

# ***Ab initio* molecular dynamics: Propagating the density matrix with Gaussian orbitals**

H. Bernhard Schlegel and John M. Millam

*Department of Chemistry, Wayne State University, Detroit, Michigan 48202-3489*

Srinivasan S. Iyengar and Gregory A. Voth

*Department of Chemistry and Henry Eyring Center for Theoretical Chemistry, 315 S. 1400 E., Room 2020, University of Utah, Salt Lake City, Utah 84112-0850*

Andrew D. Daniels and Gustavo E. Scuseria

*Department of Chemistry and Center for Nanoscale Science and Technology, Rice University, Houston, Texas 77005-1892*

Michael J. Frisch

*Gaussian, Inc., 140 Washington Ave., North Haven, Connecticut 06473-1712*

(Received 28 December 2000; accepted 27 March 2001)

We propose and implement an alternative approach to the original Car–Parrinello method where the density matrix elements (instead of the molecular orbitals) are propagated together with the nuclear degrees of freedom. Our new approach has the advantage of leading to an  $\mathcal{O}(N)$  computational scheme in the large system limit. Our implementation is based on atom-centered Gaussian orbitals, which are especially suited to deal effectively with general molecular systems. The methodology is illustrated by applications to the three-body dissociation of triazine and to the dynamics of a cluster of a chloride ion with 25 water molecules. © 2001 American Institute of Physics.  
[DOI: 10.1063/1.1372182]

## **I. INTRODUCTION**

The Car–Parrinello (CP) (Ref. 1) method has become a central tool in *ab initio* molecular dynamics (MD) simulations<sup>2–4</sup> over the last several years. The CP method is essentially an extended Lagrangian MD scheme in which the electronic degrees of freedom are not iterated to convergence at each step, but are instead treated as fictitious dynamical variables and propagated along with the nuclear degrees of freedom by a simple adjustment of time scales. The resultant energy surface remains close to a converged adiabatic electronic surface. Such an approach offers significant computational advantages for efficient MD calculations. Numerous important examples of applications with density functional theory (DFT) and the CP method are now well documented in the literature (see, e.g., Ref. 5). In the original CP approach, the Kohn–Sham molecular orbitals  $\psi_i$  were chosen as dynamical variables to represent the electronic degrees of freedom.<sup>1</sup> However, this is not the only possible choice, and in this paper we propose an alternative approach in which the individual elements of the reduced one-particle density matrix,  $\mathbf{P}$ , are explicitly treated as dynamical variables.

Car–Parrinello calculations are usually carried out in a plane wave basis. These form an orthonormal basis that is independent of the atom positions. Although most of the integrals can be calculated easily using fast Fourier transform methods, a very large basis set is needed to represent the orbitals to the desired accuracy. The problem is less severe if pseudopotentials are used to replace the core electrons, since the plane wave basis does not easily describe the cusps and high electron density near the nuclei. Recently, Gaussian or-

bitals have also become a useful adjunct to treat this problem.<sup>6</sup>

Quantum chemistry calculations for molecules typically use atom centered Gaussian basis functions. Compared to a plane wave basis, far fewer Gaussians are needed to achieve a desired accuracy, since they are localized in regions of higher electron density and follow the nuclei as they move. The necessary integrals are somewhat more difficult to calculate than for a plane wave basis, but very efficient algorithms are available to calculate these integrals and their derivatives with respect to the atom positions.<sup>7</sup> Atom centered basis functions have been used extensively in Born–Oppenheimer dynamics (for a review, see Ref. 8). Gaussian basis functions have also been used with Car–Parrinello dynamics in generalized valence bond (GVB) (Refs. 9, 10) and Hartree–Fock<sup>11</sup> calculations where orbital and wave function coefficients are propagated in a scheme that has some difficulty with energy conservation.<sup>9</sup> As shown below, energy conservation and adiabaticity are well controlled in the present method.

It is now well established that molecular calculations with atom centered functions can be carried out in computational times that scale linearly with system size in the large system limit.<sup>12–14</sup> This has been achieved through the use of Gaussian orbitals in Kohn–Sham DFT calculations, fast multipole methods for the Coulomb problem,<sup>15,16</sup> and density matrix search alternatives with sparse matrix multiplication techniques to bypass the  $\mathcal{O}(N^3)$  Hamiltonian diagonalization bottleneck.<sup>14,17,18</sup> For similar reasons, our approach using density matrices and Gaussians for *ab initio* MD has the fundamental advantage of leading to  $\mathcal{O}(N)$  scaling of com-

putational time with system size. The literature on  $\mathcal{O}(N)$  electronic structure methods using alternative approaches is extensive (see Ref. 12, and references therein).

In the present paper, we outline the first steps toward an efficient *ab initio* method for simulating molecular dynamics of general molecular systems by propagating the density represented by a Gaussian basis. This paper is organized as follows: In Sec. II, we present the theoretical framework for our method. Following this, in Sec. III, we consider two illustrative examples: the trajectories for photodissociation of 1,3,5 triazine and the solvation of a chloride ion by a cluster of 25 water molecules. In Sec. IV we present our concluding remarks.

## II. THEORY

Like density matrix search methods for calculating electronic energies,<sup>17,18</sup> the equations for propagation of the density matrix are simplest in an orthonormal basis. In an orthonormal basis, an extended Lagrangian for the system can be written as

$$\mathcal{L} = \frac{1}{2} \text{Tr}(\mathbf{V}^T \mathbf{M} \mathbf{V}) + \frac{1}{2} \mu \text{Tr}(\mathbf{W} \mathbf{W}) - E(\mathbf{R}, \mathbf{P}) - \text{Tr}[\mathbf{\Lambda}(\mathbf{P} \mathbf{P} - \mathbf{P})], \quad (1)$$

where  $\mathbf{M}$ ,  $\mathbf{R}$ , and  $\mathbf{V}$  are the nuclear masses, positions, and velocities, respectively. The density matrix, density matrix velocity, and the fictitious mass for the electronic degrees of freedom are  $\mathbf{P}$ ,  $\mathbf{W}$ , and  $\mu$ , respectively. (Note that  $\mu$  does not have the units of mass since  $\mathbf{P}$  does not have the units of distance.) Constraints on the total number of electrons,  $N_e$ , and on the idempotency of the density matrix are imposed using a Lagrangian multiplier matrix  $\mathbf{\Lambda}$ .

Although Gaussian basis functions are not orthogonal, the Hamiltonian matrix and the density matrix,  $\mathbf{P}'$ , can easily be transformed to an orthonormal basis using  $\mathbf{P} = \mathbf{U} \mathbf{P}' \mathbf{U}^T$ , where  $\mathbf{U}^T \mathbf{U} = \mathbf{S}'$ , the overlap matrix. There are a number of choices for this transformation (e.g.,  $\mathbf{U}$  can be obtained by Cholesky decomposition,<sup>19</sup> or  $\mathbf{U} = \mathbf{S}'^{1/2}$  for Löwdin symmetric orthonormalization). This matrix can also include an additional transformation so that overall rotation of the system is factored out of the propagation of the density.

As in the conjugate gradient density matrix search method<sup>17,18</sup> the energy is calculated using the McWeeny purification transformation of the density,<sup>20</sup>  $\tilde{\mathbf{P}} = 3\mathbf{P}^2 - 2\mathbf{P}^3$ . This leads to

$$E = \text{Tr} \left[ \mathbf{h}' \tilde{\mathbf{P}}' + \frac{1}{2} \mathbf{G}'(\tilde{\mathbf{P}}') \tilde{\mathbf{P}}' \right] + V_{NN} = \text{Tr} \left[ \mathbf{h} \tilde{\mathbf{P}} + \frac{1}{2} \mathbf{G}(\tilde{\mathbf{P}}) \tilde{\mathbf{P}} \right] + V_{NN}, \quad (2)$$

where  $\mathbf{h}'$  and  $\mathbf{G}'(\tilde{\mathbf{P}}')$  are the one and two electron matrices in the nonorthogonal basis and  $V_{NN}$  is the nuclear repulsion energy. In the orthonormal basis, these matrices are  $\mathbf{h} = \mathbf{U}^{-T} \mathbf{h}' \mathbf{U}^{-1}$ , etc. In light of Eq. (2), the Lagrangian in Eq. (1) is invariant with respect to unitary transformations among orthogonal bases.

Using the principle of stationary action<sup>21</sup> for the Lagrangian in Eq. (1), the Euler–Lagrange equation for the

propagation of the density (in the orthonormal basis) and the propagation of the nuclei may be obtained as

$$\mu \frac{d^2 \mathbf{P}}{dt^2} = - \left[ \frac{\partial E(\mathbf{R}, \mathbf{P})}{\partial \mathbf{P}} \Big|_{\mathbf{R}} + \mathbf{\Lambda} \mathbf{P} + \mathbf{P} \mathbf{\Lambda} - \mathbf{\Lambda} \right], \quad (3)$$

$$\mathbf{M} \frac{d^2 \mathbf{R}}{dt^2} = - \frac{\partial E(\mathbf{R}, \mathbf{P})}{\partial \mathbf{R}} \Big|_{\mathbf{P}}. \quad (4)$$

In Eq. (3),  $\mu$  is a scalar and all elements of the density matrix are weighted identically. Generalizing  $\mu$  to a matrix will be considered in a subsequent paper.<sup>22</sup> A solution to Eq. (3) represents the dynamics of the one-electron density matrix in orthonormal bases. The nonorthogonal (Gaussian) basis is used only to calculate the energy,  $E$ , and its derivatives,  $[\partial E(\mathbf{R}, \mathbf{P})/\partial \mathbf{P}]_{\mathbf{R}}$  and  $[\partial E(\mathbf{R}, \mathbf{P})/\partial \mathbf{R}]_{\mathbf{P}}$ . But, since Gaussian basis sets are atom centered, it is important to note that the transformation matrix  $\mathbf{U}$  is time dependent, i.e., the relation between the nonorthogonal and the orthonormal basis is different for each time  $t$ .

The derivative of the energy with respect to the density matrix in the orthonormal basis is

$$\frac{\partial E(\mathbf{R}, \mathbf{P})}{\partial \mathbf{P}} \Big|_{\mathbf{R}} = 3\mathbf{F} \mathbf{P} + 3\mathbf{P} \mathbf{F} - 2\mathbf{F} \mathbf{P}^2 - 2\mathbf{P} \mathbf{F} \mathbf{P} - 2\mathbf{P}^2 \mathbf{F}, \quad (5)$$

where  $\mathbf{F}$  is the Fock matrix  $[\mathbf{F} \equiv \mathbf{h} + \mathbf{G}(\tilde{\mathbf{P}})]$  in the orthonormal basis built with the McWeeny purified density.

Equations (3) and (4) may be integrated using the velocity Verlet algorithm.<sup>23</sup> The contributions for the density and density velocity are given by

$$\mathbf{P}_{i+1} = \mathbf{P}_i + \mathbf{W}_i \Delta t - \frac{\Delta t^2}{2\mu} \left[ \frac{\partial E(\mathbf{R}_i, \mathbf{P}_i)}{\partial \mathbf{P}} \Big|_{\mathbf{R}} + \mathbf{\Lambda}_i \mathbf{P}_i + \mathbf{P}_i \mathbf{\Lambda}_i - \mathbf{\Lambda}_i \right], \quad (6)$$

$$\begin{aligned} \mathbf{W}_{i+1/2} &= \mathbf{W}_i - \frac{\Delta t}{2\mu} \left[ \frac{\partial E(\mathbf{R}_i, \mathbf{P}_i)}{\partial \mathbf{P}} \Big|_{\mathbf{R}} + \mathbf{\Lambda}_i \mathbf{P}_i + \mathbf{P}_i \mathbf{\Lambda}_i - \mathbf{\Lambda}_i \right] \\ &= \frac{\mathbf{P}_{i+1} - \mathbf{P}_i}{\Delta t}, \end{aligned} \quad (7)$$

$$\begin{aligned} \mathbf{W}_{i+1} &= \mathbf{W}_{i+1/2} - \frac{\Delta t}{2\mu} \left[ \frac{\partial E(\mathbf{R}_{i+1}, \mathbf{P}_{i+1})}{\partial \mathbf{P}} \Big|_{\mathbf{R}} + \mathbf{\Lambda}_{i+1} \mathbf{P}_{i+1} \right. \\ &\quad \left. + \mathbf{P}_{i+1} \mathbf{\Lambda}_{i+1} - \mathbf{\Lambda}_{i+1} \right], \end{aligned} \quad (8)$$

where the Lagrangian constraint matrix for the  $i$ th timestep i.e.,  $\mathbf{\Lambda}_i$ , is chosen to satisfy  $\text{Tr}[\mathbf{P}_{i+1}] = N_e$  and  $\mathbf{P}_{i+1}^2 = \mathbf{P}_{i+1}$ . The constraint terms above contribute only to the occupied–occupied and virtual–virtual blocks of the density matrix, since  $\mathbf{P}_i[\mathbf{\Lambda}_i \mathbf{P}_i + \mathbf{P}_i \mathbf{\Lambda}_i - \mathbf{\Lambda}_i] \mathbf{Q}_i = \mathbf{Q}_i[\mathbf{\Lambda}_i \mathbf{P}_i + \mathbf{P}_i \mathbf{\Lambda}_i - \mathbf{\Lambda}_i] \mathbf{P}_i = 0$ , where  $\mathbf{Q} = \mathbf{I} - \mathbf{P}$  and  $\mathbf{I}$  is the identity matrix. Thus, instead of solving for  $\mathbf{\Lambda}_i$ , we can solve directly for  $\mathbf{P}_{i+1}$  by a simple iterative procedure that minimizes  $\text{Tr}[(\mathbf{P}_{i+1}^2 - \mathbf{P}_{i+1})^2]$ . Starting with  $\mathbf{P}_{i+1} = \mathbf{P}_i + \mathbf{W}_i \Delta t - [\partial E(\mathbf{R}_i, \mathbf{P}_i)/\partial \mathbf{P}]_{\mathbf{R}} (\Delta t^2/2\mu)$  we iterate

$$\mathbf{P}_{i+1} \leftarrow \mathbf{P}_{i+1} + \mathbf{P}_i \mathbf{T} \mathbf{P}_i + \mathbf{Q}_i \mathbf{T} \mathbf{Q}_i, \quad (9)$$

where  $\mathbf{T} = \tilde{\mathbf{P}}_{i+1} - \mathbf{P}_{i+1} = 3\mathbf{P}_{i+1}^2 - 2\mathbf{P}_{i+1}^3 - \mathbf{P}_{i+1}$ . The iteration converges rapidly (as will be seen in the results section) and is stopped when  $\{\text{Tr}[(\mathbf{P}_{i+1}^2 - \mathbf{P}_{i+1})^2]\}^{1/2}/N < 10^{-12}$ . The final density velocity needs to satisfy the time derivative of the idempotency condition,  $\mathbf{W}_{i+1}\mathbf{P}_{i+1} + \mathbf{P}_{i+1}\mathbf{W}_{i+1} = \mathbf{W}_{i+1}$ . This can be solved exactly to yield  $\mathbf{\Lambda}_{i+1} = [2\mathbf{P}_{i+1}\mathbf{W}_{i+1}^*\mathbf{P}_{i+1} - \mathbf{W}_{i+1}^*](2\mu/\Delta t)$ , where  $\mathbf{W}_{i+1}^* = \mathbf{W}_{i+1/2} - [\partial E(\mathbf{R}_{i+1}, \mathbf{P}_{i+1})/\partial \mathbf{P} |_{\mathbf{R}}](\Delta t/2\mu)$ . With this, Eq. (8) simplifies to

$$\mathbf{W}_{i+1} = \mathbf{W}_{i+1}^* - \mathbf{P}_{i+1}\mathbf{W}_{i+1}^*\mathbf{P}_{i+1} - \mathbf{Q}_{i+1}\mathbf{W}_{i+1}^*\mathbf{Q}_{i+1}. \quad (10)$$

In an orthonormal basis, the constraints do not depend on the nuclear coordinates; hence, there are no constraints for the propagation of the nuclear coordinates. Since the energy is calculated using the purified density, the derivatives with respect to the nuclear positions [for Eq. (4)] are also calculated using the purified density. If atom centered basis functions are used, the derivative of the energy with respect to the nuclei consists of not only the Hellmann–Feynman term but also the Pulay (or wave function derivative) contributions,<sup>24</sup>

$$\begin{aligned} \left. \frac{\partial E}{\partial \mathbf{R}} \right|_{\mathbf{P}} &= \text{Tr} \left[ \left. \frac{d\mathbf{h}'}{d\mathbf{R}} \tilde{\mathbf{P}}' + \frac{1}{2} \frac{\partial \mathbf{G}'(\tilde{\mathbf{P}}')}{\partial \mathbf{R}} \right|_{\tilde{\mathbf{P}}'} + \mathbf{F}' \left. \frac{\partial \tilde{\mathbf{P}}'}{\partial \mathbf{R}} \right|_{\tilde{\mathbf{P}}'} \right] + \frac{\partial V_{NN}}{\partial \mathbf{R}} \\ &= \text{Tr} \left[ \left. \frac{d\mathbf{h}}{d\mathbf{R}} \tilde{\mathbf{P}} + \frac{1}{2} \frac{\partial \mathbf{G}(\tilde{\mathbf{P}})}{\partial \mathbf{R}} \right|_{\tilde{\mathbf{P}}} \right] + \frac{\partial V_{NN}}{\partial \mathbf{R}}. \end{aligned} \quad (11)$$

We have used, here, the fact that  $\mathbf{h}$  and  $\mathbf{h}'$  are not functions of  $\mathbf{P}$ , and hence  $[\partial \mathbf{h}/\partial \mathbf{R}]_{\mathbf{P}} \equiv d\mathbf{h}/d\mathbf{R}$  and similarly for  $\mathbf{h}'$ .

The derivatives of the integrals in the orthonormal basis depend not only on the derivatives of the integrals in the nonorthogonal, atomic orbital basis, but also on the derivatives of the transformation to the orthonormal basis. For example, the derivative of the one electron matrix in the orthonormal basis is given by

$$\begin{aligned} \frac{d\mathbf{h}}{d\mathbf{R}} &= \frac{d\mathbf{U}^{-T}\mathbf{h}'\mathbf{U}^{-1}}{d\mathbf{R}} \\ &= \mathbf{U}^{-T} \frac{d\mathbf{h}'}{d\mathbf{R}} \mathbf{U}^{-1} + \frac{d\mathbf{U}^{-T}}{d\mathbf{R}} \mathbf{h}' \mathbf{U}^{-1} + \mathbf{U}^{-T} \mathbf{h}' \frac{d\mathbf{U}^{-1}}{d\mathbf{R}} \\ &= \mathbf{U}^{-T} \frac{d\mathbf{h}'}{d\mathbf{R}} \mathbf{U}^{-1} + \frac{d\mathbf{U}^{-T}}{d\mathbf{R}} \mathbf{U}^T \mathbf{h} + \mathbf{h} \mathbf{U} \frac{d\mathbf{U}^{-1}}{d\mathbf{R}} \\ &= \mathbf{U}^{-T} \frac{d\mathbf{h}'}{d\mathbf{R}} \mathbf{U}^{-1} - \mathbf{U}^{-T} \frac{d\mathbf{U}^T}{d\mathbf{R}} \mathbf{h} - \mathbf{h} \frac{d\mathbf{U}}{d\mathbf{R}} \mathbf{U}^{-1}. \end{aligned} \quad (12)$$

This makes use of the relation  $\mathbf{U}[d\mathbf{U}^{-1}/d\mathbf{R}] = -[d\mathbf{U}/d\mathbf{R}]\mathbf{U}^{-1}$  which can be obtained by differentiating  $\mathbf{U}\mathbf{U}^{-1} = \mathbf{I}$ . The same terms arise if we compute the density matrix derivative in the non-orthogonal basis,

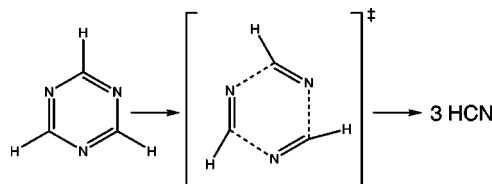


FIG. 1. Photodissociation of 1,3,5 triazine.

$$\begin{aligned} \left. \frac{\partial \tilde{\mathbf{P}}'}{\partial \mathbf{R}} \right|_{\mathbf{P}} &= \left. \frac{\partial \mathbf{U}^{-1} \tilde{\mathbf{P}} \mathbf{U}^{-T}}{\partial \mathbf{R}} \right|_{\mathbf{P}} \\ &= \frac{d\mathbf{U}^{-1}}{d\mathbf{R}} \tilde{\mathbf{P}} \mathbf{U}^{-T} + \mathbf{U}^{-1} \tilde{\mathbf{P}} \frac{d\mathbf{U}^{-T}}{d\mathbf{R}} \\ &= \frac{d\mathbf{U}^{-1}}{d\mathbf{R}} \mathbf{U} \tilde{\mathbf{P}}' + \tilde{\mathbf{P}}' \mathbf{U}^T \frac{d\mathbf{U}^{-T}}{d\mathbf{R}} \\ &= -\mathbf{U}^{-1} \frac{d\mathbf{U}}{d\mathbf{R}} \tilde{\mathbf{P}}' - \tilde{\mathbf{P}}' \frac{d\mathbf{U}^T}{d\mathbf{R}} \mathbf{U}^{-T}, \end{aligned} \quad (13)$$

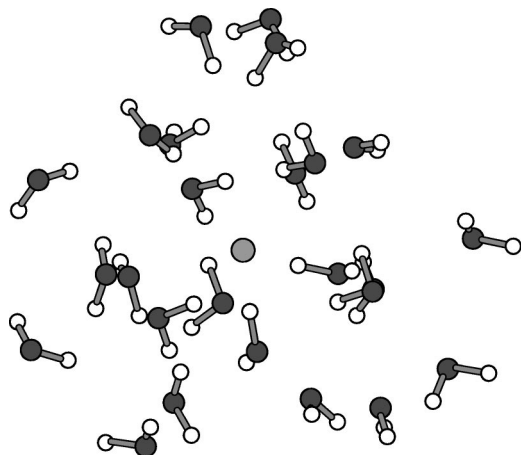
where we have substituted  $\partial \mathbf{U}/\partial \mathbf{R} |_{\mathbf{P}} = d\mathbf{U}/d\mathbf{R}$ , since  $\mathbf{U}$  is explicitly only a function of  $\mathbf{R}$ , and not of  $\mathbf{P}$ , and depends on time only through its dependence on  $\mathbf{R}$ . Inserting Eq. (12) or (13) into Eq. (11), the derivative of the energy with respect to the nuclei becomes

$$\begin{aligned} \left. \frac{\partial E}{\partial \mathbf{R}} \right|_{\mathbf{P}} &= \text{Tr} \left[ \left. \frac{d\mathbf{h}'}{d\mathbf{R}} \tilde{\mathbf{P}}' + \frac{1}{2} \frac{\partial \mathbf{G}'(\tilde{\mathbf{P}}')}{\partial \mathbf{R}} \right|_{\tilde{\mathbf{P}}'} \right] \\ &\quad - \text{Tr} \left[ \mathbf{F}' \mathbf{U}^{-1} \frac{d\mathbf{U}}{d\mathbf{R}} \tilde{\mathbf{P}}' + \tilde{\mathbf{P}}' \frac{d\mathbf{U}^T}{d\mathbf{R}} \mathbf{U}^{-T} \mathbf{F}' \right] + \frac{\partial V_{NN}}{\partial \mathbf{R}} \\ &= \text{Tr} \left[ \mathbf{U}^{-T} \frac{d\mathbf{h}'}{d\mathbf{R}} \mathbf{U}^{-1} \tilde{\mathbf{P}} + \frac{1}{2} \mathbf{U}^{-T} \frac{\partial \mathbf{G}'}{\partial \mathbf{R}} \right|_{\tilde{\mathbf{P}}} \mathbf{U}^{-1} \tilde{\mathbf{P}} \right] \\ &\quad - \text{Tr} \left[ \mathbf{F} \frac{d\mathbf{U}}{d\mathbf{R}} \mathbf{U}^{-1} \tilde{\mathbf{P}} + \tilde{\mathbf{P}} \mathbf{U}^{-T} \frac{d\mathbf{U}^T}{d\mathbf{R}} \mathbf{F} \right] + \frac{\partial V_{NN}}{\partial \mathbf{R}}. \end{aligned} \quad (14)$$

For a converged SCF calculation, the Fock matrix and the density commute in an orthonormal basis, leading to

$$\begin{aligned} \frac{dE}{d\mathbf{R}} &= \left. \frac{\partial E}{\partial \mathbf{R}} \right|_{\mathbf{P}'} \\ &= \text{Tr} \left[ \left. \frac{d\mathbf{h}'}{d\mathbf{R}} \tilde{\mathbf{P}}' + \frac{1}{2} \frac{\partial \mathbf{G}'(\tilde{\mathbf{P}}')}{\partial \mathbf{R}} \right|_{\tilde{\mathbf{P}}'} \right] \\ &\quad - \text{Tr} \left[ \tilde{\mathbf{P}}' \mathbf{F} \left( \frac{d\mathbf{U}}{d\mathbf{R}} \mathbf{U}^{-1} + \mathbf{U}^{-T} \frac{d\mathbf{U}^T}{d\mathbf{R}} \right) \right] + \frac{\partial V_{NN}}{\partial \mathbf{R}} \\ &= \text{Tr} \left[ \left. \frac{d\mathbf{h}'}{d\mathbf{R}} \tilde{\mathbf{P}}' + \frac{\partial \mathbf{G}'(\tilde{\mathbf{P}}')}{\partial \mathbf{R}} \right|_{\tilde{\mathbf{P}}'} \right] - \text{Tr} \left[ \mathbf{F}' \tilde{\mathbf{P}}' \frac{d\mathbf{S}'}{d\mathbf{R}} \tilde{\mathbf{P}}' \right], \end{aligned} \quad (15)$$

where we have used the fact that  $\tilde{\mathbf{P}} \equiv \mathbf{P}$  for converged densities and  $\mathbf{U}^{-T}[d\mathbf{U}^T/d\mathbf{R}] + [d\mathbf{U}/d\mathbf{R}]\mathbf{U}^{-1} = \mathbf{U}^{-T}[d\mathbf{S}'/d\mathbf{R}]\mathbf{U}^{-1}$ . The latter is obtained by differentiating  $\mathbf{U}^{-T}\mathbf{S}'\mathbf{U}^{-1} = \mathbf{I}$  (e.g.,  $[d\mathbf{U}^{-T}/d\mathbf{R}]\mathbf{S}'\mathbf{U}^{-1} + \mathbf{U}^{-T}[d\mathbf{S}'/d\mathbf{R}]\mathbf{U}^{-1} + \mathbf{U}^{-T}\mathbf{S}'[d\mathbf{U}^{-1}/d\mathbf{R}] = 0$ ). It may be noted that the second expression in Eq. (15) is the usual expression for the energy derivative for a converged SCF

FIG. 2. Starting geometry for the  $\text{Cl}^-(\text{H}_2\text{O})_{25}$  cluster.

calculation<sup>24</sup> and corresponds to  $[\partial\mathbf{P}'/\partial\mathbf{R}]\mathbf{P} = -\mathbf{P}'[d\mathbf{S}'/d\mathbf{R}]\mathbf{P}'$  [from Eqs. (11) and (15)].

When the SCF is not converged, the derivatives of  $\mathbf{U}$  must be computed explicitly. In the case of Löwdin orthonormalization, this can be expressed as<sup>25</sup>

$$\frac{d\mathbf{U}}{d\mathbf{R}} = \sum_{i,j} \mathbf{s}_i \frac{1}{\sigma_i^{1/2} + \sigma_j^{1/2}} \left( \mathbf{s}_i^T \frac{d\mathbf{S}'}{d\mathbf{R}} \mathbf{s}_j \right) \mathbf{s}_j^T, \quad (16)$$

where  $\mathbf{s}_i$  and  $\sigma_i$  are, respectively, the  $i$ th eigenvector and eigenvalue of the overlap matrix  $\mathbf{S}'$ . For Cholesky orthonormalization, the derivative of the transformation is an upper triangular matrix that satisfies  $[d\mathbf{U}^T/d\mathbf{R}]\mathbf{U} + \mathbf{U}^T[d\mathbf{U}/d\mathbf{R}] = [d\mathbf{S}'/d\mathbf{R}]$  and can be obtained by a simple modification of the Cholesky decomposition procedure. Alternatively, one can use the fact that  $\mathbf{U}$ ,  $\mathbf{U}^{-1}$ ,  $[d\mathbf{U}/d\mathbf{R}]$  and  $[d\mathbf{U}^{-1}/d\mathbf{R}]$  are the upper triangular, and  $\mathbf{U}^T$ ,  $\mathbf{U}^{-T}$ ,  $[d\mathbf{U}^T/d\mathbf{R}]$  and  $[d\mathbf{U}^{-T}/d\mathbf{R}]$  are the lower triangular. Thus,  $\mathbf{U}^{-T}[d\mathbf{U}^T/d\mathbf{R}] + [d\mathbf{U}/d\mathbf{R}]\mathbf{U}^{-1} = \mathbf{U}^{-T}[d\mathbf{S}'/d\mathbf{R}]\mathbf{U}^{-1}$  can be partitioned uniquely into upper and lower triangular parts such that

$$\begin{aligned} \left[ \frac{d\mathbf{U}}{d\mathbf{R}} \mathbf{U}^{-1} \right]_{\mu,\nu} &= \left( \mathbf{U}^{-T} \frac{d\mathbf{S}'}{d\mathbf{R}} \mathbf{U}^{-1} \right)_{\mu,\nu}, & \mu < \nu \\ &= \frac{1}{2} \left( \mathbf{U}^{-T} \frac{d\mathbf{S}'}{d\mathbf{R}} \mathbf{U}^{-1} \right)_{\mu,\nu}, & \mu = \nu \\ &= 0, & \mu > \nu, \end{aligned} \quad (17)$$

since  $\mathbf{U}^{-T}[d\mathbf{U}^T/d\mathbf{R}] = \{[d\mathbf{U}/d\mathbf{R}]\mathbf{U}^{-1}\}^T$ . These expressions can then be used in Eq. (14) to compute of the derivatives of the energy with respect to the nuclear coordinates in the general case of unconverged densities.

### III. APPLICATIONS

To illustrate the implementation of the approach discussed above, we have considered the dynamics of two systems: the photodissociation of 1,3,5 triazine,  $\text{C}_3\text{N}_3\text{H}_3$  (Fig. 1), and the solvation of a  $\text{Cl}^-$  ion by a cluster of 25 water molecules (Fig. 2).

Triazine can be excited into either the  ${}^1E(\pi^* \leftarrow n)$  or the  ${}^1A_2(\pi^* \leftarrow \pi)$  states; both rapidly convert to the ground state which fragments into three molecules of HCN in a concerted fashion.<sup>26</sup> Table I summarizes calculations with orthogonal semiempirical (PM3) and nonorthogonal *ab initio* Hartree–Fock (HF) and DFT methods with polarized basis sets, using a development version of the GAUSSIAN series of electronic structure codes.<sup>27</sup> Trajectories were started at the transition state with converged densities and 10 kcal/mol translational energy in the transition vector, zero point energy in the remaining vibrational modes and a rotational energy of 298 K. A total integration time of 100 fs ensured that the products were well separated. A step size of 0.1 fs was chosen so that total energy (kinetic and potential energy for the atoms plus the fictitious kinetic energy of the density matrix elements) was conserved to  $\sim 10^{-5}$  hartree. When a fictitious mass of 0.050 amu bohr<sup>2</sup> ( $\sim 91$  a.u.) is used for all of the density matrix elements in HF or DFT calculations, the periods of oscillation range from 0.2 fs for core functions to 1.5 fs for valence functions. Hence, the oscillation of the slowest elec-

TABLE I. Comparison of trajectories for triazine three-body photodissociation.

Theoretical method	Integration technique	Conservation of the total energy (in hartree) <sup>a</sup>	Path error <sup>b</sup> (in bohr)
PM3	Converged density	0.000020	0.0030
	Propagated density	0.000150	0.095
	Propagated (no projection)	0.000019	0.095
HF/ 6-31G(d)	Converged density	0.000014	0.0029
	Propagated density, Löwdin orthogonalization	0.000079	0.14
B3LYP <sup>c</sup> / 6-31G(d)	Propagated density, Cholesky orthogonalization	0.000130	0.22
	Converged density	0.000030	0.0030
	Propagated density, Löwdin orthogonalization	0.000056	0.11
	Propagated density, Cholesky orthogonalization	0.000057	0.21

<sup>a</sup>Maximum deviation of the total energy of the system (defined as the sum of the total potential energy,  $E$ , the nuclear kinetic energy and the fictitious kinetic energy of the density matrix) during the trajectory. 1 hartree (a.u. of energy) = 627.51 kcal/mol.

<sup>b</sup>Relative to a reference trajectory calculated using the Hessian method with a step size of 0.025 amu<sup>1/2</sup> bohr.

<sup>c</sup>Reference 28.

tronic modes is about an order of magnitude faster than the fastest nuclear modes (10–15 fs), ensuring very good separation between the degrees of freedom, even without thermostats.

To gauge the accuracy of these methods, reference trajectories were calculated with a Hessian based algorithm<sup>29</sup> with a small step of fixed distance (0.025 amu<sup>1/2</sup> bohr in mass-weighted coordinates), which corresponds to a maximum error of  $2.4 \times 10^{-8}$  hartree in the conservation of the total energy. When converged densities are used, the conservation of the total energy and the agreement of the trajectory with the more accurate Hessian-based path are very good. The error in the conservation of total energy is only slightly larger when the density is propagated, and the maximum deviation of the trajectory from the more accurate path is  $\sim 0.2$  bohr, for both Löwdin and Cholesky orthonormalization. For all calculations idempotency of the density matrix is maintained to  $10^{-12}$  or better through the use of Eqs. (9) and (10), and angular momentum was conserved to better than  $10^{-10}\hbar$  since projection methods were used to remove residual angular forces.<sup>30</sup> Subsequent to initial heating of the electronic degrees of freedom, the adiabaticity (i.e., the drift in the fictitious kinetic energy of the density matrix) is conserved to better than  $10^{-4}$  hartrees. (The initial heating was a result of the choice of the density matrix velocities at the initial time-step. This may be considered akin to an equilibration process where the density matrix elements “catch-up” with the nuclei.)

Solvation of ions in aqueous medium is of great practical interest and has recently attracted considerable attention from the first principles molecular dynamics community.<sup>31–36</sup> We have computed trajectories for the chloride–water cluster using the current method. Starting geometries were calculated by MM3 minimization and initial density matrices were obtained from fully converged SCF calculations (see Fig. 2). The initial nuclear velocities were chosen randomly and scaled so that the total nuclear kinetic energy was equal to 0.100 hartree ( $\approx 63$  kcal/mol). The density matrix velocity was chosen to be zero at the initial time step. Trajectories were computed by density functional theory with the PBE (Ref. 37) exchange correlation functional and using Cholesky orthonormalization. (Use of the Löwdin orthogonalization procedure did not provide any distinguishably different trajectories.) Two different sets of trajectory calculations were performed: one using an all electron double zeta basis set with polarization functions on the chlorine, the other using a valence double zeta basis set with polarization functions on the chlorine and pseudopotentials<sup>38–40</sup> to replace the oxygen 1s and the chlorine 1s, 2s, and 2p core functions. Various combinations of step sizes and fictitious masses were tested; selected results for  $\mu = 0.10$  amu bohr<sup>2</sup> ( $\approx 182$  a.u.) are shown in Table II. Angular momentum is conserved to better than  $10^{-10}\hbar$  through the use of projection methods.

Conservation of the total energy, shown in Fig. 3, reflects the accuracy of the integration procedure. Note that we are able to obtain good energy conservation while using hydrogens in the water cluster, and we did not have to substitute heavier isotopes of hydrogen, as is often done in plane-

TABLE II. Trajectories for the  $\text{Cl}^-(\text{H}_2\text{O})_{25}$  cluster with no thermostats.<sup>a</sup>

Functional <sup>b</sup> basis set	Basis set size	Time step (fs)	Trajectory time (fs)	Conservation of total energy <sup>c</sup> (hartree)	Average number of purification steps [using Eq. (9)]
PBE <sup>d</sup> /3-21G*	344	0.06	120	0.0024	2
PBE <sup>d</sup> /CEP-31G <sup>e</sup>	314	0.05	50	0.0001	2
PBE <sup>d</sup> /CEP-31G <sup>e</sup>	314	0.10	45	0.0001	3

<sup>a</sup>Value of the scalar fictitious mass,  $\mu$ , is chosen to be 0.1 amu bohr<sup>2</sup> ( $\approx 182$  a.u.).

<sup>b</sup>Exchange correlation functional used for the DFT calculation.

<sup>c</sup>Maximum deviation of the total energy of the system (defined as the sum of the total potential energy,  $E$ , the nuclear kinetic energy, and the fictitious kinetic energy of the density matrix) during the trajectory.

<sup>d</sup>Reference 37.

<sup>e</sup>Pseudopotential plus polarization basis functions added for  $\text{Cl}^-$  (Refs. 38–40).

wave Car–Parrinello calculations. The range of values used for the fictitious mass, in this study, are about five times smaller than those used in the standard plane-wave Car–Parrinello calculations.<sup>31–36</sup> The time steps used for the calculations using pseudopotentials are comparable to those commonly used for plane-wave Car–Parrinello runs.<sup>31–36</sup> As it may be anticipated, our results indicate that pseudopotential calculations permit larger time steps to be used while still maintaining energy conservation to a satisfactory degree. We also find that the adiabaticity for the pseudopotential calculations was maintained to  $10^{-3}$  hartree after the initial heating of the density matrix. This adiabatic behavior is indicated in Fig. 3, where the change in total energy, the fictitious kinetic energy, and the difference between the two are shown, all for the pseudopotential calculation. The zero of time in the trajectory was chosen after the initial increase in the fictitious kinetic energy. (Again, the initial heating of the density matrix is due to the initial density matrix velocities being chosen as equal to zero in this study.) The effect of

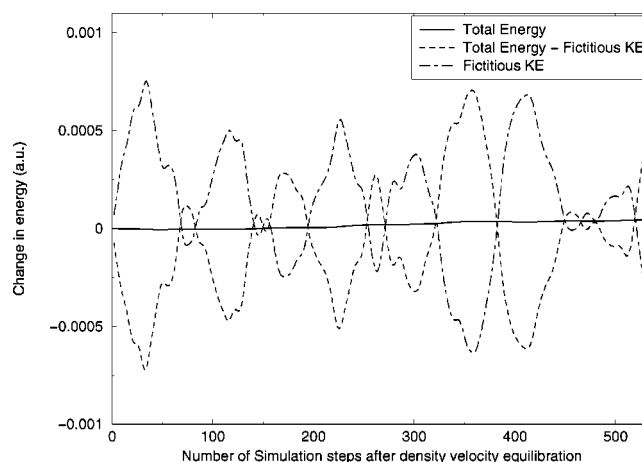


FIG. 3. Trajectory for the CEP-31G pseudopotential calculation of  $\text{Cl}^-(\text{H}_2\text{O})_{25}$  cluster;  $\Delta t = 0.05$  fs;  $\mu = 0.1$  amu bohr<sup>2</sup>. Plot shows the change in total energy, the fictitious kinetic energy of the density matrix and the difference between the two as a function of the number of simulation steps, subsequent to equilibration of the density matrix velocities (see text for details).

various choices of initial velocities will be studied in later publications. The results presented here demonstrate the applicability and accuracy of the current method. More detailed studies of adiabaticity versus time-step and fictitious mass will follow.<sup>22</sup>

#### IV. CONCLUDING REMARKS

In this paper, we have demonstrated that Car–Parrinello-type *ab initio* molecular dynamics can be carried out by propagating the density matrix rather than the orbital coefficients, and by using atom centered functions instead of plane waves. For the examples studied it was shown that reasonable energy conservation may be maintained during the trajectories, in spite of having hydrogen atoms in the system. This could potentially be useful in areas such as isotopic substitution.

Gaussian orbitals have the distinct advantage of leading to sparse matrix representations for the Hamiltonian and density matrix in the large system limit. With sparse matrix multiplication techniques, Cholesky orthonormalization, and fast multipole methods, our approach will thus lead to  $\mathcal{O}(N)$  scaling with system size. This aspect will be the subject of further study.

#### ACKNOWLEDGMENTS

This work was supported by the Natural Science Foundation (CHE-9982156 and CHE-9874005), the Office of Naval Research (GAV), and Gaussian, Inc. We would like to acknowledge Professor Jack Simons for his input on a preliminary version of this manuscript.

<sup>1</sup>R. Car and M. Parrinello, Phys. Rev. Lett. **55**, 2471 (1985).

<sup>2</sup>D. K. Remler and P. A. Madden, Mol. Phys. **70**, 921 (1990).

<sup>3</sup>M. C. Payne, M. P. Teter, D. C. Allan, T. A. Arias, and J. D. Joannopoulos, Rev. Mod. Phys. **64**, 1045 (1992).

<sup>4</sup>M. E. Tuckerman, P. J. Ungar, T. von Rosenvinge, and M. L. Klein, J. Phys. Chem. **100**, 12878 (1996).

<sup>5</sup>B. L. Trout and M. Parrinello, J. Phys. Chem. B **103**, 7340 (1999); P. L. Silvestrelli and M. Parrinello, J. Chem. Phys. **111**, 3572 (1999).

<sup>6</sup>G. Lippert, J. Hutter, and M. Parrinello, Mol. Phys. **92**, 477 (1997); Theor. Chem. Acc. **103**, 124 (1999).

<sup>7</sup>T. Helgaker and P. R. Taylor, in *Modern Electronic Structure Theory* (World Scientific, Singapore, 1995), p. 725.

<sup>8</sup>K. Bolton, W. L. Hase, and G. H. Peslherbe, in *Modern Methods for Multidimensional Dynamics Computation in Chemistry*, edited by D. L. Thompson (World Scientific, Singapore, 1998).

<sup>9</sup>D. A. Gibson, I. V. Ionova, and E. A. Carter, Chem. Phys. Lett. **240**, 261 (1995).

<sup>10</sup>B. Hartke and E. A. Carter, J. Chem. Phys. **97**, 6569 (1992).

<sup>11</sup>B. Hartke and E. A. Carter, Chem. Phys. Lett. **189**, 358 (1992).

<sup>12</sup>S. Goedecker, Rev. Mod. Phys. **71**, 1085 (1999).

<sup>13</sup>G. E. Scuseria, J. Phys. Chem. A **103**, 4782 (1999).

<sup>14</sup>X.-P. Li, W. Nunes, and D. Vanderbilt, Phys. Rev. B **47**, 10891 (1993).

<sup>15</sup>C. A. White and M. Head-Gordon, J. Chem. Phys. **101**, 6593 (1994).

<sup>16</sup>M. C. Strain, G. E. Scuseria, and M. J. Frisch, Science **271**, 51 (1996).

<sup>17</sup>J. M. Millam and G. E. Scuseria, J. Chem. Phys. **106**, 5569 (1997).

<sup>18</sup>A. D. Daniels, J. M. Millam, and G. E. Scuseria, J. Chem. Phys. **107**, 425 (1997).

<sup>19</sup>G. H. Golub and C. F. van Loan, *Matrix Computations* (The Johns Hopkins University Press, Baltimore, 1996).

<sup>20</sup>R. McWeeny, Rev. Mod. Phys. **32**, 335 (1960).

<sup>21</sup>H. Goldstein, *Classical Mechanics* (Addison–Wesley, Cambridge, 1980).

<sup>22</sup>S. S. Iyengar, H. B. Schlegel, J. M. Millam, G. A. Voth, A. D. Daniels, G. E. Scuseria, and M. J. Frisch (in preparation).

<sup>23</sup>W. C. Swope, H. C. Andersen, P. H. Berens, and K. R. Wilson, J. Chem. Phys. **76**, 637 (1982); M. E. Tuckerman and M. Parrinello, *ibid.* **101**, 1302 (1994).

<sup>24</sup>P. Pulay, Mol. Phys. **17**, 197 (1969).

<sup>25</sup>P. Jorgensen and J. Simons, J. Chem. Phys. **79**, 334 (1983).

<sup>26</sup>G. S. Ondrey and R. Bersohn, J. Chem. Phys. **81**, 4517 (1984); S. R. Goates, J. O. Chu, and G. W. Flynn, *ibid.* **81**, 4521 (1984); T. Gejo, J. A. Harrison, and J. R. Huber, J. Phys. Chem. **100**, 1391 (1996); Y. Osamura, M. Unno, and K. Hashimoto, J. Am. Chem. Soc. **109**, 1370 (1987); S. V. Pai, C. F. Chabalowski, and B. M. Rice, J. Phys. Chem. **100**, 15368 (1996); **100**, 5681 (1996).

<sup>27</sup>M. J. Frisch, G. W. Trucks, H. B. Schlegel *et al.*, GAUSSIAN 99, Development Version, Revision B.05 (Gaussian, Inc., Pittsburgh, PA, 1998).

<sup>28</sup>A. D. Becke, J. Chem. Phys. **98**, 5648 (1993).

<sup>29</sup>J. M. Millam, V. Bakken, W. Chen, W. L. Hase, and H. B. Schlegel, J. Chem. Phys. **111**, 3800 (1999).

<sup>30</sup>For the semiempirical calculations with converged densities, the angular forces originate primarily from the numerical differentiation of the integrals used in the gradients (angular momentum conservation is  $\sim 0.3\hbar$  for PM3 without projection). When the electronic structure is propagated, additional angular forces can arise from the fact that the electronic energy is not converged with respect to the density.

<sup>31</sup>C. J. Mundy, J. Hutter, and M. Parrinello, J. Am. Chem. Soc. **122**, 4837 (2000).

<sup>32</sup>F. Bruge, M. Bernasconi, and M. Parrinello, J. Am. Chem. Soc. **121**, 10883 (1999).

<sup>33</sup>L. M. Ramaniah, M. Bernasconi, and M. Parrinello, J. Chem. Phys. **111**, 1587 (1999).

<sup>34</sup>E. J. Meijer and M. Sprik, J. Phys. Chem. A **102**, 2893 (1998).

<sup>35</sup>D. Marx, M. Sprik, and M. Parrinello, Chem. Phys. Lett. **373**, 360 (1997).

<sup>36</sup>M. Tuckerman, K. Laasonen, M. Sprik, and M. Parrinello, J. Phys. Chem. **99**, 5749 (1995).

<sup>37</sup>J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996); **78**, 1396 (1997).

<sup>38</sup>W. Stevens, H. Basch, and J. Krauss, J. Chem. Phys. **81**, 6026 (1984).

<sup>39</sup>W. J. Stevens, M. Krauss, H. Basch, and P. G. Jasien, Can. J. Chem. **70**, 612 (1992).

<sup>40</sup>T. R. Cundari and W. J. Stevens, J. Chem. Phys. **98**, 5555 (1993).