Effect of time-dependent basis functions and their superposition error on atom-centered density matrix propagation (ADMP): Connections to wavelet theory of multiresolution analysis

Srinivasan S. Iyengar[a)]
Department of Chemistry and Department of Physics, Indiana University, Bloomington, Indiana 47405

Michael J. Frisch
Gaussian, Inc., North Haven, Connecticut 06473-1712

(Received 11 May 2004; accepted 16 June 2004)

We present a rigorous analysis of the primitive Gaussian basis sets used in the electronic structure theory. This leads to fundamental connections between Gaussian basis functions and the wavelet theory of multiresolution analysis. We also obtain a general description of basis set superposition error which holds for all localized, orthogonal or nonorthogonal, basis functions. The standard counterpoise correction of quantum chemistry is seen to arise as a special case of this treatment. Computational study of the weakly bound water dimer illustrates that basis set superposition error is much less for basis functions beyond the 6-31+G* level of Gaussians when structure, energetics, frequencies, and radial distribution functions are to be calculated. This result will be invaluable in the use of atom-centered Gaussian functions for *ab initio* molecular dynamics studies using Born-Oppenheimer and atom-centered density matrix propagation. © *2004 American Institute of Physics.* [DOI: 10.1063/1.1780157]

I. INTRODUCTION

Over the past several years ''on-the-fly'' approaches to dynamics of nuclei and electrons[1–6] have become popular, leading to the field of *ab initio* molecular dynamics. Here, approximations to the electronic structure are calculated as needed, as the nuclear configurations are propagated to simulate dynamics on the Born-Oppenheimer surface. In recent years the Car-Parrinello (CP) method[3,7–10] has emerged as a very powerful implementation of *ab initio* molecular dynamics.[3,4,7–10] The CP method is an extended Lagrangian molecular dynamics (MD) scheme[11,12] where the electronic degrees of freedom are treated explicitly as dynamical variables and propagated with the nuclear degrees of freedom. Such a scheme offers significant computational advantages for efficient MD calculations. Several important applications with density functional theory (DFT) and the CP method are well documented in the literature.[7,8] Traditionally, CP calculations employ Kohn-Sham orbitals expanded in plane-wave basis sets (occasionally augmented with Gaussian orbitals[13,14]), although recently[15] Wannier basis functions[16,17] have been used to represent the molecular orbitals. Gaussian basis functions within the generalized valence bond (GVB)[18,19] and Hartree-Fock[20] framework and floating Gaussians[21] have also been used within the CP scheme.

In Refs. 6 and 22–26 we have described the theory, implementation, and initial applications of an extended Lagrangian MD method that employs atom-centered Gaussian basis functions and one-particle electronic density matrix propagation. Here, the basis functions follow the nuclei and hence this method is called atom-centered density matrix propagation (ADMP). The ADMP method has many attractive features. Systems can be simulated by accurately treating all electrons or by using pseudopotentials. Through the use of smaller values for the tensorial fictitious mass, large time steps can be employed (ca. 0.25 fs in water simulations) and lighter atoms such as hydrogens are routinely used. A wide variety of exchange-correlation functionals can be utilized, including hybrid density functionals such as B3LYP. Atom-centered functions can be used with the appropriate physical boundary conditions for molecules, polymers, surfaces, and solids, without the need to treat replicated images to impose $3d$ periodicity. Consequently, charged systems and QM/MM models of biological systems[26] can be treated as easily as neutral molecules. Deviations from the Born-Oppenheimer surface and the mixing of fictitious and real kinetic energies can also be rigorously controlled on-the-fly in ADMP. Because of fewer basis functions per atom, larger time steps, and its asymptotic $O(N)$ scaling,[27] good computational efficiency can be achieved with the ADMP method. Interesting applications such as solvation of excess protons in water clusters,[28,29] hydroxyl-stretch redshifts in chloride-water clusters,[23] product distributions in gas-phase chemical reactions,[23] and proton shuttling in biological ion-channels and water wires[26] have already been studied. ADMP trajectories of the order of picoseconds show stable dynamics, and the adiabaticity can be controlled effectively in these systems without thermostats. ADMP has also been recently extended to treat nuclear quantum effects using full quantum wave packet dynamics.[30] The important conceptual and computational differences between ADMP and other Gaussian basis-set based implementations[18–21] of the Car-Parrinello formal-

[a)]Electronic mail: iyengar@indiana.edu

ism have been discussed in detail in Ref. 23.

Current implementation of the ADMP approach has been found to be computationally superior to Born-Oppenheimer dynamics.[23] This important result can be conceptualized based on the following: in Born-Oppenheimer dynamics, the density matrix is to be converged at every dynamics step. Assuming that the largest possible time step is used during dynamics, SCF convergence requires $\approx 8-12$ SCF steps. (This depends on the convergence threshold and difficult cases such as transition metal complexes may require more SCF steps.) In ADMP, on the contrary, only the equivalent of 1 SCF step is required per dynamics step; this 1 SCF step is necessary to calculate the Fock matrix required for propagating the density matrix. (A brief review of ADMP is presented in Sec. II.) Both BOMD and ADMP evaluate the gradient of energy with respect to nuclear coordinates and this calculation requires approximately the same amount of time in both methods. Note that the gradients used in ADMP are more general than those used in BOMD (Refs. 6 and 24) on account of the non-negligible magnitude of the commutator of the Fock and density matrix (see Sec. II and Ref. 24 for details). However, the additional terms require no significant computation over the standard BOMD gradient calculation. The calculation of nuclear force requires approximately three times as much computation time as a single SCF cycle. This makes ADMP faster than BOMD by over a factor of 4 per dynamics step. However, the requirement that the ADMP energies oscillate about the BO values with small amplitudes[22] implies that ADMP step sizes cannot be as large at those in BO dynamics. But good energy conservation, which applies to both methods, limits the BO steps to at most twice those of ADMP.[23] (ADMP already uses reasonably large time steps on account of smaller values for the fictitious mass and an innovative tensorial fictitious mass scheme.[22]) This allows ADMP to be over a factor of 2 superior to BOMD, but this estimate is only for cases where the SCF convergence in BOMD is not difficult.[23] The hard to converge cases would require more SCF steps thus making ADMP more efficient as compared to BOMD for these cases. Furthermore, computational improvements that speed up the gradient evaluation will tilt this comparison further towards ADMP.

In this paper we study the effect of basis-set superposition error (BSSE)[31–36] on dynamics (both BOMD and ADMP). In ADMP and in the BOMD implementation within the Gaussian suite of electronic structure programs[37] atom-centered Gaussians are used as electronic basis sets. As atoms move, the basis functions remain with the atoms. This has the important advantage that the basis functions are always present in regions where the density has important contributions, however, has the disadvantage that the vector space that represents the molecular orbitals is not constant and changes with time and could change appreciably, when small basis sets are used. In this paper we analyze this problem both analytically and computationally.

The paper is organized as follows: in Sec. II we first present a brief overview of ADMP to facilitate some of the discussion. Further details may be found in Refs. 6, 22–26, and 30. In Sec. III, we analyze the primitive Gaussian basis functions that are used in the ADMP approach. We show in Sec. III that these primitive Gaussians are very closely related to the theory of multiresolution analysis which has given rise to wavelet basis functions.[38–41] We show in Sec. III that primitive Gaussians, in fact, belong to a very general category of multiwavelets with nonintegral scale factors. This is an important connection, since wavelets are being used in digital signal processing,[39–41] computational fluid mechanics,[38] and in a number of other areas of scientific computing that involve solutions to nonlinear differential equations. Wavelets are preferred over plane waves in these areas due to their locality and efficiency in representing nonuniform functions. Quantum chemistry of the gas phase and quantum chemistry of reactions in the condensed phase involve the accurate description of wave functions that may be spatially nonuniform. As a result the use of wavelets in electronic structure theory has become important recently.[42–44] In Sec. III A we exploit this connection to analyze, formally, basis-set superposition error and we arrive at error bounds for this quantity. We also make connections to the standard counterpoise expression.[31–36] In Sec. IV we present a comparison of ADMP and BOMD calculations performed with and without counterpoise corrections, and show that for commonly used basis sets such as 6-31+G**, the effect of basis-set superposition error on observable quantities is small. In Sec. V we present our conclusions.

II. A BRIEF OVERVIEW OF ATOM-CENTERED DENSITY MATRIX PROPAGATION (ADMP)

The ADMP equations of motion for the nuclei and density matrix are

$$\mathbf{M}\frac{d^2\mathbf{R}}{dt^2} = -\left.\frac{\partial E(\mathbf{R},\mathbf{P})}{\partial \mathbf{R}}\right|_{\mathbf{P}}, \tag{1}$$

$$\mu^{1/2}\frac{d^2\mathbf{P}}{dt^2}\mu^{1/2} = -\left[\left.\frac{\partial E(\mathbf{R},\mathbf{P})}{\partial \mathbf{P}}\right|_{\mathbf{R}} + \mathbf{\Lambda P} + \mathbf{P\Lambda} - \mathbf{\Lambda}\right], \tag{2}$$

where $\mathbf{R}$, $\mathbf{V}$, and $\mathbf{M}$ are the nuclear positions, velocities, and masses, and $\mathbf{P}$, $\mathbf{W}$, and μ are the density matrix, the density matrix velocity, and the fictitious mass tensor for the electronic degrees of freedom. $\mathbf{\Lambda}$ is a Lagrangian multiplier matrix used to impose N-representability of the single particle density matrix. The energy $E(\mathbf{R},\mathbf{P})$ is calculated using McWeeny purification, $\tilde{\mathbf{P}} = 3\mathbf{P}^2 - 2\mathbf{P}^3$,

$$\begin{aligned} E &= \mathrm{Tr}[\mathbf{h}'\tilde{\mathbf{P}}' + \tfrac{1}{2}\mathbf{G}'(\tilde{\mathbf{P}}')\tilde{\mathbf{P}}'] + E_{xc} + V_{NN} \\ &= \mathrm{Tr}[\mathbf{h}\tilde{\mathbf{P}} + \tfrac{1}{2}\mathbf{G}(\tilde{\mathbf{P}})\tilde{\mathbf{P}}] + E_{xc} + V_{NN}. \end{aligned} \tag{3}$$

Here, $\mathbf{h}'$ is the one electron matrix in the nonorthogonal Gaussian basis and $\mathbf{G}'(\tilde{P}')$ is the two electron matrix for Hartree-Fock calculations, but for DFT it represents the Coulomb potential. The term E_{xc} is the DFT exchange-correlation functional (for Hartree-Fock $E_{xc}=0$), while V_{NN} represents the nuclear repulsion energy. In the orthonormal basis, these matrices are $\mathbf{h} = \mathbf{U}^{-\mathbf{T}}\mathbf{h}'\mathbf{U}^{-1}$, etc., where the overlap matrix for the nonorthogonal Gaussian basis, $\mathbf{S}'$, is factorized to yield $\mathbf{S}' = \mathbf{U}^{\mathbf{T}}\mathbf{U}$. There are a number of choices for the transformation matrix $\mathbf{U}$, e.g., $\mathbf{U}$ can be obtained from

Cholesky decomposition[45] of $\mathbf{S}'$ or $\mathbf{U}=\mathbf{S}'^{1/2}$ for Löwdin symmetric orthogonalization. The matrix $\mathbf{U}$ can also include an additional transformation so that overall rotation of the system is factored out of the propagation of the density. The density matrix in the orthonormal basis $\mathbf{P}$ is related to the density matrix in the nonorthogonal Gaussian basis $\mathbf{P}'$ by $\mathbf{P}\equiv\mathbf{U}\mathbf{P}'\mathbf{U}^{\mathbf{T}}$. The gradient terms involved in the equations of motion are

$$\left.\frac{\partial E(\mathbf{R},\mathbf{P})}{\partial \mathbf{P}}\right|_{\mathbf{R}} = 3\mathbf{F}\mathbf{P}+3\mathbf{P}\mathbf{F}-2\mathbf{F}\mathbf{P}^2-2\mathbf{P}\mathbf{F}\mathbf{P}-2\mathbf{P}^2\mathbf{F}, \tag{4}$$

where $\mathbf{F}$ is the Fock matrix and in the nonorthogonal basis,

$$\mathbf{F}'_{\nu,\sigma}\equiv\mathbf{h}'_{\nu,\sigma}+\mathbf{G}'(\tilde{\mathbf{P}}')_{\nu,\sigma}+\frac{\partial E_{xc}}{\partial \mathbf{P}'}, \tag{5}$$

while the orthogonal basis Fock matrix is $\mathbf{F}=\mathbf{U}^{-\mathbf{T}}\mathbf{F}'\mathbf{U}^{-1}$. The nuclear gradients are

$$\begin{aligned}\left.\frac{\partial E}{\partial \mathbf{R}}\right|_{\mathbf{P}} =& \left\{ \mathrm{Tr}\left[\frac{d\mathbf{h}'}{d\mathbf{R}}\tilde{\mathbf{P}}'+\frac{1}{2}\left.\frac{\partial \mathbf{G}'(\mathbf{P}')}{\partial \mathbf{R}}\right|_{\mathbf{P}'}\tilde{\mathbf{P}}'\right]\right.\\ &\left.-\mathrm{Tr}\left[\mathbf{F}'\tilde{\mathbf{P}}'\frac{d\mathbf{S}'}{d\mathbf{R}}\tilde{\mathbf{P}}'\right]+\left.\frac{\partial E_{xc}}{\partial \mathbf{R}}\right|_{\mathbf{P}}+\frac{\partial V_{NN}}{\partial \mathbf{R}}\right\}\\ &+\mathrm{Tr}\left[[\tilde{\mathbf{P}},\mathbf{F}]\left(\tilde{\mathbf{Q}}\frac{d\mathbf{U}}{d\mathbf{R}}\mathbf{U}^{-1}-\tilde{\mathbf{P}}\mathbf{U}^{-\mathbf{T}}\frac{d\mathbf{U}^{\mathbf{T}}}{d\mathbf{R}}\right)\right],\end{aligned} \tag{6}$$

where $\tilde{\mathbf{Q}}\equiv\mathbf{I}-\tilde{\mathbf{P}}$. Note that as the commutator $[\tilde{\mathbf{P}},\mathbf{F}]\rightarrow 0$, the nuclear forces tend to those used in the standard Born-Oppenheimer MD.[6,24,46] However, in ADMP, the magnitude of the commutator $[\tilde{\mathbf{P}},\mathbf{F}]$ is non-negligible and hence the general expression for the nuclear gradients[6,24] in Eq. (6) is used.

Like CP, ADMP represents fictitious dynamics where the density matrix is propagated instead of being converged. The accuracy and efficiency are governed by the choice of the fictitious mass tensor μ; hence one must be aware of the limits on this quantity. We have derived two independent criteria[22,24] that place bounds on the choice of the fictitious mass. First, the choice of the fictitious mass determines the magnitude of the commutator $[\tilde{\mathbf{P}},\mathbf{F}]$ thus determining the extent of deviation from the Born-Oppenheimer surface,[24]

$$\|[\mathbf{F},\mathbf{P}_{\mathrm{approx}}]\|_F \geqslant \frac{1}{\|[\mathbf{P}_{\mathrm{approx}},\mathbf{W}]\|_F}\left|\mathrm{Tr}\left[\mathbf{W}\mu^{1/2}\frac{d\mathbf{W}}{dt}\mu^{1/2}\right]\right|, \tag{7}$$

where $\|[\cdots]\|_F$ is the Frobenius norm[45,47] of the commutator and is defined as $\|A\|_F=\sqrt{\Sigma_{i,j}A_{i,j}^2}$. Second, the rate of change of the fictitious kinetic energy,

$$\begin{aligned}\frac{d\mathcal{H}_{\mathrm{fict}}}{dt} &= \mathrm{Tr}\left[\mathbf{W}\mu^{1/2}\frac{d^2\mathbf{P}}{dt^2}\mu^{1/2}\right]\\ &= -\mathrm{Tr}\left[\mathbf{W}\left(\left.\frac{\partial E(\mathbf{R},\mathbf{P})}{\partial \mathbf{P}}\right|_{\mathbf{R}}+\mathbf{\Lambda}\mathbf{P}+\mathbf{P}\mathbf{\Lambda}-\mathbf{\Lambda}\right)\right],\end{aligned} \tag{8}$$

is to be bounded and oscillatory and this again is determined by the choice of fictitious mass tensor. We have shown that ADMP gives results that are in good agreement with BOMD and is computationally superior to BOMD.[23] However, one must monitor the quantities in Eqs. (7) and (8) to ascertain that the ADMP dynamics is physically consistent. In all applications studied to date[6,22,23,26,28,29] these conditions are satisfied thus yielding a computationally efficient and accurate approach to model dynamics on the Born-Oppenheimer surface.

III. FORMAL ANALYSIS OF PRIMITIVE GAUSSIAN BASIS FUNCTIONS: CONNECTIONS TO MULTI-RESOLUTION ANALYSIS

The general form of the primitive cartesian Gaussian basis functions used in ADMP is

$$\chi^{\mathbf{R}}_{l,m,n}(\mathbf{r})=(x-R_x)^l(y-R_y)^m(z-R_z)^n\exp[-\alpha(\mathbf{r}-\mathbf{R})^2], \tag{9}$$

where $\mathbf{R}\equiv\{R_x,R_y,R_z\}$ is the Gaussian center and l, m, and n are integers that determine the orbital angular momentum of the basis function. The following two characteristics are evident upon inspection of Eq. (9):

(1) Translation: The basis function in Eq. (9) has a translational property which is represented by its dependence on the atom center $\mathbf{R}$.
(2) Dilation: There is a dilatory characteristic associated with Eq. (9) wherein for a given value of l, m, n, and $\mathbf{R}$, the exponent α can have multiple values. For example, many different α values may be used to construct a contracted $1s$-type or $2s$-type Gaussian basis function. The primitive Gaussians with smaller α are simply a ''dilated'' version of the original for every value of l, m, n, and $\mathbf{R}$.

Based on the above observations, we see that *the primitive Gaussian basis functions used in electronic structure theory are closely related to the theory of multiresolution analysis that has been well-studied over the last couple of decades*.[38–41] Multiresolution analysis leads to basis functions called ''wavelets'' which have, in recent times, been very popular in scientific computing.[38–41] Since wavelets share a similar property as those highlighted above, the wavelets have been broadly described as a *translation-dilation* basis set. In fact, as we will see below that Eq. (9) is an example of a *multiwavelet*.[48] To understand the similarities between the wavelet theory and the primitive Gaussian functions in Eq. (9), let us consider the following definitions for the *scaling function* and the *wavelet function* that form the basis for multiresolution analysis:

$$\xi_{i,j}(x)=a^{-j/2}\xi(xa^{+j}-i), \tag{10}$$

$$\eta_{i,j}(x)=a^{-j/2}\eta(xa^{+j}-i), \tag{11}$$

where $\xi(x)$ and $\eta(x)$ can either be orthogonal or nonorthogonal (they are generally orthogonal in all wavelet applications) and $\xi(x)$ is called the scaling function and $\eta(x)$ is called the wavelet function. In most cases $\xi(x)$ constitutes a function with average lower frequency (and hence called a low-pass filter) while the wavelet function is generally a high-pass filter since it is generally chosen to have a Fourier transform that is predominantly nonzero in the high-frequency region. Using Eqs. (10) and (11) a complete hier-

archy of basis functions are constructed and this hierarchical representation allows accurate and efficient ways to store images and solve partial differential equations when $\xi(x)$ and $\eta(x)$ are appropriately chosen. The indices i and j are generally integers although that is not a requirement of the theory and a is a scaling parameter that specifies the extent of ''dilation.'' Again, in most signal processing applications the two-scale version $(a=2)$ of these equations is common. Based on Eqs. (10) and (11) a variety of such hierarchical wavelet bases have been developed[39,40,49] and some of these have also been recently used in electronic structure calculations.[42–44]

To illustrate Eqs. (10) and (11) let us consider the example of the two-scale Haar wavelet and scaling functions. The basic Haar scaling function is a square function where $\xi(x)$ is equal to 1, for $0\leqslant x\leqslant 1$, and zero otherwise. The Haar wavelet function is the orthogonal complement of the Haar scaling function and hence $\eta(x)$ is equal to 1, for $0\leqslant x\leqslant 1/2$, and $\eta(x)$ is equal to -1 for $1/2\leqslant x\leqslant 1$, and zero otherwise. The translational and dilation properties in Eqs. (10) and (11) are then used to generate a complete hierarchy of Haar wavelet basis functions $(a=2)$. Two properties are evident upon inspection of the Haar wavelet and scaling functions: (a) localization: all the η and ξ functions are localized, (b) multiresolution: the wavelet functions with higher j values may be used to represent highly oscillatory (high frequency) functions, or highly oscillatory regions of a function, while the scaling functions with lower j values may represent flatter (low frequency) regions. Thus, if a given function displays high frequency behavior in some region and low frequency behavior in another region, the higher j wavelet functions may be used to represent the high frequency region and the low j scaling functions may be used to represent the lower frequency regions. Intermediate frequency regions may be represented using other components of the hierarchical scheme presented in Eqs. (10) and (11). In this way an accurate and efficient fit could be obtained for nonuniform functions and this is the reason wavelets have become popular in various areas of scientific computing. Compare this with a fit accomplished using a Fourier series. Here using the high-frequency Fourier functions to represent the high-frequency region leads to the Gibb's phenomenon or rapid intertwining in the low-frequency region since the Fourier functions are not local. (This aspect is further discussed later in this section.)

On comparing Eqs. (10) and (11) with Eq. (9) the following statements are apparent. The exponent α in Eq. (9) takes the role of a scaling factor similar to a^{-j} in Eqs. (10) and (11), which makes it possible to have different spreads for orbitals with different principal quantum numbers. The vector $\mathbf{R}\equiv\{R_x,R_y,R_z\}$ takes the role of translation as performed by the index i in Eqs. (10) and (11). (To be exact, we must note that in the atom-centered form of the basis sets noted in Eq. (9) the translates $\mathbf{R}$ are not uniform. This is also the case for the scaling factor. In the case of *even tempered* or *uniform* basis sets, these would be uniform and the atom-centered forms in Eq. (9) are a subset of such *even tempered* basis sets. It is advantageous to use this subset in electronic structure theory of molecular systems since the electron density is generally a maximum at the nuclear center.[50]) Due to this similarity between the Gaussian functions and wavelets, the *Mexican hat wavelet* $\{(1-x^2)\exp[-x^2/2]\}$ has been used in analyzing vision and the *Morlet wavelet* $\{\exp[-\imath\omega x]\times\exp[-x^2/2]\}$ has been used in signal processing. The Mexican hat wavelet is related to the d-type Gaussian basis derived from Eq. (9). Note the similarity between the Morlet wavelet and coherent states,[51] and also the periodic form of Eq. (9) which have been used to enforce periodic boundary conditions with Gaussian basis functions.[52]

Unlike Eqs. (10) and (11), Eq. (9) may have many basic forms [as opposed to just two forms for Eqs. (10) and (11), one scaling function and one wavelet function; for example, see the Haar wavelet discussion above]. One can have as many sets of $\{l,m,n\}$ as required for each $\mathbf{R}$, each set corresponding to a given angular momentum. In this sense, Eq. (9) is a *multi wavelet*.[48] [In fact, the Cartesian Gaussians in Eq. (9) approach a *frame* in the limit of large angular momenta.[41]]

Equations (10) and (11) can both have real values for the scaling factor a (although real values for the scaling factor are seldom used in most wavelet applications) and in this case the associated wavelet and scaling functions turn out to be nonorthogonal.[40] By contrast, the scaling parameter in Eq. (9) is always real, however, the associated nonorthogonality of Gaussian basis functions has been known for a long time, before the development of wavelets.

All this is extremely interesting since it has been recently realized that wavelets and multiwavelets have great utility in fitting diverse data and in solving partial-differential equations of various kinds. Wavelet signal processing is also replacing the standard digital Fourier transforms (used in jpeg files) to become the industry standard in image compression and data storage.[53] One of the main reasons for this is, when the scaling and wavelet functions are appropriately chosen, the multiresolution analysis theory provides a framework to accurately represent nonuniform functions that may portray a diverse spatial (or temporal) frequency dependence. When a truncated series of plane waves of some order N is used to approximate such functions, the residue is a Fourier function of order $N+1$, which is a very rapidly oscillatory function. This leads to the so-called Gibb's phenomenon[54] where the approximated function constantly intertwines around the real function. Wavelets and multiwavelets on the contrary give rise to what are known as ''well-tempered'' or smooth approximations[55,56] to fitted data and this intertwining effect is generally not present due to the localized nature of the functions chosen in Eqs. (10), (11), and (9). [An important difference between a plane-wave fit and a fit based on Eqs. (10), (11), and (9) is as follows: While an approximation to a function based on a plane-wave expansion results in a *normed* convergence to the true function, an approximation based on Eqs. (10), (11), and (9) results in a *uniform* convergence to the true function.[57] The *normed* convergence is a *weak* convergence criterion as compared to the *uniform* convergence.] Alternately, such artifacts in the plane-wave bases may be eliminated by using a ''smoothing'' or ''windowing'' process to reduce the contribution of high frequency noise.[54,58] Without the ''windowing,'' this constitutes one of

the major impediments in the accurate description of localized systems using plane-waves even when large box sizes are used.[8]

In ADMP there is an additional complication that arises from the use of Eq. (9) as the primitive basis set. The translation property of the Gaussian multiwavelet basis is time-dependent on account of moving nuclei, since basis functions in ADMP are atom-centered. This, along with the nonorthogonal nature of multiwavelets associated with nonintegral scaling factors gives rise to different levels of basis-set completeness at different nuclear geometries. *Hence the "quality" of the basis set changes as the nuclei move*. This is known as the "basis-set superposition error." However, we can now see that this problem is quite general and should occur in any multiwavelet (or wavelet) application where the wavelet centers are not stationary. In the following section we present a general treatment of such errors.

A. Basis-set superposition error (BSSE): Effect on dynamics

In order to understand the effect of "basis-set superposition error" to dynamics (both ADMP and BOMD) using basis sets such as those in Eq. (9) in particular and any multiwavelets in general, we will study how the basis set space changes during a dynamics process. Let us first introduce the set $V^{\mathbf{R}}_{\{l,m,n\}}$ to represent the vector space spanned by all the primitive Gaussian basis functions belonging to the atom centered at $\mathbf{R}$. The family of angular momentum indices $\{l,m,n\}$ takes on values depending upon the choice of Gaussian basis set. As a result of moving nuclei in ADMP and BOMD, the overlap of the vector space $V^{\mathbf{R}}_{\{l,m,n\}}$ with the vector space $V^{\mathbf{R}'}_{\{l,m,n\}}$, centered on another atom, is a time-dependent quantity. Consider the set

$$V^{\mathbf{R}'}_{\{l,m,n\}} - V^{\mathbf{R}'}_{\{l,m,n\}} \cap V^{\mathbf{R}}_{\{l,m,n\}}. \tag{12}$$

The quantity $V^{\mathbf{R}'}_{\{l,m,n\}} \cap V^{\mathbf{R}}_{\{l,m,n\}}$ is the intersection between the two vector spaces, by which is meant the orthogonal set of vectors that span the space of the overlap matrix of the vectors in $V^{\mathbf{R}}_{\{l,m,n\}}$ and $V^{\mathbf{R}'}_{\{l,m,n\}}$. The subtraction in Eq. (12) is a set-theoretic subtraction that implies the removal of the space spanned by the overlap matrix from the original space. The difference in Eq. (12) may be obtained from a Schmidt orthogonalization of the vectors in $V^{\mathbf{R}'}_{\{l,m,n\}}$ from the vectors in $V^{\mathbf{R}}_{\{l,m,n\}}$. The quantity in Eq. (12) is, however, time dependent as $\mathbf{R}$ and $\mathbf{R}'$ move with time during dynamics. The difference between the quantity in Eq. (12) at time t' and the same quantity at time t, that is

$$\{(V^{\mathbf{R}',t'}_{\{l,m,n\}} - V^{\mathbf{R}',t'}_{\{l,m,n\}} \cap V^{\mathbf{R},t'}_{\{l,m,n\}}) - (V^{\mathbf{R}',t}_{\{l,m,n\}} - V^{\mathbf{R}',t}_{\{l,m,n\}} \cap V^{\mathbf{R},t}_{\{l,m,n\}})\}, \tag{13}$$

where we have used the superscript index to identify the time dependence, represents how the basis-set space changes with time in the neighborhood of $\mathbf{R}$. In fact, Eq. (13) represents the additional basis functions that have been added (or removed) at time t', since the set-theoretic subtraction in Eq. (13) represents the removal of elements in $[V^{\mathbf{R}',t}_{\{l,m,n\}} - V^{\mathbf{R}',t}_{\{l,m,n\}} \cap V^{\mathbf{R},t}_{\{l,m,n\}}]$ from the set $[V^{\mathbf{R}',t'}_{\{l,m,n\}} - V^{\mathbf{R}',t'}_{\{l,m,n\}} \cap V^{\mathbf{R},t'}_{\{l,m,n\}}]$. It is instructive to explain in words what the expressions in Eqs. (12) and (13) mean. The subtractions used are set-theoretic subtractions and hence, if $[A-B]$ is to be obtained, as is the case in all of these equations, then this is done by removing the portion of A that also exists in B. Hence in this sense, $[A-B] \equiv [A-A\cap B]$, which is obtained by the Schmidt orthogonalization of the vectors in A from the vectors in B. This explains how the quantities in Eqs. (12) and (13) are to be viewed. (Subtracting two vectors spaces is well-known in applied mathematics and the difference between two vectors spaces is called the Bellman residual.)

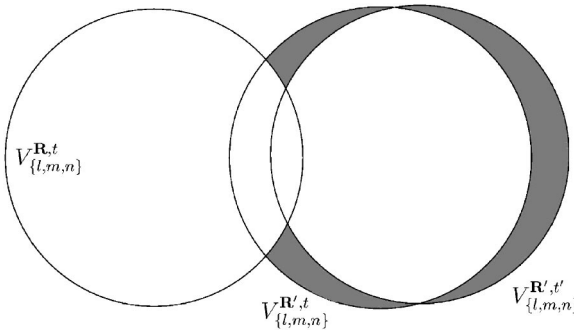


FIG. 1. Venn diagram corresponding to the set in Eq. (14), assuming that the system at $\mathbf{R}$ is stationary. The shaded area represents Eq. (14), i.e., the set comprising the basis vectors that are absent at time t and present at time t', or those that are absent at time t' but present at time t.

Since, Eq. (13) represents the additional basis functions that have been added (or removed) at time t', there must also be the opposite kind of elements, that is, those in the set, $[V^{\mathbf{R}',t}_{\{l,m,n\}} - V^{\mathbf{R}',t}_{\{l,m,n\}} \cap V^{\mathbf{R},t}_{\{l,m,n\}}]$, that are not in the set, $[V^{\mathbf{R}',t'}_{\{l,m,n\}} - V^{\mathbf{R}',t'}_{\{l,m,n\}} \cap V^{\mathbf{R},t'}_{\{l,m,n\}}]$. Hence the quantity

$$\begin{aligned}\mathcal{V}^{t,t'}_{\mathbf{R},\mathbf{R}'} \equiv &\{(V^{\mathbf{R}',t'}_{\{l,m,n\}} - V^{\mathbf{R}',t'}_{\{l,m,n\}} \cap V^{\mathbf{R},t'}_{\{l,m,n\}}) \\ &- (V^{\mathbf{R}',t}_{\{l,m,n\}} - V^{\mathbf{R}',t}_{\{l,m,n\}} \cap V^{\mathbf{R},t}_{\{l,m,n\}})\} \\ \oplus &\{(V^{\mathbf{R}',t}_{\{l,m,n\}} - V^{\mathbf{R}',t}_{\{l,m,n\}} \cap V^{\mathbf{R},t}_{\{l,m,n\}}) \\ &- (V^{\mathbf{R}',t'}_{\{l,m,n\}} - V^{\mathbf{R}',t'}_{\{l,m,n\}} \cap V^{\mathbf{R},t'}_{\{l,m,n\}})\}\end{aligned} \tag{14}$$

represents how the quality of the basis set changes in the vicinity of $\mathbf{R}$ between times t and t'. A Venn diagram representing Eq. (14) is presented in Fig. 1. (For dynamics t' may be chosen as $t+\Delta t$.) The terms in Eq. (14) are to be obtained from the direct-sum $\oplus$: The quantity $[A\oplus B]$ is simply the union of all the elements in both sets. Hence an orthogonal set belonging to $[A\oplus B]$ may be obtained from a QR decomposition of the vectors in A and B, together.

The projection of the orbitals at time t onto the vector space represented by Eq. (14) determines how the change in the quality of the basis-set expansion due to the moving nature of the atom-centered basis functions affects the physical process. This is because the basis vectors in the space $\mathcal{V}^{t,t'}_{\mathbf{R},\mathbf{R}'}$ are those vectors that are not present either at time t or at time t'. Hence the existence of a sizable component of the orbital in this vector space would ensure that the orbital is

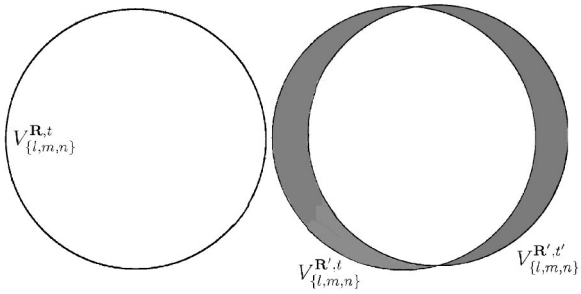


FIG. 2. Venn diagram corresponding to Eq. (15), assuming that the system at **R** is stationary. The shaded area represents the set in Eq. (15). This is similar to the case in Fig. 1 except the basis set on each atom are orthonormal as seen in the discussion in the paragraph before Eq. (15).

treated incompletely at either time t or at time t'. Thus the projection of the orbitals onto the space $\mathcal{V}^{t,t'}_{\mathbf{R},\mathbf{R}'}$ completely quantifies basis-set superposition error for a given problem.

It is interesting to consider the nature of the space $\mathcal{V}^{t,t'}_{\mathbf{R},\mathbf{R}'}$ when the basis functions centered on each atom become *large* and effectively complete. When this happens the intersection $[V^{\mathbf{R}',t}_{\{l,m,n\}} \cap V^{\mathbf{R},t}_{\{l,m,n\}} \cong V^{\mathbf{R}',t}_{\{l,m,n\}} \cong V^{\mathbf{R},t}_{\{l,m,n\}}]$ at both time t and at t'. Hence the quantity $\mathcal{V}^{t,t'}_{\mathbf{R},\mathbf{R}'}$ in Eq. (14) becomes zero (null set). In Fig. 1, this represents the fact that when all circles become suitably large, the size of the shaded portion goes to zero. This implies that basis set superposition error is completely absent for large enough atom-centered basis sets. It is also negligible when $V^{\mathbf{R}}_{\{l,m,n\}}$ and $V^{\mathbf{R}'}_{\{l,m,n\}}$ are large and not complete, but the component of an orbital in the space $\mathcal{V}^{t,t'}_{\mathbf{R},\mathbf{R}'}$ is small. In this case the vector spaces $V^{\mathbf{R}}_{\{l,m,n\}}$ and $V^{\mathbf{R}'}_{\{l,m,n\}}$ are *effectively* complete for a given problem. This is, of course, system dependent and chemistry dependent (which is why the orbital at time t enters into the discussion).

If an orthonormal multi-wavelet unlike that described in Eq. (9) were to be used, then the quantity $[V^{\mathbf{R}'}_{\{l,m,n\}} \cap V^{\mathbf{R}}_{\{l,m,n\}}]$ would be zero at all times, and, from Eq. (14), it would be the overlap of the orbitals at time t with the space,

$$[V^{\mathbf{R}',t'}_{\{l,m,n\}} - V^{\mathbf{R}',t}_{\{l,m,n\}}] \oplus [V^{\mathbf{R}',t}_{\{l,m,n\}} - V^{\mathbf{R}',t'}_{\{l,m,n\}}] \tag{15}$$

that contributes to the superposition error. This is shown in Fig. 2. Clearly, this is small when $t'-t \equiv \Delta t$ is small within a dynamics setting. [This is also the case for the quantity in Eq. (14).] However, when Δt is not negligible, this quantity remains finite, when $V^{\mathbf{R}',t}_{\{l,m,n\}}$ and $V^{\mathbf{R},t}_{\{l,m,n\}}$ are not effectively complete. It should be noted that when both $V^{\mathbf{R}',t}_{\{l,m,n\}}$ and $V^{\mathbf{R},t}_{\{l,m,n\}}$ do approach completeness, $\{[V^{\mathbf{R}',t'}_{\{l,m,n\}} - V^{\mathbf{R}',t}_{\{l,m,n\}}] \oplus [V^{\mathbf{R}',t}_{\{l,m,n\}} - V^{\mathbf{R}',t'}_{\{l,m,n\}}]\}$ is a null set as in the nonorthogonal case discussed above. This is, however, not the case when $V^{\mathbf{R}',t}_{\{l,m,n\}}$ and $V^{\mathbf{R},t}_{\{l,m,n\}}$ are not effectively complete. Thus, it is only the local and time-dependent nature of the basis functions that causes this kind of an error, and the orthogonal nature of the basis set has limited bearing on the error. This is especially interesting since recently orthonormal, but local, Wannier functions[16,17] have been used to perform Car-Parrinello dynamics.[15] A similar error may be expected in these calculations for a finite basis set, if the Wannier basis functions were time-dependent or nuclear coordinate dependent.[15]

To proceed further with our discussion of Eq. (14), we note that since the overlap of the set $\mathcal{V}^{t,t'}_{\mathbf{R},\mathbf{R}'}$ with the orbitals is crucial in determining the extent of the superposition error, an upper bound to such error for normalized orbitals may be written as

$$\|\mathcal{V}^{t,t'}_{\mathbf{R},\mathbf{R}'}\|^{C}_{2}, \tag{16}$$

where we have also used the Holder inequality and the notation $\|\cdots\|^{C}_{2}$ represents a constrained L^2 norm, where the constraint comes from the fact that we want to consider only the domain in space where a given orbital, whose basis-set superposition error we are interested in, is numerically nonzero. An unconstrained norm should present a weak upper bound.

It is interesting to now consider how the above discussion can be specialized for the case of calculating dissociation energies that have been well-studied using the counterpoise correction.[34] In such a case the two time frames being used in Eqs. (14) and (16) would be $t=0$ (the undissociated system) and $t'=\infty$ (the dissociated system). Since the overlap of the two spaces would be zero for the dissociated system, the expression in Eq. (14) becomes

$$\begin{aligned}\mathcal{V}^{0,\infty}_{\mathbf{R},\mathbf{R}'} \equiv & \{V^{\mathbf{R}',\infty}_{\{l,m,n\}} - (V^{\mathbf{R}',0}_{\{l,m,n\}} - V^{\mathbf{R}',0}_{\{l,m,n\}} \cap V^{\mathbf{R},0}_{\{l,m,n\}})\} \\ & \oplus \{(V^{\mathbf{R}',0}_{\{l,m,n\}} - V^{\mathbf{R}',0}_{\{l,m,n\}} \cap V^{\mathbf{R},0}_{\{l,m,n\}}) - V^{\mathbf{R}',\infty}_{\{l,m,n\}}\} \\ & = V^{\mathbf{R}',\infty}_{\{l,m,n\}} \oplus (V^{\mathbf{R}',0}_{\{l,m,n\}} - V^{\mathbf{R}',0}_{\{l,m,n\}} \cap V^{\mathbf{R},0}_{\{l,m,n\}}).\end{aligned} \tag{17}$$

We obtain the second expression above from the fact that the orbitals localized around **R** for the undissociated system will have no overlap with $V^{\mathbf{R}',\infty}_{\{l,m,n\}}$ (which is infinitely far from **R** at $t'=\infty$, the dissociated system). This expression implies that the dissociated system at $t=\infty$ is to be corrected by augmenting additional basis functions from the set $\mathcal{V}^{0,\infty}_{\mathbf{R},\mathbf{R}'}$ for the system at **R**.

It is useful to compare this last expression in Eq. (17) with the standard counterpoise correction[34] used in most calculations. Here, in the space $V^{\mathbf{R},\infty}_{\{l,m,n\}}$ (the dissociated system) the basis functions localized on **R**, and also basis functions localized on all other $\mathbf{R}'$ are included. Note now that Eq. (17) suggests that the orthogonal portion of $V^{\mathbf{R}'}_{\{l,m,n\}}$, that is $V^{\mathbf{R}',0}_{\{l,m,n\}} - V^{\mathbf{R}',0}_{\{l,m,n\}} \cap V^{\mathbf{R},0}_{\{l,m,n\}}$, be included along with the basis functions on **R** and $V^{\mathbf{R}',\infty}_{\{l,m,n\}}$. (The latter would have no overlap with an orbital localized at **R** and hence would not be expected to contribute much.) In this sense Eq. (17) is different from the standard counterpoise correction, in that it avoids redundancy, by including only the orthogonal component. However, it is important to note that the space spanned by adding basis functions from $V^{\mathbf{R}'}_{\{l,m,n\}}$ or by adding basis functions from $V^{\mathbf{R}',0}_{\{l,m,n\}} - V^{\mathbf{R}',0}_{\{l,m,n\}} \cap V^{\mathbf{R},0}_{\{l,m,n\}}$ is essentially the same. (Although, the former strategy may present some numerical problems where large basis sets are used for tightly bound systems.) Since a space analogous to Eq. (17) is added in as part of $V^{\mathbf{R},\infty}_{\{l,m,n\}}$ the basis-set superposition error is for-

TABLE I. Comparison of optimized geometries for water dimer.

	B3LYP			BLYP			PBE		
	6-31+G**	6-31+G*	6-31G*[a]	6-31+G**	6-31+G*	6-31G*	6-31+G**	6-31+G*	6-31G*
rms deviation (Å)[b]	0.0288	0.009	0.4605	0.0308	0.009	0.4789	0.0277	0.007	0.4601
Error in dissociation energy (kcal/mol)[c]	0.816	1.105		0.835	1.114		0.900	1.216	

[a]6-31G* basis set optimizes to the double bridge water structure with B3LYP, PBE, and BLYP in the absence of counterpoise correction. Optimizes to the correct structure when counterpoise correction is included.
[b]Distance matrix RMS difference between the optimized structures obtained with and without counterpoise correction.
[c]The difference in the dissociation energy of the water dimer with and without counterpoise correction. The dissociation energy is calculated as dimer minus two times monomer energy.

mally zero in such calculations. The corrected energy for a system comprising N monomer units within the counterpoise scheme then has the form

$$E_{\text{counterpoise}} = E(A) - \sum_{i}^{N} [E_i(A) - E_i(a_i)]. \tag{18}$$

Here $E(A)$ is the energy of the full system using a certain basis set (represented as A), $E_i(A)$ is the energy for the ith monomer using the basis set for the full system A (which is analogous to including basis functions localized on all other $\mathbf{R}'$ in the space $V^{\mathbf{R},\infty}_{\{l,m,n\}}$ as noted above) and $E_i(a_i)$ is the energy for the ith monomer using only the basis functions localized on the ith monomer. For a system containing N monomers, for example, a water cluster with N water molecules, the energy in Eq. (18) requires one to calculate $2N+1$ different energy values that go into the right-hand side of the equation to obtain a counterpoise corrected value for the total energy of the system. The $2N+1$ different energy values correspond to a single calculation for the full system and two different calculations for each monomer, one with all the basis functions and the other with only the monomer basis functions. In this fashion the total energy is corrected by introducing the over-stabilization of the monomer due to the presence of basis functions localized on other monomers in the full calculation $[E_i(A)-E_i(a_i)]$. This contribution $[E_i(A)-E_i(a_i)]$ gets smaller as the basis set localized on monomer i gets more complete and this aspect is evident from our earlier formal discussion. It is also important to note that the energy expression in Eq. (18) only depends upon the nuclear coordinates of the system and hence gradients with respect to these can be obtained to perform both geometry optimizations as well as Born-Oppenheimer dynamics on this corrected surface, which is the subject of the fllowing section.

IV. COMPUTATIONAL TESTS

To test the effect of basis-set superposition error on energy, geometry, and frequencies, we consider the weakly bound water dimer. This system is chosen due to (a) its computational simplicity, and (b) the existence of nonbonded interactions which makes this an important system for basis-set superposition error. Calculations are performed with and without counterpoise corrections at various levels of theory and basis set. The dimer geometry was optimized in two different ways: in one case the corrected expression for energy in Eq. (18) was used, and in the other case the standard energy expression, without counterpoise correction, was used. The geometries and harmonic frequencies thus obtained were compared. Dynamics calculations were also performed using ADMP, Born-Oppenheimer dynamics with (and without) counterpoise correction, and the velocity-velocity correlation function obtained from these were compared with that obtained from a Car-Parrinello calculation using plane-wave basis sets.

In Table I, the geometries and dissociation energies obtained from optimization with and without counterpoise correction at the 6-31G*, 6-31+G*, and 6-31+G** levels of basis set using B3LYP, BLYP, and PBE density functionals are compared. There is very little difference in the RMS deviations between the structures obtained beyond the 6-31+G* level of basis set. 6-31G*, on the contrary is too low a level for the water dimer system, since while the counterpoise-corrected optimization yields the correct structure for all three functionals, the optimization without counterpoise correction leads to a double-bridge water-dimer structure for all three functionals. Diffuse functions are evidently critical to get the right chemistry in this case.

The harmonic frequencies are calculated at the optimized geometries and these are presented in Fig. 3. We present here only the results from the 6-31+G* calculations. It suffices to note that above (and including) this level of basis set, the agreement is good between the counterpoise-corrected values and those without the counterpoise correction. From the figures it is seen that the low-frequency modes (less than 300 cm^{-1}) do not agree as well as the higher frequency bending, stretching, and some of the librational modes. This is because the low-frequency modes involve high-amplitude motions which cause large change in overlap and hence these are more susceptible to basis-set superposition error. All the hydrogen bond stretching modes (ca. 3500 cm^{-1}), the bending modes and some of the higher frequency librational modes are in good agreement between the calculations including counterpoise and those without, implying a lesser degree of sensitivity to basis-set superposition error.

The Fourier transform of the velocity-velocity autocorrelation function obtained from ADMP, Born-Oppenheimer, and Car-Parrinello dynamics simulations are presented in Fig. 4. Here the Born-Oppenheimer dynamics forces are derived from Eq. (18) and hence include counterpoise correction. As can be seen, the ADMP results are close to the

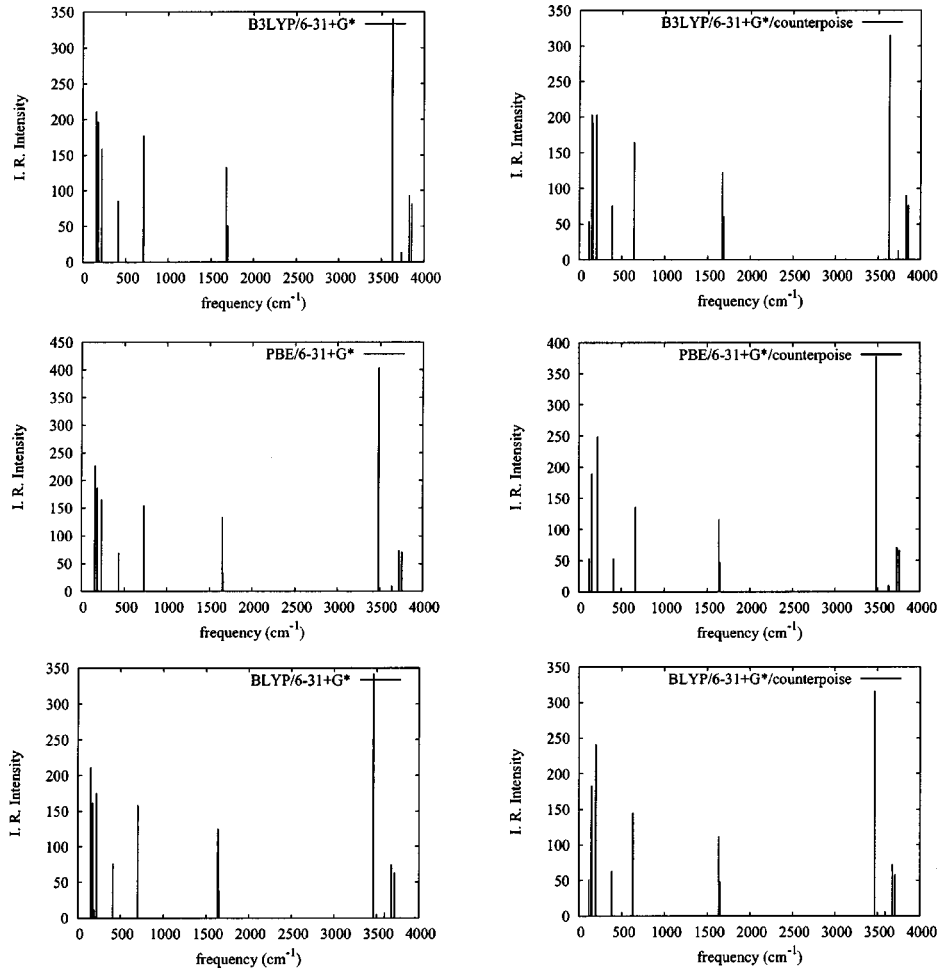


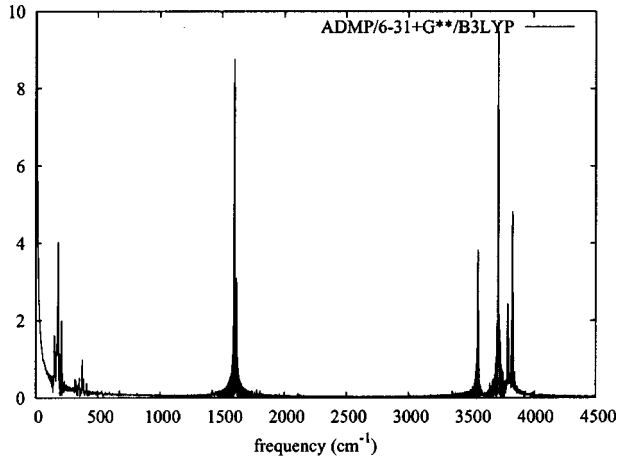
FIG. 3. Harmonic frequencies from optimized structures obtained with and without counterpoise corrections using the B3LYP, BLYP, and BPBE and the 6-31+G* basis set. For counterpoise corrections optimization was performed on the modified surface denoted by Eq. (18).

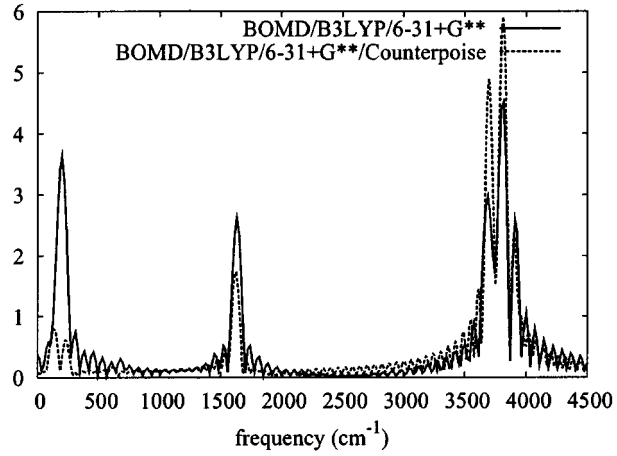
BOMD results. As a further example we have also presented a comparison between the oxygen-oxygen radial distribution functions obtained from ADMP and BOMD simulations in Fig. 5. As can be seen, these are in good agreement as well.

The Car-Parrinello calculations were performed using deuterium atoms instead of hydrogen atoms, as is standard in current literature,[8] and by using a plane-wave cutoff of 90 Ryd. The fictitious mass was chosen to be 1100 a.u. (common in current simulations[8]) with a time step of about 0.1 fs. The corresponding ADMP fictitious mass was 0.1 a.m.u. bohr$^2 \approx 180$ a.u. with a time step of 0.25 fs which is possible due to the tensorial mass-weighting scheme used in ADMP.[22] The frequencies thus obtained from the Car-Parrinello simulations were harmonically corrected (by multiplying by a factor of $\sqrt{2}$) to obtain the corresponding hydrogen spectrum. (While this is based on a harmonic bond approximation, clearly anharmonic forces are sampled during a molecular dynamics simulation. As a result this is only an estimated correction factor.) The fictitious mass was chosen based on previous studies[8] to allow for reasonably large time steps, however, recent studies[9] indicate that a lower fictitious mass would have to be chosen to allow for more accurate dynamics. The lower fictitious mass would necessitate smaller time steps in the Car-Parrinello approach. It is also seen from Fig. 4 that while the ADMP results are close to those from counterpoise-corrected BOMD, the results from the CP simulations are uniformly red-shifted. Such systematic bias in frequencies in the CP calculations has been noted previously,[9,10,23] and this is the result of the larger fictitious mass used in standard CP simulations. (Furthermore, our formal analysis[24] has shown that deviations from the Born-Oppenheimer surface is directly proportional to the size of the fictitious mass.) Nuclear mass rescaling[10] has been suggested as an approach to overcome this problem if lower values for the fictitious mass are not used,[9] however, we have not used this in the current CP simulations.

V. CONCLUSIONS

In this paper we have performed a rigorous analysis of primitive Gaussian basis sets and the effect of using such functions within an atom-centered *ab initio* dynamics scheme. Our analysis has led to important connections between the Gaussian basis functions and the wavelet theory that is now popular in computational fluid mechanics, digital signal processing, and other areas of applied mathematics and scientific computing. These connections lead to a general treatment of the basis-set superposition error; we find that





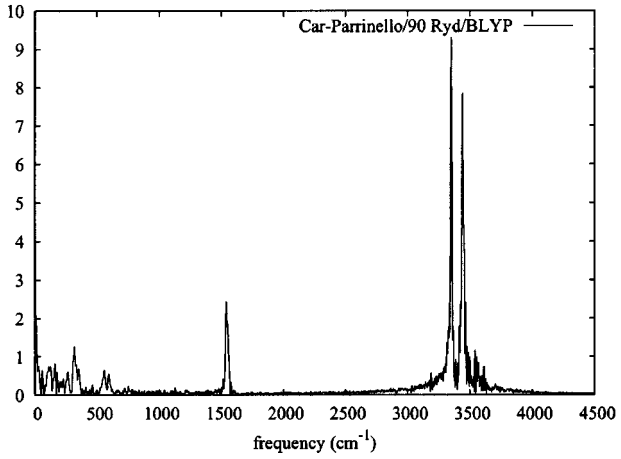


FIG. 4. Fourier transform of the velocity-velocity autocorrelation function obtained from ADMP, Car-Parrinello, and Born-Oppenheimer dynamics calculations. The Car-Parrinello calculations were performed using deuterium atoms and hence the resulting spectrum was scaled by the factor $\sqrt{2}$ to convert to the approximate proton spectrum. As can be seen the Car-Parrinello results are slightly red-shifted compared to the ADMP and Born-Oppenheimer results, while the ADMP and Born-Oppenheimer results are in reasonable agreement.

this is closely related to the time dependence of the instantaneous incompleteness of basis sets that have a multiwavelet character, of which the atom-centered Gaussian basis functions are one example. The standard counterpoise corrections appear as a special case of our discussion. Our

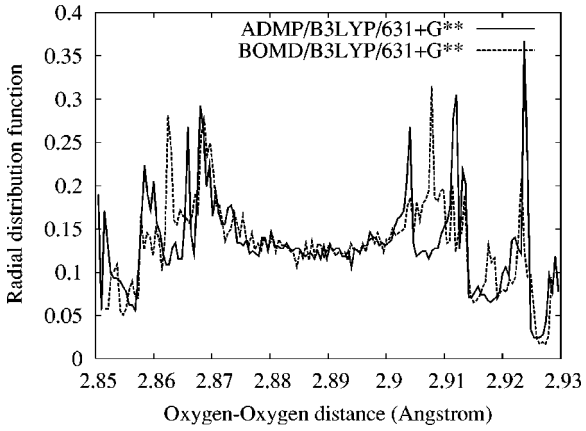


FIG. 5. Radial O-O distribution function from BOMD and ADMP simulations.

analysis also rigorously leads to the expected result that as the size of the atom-centered Gaussian basis set increases, the basis set superposition error goes to zero.

We have carried out computational tests on the weakly bound water-dimer system to critically ascertain the limits of *ab initio* dynamics approaches that use atom-centered Gaussian basis functions insofar as basis set size and superposition error are concerned. We find that beyond 6-31+G* the errors in geometry, energy, frequency, and structure distribution are within the accuracy of the density functional used. For more strongly bound chemically reactive systems one could expect to use a slightly lower level of basis set.

Our analysis also provides new ways to ascertain the quality of a basis set during dynamics. This will be used, in future publications, to adaptively control and change the quality of the basis set during dynamics.

ACKNOWLEDGMENT

S.S.I. gratefully acknowledges the Camille and Henry Dreyfus Foundation for their support.

[1] I. S. Y. Wang and M. Karplus, J. Am. Chem. Soc. **95**, 8160 (1973).
[2] C. Leforestier, J. Chem. Phys. **68**, 4406 (1978).
[3] R. Car and M. Parrinello, Phys. Rev. Lett. **55**, 2471 (1985).
[4] M. C. Payne, M. P. Teter, D. C. Allan, T. A. Arias, and J. D. Joannopoulos, Rev. Mod. Phys. **64**, 1045 (1992).
[5] E. Deumens, A. Diz, R. Longo, and Y. Öhrn, Rev. Mod. Phys. **66**, 917 (1994).
[6] H. B. Schlegel, J. M. Millam, S. S. Iyengar, G. A. Voth, A. D. Daniels, G. E. Scuseria, and M. J. Frisch, J. Chem. Phys. **114**, 9758 (2001).
[7] D. K. Remler and P. A. Madden, Mol. Phys. **70**, 921 (1990).
[8] D. Marx and J. Hutter, *Modern Methods and Algorithms of Quantum Chemistry* (John von-Neumann Institute for Computing, Julich, 2000), Vol. 1, p. 301.
[9] J. C. Grossman, E. Schwegler, E. W. Draeger, F. Gygi, and G. Galli, J. Chem. Phys. **120**, 300 (2004).
[10] P. Tangney and S. Scandolo, J. Chem. Phys. **116**, 14 (2002).
[11] H. C. Andersen, J. Chem. Phys. **72**, 2384 (1980).
[12] M. Parrinello and A. Rahman, Phys. Rev. Lett. **45**, 1196 (1980).
[13] G. Lippert, J. Hutter, and M. Parrinello, Theor. Chem. Acc. **103**, 124 (1999).
[14] G. Lippert, J. Hutter, and M. Parrinello, Mol. Phys. **92**, 477 (1997).
[15] M. Sharma, Y. Wu, and R. Car, Int. J. Quantum Chem. **95**, 821 (2003).
[16] W. Kohn, Phys. Rev. **115**, 809 (1959).
[17] N. Marzari and D. Vanderbilt, Phys. Rev. B **56**, 12847 (1997).

[18] D. A. Gibson, I. V. Ionova, and E. A. Carter, Chem. Phys. Lett. **240**, 261 (1995).
[19] B. Hartke and E. A. Carter, J. Chem. Phys. **97**, 6569 (1992).
[20] B. Hartke and E. A. Carter, Chem. Phys. Lett. **189**, 358 (1992).
[21] G. Martyna, C. Cheng, and M. L. Klein, J. Chem. Phys. **95**, 1318 (1991).
[22] S. S. Iyengar, H. B. Schlegel, J. M. Millam, G. A. Voth, G. E. Scuseria, and M. J. Frisch, J. Chem. Phys. **115**, 10291 (2001).
[23] H. B. Schlegel, S. S. Iyengar, X. Li, J. M. Millam, G. A. Voth, G. E. Scuseria, and M. J. Frisch, J. Chem. Phys. **117**, 8694 (2002).
[24] S. S. Iyengar, H. B. Schlegel, G. A. Voth, J. M. Millam, G. E. Scuseria, and M. J. Frisch, Isr. J. Chem. **42**, 191 (2002).
[25] S. S. Iyengar, H. B. Schlegel, and G. A. Voth, J. Phys. Chem. A **107**, 7269 (2003).
[26] N. Rega, S. S. Iyengar, G. A. Voth, H. B. Schlegel, and M. J. Frisch, J. Phys. Chem. B **108**, 4210 (2004).
[27] G. E. Scuseria, J. Phys. Chem. A **103**, 4782 (1999).
[28] S. S. Iyengar, M. K. Petersen, T. J. F. Day, C. J. Burnham, and G. A. Voth, J. Am. Chem. Soc. (to be published).
[29] M. K. Petersen, S. S. Iyengar, T. J. F. Day, C. J. Burnham, and G. A. Voth, J. Phys. Chem. B (to be published).
[30] S. S. Iyengar, Phys. Rev. Lett. (to be published).
[31] S. F. Boys and F. Bernardi, Mol. Phys. **19**, 553 (1970).
[32] E. Clementi, J. Chem. Phys. **46**, 3851 (1967).
[33] B. Liu and A. D. McLean, J. Chem. Phys. **59**, 4557 (1973).
[34] E. R. Davidson and S. J. Chakravorty, Chem. Phys. Lett. **217**, 48 (1994).
[35] A. K. Rappe and E. R. Bernstein, J. Phys. Chem. A **104**, 6117 (2000).
[36] F. Chen and E. R. Davidson, Chem. Phys. Lett. **360**, 99 (2002).
[37] M. J. Frisch, G. W. Trucks, H. B. Schlegel *et al.*, GAUSSIAN 03, Revision b.02, Gaussian, Inc., Pittsburgh, PA, 2003.
[38] A. Grossman and J. Morlet, SIAM J. Math. Anal. **15**, 723 (1984).
[39] G. Strang, SIAM Rev. **31**, 613 (1989).
[40] I. Daubechies, *Ten Lectures in Wavelets* (SIAM, Philadelphia, PA, 1992).
[41] G. Strang and T. Nguyen, Wavelets and Filter Banks (Wellesley-Cambridge, Wellesley, MA, 1996).
[42] T. Arias, Rev. Mod. Phys. **71**, 267 (1999).
[43] S. Goedecker, Rev. Mod. Phys. **71**, 1085 (1999).
[44] B. R. Johnson, J. P. Modisette, P. J. Nordlander, and J. L. Kinsey, J. Chem. Phys. **110**, 8309 (1999).
[45] G. H. Golub and C. F. van Loan, *Matrix Computations* (The Johns Hopkins University Press, Baltimore, 1996).
[46] P. Pulay, Mol. Phys. **17**, 197 (1969).
[47] F. Riesz and B. Sz.-Nagy, *Functional Analysis* (Dover, Inc., New York, 1990).
[48] G. Strang and V. Strela, J. Opt. Eng. **33**, 2104 (1994).
[49] D. K. Hoffman, G. W. Wei, D. S. Zhang, and D. J. Kouri, Phys. Rev. E **57**, 6152 (1998).
[50] In quantum chemistry linear combinations of Eq. (9) are used as basis functions to represent the electronic structure. Our analysis follows for these as well.
[51] J. R. Klauder and I. Daubechies, Phys. Rev. Lett. **52**, 1161 (1984).
[52] K. N. Kudin and G. E. Scuseria, Phys. Rev. B **61**, 16440 (2000).
[53] Z. Shi, G. W. Wei, D. J. Kouri, D. K. Hoffman, and Z. Bao, IEEE Trans. Image Process. **10**, 1488 (2001).
[54] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C* (Cambridge University Press, New York, 1992).
[55] D. K. Hoffman, N. Nayar, O. A. Sharafeddin, and D. J. Kouri, J. Phys. Chem. **95**, 8299 (1991).
[56] D. J. Kouri and D. K. Hoffman, Phys. Rev. Lett. **85**, 5263 (2000).
[57] C. Chandler and A. Gibson, J. Approx. Theory **100**, 233 (1999).
[58] Y. Huang, S. S. Iyengar, D. J. Kouri, and D. K. Hoffman, J. Chem. Phys. **105**, 927 (1996).