(- \frac{u^2(s-t)'\bar{J}_{ij}(s-t)}{1 + u^2 \mathrm{tr}[J_{ii}A_{kl}^{-1}]} \right)$$

the form of which is very similar to the electron repulsion integral. Using the same steps as in that integral, we get

$$NA_{kl}^{it} = S_{kl} \left(\frac{1}{(s-t)'\bar{J}_{ii}(s-t)} \right)^{1/2} \mathrm{erf} \left(\left(\frac{(s-t)'\bar{J}_{ii}(s-t)}{\mathrm{tr}[J_{ii}A_{kl}^{-1}]} \right)^{1/2} \right) \tag{134}$$

After converting the error function to the F_0 function, both of the potential energy integrals have the same general form

$$V_{kl} = \frac{2}{\pi^{1/2}} S_{kl} \alpha^{-1/2} F_0\left(\frac{\rho}{\alpha}\right) \tag{135}$$

For $V_{kl} = ER_{kl}^{ij}$, we have $\alpha = \mathrm{tr}[J_{ij}A_{kl}^{-1}]$ and $\rho = s'\bar{J}_{ij}s$; for $V_{kl} = NA_{kl}^{it}$, we have $\alpha = \mathrm{tr}[J_{ii}A_{kl}^{-1}]$ and $\rho = (s-t)'\bar{J}_{ii}(s-t)$.

The F_0 function is defined as

$$F_0(x) = \tfrac{1}{2}\sqrt{\tfrac{\pi}{x}} \, \mathrm{erf}(\sqrt{x})$$

2. *Non-Born–Oppenheimer Integrals over Correlated Gaussians*

Most of the integrals used in the non-BO calculations—the overlap integral and the potential energy integral—are similar to those presented above and will not be derived. A slight difference in the kinetic energy integral will be shown. In this case we find integrals over the operators in the non-BO Hamiltonian:

$$\hat{H} = -\nabla_r' \bar{M} \nabla_r + \sum_{i=1}^{n} \frac{q_0 q_i}{r_i} + \sum_{i<j}^{n} \frac{q_i q_j}{r_{ij}} \tag{136}$$

The kinetic energy integral is given by

$$T_{kl} = S_{kl}(4(s - s_k)'\bar{A}_k\bar{M}\bar{A}_l(s - s_l) + 6\mathrm{tr}[MA_lA_{kl}^{-1}A_k]) \tag{137}$$

This form differs from the Born–Oppenheimer form in that the constant $\frac{1}{2}$ (for electron masses) is replaced by the mass matrix, M (described above), but all other steps are similar.

3. *Electric Field Integrals*

The electric field term is most easily evaluated with one electric field component at a time. Taking the z component, we have

$$\begin{aligned} FT_{kl}^z &= \langle g_k| \sum_{i=1}^{n} \mu_z^i \varepsilon_z |g_l\rangle \\ FT_{kl}^z &= \langle g_k| \sum_{i=1}^{n} q_i r^{3i} \varepsilon_z |g_l\rangle \\ FT_{kl}^z &= \sum_{i=1}^{n} \varepsilon_z q_i \langle g_k| r^{3i} |g_l\rangle \end{aligned} \tag{138}$$

and the integral reduces to an evaluation of $\langle g_k|r_{3i}|g_l\rangle$. We will present the general case of this integral for all components, i:

$$r_{kl}^i = \langle g_k|r^i|g_l\rangle \tag{139}$$

This integral is solved by examining the expectation value of the operator:

$$o'E_{ii}\bar{A}_l^{-1}\nabla_r \tag{140}$$

where o is a vector of ones, and E_{ii} is a $3n \times 3n$ matrix as described above.

The gradient of the basis function is simply:

$$\nabla_r' g_l = g_l(2\bar{A}_l s_l - 2\bar{A}_l r) \tag{141}$$

so the expectation value of the operator is

$$\langle o'E_{ii}\bar{A}_l^{-1}\nabla_r\rangle = 2o'E_{ii}s_l S_{kl} - 2r_{kl}^i \tag{142}$$

Looking at the form of the basis function (2), it is obvious that

$$\nabla_r' g_l = -\nabla_{s_l}' g_l \tag{143}$$

so that

$$\langle o' E_{ii} \bar{A}_l^{-1} \nabla_r \rangle = -\langle o' E_{ii} \bar{A}_l^{-1} \nabla_{s_l} \rangle \tag{144}$$

The right-hand side of the above is easily evaluated by pulling the operator out of the integral:

$$\langle o' E_{ii} \bar{A}_l^{-1} \nabla_{s_l} \rangle = 2S_{kl}(o' E_{ii} s - o' E_{ii} s_l) \tag{145}$$

Setting one of the operators equal to the negative of the other, we get

$$r_{kl}^i = S_{kl} o' E_{ii} s = S_{kl} s(i) \tag{146}$$

4. *Energy Gradients*

The gradients of the molecular integrals with respect to the nonlinear variational parameters (i.e., the exponential parameters A_k and the orbital centers s_k) were derived using the methods of matrix differential calculus as introduced by Kinghorn [116]. It was shown there that the energy gradient with respect to all nonlinear variational parameters can be written as

$$\nabla_a E = \frac{1}{c' S c} \left(\frac{\partial \text{vech } H}{\partial a'} - E \frac{\partial \text{vech } S}{\partial a'} \right)' (\text{vech } [2cc' - \text{diag } cc']) \tag{147}$$

where a is a $m(n(n+1)/2 + 3n)$ vector of all the nonlinear parameters. Each function, g_k, contains the lower triangular matrix L_k with $n(n+1)/2$ nonzero elements and the $3n$ vector s_k. The vector a is made by stacking, function by function, the nonzero elements of L_k followed by the vector s_k. The matrices of derivatives $\partial \text{vech } H / \partial a'$ and $\partial \text{vech } S / \partial a'$ are sparse, since each row has at most $2(n(n+1)/2 + 3n)$ nonzero elements. The derivatives of the Hamiltonian are determined using the derivatives of the molecular integrals (Born–Oppenheimer):

$$\frac{\partial \text{vech } H}{\partial a'} = \frac{\partial \text{vech } T}{\partial a'} + \frac{\partial \text{vech } ER}{\partial a'} + \frac{\partial \text{vech } NR}{\partial a'} - \frac{\partial \text{vech } NA}{\partial a'} \tag{148}$$

where NR stands for the nuclear–nuclear repulsion energy, or (non-Born–Oppenheimer)

$$\frac{\partial \text{vech } H}{\partial a'} = \frac{\partial \text{vech } T}{\partial a'} + \frac{\partial \text{vech } V}{\partial a'} \tag{149}$$

The differential of the overlap integral with respect to s_k is straightforward,

$$\begin{aligned} d_{s_k} S_{kl} &= S_{kl} d_{s_k} (-s_k' \bar{A}_k s_k + e' \bar{A}_{kl}^{-1} e) \\ &= S_{kl} (-2 s_k' \bar{A}_k d s_k + 2 e' \bar{A}_{kl}^{-1} d e) \\ &= S_{kl} (-2 s_k' \bar{A}_k + 2 s' \bar{A}_k) d s_k \\ &= 2 S_{kl} ((s - s_k)' \bar{A}_k) d s_k \end{aligned} \tag{150}$$

The kinetic energy differential is only slightly more complicated:

$$\begin{aligned} d_{s_k} T_{kl} &= \frac{T_{kl}}{S_{kl}} d_{s_k} S_{kl} + 2 S_{kl} d_{s_k} ((s - s_k)' \bar{A}_k \bar{A}_l (s - s_l)) \\ &= \frac{T_{kl}}{S_{kl}} d_{s_k} S_{kl} + 2 S_{kl} ((s - s_k)' \bar{A}_k \bar{A}_l d s + (s - s_l)' \bar{A}_k \bar{A}_l d s - (s - s_l)' \bar{A}_k \bar{A}_l \mathrm{d} s_k) \\ &= \frac{T_{kl}}{S_{kl}} d_{s_k} S_{kl} + 2 S_{kl} ((s - s_k) \bar{A}_k \bar{A}_l \bar{A}_{kl}^{-1} \bar{A}_k + (s - s_l) \bar{A}_k \bar{A}_l \bar{A}_{kl}^{-1} \bar{A}_k \\ &\quad - (s - s_l) \bar{A}_k \bar{A}_l)' d s_k \end{aligned} \tag{151}$$

In the general form of the potential energy integrals, only ρ has any dependence on s_k. The general differential is

$$d_{s_k} V_{kl} = \frac{V_{kl}}{S_{kl}} d_{s_k} S_{kl} + \frac{2}{\pi^{1/2}} S_{kl} \alpha^{-3/2} F_0' \left(\frac{\rho}{\alpha} \right) d\rho \tag{152}$$

For $V_{kl} = ER_{kl}^{ij}$:

$$d\rho = 2 s' \bar{J}_{ij} \bar{A}_{kl}^{-1} \bar{A}_k d s_k \tag{153}$$

For $V_{kl} = NA_{kl}^{it}$:

$$d\rho = 2 (s - t)' \bar{J}_{ij} \bar{A}_{kl}^{-1} \bar{A}_k d s_k \tag{154}$$

As shown in (4), the matrices A_k are actually products of triangular matrices $L_k L_k'$ and the derivatives reflect this. The derivatives we must find are

$$\frac{\partial O_{kl}}{\partial \mathrm{vech}\ L_k}$$

In doing so, we make frequent use of the construct

$$v' \bar{M} v = \mathrm{tr}[(v \cdot v) M] \tag{155}$$

to get rid of pesky Kronecker products.

The full differential of the overlap integral with respect to L_k is

$$d_k S_{kl} = 2^{3n} (d_k (p_{kl}^{3/2}) \exp(\gamma) + p_{kl}^{3/2} \exp(\gamma) \mathrm{d}_k \gamma) \tag{156}$$

We will examine first the part containing the determinants:

$$\begin{aligned}
d_k(p_{kl})^{3/2} &= \frac{3}{2}(p_{kl})^{1/2} d_k(p_{kl}) \\
&= \frac{3}{2}(p_{kl})^{1/2} \frac{1}{|A_{kl}|^2} \left(\frac{1}{2} |A_{kl}||A_l|^{1/2} |A_k|^{1/2} \mathrm{tr}[A_k^{-1} d_k(A_k)] \right. \\
&\quad \left. - |A_{kl}||A_l|^{1/2} |A_k|^{1/2} \mathrm{tr}[A_{kl}^{-1} d_k(A_k)] \right) \\
&= \frac{3}{2}(p_{kl})^{3/2} (\mathrm{tr}[L_k' A_k^{-1} \mathrm{d}L_k] - 2\mathrm{tr}[L_k' A_{kl}^{-1} dL_k]) \\
&= \frac{3}{2}(p_{kl})^{3/2} (\mathrm{vech}\,[A_k^{-1} L_k]' - 2\,\mathrm{vech}[A_{kl}^{-1} L_k]') \mathrm{vech}\,[dL_k]. \qquad (157)
\end{aligned}$$

Next, we need to find the differential of γ,

$$\begin{aligned}
d_k\gamma &= -d_k(s_k' \bar{A}_k s_k) + d_k(e' \bar{A}_{kl}^{-1} e) \\
&= -s_k' d(\bar{A}_k) s_k + 2e' \bar{A}_{kl}^{-1} d(\bar{A}_k) s_k - e' \bar{A}_{kl}^{-1} d(\bar{A}_k) \bar{A}_{kl}^{-1} e \\
&= -\mathrm{tr}[(s_k \cdot s_k) dA_k] + 2\mathrm{tr}[(s_k \cdot s) dA_k] - \mathrm{tr}[(s \cdot s) dA_k] \\
&= (2\mathrm{vech}\,[(s_k \cdot s) L_k] + 2\mathrm{vech}\,[(s_k \cdot s)' L_k] \\
&\quad - 2\mathrm{vech}\,[(s_k \cdot s_k) L_k] - 2\mathrm{vech}\,[(s \cdot s) L_k])' \mathrm{vech}\,[dL_k] \qquad (158)
\end{aligned}$$

The differential of the kinetic energy integral is

$$d_k T_{kl} = \frac{T_{kl}}{S_{kl}} d_k S_{kl} + S_{kl}(2d_k((s - s_k)' \bar{A}_k \bar{A}_l (s - s_l)) + 3d_k \mathrm{tr}[A_k A_{kl}^{-1} A_l]) \qquad (159)$$

Let us look at the first term in the parentheses:

$$\begin{aligned}
&d_k((s-s_k)' \bar{A}_k \bar{A}_l (s-s_l)) = d_k(s-s_k)' \bar{A}_k \bar{A}_l (s-s_l) \\
&\quad + (s-s_k)' d_k(\bar{A}_k) \bar{A}_l (s-s_l) + (s-s_k)' \bar{A}_k \bar{A}_l d_k(s-s_l) \\
&= -s' d_k(\bar{A}_k) \bar{A}_{kl}^{-1} \bar{A}_k \bar{A}_l (s-s_l) + s_k' d_k(\bar{A}_k) \bar{A}_{kl}^{-1} \bar{A}_k \bar{A}_l (s-s_l) \\
&\quad + (s-s_k)' d_k(\bar{A}_k) \bar{A}_l (s-s_l) - (s-s_k)' \bar{A}_k \bar{A}_l \bar{A}_{kl}^{-1} d_k(\bar{A}_k) s \\
&\quad + (s-s_k)' \bar{A}_k \bar{A}_l \bar{A}_{kl}^{-1} d_k(\bar{A}_k) s_k \\
&= -\mathrm{tr}[A_{kl}^{-1} A_k A_l ((s-s_l) \cdot s) d_k(A_k)] + \mathrm{tr}[A_{kl}^{-1} A_k A_l ((s-s_l) \cdot s_k) d_k(A_k)] \\
&\quad + \mathrm{tr}[A_l ((s-s_l) \cdot (s-s_k)) d_k(A_k)] - \mathrm{tr}[(s \cdot (s-s_k)) A_k A_l A_{kl}^{-1} d_k(A_k)] \\
&\quad + \mathrm{tr}[(s_k \cdot (s-s_k)) A_k A_l A_{kl}^{-1} d_k(A_k)] \\
&= (-\mathrm{vech}\,[A_{kl}^{-1} A_k A_l ((s-s_l) \cdot s) L_k] - \mathrm{vech}\,[(A_{kl}^{-1} A_k A_l ((s-s_l) \cdot s))' L_k] \\
&\quad + \mathrm{vech}\,[A_{kl}^{-1} A_k A_l ((s-s_l) \cdot s_k) L_k] + \mathrm{vech}\,[(A_{kl}^{-1} A_k A_l ((s-s_l) \cdot s_k))' L_k] \\
&\quad + \mathrm{vech}\,[A_l ((s-s_l) \cdot (s-s_k)) L_k] + \mathrm{vech}\,[(A_l ((s-s_l) \cdot (s-s_k)))' L_k] \\
&\quad - \mathrm{vech}\,[(s \cdot (s-s_k)) A_k A_l A_{kl}^{-1} L_k] - \mathrm{vech}\,[((s \cdot (s-s_k)) A_k A_l A_{kl}^{-1})' L_k] \\
&\quad + \mathrm{vech}\,[(s_k \cdot (s-s_k)) A_k A_l A_{kl}^{-1} L_k] \\
&\quad + \mathrm{vech}\,[((s_k \cdot (s-s_k)) A_k A_l A_{kl}^{-1})' L_k])' \mathrm{vech}\,[dL_k] \qquad (160)
\end{aligned}$$

And finally let's look at the second term in the parentheses:

$$\begin{aligned} d_k\mathrm{tr}[A_kA_{kl}^{-1}A_l] &= \mathrm{tr}[d_kA_kA_{kl}^{-1}A_l] - \mathrm{tr}[A_kA_{kl}^{-1}d_kA_kA_{kl}^{-1}A_l] \\ &= \mathrm{tr}[A_{kl}^{-1}A_ld_kA_k] - \mathrm{tr}[A_{kl}^{-1}A_lA_kA_{kl}^{-1}\mathrm{d}_kA_k] \\ &= (\mathrm{vech}[(A_{kl}^{-1}A_l + A_lA_{kl}^{-1})L_k] \\ &\quad - \mathrm{vech}[(A_{kl}^{-1}A_lA_kA_{kl}^{-1} + A_{kl}^{-1}A_kA_lA_{kl}^{-1})L_k])'\mathrm{vech}[\mathrm{d}L_k] \end{aligned} \tag{161}$$

For evaluation of the derivatives of the potential terms, we will refer again to the general form of the potential energy and the definitions of ρ and α. The general form of the differential is

$$d_kV_{kl} = \frac{V_{kl}}{S_{kl}}d_kS_{kl} - \frac{1}{\pi^{1/2}}S_{kl}\alpha^{-3/2}F_0\left(\frac{\rho}{\alpha}\right)d\alpha + \frac{2}{\pi^{1/2}}S_{kl}\alpha^{-1/2}F_0'\left(\frac{\rho}{\alpha}\right)\left(\frac{\alpha d_k\rho - \rho d_k\alpha}{\alpha^2}\right) \tag{162}$$

So in both cases it is only nessesary to find the derivatives of ρ and α. For $V_{kl} = ER_{kl}^{ij}$ we have

$$\begin{aligned} d_k\rho &= 2s'\bar{\boldsymbol{J}}_{ij}d_ks = 2s'\bar{\boldsymbol{J}}_{ij}\mathrm{d}_k(\bar{\boldsymbol{A}}_{kl}^{-1})e + 2s'\bar{\boldsymbol{J}}_{ij}\bar{\boldsymbol{A}}_{kl}^{-1}d_k(\bar{\boldsymbol{A}}_k)s_k \\ &= -2s'\bar{\boldsymbol{J}}_{ij}\bar{\boldsymbol{A}}_{kl}^{-1}d_k(\bar{\boldsymbol{A}}_k)s + 2s'\bar{\boldsymbol{J}}_{ij}\bar{\boldsymbol{A}}_{kl}^{-1}d_k(\bar{\boldsymbol{A}}_k)s_k \\ &= -2\mathrm{tr}[(s\cdot s)J_{ij}A_{kl}^{-1}d_k(A_k)] + 2\mathrm{tr}[(s_k\cdot s)J_{ij}A_{kl}^{-1}d_k(A_k)] \\ &= (2\mathrm{vech}\ [(s_k\cdot s)J_{ij}A_{kl}^{-1}L_k] + 2\mathrm{vech}\ [((s_k\cdot s)J_{ij}A_{kl}^{-1})'L_k] \\ &\quad - 2\mathrm{vech}[((s\cdot s)J_{ij}A_{kl}^{-1} + A_{kl}^{-1}J_{ij}(s.s))L_k])'\mathrm{vech}[dL_k] \end{aligned} \tag{163}$$

and

$$\begin{aligned} d_k\alpha &= d_k\mathrm{tr}[J_{ij}A_{kl}^{-1}] = -\mathrm{tr}[J_{ij}A_{kl}^{-1}d_k(A_k)A_{kl}^{-1}] \\ &= -\mathrm{tr}[A_{kl}^{-1}J_{ij}A_{kl}^{-1}d_k(A_k)] = -2\mathrm{vech}\ [A_{kl}^{-1}J_{ij}A_{kl}^{-1}L_k]'\mathrm{vech}[dL_k] \end{aligned} \tag{164}$$

So, for $V_{kl} = NA_{kl}^{it}$ we obtain

$$\begin{aligned} d_k\rho &= 2(s-t)'\bar{\boldsymbol{J}}_{ii}d_ks = 2(s-t)'\bar{\boldsymbol{J}}_{ii}\mathrm{d}_k(\bar{\boldsymbol{A}}_{kl}^{-1})e + 2(s-t)'\bar{\boldsymbol{J}}_{ii}\bar{\boldsymbol{A}}_{kl}^{-1}d_k(\bar{\boldsymbol{A}}_k)s_k \\ &= -2(s-t)'\bar{\boldsymbol{J}}_{ii}\bar{\boldsymbol{A}}_{kl}^{-1}d_k(\bar{\boldsymbol{A}}_k)s + 2(s-t)'\bar{\boldsymbol{J}}_{ii}\bar{\boldsymbol{A}}_{kl}^{-1}d_k(\bar{\boldsymbol{A}}_k)s_k \\ &= -2\mathrm{tr}[(s\cdot(s-t))J_{ii}A_{kl}^{-1}d_k(A_k)] + 2\mathrm{tr}[(s_k\cdot(s-t))J_{ii}A_{kl}^{-1}d_k(A_k)] \\ &= (2\mathrm{vech}\ [(s_k\cdot(s-t))J_{ii}A_{kl}^{-1}L_k] + 2\mathrm{vech}\ [((s_k\cdot(s-t))J_{ii}A_{kl}^{-1})'L_k] \\ &\quad - 2\mathrm{vech}\ [(s\cdot(s-t))J_{ii}A_{kl}^{-1}L_k] - 2\mathrm{vech}\ [((s\cdot(s-t))J_{ii}A_{kl}^{-1})'L_k])'\mathrm{vech}[dL_k] \end{aligned} \tag{165}$$

and

$$d_k\alpha = d_k\mathrm{tr}[J_{ii}A_{kl}^{-1}] = -\mathrm{tr}[J_{ii}A_{kl}^{-1}\mathrm{d}_k(A_k)A_{kl}^{-1}]$$
$$= -\mathrm{tr}[A_{kl}^{-1}J_{ii}A_{kl}^{-1}d_k(A_k)] = -2\mathrm{vech}\ [A_{kl}^{-1}J_{ii}A_{kl}^{-1}L_k]'\mathrm{vech}\ [dL_k] \qquad (166)$$

which completes the gradient derivation.

5. *Non-Born–Oppenheimer Energy Gradients*

Again, these gradients differ only in the kinetic energy differential from the Born–Oppenheimer form:

$$d_{s_k}T_{kl} = \frac{T_{kl}}{S_{kl}}d_{s_k}S_{kl} + 4S_{kl}((s-s_k)\bar{A}_k\bar{M}\bar{A}_l\bar{A}_{kl}^{-1}\bar{A}_k$$
$$+ (s-s_l)\bar{A}_k\bar{M}\bar{A}_l\bar{A}_{kl}^{-1}\bar{A}_k - (s-s_l)\bar{A}_k\bar{M}\bar{A}_l)'ds_k \qquad (167)$$

The L_k differential of the kinetic energy integral, with M instead of $\frac{1}{2}$, is

$$d_kT_{kl} = \frac{T_{kl}}{S_{kl}}d_kS_{kl} + S_{kl}(4d_k((s-s_k)'\bar{A}_k\bar{M}\bar{A}_l(s-s_l)) + 6d_k\mathrm{tr}[MA_lA_{kl}^{-1}A_k]) \qquad (168)$$

Let us look at the first term in the parentheses:

$$d_k((s-s_k)'\bar{A}_k\bar{A}_l(s-s_l)) = (-\mathrm{vech}[A_{kl}^{-1}A_kMA_l((s-s_l)\cdot s)L_k]$$
$$- \mathrm{vech}[(A_{kl}^{-1}A_kMA_l((s-s_l)\cdot s))'L_k]$$
$$+ \mathrm{vech}[A_{kl}^{-1}A_kMA_l((s-s_l)\cdot s_k)L_k] + \mathrm{vech}[(A_{kl}^{-1}A_kMA_l((s-s_l)\cdot s_k))'L_k]$$
$$+ \mathrm{vech}[MA_l((s-s_l)\cdot(s-s_k))L_k] + \mathrm{vech}[(MA_l((s-s_l)\cdot(s-s_k)))'L_k]$$
$$- \mathrm{vech}[(s\cdot(s-s_k))A_kMA_lA_{kl}^{-1}L_k] - \mathrm{vech}[((s\cdot(s-s_k))A_kMA_lA_{kl}^{-1})'L_k]$$
$$+ \mathrm{vech}[(s_k\cdot(s-s_k))A_kMA_lA_{kl}^{-1}L_k]$$
$$+ \mathrm{vech}[((s_k\cdot(s-s_k))A_kMA_lA_{kl}^{-1})'L_k])'\mathrm{vech}[dL_k] \qquad (169)$$

And finally let's look at the second term in the parentheses:

$$d_k\mathrm{tr}[MA_lA_{kl}^{-1}A_k] = \mathrm{tr}[d_kA_kA_{kl}^{-1}MA_l] - \mathrm{tr}[A_kA_{kl}^{-1}\mathrm{d}_kA_kA_{kl}^{-1}MA_l]$$
$$= \mathrm{tr}[A_{kl}^{-1}MA_ld_kA_k] - \mathrm{tr}[A_{kl}^{-1}MA_lA_kA_{kl}^{-1}\mathrm{d}_kA_k]$$
$$= (2\mathrm{vech}[A_{kl}^{-1}MA_lL_k] - 2\mathrm{vech}[A_{kl}^{-1}MA_lA_kA_{kl}^{-1}L_k])'\mathrm{vech}[\mathrm{d}L_k] \qquad (170)$$

6. Energy Gradients for the Electric Field Term

The differential of the electric field term follows easily from the definition of s:

$$d_{s_k} FT^z_{kl} = \frac{FT^z_{kl}}{S_{kl}} d_{s_k} S_{kl} + S_{kl}(o' E_{ii} \bar{A}^{-1}_{kl} \bar{A}_k ds_k) \tag{171}$$

For the L_k differential, the general form is

$$d_k FT^z_{kl} = \frac{FT^z_{kl}}{S_{kl}} d_k S_{kl} + S_{kl} d(o' E_{ii} s) \tag{172}$$

The only new derivative needed is the leftmost term on the left-hand side:

$$\begin{aligned} d(o' E_{ii} s) &= o' E_{ii} ds \\ &= o' E_{ii} d(\bar{A}^{-1}_{kl} e) \\ &= -o' E_{ii} (\bar{A}^{-1}_{kl} d\bar{A}_k \bar{A}^{-1}_{kl} e - \bar{A}^{-1}_{kl} de) \\ &= -o' E_{ii} (\bar{A}^{-1}_{kl} d\bar{A}_k s - \bar{A}^{-1}_{kl} \mathrm{d}\bar{A}_k s_k) \end{aligned} \tag{173}$$

At this point it is useful to define two vectors: $v1 = E_{ii} o$ and $v2 = s_k - s$. Using these, the last equation becomes

$$\begin{aligned} d(o' E_{ii} s) &= \mathrm{tr}[(v2 \cdot v1) A^{-1}_{kl} dA_k] \\ &= \mathrm{tr}[(v2 \cdot v1) A^{-1}_{kl} (dL_k L'_k + L_k dL'_k)] \\ &= \mathrm{tr}[L'_k (v2 \cdot v1) A^{-1}_{kl} dL_k] + \mathrm{tr}[(v2 \cdot v1) A^{-1}_{kl} L_k dL'_k)] \\ &= \mathrm{tr}[L'_k (v2 \cdot v1) A^{-1}_{kl} dL_k] + \mathrm{tr}[L'_k A^{-1}_{kl} (v2 \cdot v1)' dL_k)] \\ &= (\mathrm{vech}[A^{-1}_{kl} (v2 \cdot v1)' L_k] + \mathrm{vech}[(v2 \cdot v1) A^{-1}_{kl} L_k])' \mathrm{vech}[dL_k] \end{aligned} \tag{174}$$

7. Geometry Optimization Gradients

When looking at Born-Oppenhiemer calculations, we find that additional gradients are needed in geometry optimization. The gradient of the energy with respect to nuclear coordinates is given by

$$\nabla_{r_N} E = \frac{1}{c' S c} \left(\frac{\partial \mathrm{vech} H}{\partial r'_N} \right)' (\mathrm{vech}[2cc' - \mathrm{diag} cc']). \tag{175}$$

The derivatives needed are

$$\frac{\partial \mathrm{vech}\ H}{\partial r'_N} = \frac{\partial \mathrm{vech}\ NR}{\partial r'_N} - \frac{\partial \mathrm{vech}\ NA}{\partial r'_N} \tag{176}$$

where NA stands for nuclear attraction energy and NR stands for nuclear repulsion energy. Thus we need to calculate the elements:

$$\frac{\partial NA_{kl}}{\partial r'_N} = \sum_{i=1}^{n} \sum_{t=1}^{N} \frac{\partial NA_{kl}^{it}}{\partial r'_N} \tag{177}$$

and

$$\frac{\partial NR_{kl}}{\partial r'_N} = \sum_{u>t} \frac{\partial NR_{kl}^{tu}}{\partial r'_N} \tag{178}$$

For the NA integral we cannot differentiate directly with respect to r_N, so we must differentiate with respect to t, since

$$\frac{\partial NA_{kl}^{it}}{\partial r'_{N_t}} = \frac{\partial NA_{kl}^{it}}{\partial t'} \tag{179}$$

where r_{N_t} is the three-element partition of r_N pertaining to the tth nucleus and it is understood that we include only the three nonzero elements of the vector $\partial NA_{kl}^{it}/\partial t'$ in the equality.

This is then easily differentiated using the chain rule:

$$dNA_{kl}^{it} = -\frac{4}{\pi^{1/2}} S_{kl} (\mathrm{tr}[J_{ii} A_{kl}])^{-3/2} F_0' \left(\frac{(s-t)' \bar{J}_{ii} (s-t)}{\mathrm{tr}[J_{ii} A_{kl}]} \right) (s-t)' \bar{J}_{ii} \, dt \tag{180}$$

The elements of NR can be easily differentiated directly on r_N:

$$dNR_{kl}^{tu} = \frac{Z_t Z_u S_{kl}}{(r'_N \bar{J}_{tu} r_N)^{3/2}} r'_N \bar{J}_{tu} d(r_N) \tag{181}$$

8. *Expectation Values*

For non-Born–Oppenheimer calculations we have no molecular structure, per se, since all particles including nuclei are treated quantum mechanically. Thus the only information we can obtain about the structure of these molecules are the expectation values of distances between particles (r_{ij}) and powers of these distances. We present below the integrals needed in the calculation of these expectation values.

In order to find the r_{ij} integrals, we use the substitution

$$\frac{1}{r_{ij}} = \frac{2}{\pi^{1/2}} \int_0^\infty \exp(-u^2 r_{ij}^2) \, du \tag{182}$$

Thus for the integral we have

$$\langle g_k|r_{ij}|g_l\rangle = r_{ij}^{kl}$$

$$r_{ij}^{kl} = \langle g_k|\frac{r_{ij}^2}{r_{ij}}|g_l\rangle$$

$$r_{ij}^{kl} = \frac{2}{(\pi)^{1/2}}\frac{\partial}{\partial a}\int_0^\infty du\frac{-1}{u^2}\langle g_k|\exp(-au^2r'\bar{J}_{ij}r)|g_l\rangle|_{a=1} \tag{183}$$

Expanding the basis functions, we have

$$r_{ij}^{kl} = \frac{2}{(\pi)^{1/2}}\exp(-\gamma - s'\bar{A}_{kl}s)\frac{\partial}{\partial a}\int_{-\infty}^\infty dr\int_0^\infty du\,\frac{1}{u^2}\exp[-r'(\bar{A}_{kl}+au^2\bar{J}_{ij})r+2s'\bar{A}_{kl}r]|_{a=1} \tag{184}$$

Applying the formula

$$\int_{-\infty}^\infty \exp(-r'Mr + iz'r)\ dr = \pi^{n/2}|M|^{-1/2}\exp(-z'M^{-1}z/4) \tag{185}$$

where M is an $n \times n$ matrix, and r and z are vectors of length n, we find the integral over r:

$$\begin{aligned} r_{ij}^{kl} = {} & \frac{2}{(\pi)^{1/2}}\exp(-\gamma - s'\bar{A}_{kl}s)\frac{\partial}{\partial a}\int_0^\infty du\frac{\pi^{3n/2}}{u^2}|\bar{A}_{kl}+au^2\bar{J}_{ij}|^{-1/2} \\ & \times \exp[s'\bar{A}_{kl}(\bar{A}_{kl}+au^2\bar{J}_{ij})^{-1}\bar{A}_{kl}s + 2s'\bar{A}_{kl}r]|_{a=1} \end{aligned} \tag{186}$$

which, upon further simplification, becomes

$$r_{ij}^{kl} = S_{kl}\frac{2}{(\pi)^{1/2}}\frac{\partial}{\partial a}\int_0^\infty du\ \frac{1}{u^2}\left(1+au^2\mathrm{tr}[J_{ij}A_{kl}^{-1}]\right)^{-3/2}\exp\left(\frac{-au^2s'\bar{J}_{ij}s}{1+au^2\mathrm{tr}[J_{ij}A_{kl}^{-1}]}\right)\Bigg|_{a=1} \tag{187}$$

At this point we define $\alpha = \mathrm{tr}[J_{ij}A_{kl}^{-1}]$ and $\beta = s'\bar{J}_{ij}s$. Using these definitions and differentiating with respect to a, we have

$$r_{ij}^{kl} = S_{kl}\frac{2}{(\pi)^{1/2}}\int_0^\infty du\left[\frac{3\alpha}{2}(1+au^2\alpha)^{-5/2}+\beta(1+au^2\alpha)^{-7/2}\right]\exp\left(\frac{-au^2\beta}{1+au^2\alpha)}\right)\Bigg|_{a=1} \tag{188}$$

These two integrals may be solved by clever substitution and use of the integrals:

$$\int_0^1 \exp(-bt^2)\,dt = \frac{1}{2}\left(\frac{\pi}{b}\right)^{1/2}\mathrm{erf}(b^{1/2}) \tag{189}$$

$$\int_0^1 t^2 \exp(-bt^2)\,dt = \frac{1}{4}\left(\frac{\pi}{b^3}\right)^{1/2}\mathrm{erf}(b^{1/2}) - \frac{1}{2b}\exp(-b) \tag{190}$$

and

$$\int_0^1 t^4 \exp(-bt^2)\,dt = \frac{3}{8}\left(\frac{\pi}{b^5}\right)^{1/2}\mathrm{erf}(b^{1/2}) - \frac{1}{2b}\exp(-b) - \frac{3}{4b^2}\exp(-b) \tag{191}$$

After simplification, we have

$$r_{ij}^{kl} = S_{kl}\left\{\left(\frac{1\alpha}{2\beta^{1/2}} - \beta^{1/2}\right)\mathrm{erf}\left[\left(\frac{\beta}{\alpha}\right)^{1/2}\right]\right\} - \left(\frac{\alpha}{\pi}\right)^{1/2}\exp\left(-\frac{\beta}{\alpha}\right) \tag{192}$$

The integrals for $(r^2)_{ij}^{kl}$ may be obtained as a subset of the above integral by setting it up as

$$(r^2)_{ij}^{kl} = \frac{\partial}{\partial u}\langle g_k|\exp(-ur'\bar{\boldsymbol{J}}_{ij}r)|g_l\rangle|_{u=0} \tag{193}$$

With relative ease we obtain

$$(r^2)_{ij}^{kl} = S_{kl}\left(\frac{3}{2}\alpha - \beta\right) \tag{194}$$

9. *One-Particle Densities*

Here we will present the formulae needed for calculating the reduced one-particle density matrices from the floating correlated Gaussians used in this work. The first-order density matrix for wave function $\Psi(\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_n)$ for particle 1 is defined as

$$P(\mathbf{r}_1, \mathbf{r}_1') = \int \Psi^*(\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_n)\Psi(\mathbf{r}_1', \mathbf{r}_2, \ldots, \mathbf{r}_n)d\mathbf{r}_2 \cdots d\mathbf{r}_n \tag{195}$$

Since we are using the expansion

$$\Psi(\mathbf{r}) = \sum_{k=1}^{M} c_k \hat{\boldsymbol{Y}}\phi_k(\mathbf{r}) \tag{196}$$

we need to only find the matrix elements:

$$P_{kl}(\mathbf{r}_1, \mathbf{r}'_1) = \int \phi_k(\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_n)\phi_l(\mathbf{r}'_1, \mathbf{r}_2, \ldots, \mathbf{r}_n)d\mathbf{r}_2 \cdots d\mathbf{r}_n \tag{197}$$

for basis functions ϕ_k and ϕ_l.

We will follow the derivations presented in the work by Poshusta and Kinghorn [104] for the single-center Gaussians. The present work differs from theirs in that we use multicenter Gaussians to shift density away from the origin of the coordinate system. A definition used in the integral formulas is

$$\bar{\mathbf{M}} = \mathbf{M} \otimes \mathbf{I}_3$$

where $\mathbf{M}$ is a matrix and $\mathbf{I}_3$ is the 3×3 identity. As in the previous work [104] and for the convenience of the presentation, we define the following augmented coordinate vectors:

$$\mathbf{r}^+ = \begin{pmatrix} \mathbf{r}'_1 \\ \mathbf{r} \end{pmatrix} = \begin{pmatrix} \mathbf{u} \\ \mathbf{v} \end{pmatrix} \tag{198}$$

where

$$\mathbf{u} = \begin{pmatrix} \mathbf{r}'_1 \\ \mathbf{r}_1 \end{pmatrix}, \qquad \mathbf{v} = \begin{pmatrix} \mathbf{r}_2 \\ \mathbf{r}_3 \\ \cdots \\ \mathbf{r}_n \end{pmatrix} \tag{199}$$

These vectors allow us to separate the coordinates $\mathbf{r}_1$ and $\mathbf{r}'_1$ from the coordinates over which the integration is performed ($\mathbf{v}$) in the density calculation. Likewise, the vector of the shifts needs to be augmented to match the dimension of $\mathbf{r}^+$:

$$\mathbf{s}_k^+ = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \mathbf{s}_k \end{pmatrix} \tag{200}$$

We will also need

$$\bar{\mathbf{A}}_k^+ = \begin{pmatrix} \bar{0} & \bar{0} \\ \bar{0} & \bar{\mathbf{A}}_k \end{pmatrix} \tag{201}$$

and

$$\bar{\prod} = \begin{pmatrix} \bar{0} & \bar{1} & \bar{0} \\ \bar{1} & \bar{0} & \bar{0} \\ \bar{0} & \bar{0} & \bar{\mathbf{I}}_{n-1} \end{pmatrix} \tag{202}$$

where $\bar{\mathbf{I}}_{n-1}$ is the $(n-1) \times (n-1)$ identity matrix. Using the above definitions, we can formally represent $\phi_k(\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_n)$ and $\phi_k(\mathbf{r}'_1, \mathbf{r}_2, \ldots, \mathbf{r}_n)$ as

$$\phi_k(\mathbf{r}_1, \mathbf{r}_2, \cdots \mathbf{r}_n) = \exp\left[-(\mathbf{r}^+ - \mathbf{s}_k^+)' \bar{\mathbf{A}}_k^+ (\mathbf{r}^+ - \mathbf{s}_k^+)\right] \tag{203}$$

and

$$\phi_l(\mathbf{r}'_1, \mathbf{r}_2, \cdots \mathbf{r}_n) = \exp\left[-\left(\mathbf{r}^+ - \bar{\prod}\mathbf{s}_l^+\right)' \left(\bar{\prod}\bar{\mathbf{A}}_l^+ \bar{\prod}\right)\left(\mathbf{r}^+ - \bar{\prod}\mathbf{s}_l^+\right)\right] \tag{204}$$

The product of the Gaussians in the integrand of P_{kl} may then be written (using the abbreviations $\bar{\Pi}\bar{\mathbf{A}}_l^+\bar{\Pi} = \bar{\mathbf{A}}_l^{+p}$ and $\bar{\Pi}\mathbf{s}_l^+ = \mathbf{s}_l^{+p}$) as

$$\begin{aligned} &\phi_k(\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_n)\phi_l(\mathbf{r}'_1, \mathbf{r}_2, \ldots, \mathbf{r}_n) \\ &\quad = \exp[-\mathbf{s}'_k\bar{\mathbf{A}}_k\mathbf{s}_k - \mathbf{s}'_l\bar{\mathbf{A}}_l\mathbf{s}_l - \mathbf{r}^{+\prime}(\bar{\mathbf{A}}_k^+ + \bar{\mathbf{A}}_l^{+p})\mathbf{r}^+ + 2\mathbf{r}^{+\prime}(\bar{\mathbf{A}}_k^+\mathbf{s}_k^+ + \bar{\mathbf{A}}_l^{+p}\mathbf{s}_l^{+p})] \end{aligned} \tag{205}$$

Using the definitions

$$\bar{\mathbf{Z}}^{kl} = \bar{\mathbf{A}}_k^+ + \bar{\mathbf{A}}_l^{+p} \tag{206}$$

and

$$\mathbf{z}^+ = \bar{\mathbf{A}}_k^+\mathbf{s}_k^+ + \bar{\mathbf{A}}_l^{+p}\mathbf{s}_l^{+p} \tag{207}$$

we can rewriting the above expression as

$$\phi_k(\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_n)\phi_l(\mathbf{r}'_1, \mathbf{r}_2, \ldots, \mathbf{r}_n) = \exp[-\mathbf{s}'_k\bar{\mathbf{A}}_k\mathbf{s}_k - \mathbf{s}'_l\bar{\mathbf{A}}_l\mathbf{s}_l - \mathbf{r}^{+\prime}\bar{\mathbf{Z}}^{kl}\mathbf{r}^+ + 2\mathbf{r}^{+\prime}\mathbf{z}^+] \tag{208}$$

The last two terms in the exponent above can be expanded using the augmented matrices defined above:

$$\begin{aligned} &- \mathbf{r}^{+\prime}(\bar{\mathbf{A}}_k^+ + \bar{\mathbf{A}}_l^{+p})\mathbf{r}^+ + 2\mathbf{r}^{+\prime}(\bar{\mathbf{A}}_k^+\mathbf{s}_k^+ + \bar{\mathbf{A}}_l^{+p}\mathbf{s}_l^{+p}) \\ &= \exp\left[-(\mathbf{uv})\begin{pmatrix} \bar{\mathbf{Z}}_u^{kl} & \bar{\mathbf{Z}}_{uv}^{kl} \\ \bar{\mathbf{Z}}_{vu}^{kl} & \bar{\mathbf{Z}}_v^{kl} \end{pmatrix}\begin{pmatrix} \mathbf{u} \\ \mathbf{v} \end{pmatrix} + 2(\mathbf{uv})\begin{pmatrix} \mathbf{z}^{+u} \\ \mathbf{z}^{+v} \end{pmatrix}\right] \end{aligned} \tag{209}$$

where $\bar{\mathbf{Z}}_u^{kl}$ is a 6×6 matrix , $\bar{\mathbf{Z}}_v^{kl}$ is a $(n-1) \times (n-1)$ matrix, and $\mathbf{z}^{+u}$ is a vector of length 6, while $\mathbf{z}^{+v}$ is a vector of length $(3n-3)$. Expanding the above expression again we get

$$\exp[-\mathbf{u}'\bar{\mathbf{Z}}_u^{kl}\mathbf{u} - \mathbf{v}'\bar{\mathbf{Z}}_v^{kl}\mathbf{v} - 2\mathbf{u}'\bar{\mathbf{Z}}_{uv}^{kl}\mathbf{v} + 2\mathbf{u}'\mathbf{z}^{+u} + 2\mathbf{v}'\mathbf{z}^{+v}] \tag{210}$$

Now, after substituting the above expression into the original expression, we get

$$\begin{aligned}&\phi_k(\mathbf{r}_1,\mathbf{r}_2,\ldots,\mathbf{r}_n)\phi_l(\mathbf{r}'_1,\mathbf{r}_2,\ldots,\mathbf{r}_n)\\ &\quad=\exp[-\mathbf{s}'_k\bar{\mathbf{A}}_k\mathbf{s}_k-\mathbf{s}'_l\bar{\mathbf{A}}_l\mathbf{s}_l-\mathbf{u}'\bar{\mathbf{Z}}_u^{kl}\mathbf{u}+2\mathbf{u}'\mathbf{z}^{+u}]\times\exp[-\mathbf{v}'\bar{\mathbf{Z}}_v^{kl}\mathbf{v}-2\mathbf{u}'\bar{\mathbf{Z}}_{uv}^{kl}\mathbf{v}+2\mathbf{v}'\mathbf{z}^{+v}]\end{aligned} \tag{211}$$

We can now integrate over the variables in the $\mathbf{v}$ vector and we get

$$\begin{aligned}P_{kl}(\mathbf{r}_1,\mathbf{r}'_1)=\pi^{3(n-1)/2}|\mathbf{Z}_v^{kl}|^{-3/2}&\exp[-\mathbf{s}'_k\bar{\mathbf{A}}_k\mathbf{s}_k-\mathbf{s}'_l\bar{\mathbf{A}}_l\mathbf{s}_l-\mathbf{u}'\bar{\mathbf{Z}}_u^{kl}\mathbf{u}+2\mathbf{u}'\mathbf{z}^{+u}]\\ &\times\exp[-\mathbf{u}'\bar{\mathbf{Z}}_{uv}^{kl}(\bar{\mathbf{Z}}_v^{kl})^{-1}\bar{\mathbf{Z}}_{vu}^{kl}\mathbf{u}-2\mathbf{u}'\bar{\mathbf{Z}}_{uv}^{kl}(\bar{\mathbf{Z}}_v^{kl})^{-1}\mathbf{z}^{+v}+\mathbf{z}^{+v\prime}(\bar{\mathbf{Z}}_v^{kl})^{-1}\mathbf{z}^{+v}]\end{aligned} \tag{212}$$

By grouping all the terms that do not include $\mathbf{u}$, we can define the following ''augmented'' overlap matrix element:

$$S_{kl}^{+}=\pi^{3(n-1)/2}|\mathbf{Z}_v^{kl}|^{-3/2}\exp[-\mathbf{s}'_k\bar{\mathbf{A}}_k\mathbf{s}_k-\mathbf{s}'_l\bar{\mathbf{A}}_l\mathbf{s}_l+\mathbf{z}^{+v\prime}(\bar{\mathbf{Z}}_v^{kl})^{-1}\mathbf{z}^{+v}] \tag{213}$$

Using this quantity, we get

$$P_{kl}(\mathbf{r}_1,\mathbf{r}'_1)=S_{kl}^{+}\times\exp[-\mathbf{u}'(\bar{\mathbf{Z}}_u^{kl}-\bar{\mathbf{Z}}_{uv}^{kl}(\bar{\mathbf{Z}}_v^{kl})^{-1}\bar{\mathbf{Z}}_{vu}^{kl})\mathbf{u}+2\mathbf{u}'(\mathbf{z}^{+u}-\bar{\mathbf{Z}}_{uv}^{kl}(\bar{\mathbf{Z}}_v^{kl})^{-1}\mathbf{z}^{+v})] \tag{214}$$

Now we define two new matrices:

$$\bar{\mathbf{W}}=\bar{\mathbf{Z}}_u^{kl}-\bar{\mathbf{Z}}_{uv}^{kl}(\bar{\mathbf{Z}}_v^{kl})^{-1}\bar{\mathbf{Z}}_{vu}^{kl} \tag{215}$$

and

$$\mathbf{d}=\mathbf{z}^{+u}-\bar{\mathbf{Z}}_{uv}^{kl}(\bar{\mathbf{Z}}_v^{kl})^{-1}\mathbf{z}^{+v} \tag{216}$$

and we arrive at the following final expression for $P_{kl}(\mathbf{r}_1,\mathbf{r}'_1)$:

$$P_{kl}(\mathbf{r}_1,\mathbf{r}'_1)=S_{kl}^{+}\times\exp[-\mathbf{u}'\bar{\mathbf{W}}\mathbf{u}+2\mathbf{u}'\mathbf{d}] \tag{217}$$

Thus the density matrix elements take on the familiar form of a Gaussian with a shifted center. We can now find the one-particle density matrix element by setting $\mathbf{r}_1 = \mathbf{r}_1' \equiv \begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix}$:

$$D_{kl}(\mathbf{r_1}) = P_{kl}(\mathbf{r}_1, \mathbf{r}_1) = S_{kl}^{+} \times \exp\left[-(W_{11} + W_{22} + 2W_{12})(x_1^2 + y_1^2 + z_1^2) + 2\begin{pmatrix} \mathbf{r}_1 \\ \mathbf{r}_1 \end{pmatrix}' \mathbf{d}\right] \tag{218}$$

This last expression was used in the code to calculate the densities. As derived, the formulae apply to the density of pseudoparticle 1 (and all the particles identical to it). As mentioned, the density of other particles may be obtained by permuting the desired particle to position 1 in the basis functions.

B. Spherically Symmetric Molecules

In the standard language of chemistry we describe a molecular geometry with bond lengths and bond angles. In the true quantum-mechanical picture of a molecule, though, at best we can know average values of the distances between particles in our system. We will call these two viewpoints the chemical and the physical picture, respectively, following the designations of Monkhorst [55, 56]. It should be noted that we basically arrive at the chemical picture by making the adiabatic and BO approximations in the physical picture. In the chemical picture, we may solve the electronic problem for a fixed molecular geometry—that is, for fixed bond lengths—or we may optimize the nuclear geometry in the average field of the electrons and find fictitious equilibrium bond lengths. This bond length, for a linear molecule, corresponds to the very bottom of the potential energy curve; likewise, for polyatomic molecules, it corresponds to the very bottoms of potential energy surfaces (PES's). At no point does the molecule ever reside in this point in the PES. We know that for every normal mode of vibration of the molecule $E_{v_i} = h\nu\left(v_i + \frac{1}{2}\right)$ the molecule has to be at least in the $v_i = 0$ state, and this state lies above the troughs of the PES by $\frac{1}{2}h\nu_i$.

The above is a well-understood problem of the BO approximation, and the most accurate calculations of molecular properties takes this into account. A less well understood difference between the physical and chemical pictures is that, in the physical picture, the ground state of any molecule is spherically symmetric. This may be understood in the chemical picture by noting that when all degrees of freedom are taken into account, the total wave function contains the nuclear vibratrional and rotational wave functions as well as the electronic wave function:

$$\Psi(r_N, r_e) = \psi(r_e)\psi_{vib}(r_N)\psi_{rot}(r_N) \tag{219}$$

In the ground state, the nuclear rotation state is the $J = 0$ state, which is spherically symmetric. This total wave function may be thought of in analogy to the hydrogen atom as a radial part (which is the expansion in explicitly correlated Gaussians) multiplied by the rotational function, which for the ground state is simply a constant. This point may be understood by the classical picture of the molecule rotating about the center of mass. In this picture, the time-averaged wave function will be spherically symmetric, because the molecule over time has no preferred direction in space.

This presents a problem when we discuss the dipole moment of a polar heteronuclear diatomic molecule, AX, where X will be the more electronegative. In the chemical picture, it is quite common to say that in the ground state the molecule lies along some axis and that it has a definite dipole moment. In the physical picture, we say that the molecule has no measurable dipole moment in the ground state.

When we apply an electric field to the molecule, though, say along the Z axis, the molecule will begin to align itself with the field. We call this the rotational polarization of the molecule. In this case, though, when the Hamiltonian includes a field directed along the z axis, the rotational quantum number J is no longer a good quantum number, and we now have only eigenstates of $\hat{J}_z$. This molecule is no longer spherically symmetric, and thus we may measure the dipole moment. All that is necessary for the molecule to rotationally polarize is for the energy lowering due to the interaction with the field to be greater than a rotational excitation.

Another problem comes in examining the polarizability. In the physical picture, the spherically symmetric molecule, just like an atom, has isotropic polarizability. In the chemical picture, for a diatomic molecule we have two unique polarizabilities: (1) α_{zz} and α_{xx} in the internal coordinate system or (2) $\alpha_{ZZ} = \frac{1}{3}(\alpha_{xx} + \alpha_{xx} + \alpha_{zz})$ (isotropic polarizability) and $\Delta\alpha = \alpha_{zz} - \alpha_{xx}$ [polarizability anisotropy(PA)] in the laboratory coordinate system. The latter are the values that are measured in experiments. The isotropic polarizability of the physical picture is comparable to the BO-isotropic polarizability at very low field strengths but is comparable to α_{zz} at higher field strengths. We cannot extract the PA from the physical picture, because in this ideal model of dilute gas-phase molecules the polarizability is isotropic.

One possible solution to this problem may be had from examining the time development of a molecule on the electric field. Before the field is applied, the molecule is spherically symmetric, and no PA exists. As soon as the field is applied, the molecule will distort and lose the symmetry. Early in this distortion process, though, especially if the field is small, the molecule is still symmetric. If one can calculate the polarizability at this state, and then calculate the polarizability of the state when the molecule is fully aligned with the field, then these two values can give the PA.

C. Good Quantum Numbers and the Symmetry Properties of the Basis Functions

In the absence of an electric field, the non-BO Hamiltonian commutes with the square of the angular momentum operator, $[\hat{H}, \hat{J}^2] = 0$, and so the eigenfunctions of the Hamiltonian also have to be eigenfunctions of $\hat{J}^2$. This condition is met, for example, by functions such as

$$g_k = \prod_{i=1}^{n} x_i^{p_{ki}} y_i^{l_{ki}} z_i^{m_{ki}} \exp(-\alpha_i^k (r_i)^2) \prod_{j>i}^{n} \exp(-\beta_{ij}^k (r_i - r_j)^2) \tag{220}$$

In the presence of an electric field, this commutation is no longer true: $[\hat{H}, \hat{J}^2] \neq 0$, though if the field is applied along the z axis, then the Hamiltonian does commute with the z component of angular momentum: $[\hat{H}, \hat{J}_z] = 0$. In this case the eigenfunctions of the Hamiltonian must also be eigenfunctions of $\hat{J}_z$. This is met by functions such as those used in this work:

$$g_k = \prod_{i=1}^{n} \exp(-\alpha_i^k (r_i - R_i^k)^2) \prod_{j>i}^{n} \exp(-\beta_{ij}^k (r_i - r_j)^2) \tag{221}$$

so long as the only components of R_i^k that are nonzero are the z components. Thus, the wave functions used in this work are not rigorously correct for the field-free calculations, but the energy error should be only contamination by the first few rotational states and has, in practice, never been a problem.

D. The Finite Field Method

The response of the energy of a molecule to a static electric field along the z axis may be written as

$$E(f) = E_o - \frac{\partial E}{\partial f} f - \frac{1}{2!} \frac{\partial E^2}{\partial^2 f} f^2 - \frac{1}{3!} \frac{\partial E^3}{\partial^3 f} f^3 - \frac{1}{4!} \frac{\partial E^4}{\partial^4 f} f^4 \cdots \tag{222}$$

We define these derivatives as the usual response properties:

$$E(f) = E_o - \mu f - \frac{1}{2!} \alpha f^2 - \frac{1}{3!} \beta f^3 - \frac{1}{4!} \gamma f^4 \cdots \tag{223}$$

where α is called the polarizability, β is called the first hyperpolarizability, and γ is called the second hyperpolarizability. To calculate the response of a molecule to a static electric field only requires knowledge of these properties.

We use a very simple method for calculating the response of molecules to electric fields. We calculate the wave function and energy for molecules in

electric fields of various strengths and plot this information. This plot is fitted to a polynomial, and the ith-order properties are extracted from the ith-order coefficients of the fit.

In the physical picture of a molecule, the non-BO energy at the field strength f is equal to the energy at the field $-f$, because as the direction of the field changes the the molecule orients itself with the field so as to have as low a potential energy as possible. Thus, for any ground-state spherically symmetric molecule the energy is an even function of the field, and if it is approximated by a polynomial, only even powers need to be used. This obviously results in a zero dipole, as well as zero-valued odd-ordered properties in general for any system.

An alternative approach is to apply stronger fields and only use energies calculated for positive field strengths in generating the polynomial fit. In this case the energy is a function of both odd and even powers in the polynomial fit. We will show that the dipole moments derived from our non-BO calculations with the procedure that uses only positive fields and polynomial fits with both even and odd powers match very well the experimental results. Thus in the present work we will show results obtained using interpolations with even- and odd-power polynomials. Methods other than the finite field method exist where the noise level in the numerical derivatives is smaller (such as the Romberg method), but such methods still do not allow calculation of odd-ordered properties in the non-BO model.

E. Vibrationally Corrected Electronic Values

The conventional approach used to describe the response of a molecule to a static electric field is either to perform pure electronic BO calculations or to perform calculations where the BO values are corrected for vibrational and rotational (thermal) motion of the nuclei. The vibrationally corrected polarizabilities usually do an excellent job of correcting the errors inherent in the pure electronic BO values. Bishop has written several excellent reviews on this topic [78–80].

F. Isotopomers of H_2

Please note that much of this material is reported in the recent article by Cafiero and Adamowicz [66].

There have been no previous direct non-BO studies of the response of H_2 and its isotopomers to electric fields. The ground-state dipole moment of HD has been determined experimentally by Nelson and Tabisz [81] to be 0.000345 a.u. There have been several theoretical studies of the dipole moment of HD, all within the BO approximation but including adiabatic corrections. The calculated values by Wolniewicz, 0.000329 [83], Ford and Browne, 0.000326 [82], and Thorson et al., 0.000334 [84], all agree well with the experimental value, although they are all about 5% too small. This is an extremely difficult experiment to carry out, and because all theoretical studies agree on the value, it

may be that the experimental work is slightly inaccurate. There have been several studies of the vibrational corrections to the pure electronic BO values. Adamowicz [85] used numerical MCSCF electronic functions and numerical evaluation of the vibrational equations to obtain the following ground-state (ground electronic, $v = 0$, $J = 0$) values of α_{zz}: 6.623 (H_2), 6.569 (HD), and 6.509 (D_2) a.u. Also, he found the following for γ_{zzzz}: 1088(H_2), 1074(HD), and 1052(D_2) a.u. Bishop et al. [86] employed the sum over states method and found the static limit value of α_{zz} to be 6.397 a.u. In a different study, Bishop et al. [87] found γ_{zzzz} to be 1099 including electronic and vibrational contributions. There are no non-BO investigations of the values of β_{zzz}.

The wave functions used in this work were built from noncorrelated products of simple gaussians describing the electronic and nuclear wave functions. Initial guesses for the exponential coefficients (squares of the diagonal elements of L_k) and functional centers (s_k) were generated quasi-randomly. The exponential coefficients were randomly distributed around the value 4.0 with a variance of 1. Since the electric field in our calculations was applied along the z direction, the initial guesses for the functional centers were distributed randomly along the z axis around 1.4 bohr for the H-nucleus and one of the electrons, and around 0.0 bohrs for the second electron, with a variance of 0.015. The initial guess wave functions were optimized with respect to all of the elements of L_k and s_k using analytical gradients. The process of optimization builds the required correlation into the wave function by filling in the off-diagonal elements of L_k.

The correct permutational symmetry was implemented into the wave functions by projection onto irreducible representations of the total symmetry group S_2^e for heteronuclear species and $S_2^e \otimes S_2^H$ for homonuclear species, where e refers to electron exchange and H refers to nuclear exchange. The irreducible representations chosen were singlets in all cases.

The energy values obtained for the H_2 isotopomer series were all converged until the energy change with respect to variation was in the subpicohartree range, and the squared gradient norm for the total gradient vector was at most 10^{-14}. The energy curve was fitted with a fourth-order polynomial.

The values for the dipoles, polarizabilities, and hyperpolarizabilities of the H_2 series were obtained using (a) a 16-term basis with a fourfold symmetry projection for the homonuclear species and (b) a 32-term basis with a twofold symmetry projection for the heteronuclear species. These different expansion lengths were used so that when combined with the symmetry projections the resulting wave functions were of about the same quality, and the properties calculated would be comparable. A crude analysis shows that basis set size for an n particle system must scale as κ^n, where κ is a constant. In our previous work [64, 65] we used a 244-term wave function for the five-internal-particle system LiH to obtain experimental quality results. This gives a value of $\kappa^5 = 3^5 \approx 244$. Applying this analysis to HD, which has three internal particles,

TABLE XI
Values for the Zero-Field Energies, Dipole Moments (μ), Polarizabilities (α), Hyperpolarizabilities (β), and Second Hyperpolarizabilities (γ) for the Non-BO H_2 Isotopomer Series[a]

	$\langle H \rangle$	μ	α	(β)	γ
H_2	−1.153736345	1.00×10^{-8}	6.74	(0.0360)	1062
HD	−1.156234289	3.27×10^{-4}	6.67	(0.0306)	1038
HT	−1.157152576	4.37×10^{-4}	6.65	(0.0186)	1028
D_2	−1.159178760	9.00×10^{-9}	6.59	(0.0360)	1009
DT	−1.160312051	1.09×10^{-4}	6.56	(0.0312)	999
T_2	−1.161561149	9.00×10^{-9}	6.52	(0.0360)	989

[a]All values are in atomic units.

we get $\kappa^3 = 3^3 = 27$. Thus the basis set size of 32 terms should be sufficient to give results of similar quality to what was obtained in the LiH case. For H_2, where we have additional basis functions generated by the proton exchange symmetry projection, the basis may be smaller. The calculated values are presented in Table XI.

Allowing only positive and rather significant electric field values permits the calculation of the dipole moment of the heteronuclear species. Applying the same approach to homonuclear species (H_2, D_2, and T_2) should give the dipoles identically equal to zero. In our calculations, these actually come out to 10^{-8}. This small noise that entered our calculations was due in part to the previously mentioned fact that the zero-field wave function we use is not an eigenfunction of $\hat{J}^2$ as it should be. The level of noise introduced is negligible, because 10^{-8} is four orders of magnitude smaller than the size of the dipole moments for the heteronuclear species.

The dipole moment obtained for HD, 3.27×10^{-4} a.u., is very close to the values obtained by Wolniewicz (3.29×10^{-4}) [83] and by Ford and Brown (3.26×10^{-4}) [82]. All of these values are about 5% lower than the experimentally measured value. The ratio of the dipole moment of HD to those of HT and DT may be predicted based on electronegativity arguments to be 0.75 and 3.0. The values obtained here fulfill this prediction exactly.

We see the expected trend in polarizabilities and second hyperpolarizabilities going down the series according to total mass. H_2 has the highest polarizability, 6.74 a.u., because the light, very quantum-mechanical protons delocalize more easily in the electric field than the heavier D and T nuclei. Also, the electrons in H_2 are slightly farther away from the nuclei than in the heavier isotopomers and are more polarizable. Indeed when we add weight to the nuclei, as in HD, the polarizability goes down to 6.67 a.u. Another large jump in polarizability (6.65 a.u. to 6.59 a.u) comes between HT, the last isotopomer containing a

proton, and D_2, where both nuclei are heavier. The second hyperpolarizability for H_2 agrees well with the previous values. Our value of 1062 a.u. is smaller than Bishop's 1099 and Adamowicz's 1088 a.u. The corresponding values for the other isotopomers are also smaller than the values by Adamowicz. The value of β for the homonuclear species should be identically zero. We obtain for all three homonuclear species a value of 0.0360 a.u. This very small value is consistent across all of these species, and it is larger than zero due to the very sensitive nature of this third-order property to small numerical inaccuracies. Although we show the values of β for all the heteronuclear isotopomers in Table XI (in parentheses), they are very small and too close to the numerical noise to be trusted.

G. LiH and LiD

There have been several recent attempts to find the nuclear corrections to the LiH dipole moment. Papadopoulos et al. [88] used the perturbation theory to calculate the corrections, and Tachikawa and Osamura [57] used the Dynamic Extended Molecular Orbital method to try to calculate the nonadiabatic result directly. Results for these methods are reported in Table XII. In all cases, the calculated values are outside of the range of the experimental results [89, 90], also reported in Table XII.

For LiH and LiD, 244-term non-BO wave functions were variationally optimized. The initial guess for the LiH non-BO wave function was built by multiplying a 244-term BO wave function expanded in a basis of explicitly correlated functions by Gaussians for the H nucleus centered at and around (in all three dimensions) a point separated from the origin by the equilibrium distance of 3.015 bohr along the direction of the electric field. Thus the centers

TABLE XII
Experimental (expt.) and Theoretical (calc.) Dipole Moments (μ) for LiH and LiD from the Literature[a]

m	μ
LiH	
[88], calc.	2.317
[57], calc.	2.389
[89], expt.	2.3145
[90], expt.	2.314 (0.001)
LiD	
[57], calc.	2.392
[90], expt.	2.309 (0.001)

[a]All values are in atomic units. Values given in parentheses are experimental uncertainty.

TABLE XIII
Expectation Values for Zero-Field Energies, Virial Coefficients (η), Dipole Moments (μ), and Static Polarizabilites (α) for Non-BO LiH/D for Various Expansion Lengths (m)[a]

m	$\langle H \rangle$	η	μ	α
LiH				
24	−8.0423294	1.000000	2.4047	24.42
64	−8.0592988	1.000000	2.3394	28.48
104	−8.0619267	1.000000	2.3261	29.41
144	−8.0629324	0.999999	2.3149	29.54
244	−8.0636331	0.999999	2.3140	29.57
Experimental			2.314 ± 0.001[90]	
LiD				
244	−8.0650331	1.000000	2.3088	
Experimental			2.309 ± 0.001[90]	

[a]All values are in atomic units.

corresponding to the hydrogen nucleus were scattered from about 2.9 to about 3.1 bohr. The lithium nucleus was, of course, placed at the origin of the internal coordinate system. The functional centers corresponding to the electrons were located primarily on the two nuclei, with two electrons at the origin (about 0.0 ± 0.001 bohr in all three directions) and two electrons near the H nucleus (about 3.05 ± 0.06 bohr) per basis set. This reflects the strong ionic character in the lithium–hydrogen bond. The LiD non-BO wave function was optimized starting from the converged LiH wave function. Wave functions of various smaller expansion lengths were optimized for LiH alone. Table XIII shows the convergence properties of the dipole moment for these basis sets. It can be seen that the calculated value of the dipole converges and reaches a value near that of experiment, 2.314 a.u., as larger basis sets are used. The reported results for all functions were converged to the point where the squared norm of the total gradient was at least on the order 10^{-8}, and the energy changed at most in the ninth decimal place. At this point the dipole moment was converged to seven decimal places, which is more than experimental accuracy. The total variational energy (also in Table XIII) for our 244-term wave function is −8.0636331 hartree. The most accurate non-BO energy calculated before was that of Scheu et al. [120] equal to −8.0661557 hartree. Thus our energy value for this basis set is in error by only about 0.0025 hartree. Despite this small error, we deem the obtained convergence of the dipole moment value with the basis set size quite satisfactory, and it is highly unlikely that further enlargement of the basis can change this value by an amount close to the uncertainty level of the experiment where the LiH/LiD dipole moments were measured.

The optimized basis functions show strong correlation between the nuclei, between the nuclei and the electrons, and between the electrons. The centers describing the H/D nucleus show a spread of values from about 2.9 bohr to 3.1 bohr, with the expectation value for the internuclear distance being 3.063 bohr for LiH and 3.052 bohr for LiD. These values are in good agreement with the value 3.061 bohr for LiH obtained by Scheu et al. [120] in their non-BO calculations and are, as expected, longer than the BO 'equilibrium' value. Our values are believed to be much more accurate than those of Ref. 57, which are 3.119937 bohr and 3.104819 bohr for LiH and LiD, respectively.

The value of the dipole moment of LiH obtained in this work, 2.3140 a.u., is essentially identical to the experimental value, 2.314 ± 0.001 [90]. Our calculations simulate experiment more closely than any previous calculations. The results also provide validation of the perturbation approach of Ref. 88, since the perturbation result, 2.317 a.u., is very close to our value. At the same time, our results are much more accurate than those of Ref. 57, the only other ''direct'' calculation of the LiH dipole moment. The value of the dipole moment of LiD, 2.3088, is also of good accuracy, compared to the experimental result, 2.309 ± 0.001 [90]. Again, our result is much more accurate than that of Ref. 57.

There have been a few recent studies of the corrections due to nuclear motion to the electronic diagonal polarizability (α_{zz}) of LiH. Bishop et al. [92] calculated vibrational and rotational contributions to the polarizability. They found for the ground state ($v = 0$, the state studied here) that the vibrational contribution is 0.923 a.u. Papadopoulos et al. [88] use the perturbation method to find a corrected value of 28.93 a.u. including a vibrational component of 1.7 a.u. Jonsson et al. [91] used cubic response functions to find a corrected value for α_{zz} of 28.26 a.u., including a vibrational contribution of 1.37 a.u. In all cases, the vibrational contribution is approximately 3% of the total polarizability.

The results for the non-BO diagonal polarizability are shown in Table XIII. Our best—and, as it seems, well-converged—value of α, 29.57 a.u., calculated with a 244-term wave function, is slightly larger than the previously obtained ''corrected'' electronic values, 28.93 and 28.26 a.u. [88,91]. It is believed that the non-BO correction to the polarizability will be positive and on the order of less than 1 a.u. [92], but it is not possible to say if the difference between the value obtained in this work and the previous values for polarizability are due to this effect or to other effects, such as the basis set incompleteness in the BO calculations. An effective way of testing this would be to perform BO calculations of the electronic and vibrational components of polarizability using an extended, well-optimized set of explicitly correlated Gaussian functions. This type of calculation is outside of our current research interests and is quite expensive. It may become a possibility in the future. As such, we would like the polarizability value of 29.57 a.u. obtained in this work to serve as a standard for non-BO polarizability of LiH.

VII. THE USE OF SHIFTED GAUSSIANS IN NON-BO CALCULATIONS ON POLYATOMIC MOLECULES

We have shown that in order to calculate a non-Born–Oppenheimer wave function we must use basis functions that are eigenvalues of the total angular momentum. We have used such functions in the calculation of atoms and diatomic molecules and have shown that we are working on extending these basis sets to triatomic molecules. We may also perform non-Born–Oppenheimer calculations of reasonable accuracy using basis functions which are not rigorous eigenfunctions of angular momentum but which are complete sets and may approximate such eigenfunctions.

Non-Born–Oppenheimer wave functions calculated in this way look more like their Born–Oppenheimer counterparts in the smaller basis set limits, and thus a good starting guess for these may be taken from Born–Oppenheimer calculations in the same basis. Thus we calculate the electronic part first (this requires much fewer basis functions than does a full non-Born–Oppenhimer calculation) and then form the total basis function by multiplying each electronic portion by a guess for the nuclear portion:

$$\psi_T(r_{total}) = \psi_N(r_{nuc})\psi_e(r_{elec}) \tag{224}$$

We will present below a short description of some Born–Oppenheimer calculations we have done on this basis, followed by examples of triatomic non-Born–Oppenheimer calculations on this basis.

A. Born–Oppenheimer Calculations in a Basis of Explicitly Correlated Gaussians

FSECGs have been used in BO calculations for some time, but always with quite small systems. It has been said [94] that the bottleneck for the application of ECGs to larger systems ($n > 4$) is the large amount of time spent in optimization of the wave functions. We have removed a large part of the bottleneck in these calculations by implementing analytical gradients in the optimization and parallelizing the entire code. The use of the analytical gradients speeds up optimization, because the optimization routine needs to make only one function call to calculate the energy and the total gradient vector, rather than several function calls to calculate finite-difference gradients. The analytical gradients are also more accurate and lead to faster optimization paths.

B. Test Calculations on H_3 and H_3^+

Please note that much of this work may be found in a previous article [70].

A new upper bound for the BO energy of the ground state of H_3^+ was recently obtained by Cencek et al. [48] using explicitly correlated Gaussians. Below we

TABLE XIV
Expectation Values of the Hamiltonian, $\langle H \rangle$, Virial Coefficients, η, and Squared Gradient Norms, $\| grad \|^2$, For the Ground State of H_3^+ ($R1 = R2 = R3 = 1.65$) for the 75-Term Wave Function

	Cencek et al. [48]	This Work
$\langle H \rangle$	−1.343834724	−1.343834853
η	0.9999977	0.9999978
$\| grad \|^2$	$O(10^{-8})$	$O(10^{-12})$

show how our calculations compare against those. The Young projection operator used was for the singlet state, and the D_{3h} point group symmetry was ensured by constraining the molecule to the xy plane and applying the operator:

$$\hat{P} = \hat{1} + \hat{C}_3 + \hat{C}_3^2 + \hat{\sigma}_1 + \hat{\sigma}_2 + \hat{\sigma}_3 \tag{225}$$

We did expect to gain a slight lowering of the energy due to our use of analytical gradients in our optimization as opposed to a numerical optimization used by Cencek et al., and we did. Table XIV shows the energy obtained by Cencek et al. for the 75-term wave function, as well as the squared gradient norm—that is, the sum of the squares of the gradients with respect to all the variational parameters. As can be seen, the wave function obtained by Cencek et al. [48] had considerable ''room'' for further optimization: The gradients with respect to most variational parameters were indeed quite small, but the gradients with respect to the off-diagonal exponential coefficients were somewhat larger and thus made the squared norm of the gradient overall larger. After several cycles of our unconstrained optimization in which we reoptimized all variational parameters, the energy for this basis set was lowered and the gradients with respect to all parameters, and thus the squared norm of the gradient, became much smaller. This final wave function is more tightly converged than that obtained before.

The formulas derived in Cencek et al. [49] involve a simplifying approximation of only including one correlation term per basis function; that is, only two electrons are correlated per function. In this work, all functions include correlation among all electrons. For two electrons there is no difference in the two forms of the basis, but for three or more electrons the formulas derived here should prove more efficient; that is, they should converge faster and with a smaller basis set size.

We tested a 76-term wave function for the system H_3, including permutational and point group symmetry. The initial guess for the nonlinear parameters in the ECGs were generated randomly using Matlab. The Young

TABLE XV
Expectation Values of the Hamiltonian, $\langle H \rangle$, Virial Coefficients, η, and Squared Gradient Norms, $\| grad \|^2$, for the Ground State of H_3 ($R1 = R2 = 1.75$) for the 76-Term wave Function (This Work) and the 100-Term Wave Function (Cencek and Rychlewski) [49]

	100-term [49]	76-term
$\langle H \rangle$	−1.658491	−1.658565
η	—	1.00012
$\| grad \|^2$	—	$O(10^{-13})$

projection operator used was a doublet, and the $D_{\infty h}$ point group symmetry was ensured by constraining the molecule to the z axis and applying the operator:

$$\hat{P} = \hat{1} - \hat{\sigma}_h \tag{226}$$

The energy obtained after optimization for the 76-term function was found to be lower than the energy obtained by Cencek and Rychlewski [49] for the 100-term wave function, illustrating the power of including correlation among all the electrons in each basis term (see Table XV).

C. Geometry Optimization

Please note that much of this work may be found in a previous article [69].

In the course of this research, we introduced the first geometry optimization via analytical gradients to be used in very-high-accuracy electronic structure calculations with FSECGs [69]. In this method we simultaneously optimize the nonlinear parameter of the basis functions and the molecular structure parameters. Simultaneous optimization of both types of parameter provides a unique path to the high accuracy of the calculation. The explicitly correlated Gaussians are a particularly interesting basis to use in geometry optimization since, by monitoring the magnitude of correlation coefficients (off-diagonal elements in the L_k matrix), we can see the dynamic correlations of pairs of electrons increasing or decreasing as the molecule undergoes changes in nuclear geometry. A simple example of this is found in the results on small hydrogen clusters presented here.

Sample calculations were carried out on H_2, H_3^+, and H_3. Geometry optimizations were carried out in internal coordinates. The projection operators used in the expansion (4) represented a singlet state for H_2 and H_3^+ and a doublet state for H_3. Starting geometries that were used are given in Table XVI. The initial wave functions were centered at the nuclei. For all of the initial functions the correlation parameters were set to zero (that is, the matrices A_k were

TABLE XVI
Starting Geometries (in parentheses), Energies, $\langle \hat{H} \rangle$, Virial Coefficients, η, Squared Gradient Norms, $\| \, grad \, \|^2$, and Optimized Geometries for the Hydrogen Clusters[a]

	$\langle \hat{H} \rangle$	η	$\| \, grad \, \|^2$	Geometry
H_2(R12 = 2.000)	−1.173092	1.000000	10^{-13}	R12 = 1.399
H_3^+	−1.334711	0.999999	10^{-13}	R12 = 1.650
(R12 = 1.400,				R13 = 1.650
R13 = R23 = 1.220)				R23 = 1.649
H_3(global minimum)	−1.673468	0.999999	10^{-13}	R12 = 1.400
(R12 = R23 = 1.750)				R23 = 6.442
H_3(saddle point)	−1.655734	1.000000	10^{-13}	R12 = 1.757
(R12 = R23 = 1.750)				R23 = 1.757

[a]All reported values are in atomic units.

diagonal). Expansion lengths used were $m = 16$ for H_2 and H_3^+ and $m = 64$ for H_3.

The nonlinear optimization is always sensitive to the initial guess of the parameters (nonlinear and geometrical). For each system, several initial guesses for the parameters were used, and we report the results of the calculations which produced the lowest energies. In all cases, equilibrium structures were found; additionally, the saddle point for H_3 was found using a constrained optimization. This point is defined in the work by Liu [98] as the minimum under the conditions of linearity and σ_h symmetry. We thus reduced the optimization to a single variable—that is, the bond length. It is interesting to note that for the H_3 saddle point the correlation coefficients (off-diagonal elements of A_k) were all of about the same order of magnitude (10^{-2}), while for the equilibrium point, where H_2 and H are separated, the magnitudes of the correlation coefficients between the electrons of H_2 were much larger (10^{-2}) than those between electrons on different subunits (10^{-5}). This reflects the expected result that there is strong correlation of electrons within the localized chemical bond in H_2 and much weaker correlation of the electrons that are spatially separated. In the case of the H_3 equilibrium structure, the optimization settled into the van der Waals' well corresponding to linear H_3 with the $H_2 \cdots H$ distance equal to 6.442 bohr. This structure is in good agreement with the potential surface minima for this system determined by Truhlar and Horowitz [99], Wu et al. [100], and Tang and Toennies [101]. Final energies, virial coefficients, squared gradient norms, and geometries are given in Table XVI.

We compared the van der Waals region obtained with our method to that which may be obtained with a fairly high-end standard method, CCSD(T)/cc-pVTZ, as implemented in *Gaussian* 98 [102]. The CCSD(T) energy for this system, though not variational, should provide a reasonable estimate of the

variational energy in this basis, because the single and double excitations are much more important than the triple excitations. The CCSD(T) energy was −1.6721738 hartree, about 1.3 millihartree higher than our result. The geometry obtained with CCSD(T) was $R12 = 1.403$ bohr and $R23 = 6.478$ bohr. The expected value of $R12$ is 1.401 bohr, which is slightly closer to our value of 1.400 bohr. The literature value of $R23$ is between 6.44 and 6.48 bohr. The relatively newer result of Tang and Toennies [101] predicts a value of 6.46 bohr, roughly between the values obtained using our method and CCSD(T).

D. Extension to Non-Born–Oppenheimer

We show here how we may take the information obtained above and use it as a starting point for non-Born–Oppenheimer calculations. The five-particle systems of non-Born–Oppenheimer H_3^+ and its isotopomers were transformed via separation of the center-of-mass Hamiltonian to four-pseudoparticle systems as described above. The resulting total position vector is

$$r = \begin{pmatrix} r_1 \\ r_2 \\ r_3 \\ r_4 \end{pmatrix} \tag{227}$$

For the illustrative calculations shown here, the spin-free wave functions, Ψ, for the H_3^+ isotopomers were obtained as 50-term expansions in a basis of FSECG's $g_k(r)$:

$$\Psi(r) = \sum_{k=1}^{50} c_k \hat{Y} g_k(r) \tag{228}$$

where $\hat{Y}$ is the total Young operator described above.

The FSECG basis function for four particles is

$$g_k(r) = \exp\{-(r - s_k)'[(L_k L_k') \otimes I_3](r - s_k)\} \tag{229}$$

where s_k is a 12 vector of 1''shifts'' that are variational parameters, and L_k is an upper triangular 4×4 matrix of variational exponential parameters.

The ground-state energy values and the wave functions for the considered systems were variationally optimized:

$$E = \min \frac{\langle \Psi | \hat{H} | \Psi \rangle}{\langle \Psi | \Psi \rangle} \tag{230}$$

Each wave function was optimized with respect to the parameters L_k, s_k, and c_k. This lead to $\frac{1}{2}n(n+1) + 3n + 1$ variational parameters per basis function (23 for

the H_3^+ systems). We began with an initial guess for the wave function built from the above Born–Oppenheimer calculations and then use analytical gradients in a truncated Newton-type algorithm to find the lowest value for the energy. The initial guesses for the nuclear portions of the basis functions were randomly generated with all the s_k centers lying in the *xy* plane. For small basis sets such as those used here, it has been found that placing functional centers off of the *xy* plane makes a negligible difference in the energy and structural expectation values. Because the basis set is increased and the planar optimization space becomes saturated, the placement of functions off of the *xy* plane would become more important, with the basis set limit featuring a full three-dimensional spherical symmetry for the ground state.

For the homonuclear (HON) species, the permutation–symmetry operator had the following form: $\hat{Y} = \hat{Y}_N(S_3) \otimes \hat{Y}_e(S_2)$, where $\hat{Y}_N(S_3)$ is a Young operator for the third-order symmetric group which permutes the nuclear coordinates and $\hat{Y}_e(S_2)$ is a Young operator for the second-order symmetric group which permutes the electronic coordinates. For the fermionic nuclei (H and T, spin = 1/2) the Young operators corresponded to doublet-type representations, while for the bosonic D nuclei we use operators that correspond to the totally symmetric representation. In all cases the electronic operator corresponded to a singlet representation.

For the heteronuclear (HEN) A_2B-type species, the symmetry operator was $\hat{Y} = \hat{Y}_N(S_2) \otimes \hat{Y}_e(S_2)$. In this case the nuclear operator was a singlet for H pairs and symmetric for D pairs. Finally for the HEN HDT-type isotopomer we had $\hat{Y} = \hat{Y}_e(S_2)$. Again in all cases the electron operators represented a singlet. For discussion of the construction of the operators, see, for example, the excellent work of Pauncz [73].

It is obvious that the projection operators for the different species have different numbers of terms in them. The HON species have 12 terms ($3! \times 2!$) while the A_2B-type species have four terms, and the HDT^+ isotopomer has only two terms. This results in different sizes of the spin-projected basis sets, and for this reason the properties obtained in this work are not precisely comparable between the A_3, A_2B, and ABC systems, although a very good idea of the trends may be obtained from the data in Table XVI. While all of the above are given in terms of the original particles, it should be noted that the permutations used in the internal particle basis functions are "pseudo"-permutations induced by the permutations on real particles.

The geometrical parameters reported in Table XVI include the distances from the particle at the origin of coordinates to the other two nuclei (r_1 and r_2) as well the distance between the two "loose" nuclei (r_{12}); also included are the squares of all of these distances. For the HON species, the quantity actually calculated is $\langle r_1 + r_2 + r_{12} \rangle$. The value reported in the table is this number divided by three. For the HEN A_2B species we calculate $\langle r_1 + r_2 \rangle$ and $\langle r_{12} \rangle$

TABLE XVII
Values for the Nonadiabatic Energies, Virial Coefficients (η), and Bond Lengths of DH_2^+ for Various Expansion Lengths[a]

m	$\langle H \rangle$	η	r_{ij}
50	−1.316992613	0.999999	$r_{DH} = 1.735$
			$r_{HH} = 1.746$
80	−1.318112939	0.999996	$r_{DH} = 1.731$
			$r_{HH} = 1.741$
118	−1.318845090	0.999999	$r_{DH} = 1.731$
			$r_{HH} = 1.740$
489	−1.321226255	0.999742	$r_{DH} = 1.724$
			$r_{HH} = 1.734$

[a]All values are in atomic units.

separately. Finally, for the HDT^+ isotopomer, we find each distance separately. The same patterns were used for the squares of the distances. Since the wave functions used were nonadiabatic and include both electronic and nuclear coordinates, the bond lengths calculated are not the usual equilibrium bond lengths, r_e (i.e., the very bottom of potential wells), obtained in conventional BO calculations, but rather the r_0 bond lengths (i.e., bond lengths that are more comparable to bond lengths obtained in the BO calculations by averaging the internuclear distances over the ground vibrational state of the system).

For one isotopomer, DH_2^+, two additional larger basis sets were optimized ($m = 80$, 118, 489) in order to determine the convergence of the structure parameters. This data are presented in Table XVII. As may be seen, the energies are not yet converged. Because there are no references for the fully nonadiabatic energy of DH_2^+, we may estimate it by adding the zero-point vibrational energy to the best Born–Oppenheimer energy obtained so far. These data may be found in the work by Jaquet et al. [103], and the value is −1.326672 hartree. Although our energies for DH_2^+ are above this value, the geometry is well known to converge before the energy. The difference in bond length for the DH bond decreases by 0.003 bohr, going from 50 basis functions to 80, and then decreases by less than 0.001 bohr, going from 80 to 118 basis functions. Finally, taking a very large step to 489 basis functions, the bond lengths are seen to decrease by less than 0.01 bohr. It may be noted also that this basis set produces an energy much closer to the estimated energy than the smaller basis sets, and so the geometry may be that much more reliable.

Assuming that any additional increase in basis set size will cause a decrease on the order 0.005 bohr or less, we may say that the geometries of the systems considered here obtained with 50 basis functions are fairly well optimized. The isotopic differences across the isotopomers seem to be consistant as well. It will be expected that bond lengths obtained from the HON species will be slightly

more accurate than those obtained from the HEN species due to the difference in the spin-projected basis set size. Likewise, the HDT^+ isotopomer will have the least accurate bond lengths.

E. Discussion

The commonly accepted equilibrium bond length, r_e, for H_3^+ is 1.650 bohr (see, for example, Cencek et al. [48] and references therein). This number corresponds to the very center of the potential energy well. In actual fact, molecules in their ground states reside above this trough in the zero-point vibrational state. The more anharmonic the well is, however, the more the bond length will be shifted to higher values. H_3^+ is a particularly floppy molecule, and so the r_0 bond length would be expected to be significantly larger than the r_e bond length. This is in fact the case, since the value we obtained in our nonadiabatic calculation is 1.748 bohr.

As the nuclei get heavier, the displacement of the molecule from its equilibrium structure in the zero-point vibration becomes smaller and the r_0 bond lengths should approach the r_e bond lengths. This is in fact the case, as we see from Table XVIII. T_3^+ is about 0.04 bohr closer to the expected well than H_3^+, which is about 0.1 bohr from the expected value. Another phenomenon seen in the data is that the uncertainty in the nuclear position, calculated as $\delta = (|\langle r\rangle^2 - \langle r^2\rangle|)^{1/2}$, gets progressively smaller for each HON isotopomer, going from 0.238 to 0.194 to 0.179 as we move from H_3^+ down to T_3^+.

An interesting point in the data presented here is that for the HON species and the A_2B HEN species it is impossible to determine from the r data alone if the molecule is linear or planar triangular. Even if the expectation values of the

TABLE XVIII
Values for the Nonadiabatic Expectation Values of the Ground-State Energies, Virial Coefficients (η), Interparticular Distances, and Squares of Interparticular Distances, for Some Isotopomers of H_3^+[a]

	$\langle H\rangle$	η	r_{ij}	r_{ij}^2
H_3^+	−1.314383574	0.999999	$r_{HH} = 1.748$	$r_{HH}^2 = 3.112$
D_3^+	−1.322718305	0.999999	$r_{DD} = 1.720$	$r_{DD}^2 = 2.996$
T_3^+	−1.326427799	0.999997	$r_{TT} = 1.707$	$r_{TT}^2 = 2.946$
DH_2^+	−1.316992613	0.999999	$r_{DH} = 1.735$	$r_{DH}^2 = 3.059$
			$r_{HH} = 1.746$	$r_{HH}^2 = 3.105$
HD_2^+	−1.319779541	0.999999	$r_{HD} = 1.734$	$r_{HD}^2 = 3.056$
			$r_{DD} = 1.719$	$r_{DD}^2 = 2.997$
HDT^+	−1.320907124	0.999999	$r_{TD} = 1.715$	$r_{DT}^2 = 2.976$
			$r_{TH} = 1.729$	$r_{DD}^2 = 3.034$
			$r_{HD} = 1.734$	$r_{DD}^2 = 3.053$

[a]All values are calculated for an optimized 50-term explicitly correlated Gaussian basis set and are in atomic units.

angles were calculated, we would find that due to indistinguishability of the nuclei the angles would come out equal in either configuration. The only way to predict from the nonadiabatic calculations the actual structure of the molecules and determine whether they are linear or planar triangular is to consider the average bond lengths calculated for the HDT^+ isotopomer. In HDT^+ the indistinguishability plays no role, and in this system only we find that the molecular geometry is a near-equilateral triangle, and not linear. Thus, just as in experiment, ''isotopic substitution'' is necessary to extract information about the molecular structure from the nonadiabatic calculations.

VIII. FUTURE WORK

At this stage we are at the very beginning of development, implementation, and application of methods for quantum-mechanical calculations of molecular systems without assuming the Born–Oppenheimer approximation. So far we have done several calculations of ground and excited states of small diatomic molecules, extending them beyond two-electron systems and some preliminary calculations on triatomic systems. In the non-BO works, we have used three different correlated Gaussian basis sets. The simplest one without r_{ij} premultipliers ($\phi_k = \exp[-\mathbf{r}'(A_k \otimes I_3)\mathbf{r}]$) was used in atomic calculations; the basis with premultipliers in the form of powers of r_1 ($\phi_k = r_1^{m_k} \exp[-\mathbf{r}'(A_k \otimes I_3)\mathbf{r}]$) was used in calculations for diatomic systems; and Gaussians with shifted centers ($\phi_k = \exp\left[-(\mathbf{r}-\mathbf{s})'(A_k \otimes I_3)(\mathbf{r}-\mathbf{s})\right]$) were used in non-BO calculations of diatomic molecules in the static electric field. The latter basis was also used in non-BO calculations of ground states of some simple triatomic systems (H_3, H_3^+ and their isotopomers).

At present our effort concentrates on the development of methods for non-BO calculations of excited states of molecules with three nuclei. Our aim is to match in such calculations the accuracy we have been getting for the diatomic systems. This development will open to us the possibility of studying highly vibrationally excited, charged, and neutral clusters of hydrogen and its isotopomers. H_3 and H_3^+ are among the most interesting cases in this cathegory. At present we consider two different approaches for such calculations. The first is based on implementation of correlated Gaussian basis functions with preexponential multipliers consisting of products of all three internuclear distances (r_1, r_2, and r_{12}) raised to even powers ($\phi_k = r_1^{m_k} r_2^{n_k} r_{12}^{l_k} \exp[-\mathbf{r}'(A_k \otimes I_3)\mathbf{r}]$). Such a basis should very effectively describe the coupled motions of electrons and nuclei in systems with three heavy, repelling particles. In the second approach we will use the correlated Gaussians with shifted centers ($\phi_k = \exp\left[-(\mathbf{r}-\mathbf{s})'(A_k \otimes I_3)(\mathbf{r}-\mathbf{s})\right]$). Since these functions are not rotationally invariant eigenfunctions, we will need to implement a projection procedure for separating functional manifolds corresponding to different quantum

numbers of the square of the total angular momentum operator $\hat{J}^2$. This may be done by including a penalty term in the Hamiltonian matrix ($a \times |\langle\phi_k|\hat{J}^2 - J(J-1)|\phi_k\rangle|$; a being a positive penalty factor) which will elevate the energies of functions corresponding to the rotation eigenvalues different from the one considered in the calculation (J) to higher energies effectively separating them away.

Another development that we will undertake in the near future is development of algorithms for non-BO calculations of molecules with π-electrons (the CH radical is an example of such a system). We also contemplate development of methods for describing systems where only the light nuclei (apart from electrons) are treated as quantum particles, and the other heavier nuclei are described either classically or by using a low-level approximation. This development would move us closer to cosidering the quantum dynamics of such reactions as inter- and intramolecular proton transfer.

Finally, the development of new algorithms in atomic and molecular non-BO calculations must be carried out in parallel with the development of the computer technology, particularly in parallel with the advances in parallel computing. An option that can be considered involves the so-called meta-computing. Meta-computing on a grid of distributed computing platforms connected via high-speed networks can revolutionize computational research by enabling hybrid computations that integrate multiple systems distributed over wide geographical locations [46]. The non-BO method developed in our work is well-suited for implementation on parallel computational platforms. In our laboratory we use parallel ''Beowulf'' clusters based on commercially available PC components and connected via a fast Ethernet switch. The clusters use a Unix (Linux) operational system, and the software parallelism is facilitated by MPI. For our non-BO calculations we found the Beowulf clusters to be very cost-effective since the calculations can be easily distributed over a network of processors and executed in parallel with quite little interprocessor communication and without the need to share a common operational memory. Utilization of massively parallel systems has given us momentum to proceed with the development of the non-BO method which we hope the computers of the future will allow to apply to larger systems more central to chemistry. Hence, while the applications presented in this chapter concern very small systems, the emphasis in the development we have carried out is placed on creating a general method that is applicable to molecular systems with an arbitrary number of electrons and nuclei.

References

1. M. Born and J. P. Oppenheimer, *Ann. Phys.* **84**, 457 (1927).
2. J. M. Cobes and R. Seiler, *Quantum Dynamics of Molecules*, R. G. Woolley, ed., Plenum Press, New York, 1980, p. 435.

3. M. Klein, A. Martinez, R. Seiler, and X. P. Wang, *Commun. Math. Phys.* **143**, 607 (1992).
4. A. Carington and R. A. Kennedy, *Gas Phase Ion Chemistry*, Vol. 3, M. T. Bowers, ed., Academic Press, New York, p. 393.
5. D. R. Yarkony, *J. Phys. Chem. A* **105**, 6277 (2001).
6. A. Toniolo, M. Ben-Nun, and T. J. Martinez, *J. Phys. Chem. A* **106**, 4679 (2002).
7. R. S. Friedman, T. Podzielinski, L. S. Cederbaum, et al., *J. Phys. Chem. A* **106**, 4320 (2002).
8. S. Mahapatra, *J. Chem. Phys.* **116**, 8817 (2002).
9. M. Brouard, P. O'Keeffe, and C. Vallance, *J. Phys. Chem. A* **106**, 3629 (2002).
10. M. Baer, A. M. Mebel, and R. Englman, *Chem. Phys. Lett.* **354** 243 (2002).
11. S. Matsika and D. R. Yarkony, *J. Phys. Chem. A* **106**, 2580 (2002).
12. S. Matsika and D. R. Yarkony, *J. Chem. Phys.* **116**, 2825 (2002).
13. S. Matsika and D. R. Yarkony, *J. Chem. Phys.* **115**, 5066 (2001).
14. S. Matsika and D. R. Yarkony, *J. Chem. Phys.* **115**, 2038 (2001).
15. N. Iordanova and S. Hammes-Schiffer, *J. Am. Chem. Soc.* **124**, 4848 (2002).
16. P. K. Agarwal, S. R. Billeter, and S. Hammes-Schiffer, *J. Phys. Chem. B* **106**, 3283 (2002).
17. P. K. Agarwal, S. R. Billeter, P. T. R. Rajagopalan, et al., *Proc. Natl. Acad. Sci. USA* **99**, 2794 (2002).
18. S. Hammes-Schiffer, *CHEMPHYSCHEM* **3**, 33 (2002).
19. S. R. Billeter, S. P. Webb, P. K. Agarwal, et al., *J. Am. Chem. Soc.* **123**, 11262 (2001).
20. S. Hammes-Schiffer and S. R. Billeter, *Int. Rev. Phys. Chem.* **20**, 591 (2001).
21. M. N. Kobrak and S. Hammes-Schiffer, *J. Phys. Chem. B* **105**, 10435 (2001).
22. I. Rostov and S. Hammes-Schiffer, *J. Chem. Phys.* **115**, 285 (2001).
23. T. Iordanov, S. R. Billeter, S. P. Webb, et al., *Chem. Phys. Lett.* **338**, 389 (2001).
24. N. Iordanova, H. Decornez, and S. Hammes-Schiffer, *J. Am. Chem. Soc.* **123**, 3723 (2001).
25. S. Hammes-Schiffer, *Accounts Chem. Res.* **34**, 273 (2001).
26. S. R. Billeter, S. P. Webb, T. Iordanov, P. K. Agarwal, and S. Hammes-Schiffer, *J. Chem. Phys.* **114**, 6925 (2001).
27. A. M. Frolov and V. H. Smith, *J. Chem. Phys.* **115**, 1187 (2001).
28. F. E. Harris, A. M. Frolov, and V. H. Smith, *J. Chem. Phys.* **119**, 8833 (2003).
29. A. M. Frolov and V. H. Smith, *J. Chem. Phys.* **119**, 3130 (2003).
30. A. M. Frolov, *J. Phys. B* **35**, L331 (2002).
31. A. M. Frolov, *Phys. Rev. A* **67**, 064501 (2003).
32. A. M. Frolov, *Phys. Rev. A* **69**, 062507 (2004).
33. A. M. Frolov, *J. Phys. B* **37**, 853 (2004).
34. A. M. Frolov, *Phys. Rev. A* **69**, 022505 (2004).
35. L. Adamowicz and A. J. Sadlej, *J. Chem. Phys.* **67**, 4398 (1977).
36. L. Adamowicz and A. J. Sadlej, *Chem. Phys. Lett.* **48**, 305 (1977).
37. L. Adamowicz, *J. Quant. Chem.* **13**, 265 (1978).
38. L. Adamowicz, *Acta Phys. Pol.* **A53**, 471 (1978).
39. L. Adamowicz and A. J. Sadlej, *J. Chem. Phys.* **69**, 3992 (1978).
40. L. Adamowicz and A. J. Sadlej, *Acta Phys. Pol. A* **54**, 73 (1978).
41. L. Adamowicz and A. J. Sadlej, *Chem. Phys. Lett.* **53**, 377 (1978).

42. K. Szalewicz, L. Adamowicz, and A. J. Sadlej, *Chem. Phys. Lett.* **61**, 548 (1979).
43. D. B. Kinghorn and L. Adamowicz, *J. Chem. Phys.* **106**, 4589 (1997).
44. K. R. Lykke, K. K. Murray, and W. C. Lineberger, *Phys. Rev. A* **43**, 6104 (1991).
45. W. F. Drake, *Nucl. Instr. Meth. Phys. Res. B* **31**, 7 (1988).
46. I. Forster and C. Kesselman, *The Grid: Blueprint for a New Computing Infrastructure*, Morgan-Kaufmann, San Francisco, 1999.
47. S. F. Boys, *Proc. R. Soc. London*, Ser. A **258**, 402 (1960).
48. W. Cencek, J. Komasa, and J. Rychlewski, *Chem. Phys. Lett.* **246**, 417 (1995).
49. W. Cencek and J. Rychlewski, *J. Chem. Phys.* **98**, 1252 (1993).
50. W. Cencek, J. Komasa, and J. Rychlewski, *Chem. Phys. Lett.* **304**, 293 (1999).
51. J. Komasa, *J. Chem. Phys.* **112**, 7075 (2000)
52. R. D. Poshusta, *Int. J. Quant. Chem.* **8**, 27 (1978).
53. Y. Suzuki and K. Varga, *Stochastic Variational Approach to Quantum Mechanical Few-Body Problems*, Springer-Verlag, Berlin, 1998.
54. N. C. Handy and A. M. Lee, *Chem. Phys. Lett.* **252**, 425 (1996).
55. H. J. Monkhorst, *Phys. Rev. A* **36**, 1544 (1987).
56. H. J. Monkhorst, *Int. J. Quantum Chem.* **72**, 281 (1999).
57. M. Tachikawa, K. Mori, K. Suzuki, and K. Iguchi, *Int. J. Quantum Chem.* **70**, 491 (1998).
58. H. Nakai, K. Sodeyama, and M. Hoshino, *Chem. Phys. Lett.* **345**, 118 (2001).
59. Y. Shigeta, H. Nagao, K. Nishikawa, and K. Yamaguchi, *J. Chem. Phys.* **111**, 6171 (1999).
60. D. B. Kinghorn and L. Adamowicz, *J. Chem. Phys.* **113**, 1203 (2000).
61. P. Kozlowski and L. Adamowicz, *J. Comp. Chem.* **13**, 100 (1992).
62. P. Kozlowski and L. Adamowicz, *J. Chem. Phys.* **96**, 9013 (1992).
63. D. B. Kinghorn and R. D. Poshusta, *Phys. Rev. A* **47**, 3671 (1996).
64. M. Cafiero and L. Adamowicz, *Phys. Rev Lett.* **88**, 33002 (2002).
65. M. Cafiero and L. Adamowicz, *J. Chem. Phys.* **116**, 5557 (2002).
66. M. Cafiero and L. Adamowicz, *Phys. Rev Lett.* in press.
67. V. I. Korobov, *Phys. Rev. A* **63**, 044501 (2001).
68. D. M. Bishop and S. A. Solunac, *Phys. Rev. Lett.* **55**, 1986 (1985).
69. M. Cafiero and L. Adamowicz, *Chem. Phys. Lett.* **335**, 404 (2001).
70. M. Cafiero and L. Adamowicz, *Int. J. Quantum Chem.* **82**, 151 (2001).
71. G. Schatz and M. Ratner, *Quantum Mechanics in Chemistry*, Dover, Mineola, NY, 2002.
72. F. A. Cotton, *Chemical Applications of Group Theory*, Wiley, New York, 1990.
73. R. Pauncz, *Spin Eigenfunctions*, Plenum, New York, 1979
74. M. Hamermesh, *Group Theory and its Application to Physical Problems*, Dover, Mineola, NY, 1989.
75. K. Singer, *Proc. R. Soc. London*, Ser. A **258**, 412 (1960).
76. S. G. Nash, *SIAM J. Num. Anal.* **21**, 770 (1984).
77. M. Snir, S. Otto, S. Huss-Lederman, D. Walker, and J. Dongarra, *MPI—The Complete Reference*, Vol. 1, The MIT Press, Cambridge, MA, 1998.
78. D. M. Bishop, *Rev. Mod. Phys.* **62**, 343 (1990).
79. D. M. Bishop, *Adv. Quant. Chem.* **25**, 1 (1994).

80. D. M. Bishop, *Adv. Chem. Phys.* **104**, 1 (1998).
81. J. B. Nelson and G. C. Tabisz, *Phys. Rev. A* **28**, 2157 (1983).
82. A. L. Ford and J. C. Browne, *Phys. Rev. A* **16**, 1992 (1977).
83. L. Wolniewicz, *Can. J. Phys.* **54**, 672 (1976).
84. W. R. Thorson, J. H. Choi, and S. K. Knudson, *Phys. Rev. A* **31**, 34 (1985).
85. L. Adamowicz, *Mol. Phys.* **65**, 1047, (1988).
86. D. M. Bishop, J. Pipin, and S. M. Cybulski, *Phys. Rev. A* **43**, 4845 (1991).
87. D. M. Bishop and B. Lam, *Chem. Phys. Lett.* **134**, 283 (1986).
88. M. G. Papadopoulos, A. Willetts, N. C. Handy, and A. E. Underhill, *Mol. Phys.* **88**, 1063 (1996).
89. F. J. Lovas, and E. Tiemann, *J. Phys. Chem. Ref. Data* **3**, 609 (1974).
90. L. Wharton, L. P. Gold, and W. Klemperer, *J. Chem. Phys.* **37**, 2149 (1962).
91. D. Jonsson, P. Norman, and H. Agren, *J. Chem. Phys* **105**, 6401 (1996).
92. D. M. Bishop, B. Lam, and S. T. Epstein, *J. Chem. Phys* **88**, 337 (1988).
93. P. Kozlowski and L. Adamowicz, *J. Comp. Chem* **13**, 602 (1992).
94. J. Rychlewski, *Adv. Quantum Chem.* **31**, 173 (1998).
95. D. W. Gilmore, P. M. Kozlowski, D. B. Kinghorn, and L. Adamowicz, *Int. J. Quant. Chem.* **63**, 991 (1997).
96. P. Pulay, *Mol. Phys.* **17**, 197 (1969).
97. H. B. Schlegel, *Theor. Chem. Acc.* **103**, 294 (2000).
98. B. Liu, *J. Chem. Phys.* **58**, 1925 (1973).
99. D. G. Truhlar and C. J. Horowitz, *J. Chem. Phys.* **68**, 2466 (1978).
100. Y.-S. M. Wu, A. Kuppermann, and J. B. Anderson *Phys. Chem. Chem. Phys.* **1**, 929 (1999).
101. K. T. Tang and J. P. Toennies *Chem. Phys. Lett.*, **151**, 301 (1988).
102. Gaussian 98 (Revision A.7), M. J. Frisch, et al., Gaussian, Inc., Pittsburgh, PA, 1998.
103. R. Jaquet, W. Cencek, W. Kutzelnigg, and J. Rychlewski, *J. Chem. Phys.* **108**, 2837 (1998).
104. R. Poshusta and D. Kinghorn, *Int. J. Quantum Chem.* **60**, 213 (1996).
105. J. R. Magnus and H. Neudecker, *Matrix Differential Calculus with Applications in Statistics and Econometrics*, Wiley, Chichester, 1988.
106. W. Heitler and F. London, *Z. Phys.* **44**, 455 (1927).
107. L. Wolniewicz, private communication.
108. L. Wolniewicz, *J. Chem. Phys.* **103**, 1792 (1995).
109. I. Ben-Itzhak, E. Wells, K. D. Carnes, V. Krishnamurthi, O. L. Weaver, and B. D. Esry, *Phys. Rev. Lett.* **85**, 58 (2000).
110. A. Carrington, I. R. McNab, C. A. Montgomerie–Leach, and R. A. Kennedy, *Mol. Phys.* **72**, 735 (1991).
111. R. E. Moss and I. A. Sadler, *Mol. Phys.* **61**, 905 (1987).
112. R. E. Moss, *Mol. Phys.* **78**, 371 (1993).
113. B. D. Esry and H. R. Sadeghpour, *Phys. Rev. A* **60**, 3604 (1999).
114. R. E. Moss and L. Valenzano, *Mol. Phys.* **100**, 649 (2002).
115. M. Cafiero, S. Bubin, and L. Adamowicz, *Phys. Chem. Chem. Phys.* **5**, 1491 (2003).
116. D. B. Kinghorn, *Int. J. Quantum Chem.* **57**, 141 (1996).
117. D. B. Kinghorn and R. D. Poshusta, *Int. J. Quantum Chem.* **62**, 223 (1997).

118. D. B. Kinghorn and L. Adamowicz, *J. Chem. Phys.* **110**, 7166 (1999).
119. D. B. Kinghorn and L. Adamowicz, *Phys. Rev. Lett.* **83**, 2541 (1999).
120. C. E. Scheu, D. B. Kinghorn, and L. Adamowicz, *J. Chem. Phys.* **114**, 3393 (2001).
121. S. Bubin and L. Adamowicz, *J. Chem. Phys.* **118**, 3079 (2003).
122. S. Bubin and L. Adamowicz, *J. Chem. Phys.* **120**, 6051 (2004).
123. S. Bubin and L. Adamowicz, *J. Chem. Phys.* **121**, 6249 (2004).
124. S. Bubin, E. Bednarz, and L. Adamowicz, submitted for publication.
125. I. Dabrowski, *Can. J. Phys.* **62**, 1639 (1984).
126. H. W. Sarkas, J. H. Hendricks, S. T. Arnold, and K. H. Bowen, *J. Chem. Phys.* **100**, 1884 (1994).
127. D. Chang, K. Reimann, G. Surratt, G. Gellene, P. Lin, and R. Lucchese, *J. Chem. Phys.* **117**, 5757 (2002).
128. A. M. Frolov and V. H. Smith, *Phys. Rev. A* **55**, 2662 (1997).
129. J. Usukura, K. Varga, and Y. Suzuki, *Phys. Rev. A* **58**, 1918 (1998).
130. Z.-C. Yan and Y. K. Ho, *Phys. Rev. A* **60**, 5098 (1999).
131. M. Mella, G. Morosi, D. Bressanini, and S. Elli, *J. Chem. Phys.* **113**, 6154 (2000).
132. M. Mella, S. Chiesa, and G. Morosi, *J. Chem. Phys.* **116**, 2852 (2002).
133. K. Strasburger and H. Chojnacki, *J. Chem. Phys.* **108**, 3218 (1998).
134. K. Strasburger, *J. Chem. Phys.* **111**, 10555 (1999).
135. K. Strasburger, *J. Chem. Phys.* **114**, 615 (2001).
136. J. Mitroy and G. G. Ryzhikh, *J. Phys. B* **33**, 3495 (2000).
137. W. Cencek and J. Rychlewski, *Chem. Phys. Lett.* **320**, 549 (2000).
138. Z.-C. Yan and G. W. F. Drake, *Phys. Rev. A* **61**, 022504 (2000).
139. X. Li and J. Paldus, *J. Chem. Phys.* **118**, 2470 (2003).
140. W. C. Stwalley and W. T. Zemke, *J. Phys. Chem. Ref. Data* **22**, 87 (1993).
141. Z.-C. Yan and Y. K. Ho, *Phys. Rev. A* **60**, 5098 (1999).