ARTICLE

# Computer program ATOM-MOL-nonBO for performing calculations of ground and excited states of atoms and molecules without assuming the Born-Oppenheimer approximation using all-particle complex explicitly correlated Gaussian functions

# Cite as: J. Chem. Phys. 152, 204102 (2020); doi: 10.1063/1.5144268 Submitted: 27 December 2019 • Accepted: 30 April 2020 • Published Online: 26 May 2020

View Online Export Citation CrossMark

Sergiy Bubin<sup>1,a)</sup> (b) and Ludwik Adamowicz<sup>2,b)</sup> (b)

# AFFILIATIONS

<sup>1</sup>Department of Physics, School of Sciences and Humanities, Nazarbayev University, Nur-Sultan 010000, Kazakhstan <sup>2</sup>Department of Chemistry and Biochemistry and Department of Physics, University of Arizona, Tucson, Arizona 85721, USA and Interdisciplinary Center for Modern Technologies, Nicolaus Copernicus University, Toruń PL 87-100, Poland

Note: This article is part of the JCP Special Topic on Electronic Structure Software. <sup>a)</sup>Electronic mail: sergiy.bubin@nu.edu.kz <sup>b)</sup>Author to whom correspondence should be addressed: ludwik@arizona.edu

# ABSTRACT

In this work, we describe a computer program called ATOM-MOL-nonBO for performing bound state calculations of small atoms and molecules without assuming the Born–Oppenheimer approximation. All particles forming the systems, electrons and nuclei, are treated on equal footing. The wave functions of the bound states are expanded in terms of all-particle one-center complex explicitly correlated Gaussian functions multiplied by Cartesian angular factors. As these Gaussian functions are eigenfunctions of the operator representing the square of the total angular momentum of the system, the problem separates and calculations of states corresponding to different values of the total rotational quantum number can be solved independently from each other. Due to thorough variational optimization of the Gaussian exponential parameters, the method allows us to generate very accurate wave functions. The optimization is aided by analytically calculated energy gradient determined with respect to the parameters. Three examples of calculations performed for diatomic and triatomic molecules are shown as an illustration of calculations that can be performed with this program. Finally, we discuss the limitations, applicability range, and bottlenecks of the program.

Published under license by AIP Publishing. https://doi.org/10.1063/1.5144268

# I. INTRODUCTION

In recent years, there has been increasing interest in quantum mechanical calculations of molecular bound ground and excited states without assuming the Born–Oppenheimer (BO) approximation.<sup>1–3</sup> There are at least two major reasons why such calculations are interesting and can provide useful information. The first reason is related to higher accuracy in calculating the molecular rovibrational and electronic spectra, as well as the electronic spectra of atoms. When the BO approximation is not assumed, the wave functions and the corresponding nonrelativistic energy levels of the atom or molecule explicitly include effects originating from the finite mass of the nuclei and from the coupling of the motions of the nuclei and electrons. If such non-BO wave functions are then used to calculate relativistic and quantum electrodynamics (QED) corrections, these corrections will directly include the finite-nuclear-mass (FNM) effects, i.e., the so-called recoil effects. We have showed that such an approach can produce results whose accuracy match the accuracy of the most accurate experimental measurements.<sup>3</sup> The second reason for carrying out the non-BO atomic and molecular calculations stems from the interest in describing properties and structures of molecules and atoms by calculations where all particles forming the system are treated on equal footing. With that, the molecular structure, dipole moment, polarizabilities, etc., are obtained as expectation values of operators representing these properties, and for molecules containing identical nuclei, the indistinguishability principle leads to interesting effects not present when the BO approximation is assumed.

Despite the fact that both BO and non-BO approaches are based on the principles of quantum mechanics, there are significant differences between them when it comes to calculating the energies and the corresponding wave functions. In order to describe a molecular rovibrational spectrum within the BO approximation, one needs to perform separate calculations of the electronic wave functions and the corresponding energies at some selected configurations of the nuclei placed in different fixed positions in space. These calculations yield the so-called potential energy surface (PES), which is used in the subsequent calculation of bound rovibrational states of the molecule. In non-BO calculations, as the nuclei and the electrons forming the molecule are treated on an equal footing, the calculations yield the total energies and the corresponding total wave functions, which explicitly depend on the coordinates of both the nuclei and the electrons. Both the energies and wave functions directly include all effects (including the high-order ones) that originate from the coupling of the motion of the nuclei and electrons, i.e., adiabatic and non-adiabatic effects.

At the nonrelativistic level, the Hamiltonian used to calculate the internal states of an atomic or molecular system commutes with the operators representing the square of the total orbital angular momentum operator and its projection on some axis (usually chosen to be the z-axis). Therefore, the corresponding quantum numbers, which can be used to label the eigenstates of the Hamiltonian, are good quantum numbers. As the Hamiltonian is isotropic (i.e., invariant with respect to rotations about the center of the internal coordinate system described later in this work), the wave functions are atom-like [i.e., they transform according to the irreducible representations of the SO(3) rotation group]. Thus, if basis functions that are eigenfunctions of the square of the total orbital angular momentum and its projection on the z-axis are used in the calculations, the problem of calculating atomic or molecular bound states separates into independent calculations that are performed for groups of states, each corresponding to a particular set of the total orbital angular momentum quantum numbers. This applies equally to atoms and molecules, as the non-BO Hamiltonians for both types of systems have the same spherical symmetry.

The starting point of the approach we have developed for non-BO atomic and molecular calculations is the transformation of the nonrelativistic Hamiltonian, which depends on laboratoryframe Cartesian coordinates of all particles forming the system. This Hamiltonian describes both the relative motion of the particles around the center of mass of the system (here, we termed it the internal motion) and the translational motion of the center of mass in 3D space. The latter corresponds to a free motion of the system as a whole, which has continuous (i.e., non-discrete) energy spectrum, and thus has to be separated. After this separation, the calculation can only focus on system's bound states, which correspond to the internal motion. The internal bound states are eigenstates of the Hamiltonian (we call it the internal Hamiltonian), which is obtained by separating out the operator representing the motion of the center of mass from the laboratory Hamiltonian. The internal Hamiltonian used in our atomic and molecular non-BO calculations is given in Sec. II.

One of the central issues in non-BO calculations is the selection of appropriate basis functions for expanding the spatial part of the wave function. In this selection, several points have to be considered. They are related to the description of the interparticle correlation effects and to the description of the radial and angular nodes in the wave functions of the ground and excited states.

If the nuclei and electrons in a non-BO calculation are treated on equal footing, the nucleus–nucleus (n–n), electron–electron (e– e), and nucleus–electron (n–e) correlation effects need to be represented in the wave function of system. A natural and very effective form of this representation, at least for the case of pairwise interaction between particles, involves expanding the wave function in terms of basis functions, which explicitly depend on the inter-particle distances. Such functions are called explicitly correlated functions. Various types of explicitly correlated functions have been adopted in quantum mechanical calculations of atomic and molecular systems within the BO approximation. Most commonly, these functions are taken in some form of multi-particle Gaussian functions.

Atomic and molecular calculations where the BO approximation is not assumed represent a more recent development in the field of quantum molecular science. Such calculations require some modifications of the explicitly correlated basis functions, as the correlation effects, i.e., the e-e, n-e, and n-n (n-n is only relevant for molecules) correlation effects, that need to be described are more intricate than in the calculations concerning purely electronic states. As electrons are lighter and more delocalized in space, their wave functions (if we could talk about independent singleparticle orbitals) overlap more strongly than the wave functions of nuclei. Thus, the correlation in the motion of electrons, on the relative scale, is smaller than the correlation in the motion of nuclei.

As the e-e, n-e, and n-n correlation effects are due to the electrostatic interactions between the particles, they depend on the charges of the particles. However, they also depend on particles' masses. The mass dependence appears due to the following effect. For the light electrons, there is non-negligible probability of finding two of them (with opposite spins) at the same point in space. On the other hand, due to significantly larger masses of the nuclei, the internuclear correlation is stronger because the nuclei in a molecule avoid each other almost completely, even when their spins are opposite. The third type of correlation, the n-e correlation, can be called anticorrelation, as it is arises due to attraction and describes an effect of electrons, particularly the core electrons, following the nuclei very closely. The direct dependence of the basis functions on the electron–nucleus distances is key to accurately describing the latter correlation.

The explicitly correlated Gaussian functions that have the dependency on the interparticle distance only in the exponential factor are usually quite effective in describing the n-e and e-e correlation effects. However, as the probability of finding two nuclei at the same spatial point is several orders of magnitude smaller than finding two electrons in the same point, additional correlation factors need to be included in the explicitly correlated Gaussians for non-BO calculations. These additional factors are also needed to describe the radial and angular nodes of the wave function that appear due to ro-vibrational (and also electronic) excitations in molecular non-BO wave functions and due to electronic excitations in atomic non-BO wave functions. The need for such factors is due to Gaussians at coalescent points  $(r_{ii} \rightarrow 0)$  having maxima, which are natural and desirable in describing the n-e correlation, but not desirable in describing the n-n and, partially, the e-e correlations. The factors can be chosen in different forms, but their role is to lower (or even zeroing) the amplitude of the wave function at  $r_{ij} = 0$ . The factors should also allow us to generate radial oscillations of the wave functions for "vibrationally excited" states. The angular oscillations of the wave functions, which are also needed to describe atomic and molecular excited states, can be achieved by multiplication of the Gaussians by appropriate Cartesian spherical harmonics. Finally, it is desirable to use a common set of basis functions in both atomic and molecular non-BO calculations, as the general structures and the symmetries of the Hamiltonians for these two types of systems are the same. Such a common basis set that can be used in very accurate atomic and molecular non-BO calculations is implemented in the computer code, which is described in this work.

Non-BO calculations of ground and excited states of atomic and molecular systems are interesting and useful for the following reasons: (i) treating nuclei and electrons of the system on equal footing is approximately equivalent to computing the entire molecular potential energy surface and determining rovibrational states of the system in a single calculation, (ii) the non-BO approach enables a direct account of the coupling of the vibrational and electronic degrees of freedom and can provide a very accurate description of this effect, (iii) non-BO calculations when performed with extended well optimized basis sets and with the inclusion of leading relativistic and quantum electrodynamics (QED) effects are capable of predicting spectral transitions with accuracy that matches or even exceeds the accuracy of the most accurate high-resolution experiments, and (iv) non-BO calculations provide a unique description of chemical and physical properties of atoms and molecules outside the realm of the BO approximation.

The non-BO effects can be determined directly by an approach such as the one presented in this work or by calculating these effects using the perturbation theory, where the BO solution is taken to be the zeroth-order approximation. A question that one can asks is which of the two approaches is more accurate and convenient to apply. An example that sheds some light on this point can be found in the paper by Pachucki and Komasa<sup>4</sup> where one finds a comparison of the non-adiabatic corrections to the adiabatic energy levels of the H<sub>2</sub> molecule obtained in direct non-BO calculations by Wolniewicz, by our group, and by a perturbation-theory method derived in that work (see Table I in Ref. 4) based on the BO approach. While the agreement between our results and the results of Wolniewicz is very good, it is slightly worse when comparing with the results of Pachucki and Komasa. It seems that higher order corrections, which are more cumbersome to calculate, are needed to obtain a better agreement of the results obtained with the direct non-BO method and the perturbation theory results.

# **II. HAMILTONIAN**

Let us consider an isolated atom or molecule formed by N particles with masses  $\{M_i\}$  and charges  $\{Q_i\}$ . We start with the positions of the particles first described using the Cartesian coordinates in the laboratory frame,  $\{\mathbf{R}_i\}$ . The laboratory coordinates and the corresponding linear momenta of the particles are

$$\mathbf{R} = \begin{bmatrix} \mathbf{R}_1 \\ \mathbf{R}_2 \\ \dots \\ \mathbf{R}_N \end{bmatrix} = \begin{bmatrix} X_1 \\ Y_1 \\ Z_1 \\ \vdots \\ Z_N \end{bmatrix}, \qquad \mathbf{P} = \begin{bmatrix} \mathbf{P}_1 \\ \mathbf{P}_2 \\ \dots \\ \mathbf{P}_N \end{bmatrix} = \begin{bmatrix} P_{x1} \\ P_{y1} \\ P_{z1} \\ \vdots \\ P_{zN} \end{bmatrix}. \qquad (1)$$

The laboratory frame nonrelativistic Hamiltonian of the system is

$$H_{\rm nr}(\mathbf{R}) = \sum_{i=1}^{N} \frac{\mathbf{P}_i^2}{2M_i} + \sum_{i=1}^{N} \sum_{j>i}^{N} \frac{Q_i Q_j}{|\mathbf{R}_i - \mathbf{R}_j|}.$$
 (2)

Next, the 3N-dimensional problem represented by the above Hamiltonian is separated into two problems. The first is a 3D problem of the motion of the center of mass of the system in space. The second problem is an (3N - 3)-dimensional internal problem of the motion of the particles forming the atom or a molecule with respect to each other. The separation of the two problems is rigorous and is achieved by transforming Hamiltonian (2) to a new coordinate system. The first three coordinates,  $\mathbf{r}_0$ , in this transformation are the coordinates of the center of mass in the laboratory coordinate frame and the remaining 3N - 3 coordinates are the internal Cartesian coordinates. These coordinates, denoted as  $\mathbf{r}_i$  (*i* = 1, ..., *N* - 1), are the position vectors of particles 2 to *N* with respect to particle 1. Particle 1 is called the reference particle. While any particle in the system can be chosen to be the reference particle, it is natural to assign this role to the heaviest one. The approach is a generalization of the textbook approach applied to solve the time-independent Schrödinger equation for the hydrogen atom. Usually in an atomic calculation, the reference particle is the atomic nucleus, and in a molecular calculation, it is the heaviest nucleus. If we denote  $n \equiv N - 1$ , then, in the new coordinate system, Hamiltonian (2) becomes

$$H_{\mathrm{nr}}^{\mathrm{tot}}(\mathbf{r}_{0},\mathbf{r}) = \left(-\frac{1}{2}\frac{1}{M_{\mathrm{tot}}}\nabla_{\mathbf{r}_{0}}^{2}\right) + \left(-\frac{1}{2}\sum_{i}^{n}\frac{1}{\mu_{i}}\nabla_{\mathbf{r}_{i}}^{2}\right) - \frac{1}{2}\sum_{i\neq j}^{n}\frac{1}{m_{0}}\nabla_{\mathbf{r}_{i}}'\nabla_{\mathbf{r}_{j}} + \sum_{i< j}^{n}\frac{q_{i}q_{j}}{r_{ij}} + \sum_{i=1}^{n}\frac{q_{0}q_{i}}{r_{i}}\right), \quad (3)$$

where  $q_i = Q_{i+1}$  (i = 0, ..., n) are the charges,  $\mu_i = \frac{m_0 m_i}{m_0 + m_i}$  are the reduced masses,  $M_{\text{tot}}$  is the total mass of the system,  $m_0$  is the mass of the reference particle, and  $m_i = M_{i+1}$ .  $\nabla_{\mathbf{r}_i}$  is the gradient vector expressed in terms of the  $x_i$ ,  $y_i$ , and  $z_i$  coordinates of vector  $\mathbf{r}_i$ ,  $r_{ij} = |\mathbf{r}_i - \mathbf{r}_j| = |\mathbf{R}_{i+1} - \mathbf{R}_{j+1}|$ , and  $r_{0i} \equiv r_i = |\mathbf{r}_i| = |\mathbf{R}_{i+1} - \mathbf{R}_1|$ . The prime symbol is used to denote the vector/matrix transposition.

In the new coordinate system, the laboratory frame Hamiltonian (3) becomes a sum of the operators representing the kinetic energy of the center-of-mass motion,  $H_{nr}^{cm}(\mathbf{r}_0)$ , and the internal Hamiltonian,  $H_{nr}^{int}(\mathbf{r})$ ,

$$H_{\rm nr}^{\rm tot}(\mathbf{r}_0, \mathbf{r}) = H_{\rm nr}^{\rm cm}(\mathbf{r}_0) + H_{\rm nr}^{\rm int}(\mathbf{r}), \qquad (4)$$

where **r** is a 3*n*-component column vector. Its first three components are coordinates  $\mathbf{r}_1$ , the next three are  $\mathbf{r}_2$ , etc. In the present work, we are only concerned with the internal bound states of the system. Thus, only the eigenvalues and eigenfunctions of the internal Hamiltonian are calculated. The variational calculation in the framework of the Rayleigh–Ritz approach involves solving the generalized eigenvalue problem (finding eigenvalues  $\epsilon$  and eigenvectors C),

$$Hc = \epsilon Sc,$$
 (5)

with the (internal) Hamiltonian and overlap matrices H and S. The elements of the Hamiltonian and overlap matrices,  $H_{kl} = \langle \phi_k | H_{nr}^{int} | \phi_l \rangle$  and  $S_{kl} = \langle \phi_k | \phi_l \rangle$ , are calculated in the chosen set of explicitly correlated basis functions,  $\{\phi_k\}$ .

The computational time for calculating a matrix element of each of the two matrices is, in general, proportional to  $n^3$ , as it involves matrix operations such as inversion and multiplication. The time is also proportional to the product of the factorials of the numbers of particles in the groups of identical particles (for example, in the case of the  $H_3^+$  ion, this product of factorials is  $3! \times 2!$ as there are three identical nuclei and two identical electrons). This later dependency results from the number of permutation symmetry operators that need to be applied to the wave function to impose the proper permutational symmetry. When all matrix elements are computed, solving the generalized eigenvalue problem (5) also makes a sizable contribution to the overall cost of the calculation despite using a quick solver involving finding only a single eigenvalue and the corresponding eigenfunction. It should be noted that both the Hamiltonian and overlap matrices are complex Hermitian matrices, while the overlap matrix is also positive definite.

Upon examining the form of the internal Hamiltonian, one can see that it describes a system of *n* pseudoparticles with the masses equal to reduced masses  $\mu_i$  and charges  $q_i$  (i = 1, ..., n) moving in the central field of the charge of the reference particle,  $q_0$ . The pseudoparticles interact with each other via the Coulomb potential. In addition, their motions are coupled (correlated) through the mass-polarization terms,  $-\frac{1}{2}\sum_{i\neq j}^{n}\frac{1}{m_0}\nabla'_{r_i}\nabla_{r_j}$ . As mentioned, (3) has the symmetry of an atomic Hamiltonian. Thus, the spatial basis functions used to expand the wave function of such a Hamiltonian need to be one-center functions that provide bases for irreducible representations of the SO(3) group of rotations.

# **III. BASIS FUNCTIONS**

As nuclei (nucleus in an atomic calculation) and electrons in the calculation are treated on equal footing, the basis functions need to properly represent this egalitarian approach. In this work, we use all-particle explicitly correlated Gaussian functions as basis functions for expanding the spatial component of the non-BO wave function. There are various forms of explicitly correlated one-center n-particle Gaussian functions (ECGs). The Gaussian exponent of these functions can be represented in the following form that shows its explicit dependence on the inter-particle distances,  $r_{ij}$ :

$$\phi_{k} = \exp\left[-\alpha_{1k}r_{1}^{2} - \alpha_{2k}r_{2}^{2} - \dots - \alpha_{nk}r_{n}^{2} - \beta_{12,k}r_{12}^{2} - \beta_{13,k}r_{13}^{2} - \dots - \beta_{n-1n,k}r_{n-1n}^{2}\right],$$
(6)

where  $r_i$  is the distance between particle *i* and the center of the Gaussian and  $\alpha_{ik}$  and  $\beta_{ij,k}$  are the non-linear parameters of the Gaussian. The above function can be represented in the following alternative compact form:

$$\phi_k(\mathbf{r}) = \exp\left[-\mathbf{r}' \, \bar{\mathbf{A}}_k \, \mathbf{r}\right],\tag{7}$$

where  $\mathbf{A}_k$  is a  $3n \times 3n$  real symmetric matrix of exponential parameters.  $\mathbf{A}_k$  can be written as  $\mathbf{A}_k = \mathbf{A}_k \otimes \mathbf{I}_3$ , where  $\mathbf{I}_3$  is the  $3 \times 3$  unit matrix,  $\otimes$  denotes the Kronecker product, and  $\mathbf{A}_k$  is a  $n \times n$  symmetric matrix. To ensure square integrability of function  $\phi_k(\mathbf{r})$ , matrix  $\mathbf{A}_k$  must be positive definite. This happens automatically if  $\mathbf{A}_k$  is represented in the Cholesky factored form as  $\mathbf{A}_k = \mathbf{L}_k \mathbf{L}'_k$ , where  $\mathbf{L}_k$ is an  $n \times n$ , rank n, lower triangular matrix.  $\phi_k(\mathbf{r})$  is automatically square-integrable for the  $\mathbf{L}_k$  matrix elements being any real numbers.

As mentioned, the description of the n-n correlation and the nodes of the wave function of excited states requires including additional factors in the basis functions. It was shown that, in the case of a diatomic molecule, an effective factor can be a pre-exponential multiplier in the form of a non-negative power of the inter-nuclear distance.<sup>3,5,6</sup> For a molecule with more than two nuclei, the factor should be a product of non-negative powers of all intermolecular distances. Let us call the Gaussians with such multipliers "power Gaussians". For a diatomic system, the power Gaussian has the following form:

$$\phi_k(\mathbf{r}) = r_1^{2p_k} \exp\left[-\mathbf{r}' \, \bar{\mathbf{A}}_k \, \mathbf{r}\right],\tag{8}$$

where  $r_1$  is the distance between the reference nuclei and particle *i* and  $2p_k$  is its integer non-negative power (ranging from 0 to 250 in our calculations; we only use even powers as this considerably simplifies the calculation of the Hamiltonian integrals but has almost no effect on the accuracy of the calculation). For a triatomic molecule, the power Gaussian has the following form:<sup>7</sup>

$$\phi_k(\mathbf{r}) = r_1^{2p_{1,k}} r_2^{2p_{2,k}} r_{12}^{2p_{12,k}} \exp[-\mathbf{r}' \, \bar{\mathbf{A}}_k \, \mathbf{r}], \tag{9}$$

where  $r_1$ ,  $r_2$ , and  $r_{12}$  are internuclear distances and  $2p_{1,k}$ ,  $2p_{2,k}$ , and  $2p_{12,k}$  are their respective non-negative even-integer powers.

The larger is the  $2p_k$  power value at fixed  $\mathbf{A}_k$ , the more the two nuclei are separated from each other [i.e., the maximum of the Gaussian, which is located, for example, for Gaussian (8) on a sphere, shifts to a sphere with a larger radius]. Gaussians with the zero value of the powers need to be included to assure that the probability of finding two or more nuclei in a single point in space may not be exactly zero. In a calculation of an excited vibrational state, Gaussians (8) with different values of the  $2p_k$  powers are combined to generate the radial oscillations and the nodes in the wave function. As the non-BO wave function for an excited ro-vibrational state of, for example, a diatomic system oscillates in terms of  $r_1$ , it also oscillates in terms of the distances between the electrons and the nuclei because the maxima of the electron densities appear in the same regions where there are maxima of the densities of the nuclei. In the non-BO wave function of a diatomic molecule, where the internal coordinate  $\mathbf{r}_1$  describes the relative position of the second nucleus with respect to the first (reference) nucleus, the internal coordinates  $\mathbf{r}_2$ ,  $\mathbf{r}_3$ , etc., describe the relative positions of the electrons with respect to the reference nucleus. Thus, the oscillations of the wave functions in terms of  $\mathbf{r}_1$  are accompanied by similar oscillations of the wave function in terms of  $\mathbf{r}_2$ ,  $\mathbf{r}_3$ , etc. However, as the wave functions do not have nodes in terms of these latter coordinates, there is less need to include in the basis functions pre-exponential multipliers being powers of the  $r_2$ ,  $r_3$ , etc. distances. The dependency of the exponents of the Gaussians on these distances usually suffices in this case. This explains the effectiveness of the basis set of Gaussians (8) in the non-BO calculations of diatomic molecules.<sup>3,5,6</sup> However, this effectiveness can be increased by making the pre-exponential multipliers also dependent on powers of  $r_2$ ,  $r_3$ , etc. This point is discussed further when the basis functions used in the present work are introduced.

In recent years, we have developed and tested an alternative form of ECGs for molecular non-BO calculations.<sup>8–10</sup> These alternative basis functions are all-particle explicitly correlated Gaussians with complex exponential parameters (complex ECGs or CECGs, for short). The general form of such functions for states with the rotation quantum number of zero (L = 0) is<sup>8</sup>

$$\phi_k(\mathbf{r}) = \exp\left[-\mathbf{r}' \ \tilde{\mathbf{C}}_k \ \mathbf{r}\right] = \exp\left[-\mathbf{r}' \left(\bar{\mathbf{A}}_k + \mathrm{i}\bar{\mathbf{B}}_k\right) \mathbf{r}\right], \tag{10}$$

where  $\mathbf{\tilde{A}}_k$ , as in (7),  $\mathbf{\tilde{B}}_k$  is a real symmetric matrix of variational parameters, and  $\mathbf{i} = \sqrt{-1}$ .  $\mathbf{\tilde{A}}_k$  and  $\mathbf{\tilde{B}}_k$  can be written as  $\mathbf{\tilde{A}}_k = \mathbf{A}_k \otimes \mathbf{I}_3$  and  $\mathbf{\tilde{B}}_k = \mathbf{B}_k \otimes \mathbf{I}_3$ . For L = 1 ( $M_L = 0$ ) states, the appropriate form of the CECGs becomes<sup>10</sup>

$$\phi_k(\mathbf{r}) = z_{i_k} \exp\left[-\mathbf{r}' \, \tilde{\mathbf{C}}_k \, \mathbf{r}\right] = z_{i_k} \exp\left[-\mathbf{r}' \left(\bar{\mathbf{A}}_k + \mathrm{i}\bar{\mathbf{B}}_k\right) \, \mathbf{r}\right], \qquad (11)$$

where  $z_{i_k}$  is the *z* internal coordinate of particle  $i_k$ . Note that the value of the integer parameter  $i_k$  can be varied from 1 to *n* for each basis function and, thus, it has the corresponding index *k*. Basis functions (10) and (11) have been implemented.<sup>8–10</sup> Work on implementation of L = 2 ( $M_L = 0$ ) CECGs,

$$\phi_{k}(\mathbf{r}) = (2z_{i_{k}}z_{j_{k}} - x_{i_{k}}x_{j_{k}} - y_{i_{k}}y_{j_{k}})\exp\left[-\mathbf{r'}\ \tilde{\mathbf{C}}_{k}\ \mathbf{r}\right]$$
$$= \exp\left[-\mathbf{r'}\ \left(\tilde{\mathbf{A}}_{k} + i\tilde{\mathbf{B}}_{k}\right)\mathbf{r}\right], \tag{12}$$

is currently in progress. The limitation of L = 0, 1, or 2 in the non-BO atomic and molecular calculations means that only the states where the electronic excitations are limited to  $\Sigma$ ,  $\Pi$ , and  $\Delta$  states, states where the total rotational quantum number does not exceed three, and states of the mixed nature where the total angular momentum quantum number does not exceed three can be considered. Future developments may go beyond these constraints.

As mentioned, to ensure the square integrability of  $\phi_k(\mathbf{r})$ , matrix  $\mathbf{A}_k$  in (7) is represented in the Cholesky factored form as  $\mathbf{A}_k = \mathbf{L}_k \mathbf{L}'_k$ . The  $\mathbf{L}_k$  matrix elements, which are variational parameters in the calculation, can take any real values, positive or negative. In our recent works, we showed that functions (10) can be contracted to form pre-exponential multipliers being products of sin and cos functions dependent on squares of the inter-particle distances.<sup>11</sup> Such products multiplied by the real part of the exponent of the Gaussian, which "dumps" the sin/cos oscillations at longer distances, are capable of very effectively describing the e-e, n-n, and n-e correlation effects. They can also describe the radial oscillations and the radial nodes of the wave functions of excited states as effectively as the power Gaussians for both atomic and molecular systems. Finally, the CECG basis functions are universal and can be used in non-BO calculations of atoms and diatomic molecules, as well as in calculations of molecules with more than two nuclei.

The exact wave functions of states of atoms and molecules corresponding to the projection of the total orbital angular momentum vector on the *z* axis equal to zero are functions that are not complex. Even though complex Gaussians can be contracted into sin/cos linear combinations that are real functions, there is some value in using CECGs in the uncontracted form in the non-BO calculations. If the complex Gaussians are uncontracted and both  $A_k$  (or  $L_k$ ) and  $B_k$ matrix elements are variationally optimized as independent parameters, it gives the basis more flexibility and may open some additional trajectories in the process of optimizing the nonlinear variational parameters, which in the end can lead to more compact expansion of the wave functions in terms of the basis functions. At the same time, the complex-space optimization approach may produce a wave function with a small imaginary component. This is acceptable as the function is variational.

The angular excitations of the system, for both the atomic and molecular cases, are described by including angular functions as pre-exponential Gaussian multipliers. As mentioned, for L = 1, the Gaussian should be multiplied by the  $z_{i_k}$  factor, where  $i_k = 1$ , ..., *n*. The value of  $i_k$  in our calculations is treated as an additional variational parameter and is partially optimized. While in the atomic non-BO calculations, *i* is an electron index, in molecular calculations, it may correspond to either an electron label or a label of a nucleus. Thus, in a calculation of an excited rovibrational state, along with the main contributions to the non-BO wave function that describe angular excitations of the nuclei, there might be some minor contributions corresponding to angular excitations of the electrons. We showed that such contributions are indeed small, yet not negligible.<sup>12</sup> For rovibrational states where more than one particle is angularly excited (we calculated such states for atoms<sup>13</sup>), there can be contributions to the wave function from Gaussians where both nuclei (nucleus for an atom) and electrons are simultaneously angularly excited. The CECGs with appropriate mixed Cartesian angular factors can effectively describe such excitations.

#### **IV. COMPUTATIONAL IMPLEMENTATION**

The algorithms for performing non-BO calculations with CECGs are implemented in a computer program written in FOR-TRAN and employ MPI (Message Passing Interface) protocol for parallelization on distributed memory systems. The implementation includes algorithms for calculating the matrix elements of the internal-Hamiltonian matrix and the overlap matrix. At present, the calculations can be performed with basis functions (10) and (11). The implementation also includes the matrix elements of the analytical energy gradient determined with respect to the elements of  $L_k$ and  $B_k$  matrices, i.e., the Gaussian non-linear parameters. The procedure for calculating the gradient is described in Refs. 8 and 14. The implementation is general and can be applied to an arbitrary number and types of particles. As the calculations of matrix elements requires very little communication between different MPI processes, this part of the code is highly scalable.

The scheme implemented in the computer program is fairly straightforward. The major goal of the calculation is to generate a CECG basis set for expanding the wave function of the desired state of the system and to calculate its energy and other properties. The basis set generation is based on variational minimization of the total nonrelativistic internal non-BO energy of the considered state of the system. The calculation of the energy involves constructing the Hamiltonian and overlap matrices and solving the generalized eigenvalue problem. The variational energy minimization employs the energy gradient to navigate the search for an energy minimum (or at least a low enough point in the space of nonlinear variational parameters). The gradient is dependent on the derivatives of the Hamiltonian and overlap matrix elements determined with respect to the matrix elements of the  $L_k$  and  $B_k$  matrices.

The basis set is generated and tuned independently for each considered state of the system. The basis set is grown from a small set of Gaussians that can be obtained by a very thorough optimization, constructing it from a set of single-particle Gaussian orbitals, or taken from an already calculated neighboring state of the system. Using a small set of Gaussians whose parameters are obtained from a random number generator is also a possibility. The growing process involves adding basis functions one by one or in small groups of functions. The initial guess for a new added basis function is obtained by randomly perturbing the non-linear parameters of a subset of some function already included in the basis set and then by selecting the candidate that lowers the energy the most. For that function, the particle indices in the angular pre-exponential multiplier [for example, the  $i_k$  index in the  $z_{i_k}$  pre-exponential angular multiplier in basis function (11) or the  $i_k$  and  $j_k$  indices in function (12)] are optimized first. This is followed by an optimization of the function's  $\mathbf{L}_k$  and  $\mathbf{B}_k$  parameters. At this point, the overlap integrals between the optimized function and the functions already included in the basis set are calculated to check for linear dependencies. If any linear dependencies exceeding a predefined threshold are detected, the function is discarded and a new function is selected and optimized. It is undesirable to have linearly dependent functions in the basis set, as this may lead to severe numerical instabilities and loss of precision in the calculation. After addition of a certain number of new functions to the basis set, the whole basis set is reoptimized by cycling over all functions one by one and reoptimizing their  $L_k$ and  $\mathbf{B}_k$  exponential parameters (no reoptimization of the *i* index is done at this point). The growing of the basis set continues until the desired energy convergence is reached.

It should be noted that in our code, there is also an alternative option of reoptimization of all parameters of all basis functions simultaneously. In this case, in order to prevent the optimization process from yielding a set of highly linearly dependent functions, we add a penalty function as described in Ref. 15. In the approach, a positive-valued penalty term,

$$\mathcal{P} = \sum_{i,j} \beta(|S_{ij}|^2 - t^2), \quad \forall \ i,j : |S_{ij}| > t,$$
(13)

is added to the total variational energy of the considered state whenever two or more basis functions have the normalized overlap,  $S_{ij}$ , higher than the predefined threshold, *t*. After the addition, instead of minimizing the energy, the sum of the energy and the penalty term is minimized.  $\beta$  in (13) is a scaling factor. The penalty term depends on the overlap integrals between the basis functions and increases as these overlaps become closer to one (the basis functions are normalized). In the calculation, the penalty threshold is set to a value slightly smaller than one (e.g., t = 0.99). This ensures that during the optimization no two basis functions become almost linearly dependent.

The code, which contains a makefile that can be used to compile it, is placed in a GitHub repository and can be downloaded freely.<sup>16</sup> A description of how to generate an input file is also supplied (some input parameters and variables used in the code are described in the comments of the FORTRAN source). We also provide examples of the input file (see inout.txt) and the corresponding output files. The examples concern calculations of the ground and the first excited states of the HD<sup>+</sup> ion. The structure of the code reflects the purpose of the calculation, which achieves the lowest possible total nonrelativistic energy of the considered state of the system. The code contains subroutine DRMNG, modified by us, which realizes unconstrained optimization and is implemented via reverse communication.<sup>17</sup> During the minimization process, DRMNG calls the procedure that calculates the energy and the energy gradient, which in turn calls the procedure to calculate the Hamiltonian and overlap matrix elements and their derivatives with respect to the non-linear parameters of the bra and ket basis functions. Thus, a significant amount of the computational time goes into calculating the matrix elements. The part of the code that does that is fully parallelized. To illustrate the efficiency of the parallelization, calculations are performed where the Hamiltonian and overlap matrices are generated for the basis set of 1500 CECGs for HD<sup>+</sup> using different numbers of MPI processes (i.e., 28, 14, 7, 3, and 1). The timings for the calculations of matrices are shown in Table I. They were obtained on a server with Intel Xeon E5-2695v3 central processing unit (CPU) (14 cores/28 threads and 2.3 GHz base frequency). The data show that, until the number of MPI processes reaches 14, the

**TABLE I.** Timings (in milliseconds) for computing the Hamiltonian and overlap matrices in the basis of CECGs functions for the ground state of HD<sup>+</sup>. In the first half of the table, the basis has a fixed size of K = 1500 CECGs, while the number of MPI processes executed changes. In the second half, the number of MPI processes, 14, is fixed, while K changes.

Number of cores	Time (ms)		
1	1050		
3	422		
7	245		
14	190		
28	168		
Number of CECGs	Time (ms)		
188	4		
375	12		
750	44		
1500	188		

**TABLE II.** Wall time (minutes:seconds) used to perform HD<sup>+</sup> calculations using seven MPI processes that involve increasing the basis set size from  $K_{\text{start}}$  to  $K_{\text{stop}}$  and performing 20 optimization cycles in which all CECGs are optimized one by one. All energies are in a.u.

K <sub>start</sub>	K <sub>stop</sub>	Energy (K <sub>start</sub> )	Energy $(K_{stop})$	Wall time
100	150	-0.597 878 500 66	-0.59789557766	1:19
150	200	-0.59789557766	-0.59789735741	3:25
200	250	-0.59789735741	-0.59789776487	6:53
250	300	-0.59789776487	-0.59789789575	12:47
300	1500	-0.59789789575	-0.59789796858	(Several months)

computational time scales nearly linearly with that number. At 28 MPI processes, the scaling becomes less linear, which is partially due to the fact that hyper-threading in many modern CPUs is rarely capable of "doubling" the actual performance (gains due to enabling hyper-threading in most float-point intense applications are relatively small). In the table, we also include timings for calculations performed using 14 MPI processes using different numbers of basis functions. The numbers increase by a factor of two and the timings for calculating the matrix elements approximately increases, as expected, by a factor of four.

Finally, in order to illustrate how much time a typical calculation of a state of a molecular system takes, we performed a series of non-BO calculations for the ground state of the  $\mathrm{HD}^+$  ion in which the CECG basis set is incrementally increased from 100 CECGs to 300 CECGs. Each time, the increment of the increase is equal to 50 CECGs. After each incremental increase of the basis set, 20 optimization cycles are performed for the basis functions, where in each cycle, all basis functions in the set are variationally reoptimized one by one using the gradient-aided optimization approach. The calculations were performed using seven MPI processes. The energies and the timings corresponding to each incremental increase of the basis set size are shown in Table II.

# **V. NUMERICAL ILLUSTRATION**

All three illustrative examples presented in this work concern molecular systems. For atoms, we used Gaussians (10), (11), and (12) without the imaginary components in the exponents (see our recent benchmark calculations of the four lowest <sup>1</sup>S states of the beryllium atom<sup>18</sup> where we showed that the use of these functions can produce extremely accurate total energies for all Be isotopes). As real Gaussians, such as those used in the above-mentioned work on the <sup>1</sup>S spectrum of beryllium, are a partial case of the complex Gaussian basis set (10), the inclusion of the imaginary components in the Gaussian exponents would make these basis functions even better. However, for molecules, the effectiveness of the use of CECGs in ro-vibrational calculations has not been as well documented as for atoms. Thus, in the test calculations described below, we only consider molecules.

The first test concerns the complete vibrational spectrum of the HD<sup>+</sup> ion corresponding to zero total orbital angular momentum. The spectrum consists of 23 bound states (v = 0, ..., 22). The total energies of the ro-vibrational states obtained in the calculations performed with CECGs are shown in Table III. The results are compared with the energies obtained in the calculations performed with power Gaussians.<sup>19</sup> The number of CECGs for each state is significantly smaller than the number of power Gaussians. For example, for the lowest state, the largest number of CECGs is 1500, but there were 4000 power Gaussians in the largest basis set of this kind. However, as one can see, already with 1500 CECGs, the total energy of -0.597 897 968 588 a.u. is slightly lower than the best power Gaussian result of -0.597 897 968 577 a.u. For the first excited state, the 1500 CECG result is slightly higher than the 5000 power-Gaussian result. This trend continues as the size of the power-Gaussian basis set increases for higher states to reach 7000 for the top v = 22 state, while the number of CECGs increases to only 1700. The reason we are only able to show CECG results up to 1700 functions are due to a yet unresolved issue with developing a fast and stable (for highly excited states) eigenvalue solver that can update the solution of the complex generalized eigenvalue problem at the cost of  $\mathcal{O}(K^2)$  arithmetic operations (K is the size of the basis). Therefore, at large Kvalues that exceed ~1000, solving the generalized eigenvalue problem becomes a bottleneck in the current implementation. However, the analysis of the energy convergence indicates that is it likely that the CECG energies for all 23 states will eventually become lower than the power-Gaussian results as the number of CECGs keeps increasing.

As the level of vibrational excitation increases and the wave function acquires more radial nodes, more basis functions are needed to describe the oscillations of the wave function. The primary vibration oscillations appear in the HD<sup>+</sup> non-BO wave function in terms of the coordinate representing the internuclear distance (the length of vector  $\mathbf{r}_1$  in the present calculations). However, the vibration quantum number v is only approximately a good quantum number, as the vibrational motion is coupled in HD<sup>+</sup> with the motion of the electron. The coupling increases with the vibrational excitation. Also, as more oscillation appears in the wave function in terms of  $r_1$  with vibrational excitation, oscillations also appear in terms of the  $r_2$  coordinate, which represents the distance of the electron from the reference nucleus. As mentioned, this happens because the wave-function maxima in terms of  $r_1$  coincide with the maxima in terms of  $r_2$  as the electron follows the moving nucleus (pseudonucleus). In the power Gaussians, there is no term that can directly represent the  $r_2$  oscillations and they have to be described "indirectly" through the exponential correlation terms. On the other hand, appropriate linear combination of CECGs generates real Gaussians with sin/cos preexponential multipliers dependent on  $r_2^2$ , which can directly describe the needed oscillation of the wave

v	Basis	Energy	v	Basis	Energy	v	Basis	Energy	v	basis	Energy
0	1300	-0.597 897 968 559	1	1300	-0.589181829438	2	1400	-0.580903700073	3	1300	-0.573050545883
	1400	-0.597897968586		1400	-0.589181829504		1500	-0.580903700105		1400	-0.573050546218
	1500	-0.597897968588		1500	-0.589181829513		1600	-0.580903700125		1500	-0.573050546313
	*4000	-0.597 897 968 577		*5000	-0.589 181 829 541		*4000	-0.580903700201		*4000	-0.573050546451
4	1400	-0.565611041122	5	1500	-0.558575519512	6	1500	-0.551935946669	7	1500	-0.545685911351
	1500	-0.565611041490		1600	-0.558575520053		1600	-0.551935947550		1600	-0.545685913104
	1600	-0.565611041616		1700	-0.558575520197		1700	-0.551935947844		1700	-0.545685913571
	*4000	-0.565611042015		*4000	-0.558 575 520 667		*4000	-0.551935948624		*4000	-0.545685914996
8	1500	-0.539820634742	9	1500	-0.534337003694	10	1500	-0.529233620505	11	1500	-0.524510889995
	1600	-0.539820637628		1600	-0.534337007333		1600	-0.529233625616		1600	-0.524510896639
	1700	-0.539820638677		1700	-0.534337009157		1700	-0.529233628500		1700	-0.524510900413
	*4000	-0.539820640553		*4000	-0.534337013108		*4000	-0.529233634746		*4000	-0.524510909642
12	1500	-0.520171120258	13	1400	-0.516218648376	14	1400	-0.512660105161	15	1400	-0.509504536602
	1600	-0.520171129340		1500	-0.516218671673		1500	-0.512660139336		1500	-0.509504574611
	1700	-0.520171133163		1600	-0.516218684234		1600	-0.512660156988		1600	-0.509504603521
	*4000	-0.520 171 143 835		*4000	-0.516218708878		*5000	-0.512660191254		* 5000	-0.5095046474123
16	1400	-0.506763735683	17	1400	-0.504452529860	18	1400	-0.502589046694	19	1400	-0.501194606434
	1500	-0.506763780540		1500	-0.504452583906		1500	-0.502589109245		1500	-0.501194671839
	1600	-0.506763816149		1600	-0.504452626599		1600	-0.502589154439		1600	-0.501194710606
	*6000	-0.506 763 873 837		*6000	-0.504452691747		*6000	-0.5025892273423		*7000	-0.501194794224
20	1400	-0.500292256373	21	1500	-0.499910306544	22	1500	-0.499865761719			
	1500	-0.500292320388		1600	-0.499910325099		1600	-0.499865767252			
	1600	-0.500292362144		1700	-0.499910332875		1700	-0.499865770261			
	*7000	-0.500292453636		*7000	-0.499910359483		*7000	-0.499865778308			

TABLE III. The convergence of the total nonrelativistic non-BO energies of the pure vibrational states of HD<sup>+</sup>. The result marked with \* are taken from Ref. 19. All energy values are given in a.u.

function. This explains the effectiveness of CECGs in molecular non-BO ro-vibrational calculations. This also suggests that CECGs should be equally effective in non-BO ro-vibrational calculations of molecules with more than two nuclei as they are in the calculations of diatomics, such as the HD<sup>+</sup> molecular ion.

The HD<sup>+</sup> system is a good example where the Born-Oppenheimer approximation leads to unsatisfactory results as it fails to describe the asymmetry of the charge distribution as the vibrational excitation increases. In BO nonrelativistic electronic calculations, the average proton-electron and deuteron-electron distances are the same regardless of the level of the vibrational excitation (if no adiabatic corrections to the wave functions are included). However, when the non-BO approach is used in the calculations, the densities of the particles forming the system show charge asymmetry that increases with the level of the vibrational excitation. The values of the average interparticle distances calculated using the non-BO wave functions generated with the code described in this work presented in Table IV show that, while for lower vibrational states, the asymmetry (as measured by the difference between the protonelectron and deuteron-electron distances) is small, for the last two states before dissociation, the electron is almost entirely localized at the deuteron. Simultaneously, the character of the bond in the ion changes from being basically covalent at lower vibrational states to ionic at the top states.

The second test concerns a triatomic molecule, the  $D_3^+$  ion. It is chosen because, due to its nuclei being bosons, the permutational symmetry of the nuclei is easier to implement in the wave function than in the case of H<sub>3</sub><sup>+</sup> with fermionic nuclei (H<sub>3</sub><sup>+</sup> non-BO calculations are now possible with the newly implemented algorithm<sup>10</sup> and will be pursued in the near future). When Gaussians (10) with  $\mathbf{B}_k$  set to zero (real Gaussians) are used in a molecular calculation, for example, for H<sub>3</sub><sup>+</sup>, the natural occurrence of linear dependencies between the basis functions prevents the calculation from converging. The reason for this behavior can be explained by the optimization of the basis functions trying to describe the value of the wave function becoming equal to zero as  $r_1 \rightarrow 0$ ,  $r_2 \rightarrow 0$ , or  $r_{12} \rightarrow 0$ . In order to achieve that for real Gaussians [i.e., Gaussians (10) with  $\mathbf{B}_k = 0$ ], the optimization has to effectively generate pre-exponential factors in the Gaussians that are products of powers of  $r_1$ ,  $r_2$ , and  $r_{12}$ . This can be achieved by "differentiating" the real Gaussians with respect to appropriate matrix elements of  $L_k$ . During the optimization, this differentiation

**TABLE IV**. Charge asymmetry in pure vibrational states of HD<sup>+</sup>. Average protondeuteron ( $r_{pd}$ ), proton–electron ( $r_{pe}$ ), and deuteron–electron ( $r_{de}$ ) distances in a.u. calculated for all 23 vibrational states listed in Table III. The calculation for each state is performed with the larges CECG basis set shown for that state in Table III (\*D atom in the ground state).

v	$r_{pd}$	r <sub>pe</sub>	r <sub>de</sub>	
0	2.0548	1.6884	1.6877	
1	2.1713	1.7504	1.7495	
2	2.2918	1.8143	1.8134	
3	2.4167	1.8807	1.8796	
4	2.5467	1.9497	1.9484	
5	2.6825	2.0217	2.0202	
6	2.8250	2.0972	2.0955	
7	2.9754	2.1769	2.1748	
8	3.1349	2.2613	2.2588	
9	3.3054	2.3515	2.3484	
10	3.4891	2.4485	2.4447	
11	3.6889	2.5540	2.5491	
12	3.9087	2.6699	2.6636	
13	4.1539	2.7992	2.7907	
14	4.4320	2.9458	2.9340	
15	4.7542	3.1159	3.0987	
16	5.1377	3.3191	3.2924	
17	5.6112	3.5723	3.5270	
18	6.2274	3.9098	3.8209	
19	7.0989	4.4208	4.1976	
20	8.5499	5.5162	4.5693	
21	12.9501	12.1913	2.3060	
22	28.5345	28.4655	1.6004	
D*			1.5000	

effectively occurs by making two or more Gaussians almost identical and having them contribute to the wave function with a linear coefficient with equal absolute values, but opposite signs. This effectively generates a "numerical" derivative that approximates the analytical one. This behavior explains the tendency of the optimization to create the linear dependencies, which make the optimization ineffective and even numerically unstable. For example, we observe such behavior in the optimization of real Gaussians for vibrational states of the  $H_2^+$  molecule. By adding the imaginary components to the real Gaussians [i.e., by making  $\mathbf{B}_k \neq 0$  in Gaussians (10)], thus effectively allowing the sin/cos pre-exponential factors to appear in the Gaussians, the zeroing of the Gaussians at  $r_1 \rightarrow 0$ ,  $r_2 \rightarrow 0$ , or  $r_{12} \rightarrow 0$ is achieved. Then, there is no need for the linear dependencies to appear in the calculation. This feature is tested for  $D_3^+$  in calculations of its ground and first-excited vibrational states corresponding to the total rotation quantum number of zero. In the test, we use 400 CECGs for each state. The basis set is separately optimized for each state using the gradient-based variational energy minimization. The all-parameter optimization approach is employed. The linear dependencies appear not to be an issue in the calculation. The energies of the lowest two states obtained in the calculations are -1.328 377 6 a.u. and -1.3168588 a.u. These two results are preliminary. To obtain good-quality results, the basis sets for the two states have to be significantly increased. Also, more ro-vibrational states need to be

**TABLE V**. Convergence of the total energy for the ground state of the BH molecule (2 nuclei + 6 electrons) in calculations with explicitly correlated Gaussian basis functions.

Basis size	Energy (a.u.)	
500 1000 2000 3000 4000	$\begin{array}{r} -25.270\ 646\\ -25.277\ 313\\ -25.280\ 280\\ -25.281\ 146\\ -25.281\ 518\end{array}$	

calculated to fully assess the utility of the computer code presented in this work.

The third example of non-BO calculations concerns the largest system we have ever calculated with the non-BO approach using all-particle explicitly correlated Gaussian functions. It is the eightparticle (two nuclei plus six electrons) boronmonohydride molecule, BH. The calculations are performed for the lowest energy state of the system. As the above-described HD<sup>+</sup> calculations showed, CECGs can be equally effective in non-BO molecular calculations as explicitly correlated power Gaussians,19 which are products of CECG Gaussian exponents with  $\mathbf{B}_k = 0$  and even non-negative powers of the internuclear distance. Thus, in the BH calculations, which started long time before the present work was initiated (the calculations are still running), we use the power Gaussians. At present, the basis set size reached over 4000 functions. This required about two years of continuous calculations on a 32-core computer. As for the memory requirements, for relatively large basis sets, the bulk of allocated RAM is used for storing two square matrices of size K. When double precision is used in the calculation, it translates into  $\sim 16 \times K^2$  bytes. Only a small amount of disk space is used. The majority of CPU time goes into calculating the matrix elements of the Hamiltonian and overlap matrices and those of the energy gradient determined with respect to the Gaussian exponential parameters. Because there are six identical particles (electrons) in the BH molecule, each matrix element requires the calculation of 6! = 720 terms. As each Gaussian in the BH calculations contains 28 independent parameters, for 4000 Gaussians, the total number of parameters that are optimized is 112 000.

The results of the BH calculations are shown in Table V. They include the energy values obtained at different stages of the basisset growing process. The reported energy values correspond to 500, 1000, 2000, 3000, and 4000 Gaussians. As one can see, the energy convergence is good, but with 4000 basis functions in the basis set, the energy is converged to about the fifth significant figure only. The BH example provides a good illustration of the effort involved in non-BO molecular calculations of larger molecules.

### **VI. SUMMARY**

In this work, we describe our new computer code ATOM-MOL-nonBO for performing non-Born–Oppenheimer calculations of atoms and molecules with an arbitrary number of electrons and for molecules with an arbitrary number of nuclei. The nuclei and electrons are treated on equal footing. The basis sets of all-particle explicitly correlated complex Gaussians are employed in the calculations. The basis set is universal and can be used to calculate ground and excited electronic states of atoms and ro-vibrational states of molecules. The code is tested in the calculations of the complete vibrational spectrum of the  $HD^+$  ion in the ground rotational state and the lowest two vibrational states of the  $D_3^+$  ion. Some results are also shown from non-BO calculations of the ground state of the eight-particle BH molecule, which is the largest system we have ever calculated within the non-BO approach.

The code continues to be developed and new features are being added. Currently, it allows us to calculate states corresponding to the total rotation quantum number of zero and one. One of the most interesting applications of the code, which will be performed in the future, is calculations of the ro-vibrational spectra of  $H_3^+$  and its isotopologues. To increase the accuracy of the calculations, we are planning to add algorithms for computing the leading relativistic corrections (the algorithm for calculating these corrections for the ground rotational state are already implemented<sup>20</sup>).

# SUPPLEMENTARY MATERIAL

See supplementary material for the source code used in the present calculations can be downloaded/cloned from a git repository.<sup>16</sup> It contains the FORTRAN source (files with extensions .f90 and .f), a UNIX makefile, and several example input files, which set up calculations for the ground and first excited states of the HD<sup>+</sup> ion with 500, 1000, and 1500 basis functions. Example output files are also provided. Users need to make sure that a working installation of the MPI library is available (e.g., Open MPI<sup>21</sup>) and adjust the makefile according to their local configuration.

# ACKNOWLEDGMENTS

This work was supported by a grant from the National Science Foundation (Grant No. 1856702). S.B. also acknowledges funding from MES RK state-targeted Program (No. BR05236454) and NU faculty development (Grant No. 090118FD5345).

## REFERENCES

- <sup>1</sup>T. Yonehara, K. Hanasaki, and K. Takatsuka, Chem. Rev. 112, 499 (2012).
- <sup>2</sup>E. Mátyus, Mol. Phys. 117, 590 (2019).
- <sup>3</sup>S. Bubin, M. Pavanello, W.-C. Tung, K. L. Sharkey, and L. Adamowicz, Chem. Rev. **113**, 36 (2013).
- <sup>4</sup>K. Pachucki and J. Komasa, J. Chem. Phys. 129, 034102 (2008).

<sup>5</sup>S. Bubin, M. Cafiero, and L. Adamowicz, Adv. Chem. Phys. **131**, 377 (2005).

<sup>6</sup>J. Mitroy, S. Bubin, W. Horiuchi, Y. Suzuki, L. Adamowicz, W. Cencek, K. Szalewicz, J. Komasa, D. Blume, and K. Varga, Rev. Mod. Phys. **85**, 693 (2013).

<sup>7</sup>E. Bednarz, S. Bubin, and L. Adamowicz, Mol. Phys. 103, 1169 (2005).

<sup>8</sup>S. Bubin and L. Adamowicz, J. Chem. Phys. 124, 224317 (2006).

<sup>9</sup>S. Bubin, M. Formanek, and L. Adamowicz, Chem. Phys. Lett. **647**, 122 (2016).

<sup>10</sup>E. M. Chavez, S. Bubin, and L. Adamowicz, Chem. Phys. Lett. 717, 147 (2019).

<sup>11</sup> M. Formanek, K. L. Sharkey, N. Kirnosov, and L. Adamowicz, J. Chem. Phys. 141, 154103 (2014).

<sup>12</sup>K. L. Sharkey, N. Kirnosov, and L. Adamowicz, J. Chem. Phys. **139**, 164119 (2013).

<sup>13</sup>K. Jones, M. Formanek, and L. Adamowicz, Chem. Phys. Lett. 669, 188 (2017).
<sup>14</sup>S. Bubin and L. Adamowicz, J. Chem. Phys. 128, 114107 (2008).

<sup>15</sup>W.-C. Tung, M. Pavanello, and L. Adamowicz, J. Chem. Phys. **133**, 124106 (2010).

<sup>16</sup>See https://github.com/sbubin/ATOM-MOL-nonBO for the ATOM-MOLnonBO source code.

 $^{17}\mbox{See https://www.netlib.org/port/dmng.f}$  for the original source of subroutine DRMNG.

<sup>18</sup>I. Hornyák, L. Adamowicz, and S. Bubin, Phys. Rev. A 100, 032504 (2019).

<sup>19</sup>M. Stanke and L. Adamowicz, Phys. Rev. A **89**, 032503 (2014).

<sup>20</sup>S. Bubin, M. Stanke, and L. Adamowicz, Phys. Rev. A **95**, 062509 (2017).

<sup>21</sup> See https://www.open-mpi.org for the Open MPI library project.