

Ab Initio Molecular Dynamics: Propagating the Density Matrix with Gaussian orbitals. IV. Formal Analysis of the deviations from Born-Oppenheimer dynamics

Srinivasan S. Iyengar,^a H. Bernhard Schlegel,^b Gregory A. Voth,^a John M. Millam^b

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

*^bDepartment of Chemistry,
Wayne State University,
Detroit, MI 48202-3489*

Gustavo E. Scuseria,^c

*^cDepartment of Chemistry and Center for Nanoscale Science and Technology,
Rice University, Houston, TX 77005-1892*

Michael J. Frisch^d

*^dGaussian, Inc.,
140 Washington Ave.,
North Haven, CT 06473-1712*

(Dated: January 7, 2003)

Abstract

In the context of the recently developed Atom-centered Density Matrix Propagation (ADMP) approach to *ab initio* molecular dynamics, a formal analysis of the deviations from the Born-Oppenheimer surface is conducted. These deviations depend on the fictitious mass and on the magnitude of the commutator of the Fock and density matrices. These quantities are found to be closely inter-related and the choice of the fictitious mass provides a lower bound on the deviations from the Born-Oppenheimer surface. The relations are illustrated with an example calculation for the $\text{Cl}^-(\text{H}_2\text{O})_{25}$ cluster. We also show that there exists a direct one-to-one correspondence between approximate Born-Oppenheimer dynamics, where SCF convergence is restricted by a chosen threshold value for the commutator of the Fock and density matrices, and an extended Lagrangian dynamics performed using a finite value for the fictitious mass. The analysis is extended to the nuclear forces used in the ADMP approximation. The forces are shown to be more general than those standardly used in Born-Oppenheimer dynamics, with the addition terms in the nuclear forces depending on the commutator mentioned above.

I. INTRODUCTION

The Born-Oppenheimer (BO) dynamics methods¹⁻⁴ and extended Lagrangian (EL) approach⁵⁻¹⁰ constitute two major categories of direct classical trajectory calculations for the simulation of molecular dynamics.^{3,4} For BO methods, the electronic structure calculation is converged at each time step in the propagation. In the EL approach, the wavefunction is propagated along with the classical nuclear degrees of freedom by using an extended Lagrangian procedure^{11,12} and by adjusting the relative time-scales of the electronic and nuclear motions. The Car-Parrinello (CP) method⁵ is the prime example of the extended Lagrangian approach. Traditionally, CP calculations employ Kohn-Sham orbitals expanded in a plane-wave basis set (occasionally augmented with Gaussian orbitals.^{13,14}) Recently⁸⁻¹⁰ we have described the theory, implementation and initial applications of an extended Lagrangian molecular dynamics method that employs atom-centered Gaussian basis functions and density matrix propagation. This method is called Atom-centered Density Matrix Propagation (ADMP). The dynamics of chemical systems such as clusters and gas-phase reactions are more readily described with atom centered basis sets than with plane waves. Furthermore, the calculations will scale as $O(N)$ for large systems, where N is the number of electrons, because ADMP is based on linear scaling DFT code.¹⁵ The ADMP method is being extended to QM/MM treatments for biological systems,¹⁶ and to calculations on periodic systems employing atom centered functions.¹⁷ Interesting applications such as solvation of excess protons in water clusters,¹⁸ hydroxyl-stretch red-shifts in chloride-water clusters,¹⁰ and product distributions in gas-phase chemical reactions¹⁰ have already been studied. ADMP trajectories of the order of picoseconds show stable dynamics and adiabaticity can be controlled effectively in these systems without thermostats.

The ADMP method has a number of attractive features. Systems can be simulated by accurately treating all electrons or by using pseudopotentials. Through the use of a tensorial fictitious mass and smaller values of the mass, reasonably large time-steps can be employed and lighter atoms such as hydrogens need not be replaced with heavier isotopes. A wide variety of exchange-correlation functionals can be utilized, including hybrid density functionals. Atom centered functions can be used with the correct physical boundary conditions for molecules, polymers, surfaces and solids, without the need to treat replicated images in order to impose 3d periodicity. Consequently, charged systems and QM/MM models of biological

systems can be treated as easily as neutral molecules, whereas special handling is needed for charged systems in most implementations of the plane-wave Car-Parrinello method because they are treated as periodic systems. Deviation from the Born-Oppenheimer surface and the mixing of fictitious and real kinetic energies can be rigorously controlled on-the-fly. Because of fewer basis functions per atom, larger time steps, and asymptotic $O(N)$ scaling using established techniques, good computational efficiency can be achieved with the ADMP method.

In the current paper we have undertaken a formal analysis of several issues involved in the ADMP methodology. This paper is organized as follows. In Section II we present a brief overview of the ADMP approach. In Section III we analyze the conceptual similarities between an extended Lagrangian molecular dynamics scheme and an approximate Born-Oppenheimer molecular dynamics which leads us to interesting relations between the fictitious mass in the extended Lagrangian formalism and the commutator of the Fock and density matrices in the Born-Oppenheimer formalism. We show that the fictitious mass fully determines the minimum deviation of the trajectory from the Born-Oppenheimer surface by limiting the magnitude of the commutator of the Fock and density matrices. Furthermore, we also show that for every approximate Born-Oppenheimer dynamics trajectory, an extended Lagrangian can be constructed with an with appropriately chosen magnitude of the fictitious mass, such that both have identical deviations from the Born-Oppenheimer surface. In Section IV we show that the nuclear forces used in ADMP are, in fact, more general than those standardly used¹⁹ in Born-Oppenheimer dynamics since they include additional terms that dependent on the commutator of the Fock and density matrices (which address the limited convergence of the density matrix required within the extended Lagrangian formalism). The nuclear forces used in ADMP reduce to the standard form¹⁹ in the limit as the commutator of the Fock and density matrix goes to zero. In Section V we present our conclusions.

II. ATOM-CENTERED DENSITY MATRIX PROPAGATION (ADMP) USING AN EXTENDED LAGRANGIAN APPROACH

To assist in subsequent discussions, a brief outline of the ADMP method is given in this section. Details of the ADMP method have been described in our earlier papers.^{8,9} The

Lagrangian for propagation of the density matrix is simplest in an orthonormal basis and can be written as

$$\mathcal{L}_{ADMP} = \frac{1}{2} Tr(\mathbf{V}^T \mathbf{M} \mathbf{V}) + \frac{1}{2} Tr \left(\left[\underline{\mu}^{1/4} \mathbf{W} \underline{\mu}^{1/4} \right]^2 \right) - E(\mathbf{R}, \mathbf{P}) - Tr[\mathbf{\Lambda}(\mathbf{P}\mathbf{P} - \mathbf{P})], \quad (1)$$

where \mathbf{R} , \mathbf{V} and \mathbf{M} are the nuclear positions, velocities and masses, respectively; and \mathbf{P} , \mathbf{W} and $\underline{\mu}$ are the density matrix, the density matrix velocity and the fictitious mass tensor for the electronic degrees of freedom, respectively. The Lagrangian multiplier matrix $\mathbf{\Lambda}$ is used to impose N-representability (*i.e.*, constraints on the total number of electrons and the idempotency of the one particle density matrix). The energy is calculated using the McWeeny purification of the density, $\tilde{\mathbf{P}} = 3\mathbf{P}^2 - 2\mathbf{P}^3$. The Euler-Lagrange equations of motion for the nuclei and density matrix are:

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

and

$$\underline{\mu}^{1/2} \frac{d^2 \mathbf{P}}{dt^2} \underline{\mu}^{1/2} = - \left[\left. \frac{\partial E(\mathbf{R}, \mathbf{P})}{\partial \mathbf{P}} \right|_{\mathbf{R}} + \mathbf{\Lambda} \mathbf{P} + \mathbf{P} \mathbf{\Lambda} - \mathbf{\Lambda} \right]. \quad (3)$$

These can be integrated using the velocity Verlet algorithm:^{20,21}

$$\mathbf{P}_{i+1} = \mathbf{P}_i + \mathbf{W}_i \Delta t - \frac{\Delta t^2}{2} \underline{\mu}^{-1/2} \left[\left. \frac{\partial E(\mathbf{R}_i, \mathbf{P}_i)}{\partial \mathbf{P}} \right|_{\mathbf{R}} + \mathbf{\Lambda}_i \mathbf{P}_i + \mathbf{P}_i \mathbf{\Lambda}_i - \mathbf{\Lambda}_i \right] \underline{\mu}^{-1/2}, \quad (4)$$

$$\mathbf{W}_{i+1/2} = \mathbf{W}_i - \frac{\Delta t}{2} \underline{\mu}^{-1/2} \left[\left. \frac{\partial E(\mathbf{R}_i, \mathbf{P}_i)}{\partial \mathbf{P}} \right|_{\mathbf{R}} + \mathbf{\Lambda}_i \mathbf{P}_i + \mathbf{P}_i \mathbf{\Lambda}_i - \mathbf{\Lambda}_i \right] \underline{\mu}^{-1/2} = \frac{\mathbf{P}_{i+1} - \mathbf{P}_i}{\Delta t}, \quad (5)$$

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

The Lagrangian multiplier matrices are determined by an iterative scheme^{8,9} so that \mathbf{P}_{i+1} and \mathbf{W}_{i+1} satisfy the idempotency constraint, $\mathbf{P}^2 = \mathbf{P}$, and its time derivative, $\mathbf{P}\mathbf{W} + \mathbf{W}\mathbf{P} = \mathbf{W}$.

III. A UNIFIED FRAMEWORK FOR BORN-OPPENHEIMER AND EXTENDED LAGRANGIAN MOLECULAR DYNAMICS

In Born-Oppenheimer (BO)^{2,22,23} molecular dynamics the electronic structure must be converged at every nuclear configuration. A measure of this convergence is obtained from the commutator of the Fock and density matrices, *i.e.*, $[\mathbf{F}, \mathbf{P}]$. The smaller the magnitude of this commutator (magnitude being defined in terms of the L²-norm or the Frobenius

norm of $[\mathbf{F}, \mathbf{P}]$) the closer the resultant electronic structure is to the Born-Oppenheimer surface. Using the converged wavefunction (or the associated single-particle density matrix) the forces on the nuclei, $\partial E(\mathbf{R}, \mathbf{P})/\partial \mathbf{R}$, are determined using standard techniques.¹⁹

Consider an approximate BO dynamics methodology where the requirement of convergence of the electronic structure to the BO surface is relaxed. Specifically the electronic structure is partially converged up to a pre-defined magnitude of the commutator. In this case, the wavefunction (and hence the associated electronic structure) is a fictitious state, that is off the Born-Oppenheimer surface, and the chosen value of commutator, $[\mathbf{F}, \mathbf{P}]$ (or its norm) represents, in some sense, how far this approximate state is from the true Born-Oppenheimer wavefunction. If this state were to be used in the dynamics procedure, the force on the nuclei at the given configuration would be different from that used in standard BO dynamics, since the wavefunction is not converged. An alternative expression for the force may be obtained (see Eqs. (14) and (15) in Ref. 8, or Section IV in the current paper for a more detailed exposition) that does not require the commutator $[\mathbf{F}, \mathbf{P}]$ to be zero. Using this force, the nuclei may be propagated. This scheme then provides an alternative approach to approximate Born-Oppenheimer dynamics.

In this section we will analyze the similarities between the aforementioned approximate Born-Oppenheimer dynamics scheme and the Extended Lagrangian approach to molecular dynamics described in Refs. 5,8–10. Consider the total Hamiltonian within a Born-Oppenheimer dynamics framework:

$$\mathcal{H} = \sum_i \sum_{j=1}^3 \frac{\mathcal{V}_{i,j}^2}{2\mathbf{M}_i} + E(\mathbf{R}, \mathbf{P}), \quad (7)$$

where \mathbf{M} , \mathbf{R} and \mathcal{V} are the nuclear masses, positions, and conjugate momenta respectively, while \mathbf{P} is the electronic single particle density matrix. $\mathcal{V}_{i,j}$ represents j -th momentum component for the i -th nucleus. $E(\mathbf{R}, \mathbf{P})$ is the Hartree-Fock, DFT or semi-empirical energy functional. The change in the total derivative of the Hamiltonian for an approximate Born-Oppenheimer methodology with respect to the true Born-Oppenheimer dynamics trajectory is then given by

$$\begin{aligned} \frac{d[\mathcal{H}_{approx.BO} - \mathcal{H}_{BO}]}{dt} = & \left\{ \sum_i \sum_{j=1}^3 \left[\mathcal{V}_{i,j}^{approx.BO} \frac{d^2 \mathbf{R}_{i,j}^{approx.BO}}{dt^2} \right] - \sum_i \sum_{j=1}^3 \left[\mathcal{V}_{i,j}^{BO} \frac{d^2 \mathbf{R}_{i,j}^{BO}}{dt^2} \right] \right\} + \\ & Tr \left[\left. \frac{\partial E_{approx.BO}(\mathbf{R}, \mathbf{P}_{approx})}{\partial \mathbf{P}} \right|_{\mathbf{R}} \frac{d\mathbf{P}_{approx}}{dt} - \left. \frac{\partial E_{BO}(\mathbf{R}, \mathbf{P})}{\partial \mathbf{P}} \right|_{\mathbf{R}} \frac{d\mathbf{P}}{dt} \right] + \end{aligned}$$

$$\left\{ \sum_i \sum_{j=1}^3 \left[\frac{\partial E_{approx.BO}(\mathbf{R}, \mathbf{P}_{approx})}{\partial \mathbf{R}_{i,j}} \Big|_{\mathbf{P}_{approx}} \frac{d\mathbf{R}_{i,j}^{approx.BO}}{dt} \right] - \sum_i \sum_{j=1}^3 \left[\frac{\partial E_{BO}(\mathbf{R}, \mathbf{P})}{\partial \mathbf{R}_{i,j}} \Big|_{\mathbf{P}} \frac{d\mathbf{R}_{i,j}}{dt} \right] \right\}, \quad (8)$$

where $E_{BO}(\mathbf{R}, \mathbf{P})$ is the total energy of the system obtained using exact BO dynamics (*i.e.*, the electronic structure is exactly converged at every instant), while $E_{approx.BO}(\mathbf{R}, \mathbf{P}_{approx})$ represents the total energy of the system using an approximate BO dynamics scheme like the one described above. The set $\{\mathbf{R}_{i,j}\}$ represents the set of nuclear coordinates (for example, the j -th coordinate of the i -th atom may be represented as $\mathbf{R}_{i,j}$). Equation (8) represents how the physical nature of the trajectory obtained from the approximate BO dynamics scheme deviates from that obtained from a *true* BO dynamics calculation. Clearly if the total derivative in Eq. (8) is zero, all physical aspects involved in the two trajectories are identical, even if the magnitude of $[E_{approx.BO} - E_{BO}]$ is large.

Invoking the classical Hamilton's equations of motion for both approximate and true BO dynamics, *i.e.*,

$$\frac{\partial \mathcal{H}}{\partial \mathcal{V}_{i,j}} = \frac{d\mathbf{R}_{i,j}}{dt} \quad (9)$$

and

$$\frac{\partial \mathcal{H}}{\partial \mathbf{R}_{i,j}} = -\frac{d\mathcal{V}_{i,j}}{dt}, \quad (10)$$

one obtains

$$\frac{d[\mathcal{H}_{approx.BO} - \mathcal{H}_{BO}]}{dt} = Tr \left[\frac{\partial E_{approx.BO}(\mathbf{R}, \mathbf{P}_{approx})}{\partial \mathbf{P}} \Big|_{\mathbf{R}} \frac{d\mathbf{P}_{approx}}{dt} - \frac{\partial E_{BO}(\mathbf{R}, \mathbf{P})}{\partial \mathbf{P}} \Big|_{\mathbf{R}} \frac{d\mathbf{P}}{dt} \right] \quad (11)$$

It is interesting to note here that first and third terms in Eq. (8) cancel out upon invoking the equations of motion. This indicates that the expressions for forces on the nuclei for exact and approximate BO dynamics are not identical, as discussed above, on account of the non-zero commutator, $[\mathbf{F}, \mathbf{P}]$, in the case of the latter. This aspect is further substantiated by the two different expressions obtained in Ref. 8 for the forces and is also discussed in Section IV, in more detail.

Proceeding further with our analysis of Eq. (11), we simplify the expression by using the gradient terms provided in Ref. 8, *i.e.*,

$$\frac{\partial E_{approx.BO}}{\partial \mathbf{P}} \Big|_{\mathbf{R}} = 3\mathbf{F}\mathbf{P}_{approx} + 3\mathbf{P}_{approx}\mathbf{F} - 2\mathbf{F}\mathbf{P}_{approx}^2 - 2\mathbf{P}_{approx}\mathbf{F}\mathbf{P}_{approx} - 2\mathbf{P}_{approx}^2\mathbf{F}, \quad (12)$$

where \mathbf{F} is the Fock matrix. For DFT and Hartree-Fock calculations, \mathbf{F} is constructed using \mathbf{P}_{approx} . Since the density and Fock matrices commute for a converged BO calculation,

$$\left. \frac{\partial E_{BO}}{\partial \mathbf{P}} \right|_{\mathbf{R}} = 0. \quad (13)$$

Hence, for an idempotent density matrix

$$\begin{aligned} \frac{d[\mathcal{H}_{approx.BO} - \mathcal{H}_{BO}]}{dt} &= Tr \left[[\mathbf{F}, \mathbf{P}_{approx}] \{2\mathbf{P}_{approx} - \mathbf{I}\} \frac{d\mathbf{P}_{approx}}{dt} \right] \\ &= Tr \left\{ [\mathbf{F}, \mathbf{P}_{approx}] \left[\mathbf{P}_{approx}, \frac{d\mathbf{P}_{approx}}{dt} \right] \right\}. \end{aligned} \quad (14)$$

Clearly, in the limit as $[\mathbf{F}, \mathbf{P}_{approx}] \rightarrow 0$, the right hand side above goes to zero, thus minimizing any differences between the approximate and exact Born-Oppenheimer dynamics procedures. The term $\{d\mathbf{P}_{approx}/dt\}$, in general, contains the non-Hellmann-Feynman contributions to the dynamics. However, an approximation to this maybe obtained from finite-difference, or as will be seen later, by drawing the analogy to an extended Lagrangian dynamics scheme.

An alternative procedure to approximate BO dynamics was introduced by Car and Parrinello,⁵ based on an extended Lagrangian procedure. Here, the introduction of a fictitious wavefunction mass and a fictitious wavefunction velocity allows for the electronic and nuclear degrees of freedom to be propagated together, by an adjustment of the relative electronic and nuclear time-scales. In essence, the Car-Parrinello “wavefunction” is rarely on the BO dynamics surface, but oscillates about it, so that long time average quantities obtained from this scheme provide good agreement with more traditional methods.⁷ In this sense, the Car-Parrinello “wavefunction” is quite similar to the approximate Born-Oppenheimer state described above. Our alternate ADMP scheme⁸⁻¹⁰ to extended Lagrangian dynamics of electrons and nuclei is based on the use of the single particle density matrix and atom-centered Gaussian basis functions. This method has been described in Refs. 8-10 and a brief discussion may be found in Section II.

In Ref. 9 we have also analyzed the conjugate Hamiltonian of the ADMP Lagrangian to obtain deviations from Born-Oppenheimer dynamics. Briefly, the conjugate ADMP Hamiltonian, obtained from a Legendre transformation²⁴ of the ADMP Lagrangian in Eq. (1), is given by

$$\mathcal{H}_{ADMP} = Tr(\mathcal{W}\mathbf{W}) + Tr(\mathcal{V}^T\mathbf{V}) - \mathcal{L}_{ADMP}$$

$$\begin{aligned}
&= \frac{1}{2} \text{Tr}(\mathcal{V}^T \mathbf{M}^{-1} \mathcal{V}) + \frac{1}{2} \text{Tr}(\mathcal{W} \underline{\mu}^{-1/2} \mathcal{W} \underline{\mu}^{-1/2}) + E(\mathbf{R}, \mathbf{P}) + \\
&\quad \text{Tr}[\mathbf{\Lambda}(\mathbf{P}\mathbf{P} - \mathbf{P})]
\end{aligned} \tag{15}$$

where \mathcal{W} is the conjugate momentum of \mathbf{P} :

$$\mathcal{W} = \frac{\partial \mathcal{L}_{ADMP}}{\partial \mathbf{W}} = \underline{\mu}^{1/2} \mathbf{W} \underline{\mu}^{1/2}. \tag{16}$$

Using the result⁹ that the Hamiltonian in Eq. (15) is conservative we have shown that the fictitious dynamics obtained from the ADMP scheme converges to the real BO dynamics trajectory in the limit that the magnitude of the fictitious mass tends to zero. Specifically, the deviation of the extended Lagrangian trajectory from the BO dynamics trajectory is bounded and given by⁹

$$\begin{aligned}
\frac{d\mathcal{H}_{fict}}{dt} &= \text{Tr} \left[\frac{\partial \mathcal{H}_{fict}}{\partial \mathcal{W}} \frac{d\mathcal{W}}{dt} \right] = \text{Tr} \left[\underline{\mu}^{-1/2} \mathcal{W} \underline{\mu}^{-1/2} \frac{d\mathcal{W}}{dt} \right] = \text{Tr} \left[\mathbf{W} \underline{\mu}^{1/2} \frac{d\mathbf{W}}{dt} \underline{\mu}^{1/2} \right] \\
&= - \text{Tr} \left[\mathbf{W} \left(\frac{\partial E(\mathbf{R}, \mathbf{P})}{\partial \mathbf{P}} \Big|_{\mathbf{R}} + \mathbf{\Lambda} \mathbf{P} + \mathbf{P} \mathbf{\Lambda} - \mathbf{\Lambda} \right) \right],
\end{aligned} \tag{17}$$

where

$$\begin{aligned}
\mathcal{H}_{fict} &= \frac{1}{2} \text{Tr} [\mathcal{W} \underline{\mu}^{-1/2} \mathcal{W} \underline{\mu}^{-1/2}] \\
&= \text{Tr} \left(\left[\underline{\mu}^{1/4} \mathbf{W} \underline{\mu}^{1/4} \right]^2 \right)
\end{aligned} \tag{18}$$

is the fictitious kinetic energy. Note that this expression is also valid for a plane-wave Car-Parrinello implementation with density matrices. Also note that Tangney and Scandolo²⁵ have subsequently derived a similar expression for the Car-Parrinello scheme, by using the fact that additional terms (non-Hellmann-Feynman) should be included in the forces, since the wavefunction is not exactly converged. In the standard Car-Parrinello method, these additional terms are not included.

We will now compare the two bounding expressions Eqs. (17) and (14) obtained for the two different approximations to Born-Oppenheimer dynamics. Since Eqs. (17) and (14) measure the deviations of two different approximate Born-Oppenheimer dynamics methodologies, *namely* the extended Lagrangian and the previously discussed *approximate* BO approach, from the true Born-Oppenheimer dynamics trajectory, we can conclude that for the two approximate trajectories to deviate identically from the *true* BO dynamics trajectory the quantities on the right side of Eqs. (17) and (14) must be equal. Hence,

$$\text{Tr} \left[\mathbf{W} \underline{\mu}^{1/2} \frac{d\mathbf{W}}{dt} \underline{\mu}^{1/2} \right] = \text{Tr} \left\{ [\mathbf{F}, \mathbf{P}_{approx}] \left[\mathbf{P}_{approx}, \frac{d\mathbf{P}_{approx}}{dt} \right] \right\}. \tag{19}$$

Note that $[\mathbf{F}, \mathbf{P}_{approx}]$ and $[\mathbf{P}_{approx}, d\mathbf{P}_{approx}/dt]$ are anti-Hermitian and hence the trace of their product is purely real, as is required from the fact that the left hand side in the expression above is purely real. Equation (19) yields a relation between the fictitious mass $\underline{\mu}$, the commutator $[\mathbf{F}, \mathbf{P}_{approx}]$ and the term $\{d\mathbf{P}_{approx}/dt\}$. The latter, as stated earlier may be obtained from a finite difference approximation, for example, $\{d\mathbf{P}_{approx}/dt\} \approx [\mathbf{P}_{approx}(t) - \mathbf{P}_{approx}(t - \Delta t)]/\Delta t$. Alternately, we make the connection that since the two approximate BO methodologies deviate identically from the true BO trajectory, $\{d\mathbf{P}_{approx}/dt\} \approx \mathbf{W}$. In either cases the quantity is well defined.

It must, however, be noted that Eq. (19) represents only one equation. Hence, the matrix $\underline{\mu}$ is not uniquely determined by Eq. (19). If the matrix $\underline{\mu}$ can however be written as a product of a scalar times a matrix, *i.e.* $\underline{\mu} \equiv \mu \mathbf{A}$ (note that for $\mathbf{A} \equiv \mathbf{I}$, the identity matrix, this reduces to the scalar fictitious mass case treated in Refs. 8 and 9, or more generally, the matrix \mathbf{A} may represent a mass-weighting scheme, like the one used in Ref. 9), then

$$\mu = \frac{1}{Tr \left[\mathbf{W} \mathbf{A}^{1/2} \frac{d\mathbf{W}}{dt} \mathbf{A}^{1/2} \right]} Tr \left\{ [\mathbf{F}, \mathbf{P}_{approx}] \left[\mathbf{P}_{approx}, \frac{d\mathbf{P}_{approx}}{dt} \right] \right\}. \quad (20)$$

Equation (20) provides a rigorous connection between the two approximate BO dynamics methodologies. This also provides a further validation of our previous study⁹ showing that the magnitude of the $\underline{\mu}$ presents an estimate of the deviation from the BO dynamics trajectory. Here, however, we show that the $\underline{\mu}$ is in fact related to the commutator, $[\mathbf{F}, \mathbf{P}_{approx}]$ by Eqs. (19) and (20). This implies that for a given approximate BO dynamics trajectory there always exists an extended Lagrangian trajectory where the corresponding fictitious mass is given by Eq. (20). However, since the quantities on the right hand side of Eq. (20) are time-dependent, it is possible that for an approximate BO trajectory with a fixed criteria of convergence for the value of the commutator, $[\mathbf{F}, \mathbf{P}_{approx}]$, the corresponding extended Lagrangian dynamics, in fact has a time-dependent fictitious mass. By corollary, an extended Lagrangian dynamics trajectory with a fixed fictitious mass should correspond to an approximate BO trajectory that has the criteria $[\mathbf{F}, \mathbf{P}_{approx}]$ to be a time-dependent quantity, that is, *a constant fictitious mass extended Lagrangian trajectory corresponds to having different values for the convergence criteria at different portions of the potential energy surface*. This could be an extremely powerful distinguishing feature between the two methods.

The preceding analysis also suggests that there may exist two different ways to perform an extended Lagrangian dynamics: (a) the standard Car-Parrinello or ADMP approach, or

(b) the standard approximate BO approach but now by letting the convergence criterion $[\mathbf{F}, \mathbf{P}_{approx}]$ change depending upon the region of the potential energy surface.³⁷ In the second case, proper choice of the convergence criteria could lead to effective sampling of the surface, as should, by extension, be the case for the first strategy. Therefore, we may proceed to state that for every approximate BO dynamics trajectory, there exists an ADMP trajectory with scalar fictitious mass given by

$$\mu = \frac{1}{Tr \left[\frac{d\mathbf{P}_{approx}}{dt} \frac{d^2\mathbf{P}_{approx}}{dt^2} \right]} Tr \left\{ [\mathbf{F}, \mathbf{P}_{approx}] \left[\mathbf{P}_{approx}, \frac{d\mathbf{P}_{approx}}{dt} \right] \right\} \quad (21)$$

where the first and second derivatives maybe obtained from a finite difference approximation.

For an ADMP or a Car-Parrinello dynamics trajectory, the value of the fictitious mass is known. In addition the Fock matrix is also known at every point on the trajectory and hence the value of the commutator is exactly known. Hence, an ADMP trajectory corresponds to a BO dynamics trajectory with the appropriate value for the commutator. However, Eq. (19) also allows one to estimate this value. We will now proceed to estimate the commutator $[\mathbf{F}, \mathbf{P}_{approx}]$ corresponding to the chosen fictitious mass as allowed by Eq. (19). Equation (19) readily simplifies to

$$\left| Tr \left[\mathbf{W} \underline{\mu}^{1/2} \frac{d\mathbf{W}}{dt} \underline{\mu}^{1/2} \right] \right| \leq \|[\mathbf{F}, \mathbf{P}_{approx}]\|_F \left\| \left[\mathbf{P}_{approx}, \frac{d\mathbf{P}_{approx}}{dt} \right] \right\|_F \quad (22)$$

where $\|[\dots]\|_F$ is the Frobenius norm^{26,27} of the commutator and is defined as

$$\|A\|_F = \sqrt{\sum_{i,j} A_{i,j}^2}. \quad (23)$$

Equation (22) uses the Cauchy-Schwarz inequality.²⁶ Hence,

$$\|[\mathbf{F}, \mathbf{P}_{approx}]\|_F \geq \frac{1}{\|[\mathbf{P}_{approx}, \mathbf{W}]\|_F} \left| Tr \left[\mathbf{W} \underline{\mu}^{1/2} \frac{d\mathbf{W}}{dt} \underline{\mu}^{1/2} \right] \right| \quad (24)$$

where we have made the connection that $\{d\mathbf{P}_{approx}/dt\} \approx \mathbf{W}$ as discussed above. A generalization to Eq. (24) can be readily obtained by exploiting the Holder inequality.²⁶

$$\left| Tr \left[\mathbf{W} \underline{\mu}^{1/2} \frac{d\mathbf{W}}{dt} \underline{\mu}^{1/2} \right] \right| \leq \|[\mathbf{F}, \mathbf{P}_{approx}]\|_p \left\| \left[\mathbf{P}_{approx}, \frac{d\mathbf{P}_{approx}}{dt} \right] \right\|_q \quad (25)$$

where the quantity $\|[\dots]\|_p$ is defined as

$$\|A\|_p = \left\{ \sum_{i,j} A_{i,j}^p \right\}^{1/p} \quad (26)$$

and is not to be confused with the L^p -norm of the matrix A . Equation (26) may in fact be viewed as the L^p -norm of the vector created from all the elements in the matrix A . Equation (25) holds true for all p and q that satisfy

$$\frac{1}{p} + \frac{1}{q} = 1. \quad (27)$$

It is readily seen that Eq. (22) is a special case of Eq. (25) for $p = q = 2$. Equation (25) maybe used to obtain alternate bounds on the commutator,

$$\|[\mathbf{F}, \mathbf{P}_{approx}]\|_p \geq \frac{1}{\|[\mathbf{P}_{approx}, \mathbf{W}]\|_q} \left| \text{Tr} \left[\mathbf{W} \underline{\mu}^{1/2} \frac{d\mathbf{W}}{dt} \underline{\mu}^{1/2} \right] \right|. \quad (28)$$

It is important to realize the implications of Eqs. (24) and (28). We already know that for a given ADMP trajectory with a given fictitious mass, the value of the commutator at each point on the trajectory is known. However, Eqs. (24) and (28) only provides a lower bound to such a value as a function of the fictitious mass. This implies that, in ADMP and Car-Parrinello dynamics, the instantaneous commutator is not fully determined by the value of the fictitious mass (as is obvious from the fact that the commutator does depend upon other quantities such as nuclear coordinates which are completely independent of the fictitious mass), but the fictitious mass only provides a lower bound to the commutator; that is the fictitious mass restricts the dynamics to be at least a certain “distance” from the true BO dynamics. This relation provides a “physical” meaning for the “fictitious” mass μ .

Finally, it should be noted that the equations derived in this section hold for any single-particle density matrix representation within the extended Lagrangian formalism regardless of basis set choice, *i.e.* for Gaussian basis sets and plane-waves. For an occupied orbital representation of electronic structure, as is the case in the extended Lagrangian Car-Parrinello approach, similar relations can be easily derived.

Before we conclude this section we provide a computational example in Fig. 1 to illustrate the lower bound presented in Eq. (24). In Fig. 1 we present both the lower bound in Eq. (24) and the Frobenius norm of the commutator $[\mathbf{F}, \mathbf{P}_{approx}]$ obtained from an ADMP simulation of a $\text{Cl}^-(\text{H}_2\text{O})_{25}$ cluster. The trajectory was obtained with the mass-tensor scheme provided in Ref. 9 using a valence fictitious mass value of 0.1 amu bohr² \approx 180 a.u. and an integration time-step of 0.25 fs. The total energy and adiabaticity were well controlled during the dynamics and the simulation was performed using the 3-21G* Gaussian basis set and the PBE exchange-correlation functional²⁸ in a development version of the *Gaussian* series of

electronic structure codes.²⁹ As is seen from the figure, Eq. (24) indeed provides a lower bound to the commutator. (It is important to note here that, generally, for a converged electronic structure calculation, the value for the Frobenius norm of $[\mathbf{F}, \mathbf{P}_{approx}]$ is of the order of 10^{-14} as is the case here for the first step of the dynamics where a full electronic structure calculation was performed.) In Fig. 2 we provide the progress in the adiabaticity index derived in Ref. 9, *i.e.* Eq. (17). It is seen that the index satisfies both conditions required for adiabaticity as derived in Ref. 9, *i.e.*, (a) the index should be oscillatory and (b) the index should be bounded in magnitude.

This proves that there is indeed a direct one-to-one correspondence between trajectories obtained from an approximate Born-Oppenheimer dynamics procedure and those obtained from extended Lagrangian procedures such as those outlined in Refs. 5,8,9. Hence chemically useful dynamics information should be obtainable from both procedures up to similar extents as has been shown in Ref. 10, although it has also been shown in Ref. 10 that the extended Lagrangian scheme is computational advantageous. It is also seen that the extended Lagrangian method provides the added feature that the associated convergence criterion is a time-dependent quantity which is not the case in standard Born-Oppenheimer dynamics implementations.

It is also useful to add here that in practice deviations from the Born-Oppenheimer surface occur due to non-adiabatic coupling between the adiabatic electronic states,³⁰⁻³⁵ and a measure of this is obtained from

$$F_{jk} \equiv \langle \phi_j | \nabla_{\mathbf{R}} | \phi_k \rangle = \frac{\langle \phi_j | \nabla_{\mathbf{R}} H_{el} | \phi_k \rangle}{\epsilon_j - \epsilon_k}, \quad (j \neq k) \quad (29)$$

where $\{\phi_i\}$ are the set of adiabatic electronic states with energies $\{\epsilon_i\}$, and H_{el} is the electronic Hamiltonian (or Fock matrix in our earlier discussion). Since both extended Lagrangian and approximate BO have states that are off the Born-Oppenheimer surface, it is possible to expand each of these in terms of the adiabatic electronic states (which form a complete set):

$$\chi \equiv \sum_i c_i \phi_i, \quad (30)$$

where χ is the approximate state obtained from approximate BO or extended Lagrangian dynamics. For the case where non-adiabatic coupling is small, the quantity in Eq. (29) is small, and the expansion in Eq. (30) is dominated by the lowest energy state. This is, of course, the case only when the approximate state χ is well converged for BO dynamics

(*i.e.*, when the corresponding $\|[\mathbf{F}, \mathbf{P}_{approx}]\|_F$ is small) and the quantity in Eq. (17) (or the fictitious kinetic energy) is small with respect to the HOMO-LUMO gap, as is required from Eq. (29), for extended Lagrangian ADMP dynamics. In fact, for ADMP, the ratio of the fictitious kinetic energy and the HOMO-LUMO gap may be used as an estimate of the level of dominance of an excited state in the expansion in Eq. (30). When this ratio is small, clearly, the dynamics is strictly adiabatic. However, in the event that this ratio becomes large, as will be the case when an unnecessarily large value for the fictitious mass is chosen or in the vicinity of a conical intersection,³⁰⁻³⁵ the approximate state χ will be a linear combination of two or more states (comprising of ground and one or more of the excited states). To maintain reasonable dynamics in this region, it will be necessary to reduce the fictitious mass in ADMP calculations so as to maintain the aforementioned ratio to within reasonable limits. This will also be the case for approximate Born-Oppenheimer dynamics where the value of the $\|[\mathbf{F}, \mathbf{P}_{approx}]\|_F$ will need to be maintained small in the vicinity of a conical intersection. The behavior and generalization of ADMP for such cases will be the subject of future publications.

IV. A DISCUSSION OF THE NUCLEAR FORCES IN THE ADMP FORMALISM

As discussed in Section III, the nuclear forces for the approximate and true BO dynamics schemes should be different because the commutator is not zero in the case of the former. This is also seen from the fact that in Ref. 8 the nuclear forces derived for the purpose of ADMP are different from those standardly used,¹⁹ and are given by

$$\begin{aligned} \left. \frac{\partial E_{approx.BO}}{\partial \mathbf{R}_{i,j}} \right|_{\mathbf{P}_{approx}} &= Tr \left[\left. \frac{d\mathbf{h}'}{d\mathbf{R}_{i,j}} \tilde{\mathbf{P}}'_{approx} + \frac{1}{2} \frac{\partial \mathbf{G}'(\mathbf{P}'_{approx})}{\partial \mathbf{R}_{i,j}} \right|_{\mathbf{P}'_{approx}} \tilde{\mathbf{P}}'_{approx} \right] - \\ &Tr \left[\mathbf{F}' \mathbf{U}^{-1} \frac{d\mathbf{U}}{d\mathbf{R}_{i,j}} \tilde{\mathbf{P}}'_{approx} + \tilde{\mathbf{P}}'_{approx} \frac{d\mathbf{U}^T}{d\mathbf{R}_{i,j}} \mathbf{U}^{-T} \mathbf{F}' \right] + \\ &\left. \frac{\partial E_{xc}}{\partial \mathbf{R}_{i,j}} \right|_{\mathbf{P}_{approx}} + \frac{\partial V_{NN}}{\partial \mathbf{R}_{i,j}}. \end{aligned} \quad (31)$$

Here $\tilde{\mathbf{P}} \equiv 3\mathbf{P}^2 - 2\mathbf{P}^3$ is the McWeeny purified³⁶ density matrix. For both Hartree-Fock and DFT, \mathbf{h}' is the one electron matrix in the non-orthogonal Gaussian basis. The matrix $\mathbf{G}'(\tilde{\mathbf{P}}')$ represents the two electron matrix in the non-orthogonal Gaussian basis 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 non-orthogonal Gaussian basis, \mathbf{S}' , is factorized to yield $\mathbf{S}' = \mathbf{U}^{\mathbf{T}}\mathbf{U}$. There are a number of choices for this transformation (e.g., \mathbf{U} can be obtained by Cholesky decomposition,²⁷ or $\mathbf{U} = \mathbf{S}'^{1/2}$ for Löwdin symmetric orthonormalization). 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 non-orthogonal Gaussian basis, \mathbf{P}' , by $\mathbf{P} \equiv \mathbf{U}\mathbf{P}'\mathbf{U}^{\mathbf{T}}$.

From Eq. (31) it follows that

$$\begin{aligned} \left. \frac{\partial E_{approx.BO}}{\partial \mathbf{R}_{i,j}} \right|_{\mathbf{P}_{approx}} &= Tr \left[\left. \frac{d\mathbf{h}'}{d\mathbf{R}_{i,j}} \tilde{\mathbf{P}}'_{approx} + \frac{1}{2} \frac{\partial \mathbf{G}'(\mathbf{P}'_{approx})}{\partial \mathbf{R}_{i,j}} \right|_{\mathbf{P}'_{approx}} \tilde{\mathbf{P}}'_{approx} \right] - \\ &Tr \left[\mathbf{F} \tilde{\mathbf{P}}'_{approx} \left(\frac{d\mathbf{U}}{d\mathbf{R}_{i,j}} \mathbf{U}^{-1} + \mathbf{U}^{-\mathbf{T}} \frac{d\mathbf{U}^{\mathbf{T}}}{d\mathbf{R}_{i,j}} \right) \right] + \\ &Tr \left([\tilde{\mathbf{P}}'_{approx}, \mathbf{F}] \frac{d\mathbf{U}}{d\mathbf{R}_{i,j}} \mathbf{U}^{-1} \right) + \\ &\left. \frac{\partial E_{xc}}{\partial \mathbf{R}_{i,j}} \right|_{\mathbf{P}_{approx}} + \frac{\partial V_{NN}}{\partial \mathbf{R}_{i,j}}. \end{aligned} \quad (32)$$

For an idempotent density matrix, the second term on the right hand side above can be further simplified to obtain

$$\begin{aligned} \left. \frac{\partial E_{approx.BO}}{\partial \mathbf{R}_{i,j}} \right|_{\mathbf{P}_{approx}} &= \left\{ Tr \left[\left. \frac{d\mathbf{h}'}{d\mathbf{R}_{i,j}} \tilde{\mathbf{P}}'_{approx} + \frac{1}{2} \frac{\partial \mathbf{G}'(\mathbf{P}'_{approx})}{\partial \mathbf{R}_{i,j}} \right|_{\mathbf{P}'_{approx}} \tilde{\mathbf{P}}'_{approx} \right] - \right. \\ &Tr \left[\mathbf{F}' \tilde{\mathbf{P}}'_{approx} \frac{d\mathbf{S}'}{d\mathbf{R}_{i,j}} \tilde{\mathbf{P}}'_{approx} \right] + \left. \frac{\partial E_{xc}}{\partial \mathbf{R}_{i,j}} \right|_{\mathbf{P}_{approx}} + \frac{\partial V_{NN}}{\partial \mathbf{R}_{i,j}} \left. \right\} + \\ &Tr \left[[\tilde{\mathbf{P}}'_{approx}, \mathbf{F}] \left(\tilde{\mathbf{Q}}_{approx} \frac{d\mathbf{U}}{d\mathbf{R}_{i,j}} \mathbf{U}^{-1} - \tilde{\mathbf{P}}'_{approx} \mathbf{U}^{-\mathbf{T}} \frac{d\mathbf{U}^{\mathbf{T}}}{d\mathbf{R}_{i,j}} \right) \right] \end{aligned} \quad (33)$$

where $\tilde{\mathbf{Q}}_{approx} \equiv \mathbf{I} - \tilde{\mathbf{P}}'_{approx}$. For the case where the Fock and density matrices commute (i.e., for converged Born-Oppenheimer calculations), the last term in Eq. (33) is zero which yields

$$\begin{aligned} \left. \frac{\partial E_{BO}}{\partial \mathbf{R}_{i,j}} \right|_{\mathbf{P}} &= Tr \left[\left. \frac{d\mathbf{h}'}{d\mathbf{R}_{i,j}} \tilde{\mathbf{P}}' + \frac{1}{2} \frac{\partial \mathbf{G}'(\tilde{\mathbf{P}}')}{\partial \mathbf{R}_{i,j}} \right|_{\mathbf{P}'} \tilde{\mathbf{P}}' \right] - Tr \left[\mathbf{F}' \tilde{\mathbf{P}}' \frac{d\mathbf{S}'}{d\mathbf{R}_{i,j}} \tilde{\mathbf{P}}' \right] + \\ &\left. \frac{\partial E_{xc}}{\partial \mathbf{R}_{i,j}} \right|_{\mathbf{P}} + \frac{\partial V_{NN}}{\partial \mathbf{R}_{i,j}}. \end{aligned} \quad (34)$$

It is Eq. (34) that is commonly used for the energy derivative in converged SCF calculations.¹⁹ Equation (33), of course, holds even in the general case where the Fock and density matrices do not commute as in the extended Lagrangian formalism^{5,8-10} and the approximate BO methodology discussed in the previous section. The difference between the two is simply given by the commutator:

$$\begin{aligned}
& \left\{ \left. \frac{\partial E_{approx. BO}}{\partial \mathbf{R}_{i,j}} \right|_{\mathbf{P}_{approx}} - \left. \frac{\partial E_{BO}}{\partial \mathbf{R}_{i,j}} \right|_{\mathbf{P}_{approx}} \right\} \\
&= Tr \left[[\tilde{\mathbf{P}}_{\mathbf{approx}}, \mathbf{F}] \left(\tilde{\mathbf{Q}}_{approx} \frac{d\mathbf{U}}{d\mathbf{R}_{i,j}} \mathbf{U}^{-1} - \tilde{\mathbf{P}}_{approx} \mathbf{U}^{-\mathbf{T}} \frac{d\mathbf{U}^{\mathbf{T}}}{d\mathbf{R}_{i,j}} \right) \right] \\
&= Tr \left[[\mathbf{U}, \tilde{\mathbf{P}}'_{\mathbf{approx}}, \mathbf{F}', \mathbf{U}^{-1}] \left(\tilde{\mathbf{Q}}_{approx} \frac{d\mathbf{U}}{d\mathbf{R}_{i,j}} \mathbf{U}^{-1} - \tilde{\mathbf{P}}_{approx} \mathbf{U}^{-\mathbf{T}} \frac{d\mathbf{U}^{\mathbf{T}}}{d\mathbf{R}_{i,j}} \right) \right], \quad (35)
\end{aligned}$$

where we have introduced the notation for higher-order commutators:

$$[\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}] = \mathbf{ABCD} - \mathbf{D}^{\mathbf{T}} \mathbf{C}^{\mathbf{T}} \mathbf{B}^{\mathbf{T}} \mathbf{A}^{\mathbf{T}}. \quad (36)$$

The second part of Eq. (35) will be used later in the appendix.

From the first equality in Eq. (35) it is seen that the difference between the force on the nuclei is proportional to the commutator, which is a measure of the convergence of the electronic subsystem. Furthermore, since the commutator is related to the fictitious mass chosen in the extended Lagrangian formalism as given in Eq. (19), it is also seen that the difference in Eq. (35) is in fact proportional to the magnitude of the fictitious mass. This is particularly interesting since the greater the magnitude of the fictitious mass, the farther is the approximate trajectory (approximate Born-Oppenheimer or extended Lagrangian) from the true Born-Oppenheimer result and the greater is the difference between the nuclear force for the true Born-Oppenheimer case from the approximate methods. Hence, the force in Eq. (33) has the important feature that it adapts to any deviations from the Born-Oppenheimer result.

The presence of the transformation matrix \mathbf{U} in all of the above equations for the forces raises the important question of dependence of the nuclear forces on choice of orthonormal basis. Consider a *time-independent* unitary transformation matrix, \mathbf{G} , such that

$$\mathbf{U}' = \mathbf{G}\mathbf{U}. \quad (37)$$

The nuclear forces are, however, invariant to such a transformation as is revealed by inspec-

tion of the second term in Eq. (31) where

$$\begin{aligned} \text{Tr} \left[\mathbf{F}' \mathbf{U}'^{-1} \frac{d\mathbf{U}'}{d\mathbf{R}_{i,j}} \tilde{\mathbf{P}}'_{\text{approx}} + \tilde{\mathbf{P}}'_{\text{approx}} \frac{d\mathbf{U}'^T}{d\mathbf{R}_{i,j}} \mathbf{U}'^{-T} \mathbf{F}' \right] = \\ \text{Tr} \left[\mathbf{F}' \mathbf{U}^{-1} \frac{d\mathbf{U}}{d\mathbf{R}_{i,j}} \tilde{\mathbf{P}}'_{\text{approx}} + \tilde{\mathbf{P}}'_{\text{approx}} \frac{d\mathbf{U}^T}{d\mathbf{R}_{i,j}} \mathbf{U}^{-T} \mathbf{F}' \right]. \end{aligned} \quad (38)$$

This proves the invariance of the nuclear forces in Eq. (33) with respect to a time-independent unitary transformations. The issue of time-dependent unitary transformations is considered in Appendix A.

V. CONCLUSIONS

In this paper, we have analyzed the relation between the fictitious mass in the newly developed ADMP molecular dynamics approach and the commutator of the Fock and density matrices. We find that the fictitious mass is proportional to the commutator and hence determines the deviations of the ADMP trajectory from the true Born-Oppenheimer trajectory. Furthermore the choice of the fictitious mass limits the magnitude of the commutator, hence providing a lower bound on the deviations from the Born-Oppenheimer surface. We also note that for every approximate Born-Oppenheimer dynamics trajectory there exists an extended Lagrangian or ADMP trajectory that deviates identically from the Born-Oppenheimer surface and hence contains the same information as the approximate Born-Oppenheimer dynamics trajectory. However, standard approaches to the Born-Oppenheimer dynamics limit the size of the convergence criteria (the commutator of the Fock and density matrix) uniformly over the entire surface. This corresponds to an ADMP trajectory with time-dependent fictitious mass. By corollary, a constant (time-independent) fictitious mass ADMP trajectory corresponds to having different values for the convergence criterion at different points in the potential energy surface. This is an important property in extended Lagrangian approaches to molecular dynamics (both CP and ADMP) since it is generally not necessary to converge the electronic structure to the same extent at different areas in the potential energy surface.

We have also analyzed the forces used in ADMP and shown that these are more general than those commonly used in Born-Oppenheimer dynamics on account of terms included that are non-zero when the commutator of the Fock and density matrices is non-zero (as is the case in ADMP and other approximations to Born-Oppenheimer dynamics).

Using the analyses in the current paper it may be possible to derive a new method to perform *ab initio* molecular dynamics wherein the computational effectiveness of ADMP (or CP extended Lagrangian dynamics) is further enhanced by partial convergence of the density matrix. The extent of partial convergence may be determined based on the relations between the fictitious mass and the commutator derived here. The generality of the ADMP nuclear forces due to inclusion of additional terms that depend on the commutator (and hence the fictitious mass) allows for correction due to such partial convergence and may in fact facilitate larger time-steps to be used within a modified formalism. This aspect will be studied in future publications.

Acknowledgments

This work was supported by the Office of Naval Research (GAV), the Natural Science Foundation (CHE-9982156 and CHE-0131157), and Gaussian, Inc. An allocation of computer time from the Center of High Performance Computing at the University of Utah is gratefully acknowledged.

APPENDIX A: UNITARITY OF PROPAGATION SCHEME FOR LÖWDIN AND CHOLESKY ORTHONORMAL BASES

In this section we evaluate the behavior of the propagation scheme under unitary transformations. In particular we will concentrate on the effect of unitary transformations on Eqs. (4), (12) and (33).

1. Nuclear Forces

The unitary transformation, \mathbf{G} , defined in Eq. (37) is coordinate-dependent (and hence time-dependent) in the case for transformations between Löwdin and Cholesky bases where

$$\mathbf{G} = \mathbf{S}'^{1/2}\mathbf{U}^{-1} = \mathbf{S}'^{-1/2}\mathbf{U}^T \tag{A1}$$

for $\mathbf{U}' = \mathbf{S}'^{1/2}$. To analyze the effect of this on the nuclear forces for Löwdin and Cholesky propagation schemes we consider the difference in forces between the two separate represen-

tations using Eqs. (33) and (35):

$$\begin{aligned}
& Tr \left[[\mathbf{U}, \tilde{\mathbf{P}}'_{approx}, \mathbf{F}', \mathbf{U}^{-1}] \left(\tilde{\mathbf{Q}}_{approx} \frac{d\mathbf{U}}{d\mathbf{R}_{i,j}} \mathbf{U}^{-1} - \tilde{\mathbf{P}}_{approx} \mathbf{U}^{-\mathbf{T}} \frac{d\mathbf{U}^{\mathbf{T}}}{d\mathbf{R}_{i,j}} \right) \right] - \\
& Tr \left[[\mathbf{S}'^{1/2}, \tilde{\mathbf{P}}'_{approx}, \mathbf{F}', \mathbf{S}'^{-1/2}] \left(\tilde{\mathbf{Q}}_{approx} \frac{d\mathbf{S}'^{1/2}}{d\mathbf{R}_{i,j}} \mathbf{S}'^{-1/2} - \tilde{\mathbf{P}}_{approx} \mathbf{S}'^{-1/2} \frac{d\mathbf{S}'^{1/2}}{d\mathbf{R}_{i,j}} \right) \right] \\
& = Tr \left[[\tilde{\mathbf{P}}_{approx}, \mathbf{F}] \left(\frac{d\mathbf{S}'^{1/2} \mathbf{U}^{-1}}{d\mathbf{R}_{i,j}} \mathbf{U}^{-\mathbf{T}} \mathbf{S}'^{1/2} \right) \right] \tag{A2}
\end{aligned}$$

This implies that the difference in the forces between Löwdin and Cholesky bases for ADMP are proportional to the commutator and hence proportional to the fictitious mass as seen in Section III. However, since the dynamics in both orthonormal bases are only approximations to Born-Oppenheimer dynamics (which is independent of choice of orthonormal bases as is clear from Eq. (34)) and since these approximations can be adaptively controlled by the choice of fictitious mass as seen in Eq. (20) to improve their “closeness” to the Born-Oppenheimer surface, this difference does not present a computational problem in the limit of acceptable values for the fictitious mass. This aspect has been well-demonstrated in Ref. 10 for a wide range of chemical applications.

2. Density matrix Propagation scheme

Consider the Cholesky form of the density propagation scheme in Eq. (4):

$$\mathbf{P}_{i+1}^{\mathbf{C}} = \mathbf{P}_i^{\mathbf{C}} + \mathbf{W}_i^{\mathbf{C}} \Delta t - \frac{\Delta t^2}{2} \underline{\mu}_{\mathbf{C}}^{-1/2} \left[\left. \frac{\partial E(\mathbf{R}_i, \mathbf{P}_i^{\mathbf{C}})}{\partial \mathbf{P}} \right|_{\mathbf{R}} + \underline{\Lambda}_i^{\mathbf{C}} \mathbf{P}_i^{\mathbf{C}} + \mathbf{P}_i^{\mathbf{C}} \underline{\Lambda}_i^{\mathbf{C}} - \underline{\Lambda}_i^{\mathbf{C}} \right] \underline{\mu}_{\mathbf{C}}^{-1/2}, \tag{A3}$$

where the letter \mathbf{C} in subscript (and superscript) is used to designate the Cholesky form. Since, $\mathbf{P}_i^{\mathbf{L}} = \mathbf{G}_i \mathbf{P}_i^{\mathbf{C}} \mathbf{G}_i^{\mathbf{T}}$ where \mathbf{G}_i is the transformation matrix of Eq. (A1) at time step i , unitary transformation of Eq. (A3) yields

$$\begin{aligned}
\mathbf{G}_i \mathbf{P}_{i+1}^{\mathbf{C}} \mathbf{G}_i^{\mathbf{T}} &= \mathbf{P}_i^{\mathbf{L}} + \mathbf{W}_i^{\mathbf{L}} \Delta t - \\
& \frac{\Delta t^2}{2} \left\{ \mathbf{G}_i \underline{\mu}_{\mathbf{C}}^{-1/2} \mathbf{G}_i^{\mathbf{T}} \right\} \left[\mathbf{G}_i \left. \frac{\partial E(\mathbf{R}_i, \mathbf{P}_i^{\mathbf{C}})}{\partial \mathbf{P}^{\mathbf{C}}} \right|_{\mathbf{R}} \mathbf{G}_i^{\mathbf{T}} + \left\{ \mathbf{G}_i \underline{\Lambda}_i^{\mathbf{C}} \mathbf{G}_i^{\mathbf{T}} \right\} \left\{ \mathbf{G}_i \mathbf{P}_i^{\mathbf{C}} \mathbf{G}_i^{\mathbf{T}} \right\} + \right. \\
& \left. \left\{ \mathbf{G}_i \mathbf{P}_i^{\mathbf{C}} \mathbf{G}_i^{\mathbf{T}} \right\} \left\{ \mathbf{G}_i \underline{\Lambda}_i^{\mathbf{C}} \mathbf{G}_i^{\mathbf{T}} \right\} - \left\{ \mathbf{G}_i \underline{\Lambda}_i^{\mathbf{C}} \mathbf{G}_i^{\mathbf{T}} \right\} \right] \left\{ \mathbf{G}_i \underline{\mu}_{\mathbf{C}}^{-1/2} \mathbf{G}_i^{\mathbf{T}} \right\} \\
& = \mathbf{P}_i^{\mathbf{L}} + \mathbf{W}_i^{\mathbf{L}} \Delta t - \frac{\Delta t^2}{2} \underline{\mu}_{\mathbf{L}}^{-1/2} \left[\left. \frac{\partial E(\mathbf{R}_i, \mathbf{P}_i^{\mathbf{L}})}{\partial \mathbf{P}^{\mathbf{L}}} \right|_{\mathbf{R}} + \underline{\Lambda}_i^{\mathbf{L}} \mathbf{P}_i^{\mathbf{L}} + \mathbf{P}_i^{\mathbf{L}} \underline{\Lambda}_i^{\mathbf{L}} - \underline{\Lambda}_i^{\mathbf{L}} \right] \underline{\mu}_{\mathbf{L}}^{-1/2} \tag{A4}
\end{aligned}$$

where, $\{\dots\}^C$ and $\{\dots\}^L$ represent Cholesky and Löwdin forms of the respective operators and

$$\left. \frac{\partial E(\mathbf{R}_i, \mathbf{P}_i^L)}{\partial \mathbf{P}^C} \right|_{\mathbf{R}} = \mathbf{G}_i \left. \frac{\partial E(\mathbf{R}_i, \mathbf{P}_i^C)}{\partial \mathbf{P}^L} \right|_{\mathbf{R}} \mathbf{G}_i^T, \quad (\text{A5})$$

which is clear from Eq. (13),

$$\Lambda_i^L = \mathbf{G}_i \Lambda_i^C \mathbf{G}_i^T \quad (\text{A6})$$

and

$$\underline{\mu}_L^{-1/2} = \mathbf{G}_i \underline{\mu}_C^{-1/2} \mathbf{G}_i^T. \quad (\text{A7})$$

For the case of the non-mass-weighted dynamics

$$\underline{\mu}_C^{-1/2} = \underline{\mu}_L^{-1/2}. \quad (\text{A8})$$

The left hand side in Eq. (A4) may be used to obtain the Löwdin density matrix at time-step $(i + 1)$ if $\mathbf{G}_{i+1} \approx \mathbf{G}_i$ which is the case for suitably small time-steps. Hence the density matrices do satisfy the conditions of unitarity.

-
- ¹ H. B. Schlegel, *J. Comp. Chem.* (manuscript accepted).
- ² K. Bolton, W. L. Hase, and G. H. Peslherbe, *Modern Methods for Multidimensional Dynamics Computation in Chemistry* (World Scientific, Singapore, 1998), chap. Direct Dynamics of Reactive Systems, p. 143.
- ³ M. C. Payne, M. P. Teter, D. C. Allan, T. A. Arias, and J. D. Joannopoulos, *Rev. Mod. Phys.* **64**, 1045 (1992).
- ⁴ E. Deumens, A. Diz, R. Longo, and Y. Ohrn, *Rev. Mod. Phys.* (1994).
- ⁵ R. Car and M. Parrinello, *Phys. Rev. Lett.* **55**, 2471 (1985).
- ⁶ D. K. Remler and P. A. Madden, *Mol. Phys* **70**, 921 (1990).
- ⁷ D. Marx and J. Hutter, *Modern Methods and Algorithms of Quantum Chemistry* (John von-Neumann Institute for Computing, Julich, 2000), vol. 1, chap. Ab Initio Molecular Dynamics: Theory and Implementation, p. 301.
- ⁸ 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).
- ⁹ 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).
- ¹⁰ H. B. Schlegel, S. S. Iyengar, X. Li, J. M. Millam, G. A. Voth, G. E. Scuseria, and M. J. Frisch, *J. Chem. Phys.* (Manuscript Accepted).
- 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 G. E. Scuseria, *J. Phys. Chem. A* **103**, 4782 (1999).
- ¹⁶ N. Rega, S. S. Iyengar, G. A. Voth, H. B. Schlegel, G. E. Scuseria, and M. J. Frisch, manuscript in preparation.
- ¹⁷ S. S. Iyengar, K. N. Kudin, H. B. Schlegel, G. A. Voth, G. E. Scuseria, and M. J. Frisch, manuscript in preparation.
- ¹⁸ S. S. Iyengar, T. J. F. Day, and G. A. Voth, *J. Am. Chem. Soc.* (Submitted).
- ¹⁹ P. Pulay, *Mol. Phys.* **17**, 197 (1969).

- ²⁰ W. C. Swope, H. C. Andersen, P. H. Berens, and K. R. Wilson, *J. Chem. Phys.* **76**, 637 (1982).
- ²¹ M. E. Tuckerman and M. Parrinello, *J. Chem. Phys.* (1994).
- ²² J. M. Millam, V. Bakken, W. Chen, W. L. Hase, and H. B. Schlegel, *J. Chem. Phys.* **111**, 3800 (1999).
- ²³ X. Li, J. M. Millam, and H. B. Schlegel, *J. Chem. Phys.* **113**, 10062 (2000).
- ²⁴ H. Goldstein, *Classical Mechanics* (Addison-Wesley Press, Cambridge, Massachusetts, 1980).
- ²⁵ P. Tangney and S. Scandolo, *J. Chem. Phys.* **116**, 14 (2002).
- ²⁶ F. Riesz and B. Sz.-Nagy, *Functional Analysis* (Dover Publications, Inc., New York, 1990).
- ²⁷ G. H. Golub and C. F. van Loan, *Matrix Computations* (The Johns Hopkins University Press, Baltimore, 1996).
- ²⁸ J. P. Perdew, K. Burke, and M. Ernzerhof, *Phys. Rev. Lett.* **77**, 3865 (1996).
- ²⁹ M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. A. Montgomery, Jr., T. Vreven, K. N. Kudin, et al., *Gaussian 01, development version (revision b.01)*, Gaussian, Inc., Pittsburgh PA, 2001.
- ³⁰ B. K. Kendrick, C. A. Mead, and D. G. Truhlar, *Chem. Phys.* **277**, 31 (2002).
- ³¹ H. C. Longuet-Higgins, U. Opik, H. L. Pryce, and R. A. Sack, *Proc. R. Soc. London Ser. A* **244**, 1 (1958).
- ³² M. V. Berry, *Proc. R. Soc. London Ser. A* **392**, 45 (1984).
- ³³ C. A. Mead and D. G. Truhlar, *J. Chem. Phys.* **70**, 2284 (1979).
- ³⁴ A. Kuppermann, *Dynamics of Molecules and Chemical Reactions* (Marcel Dekker Inc., New York, New York, 1996), chap. The Geometric Phase in Reaction Dynamics, pp. 411–472.
- ³⁵ D. R. Yarkony, *Rev. Mod. Phys.* **68**, 985 (1996).
- ³⁶ R. McWeeny, *Rev. Mod. Phys.* **32**, 335 (1960).
- ³⁷ It is important to note that partial convergence as outlined in (b) has been exploited by a number of authors previously, as for example in the conjugate gradient methodology employed in Ref.³. In the present discussion, however the convergence criterion is time-dependent. Furthermore the use of plane-waves as basis sets in Ref.³ ensures that the approximate wavefunction always lags behind its exact counterpart. This aspect may be generally circumvented by the use of atom-centered basis functions that do not allow for charge “sloshing” as they move with the nuclei.

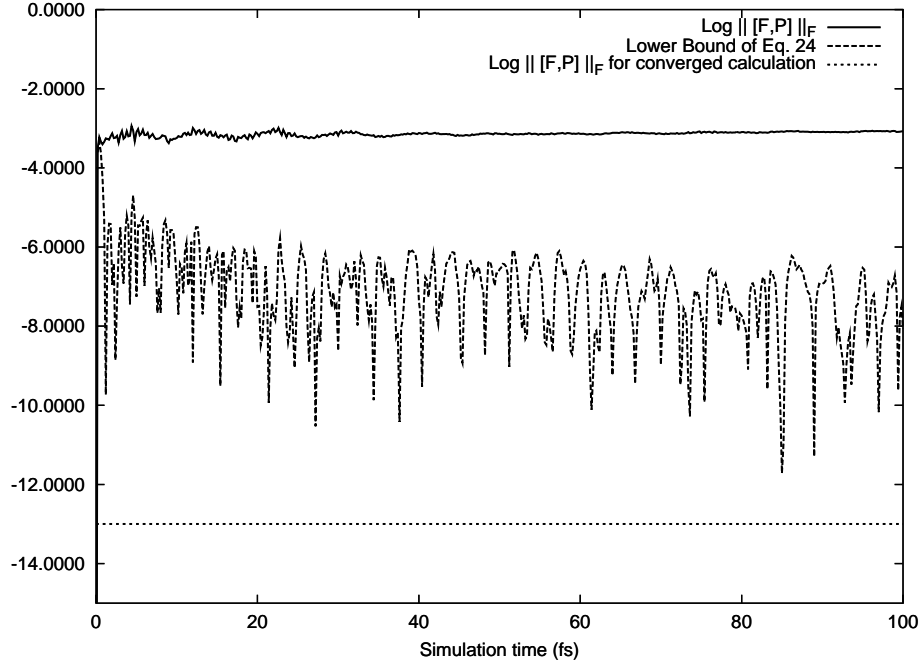


FIG. 1: Time evolution of $\|[\mathbf{F}, \mathbf{P}_{approx}]\|_F$ and the lower bound in Eq. (24) for the $\text{Cl}^-(\text{H}_2\text{O})_{25}$ cluster. Note that the Frobenius norm of the commutator for a converged calculation is approximately 10^{-14} .

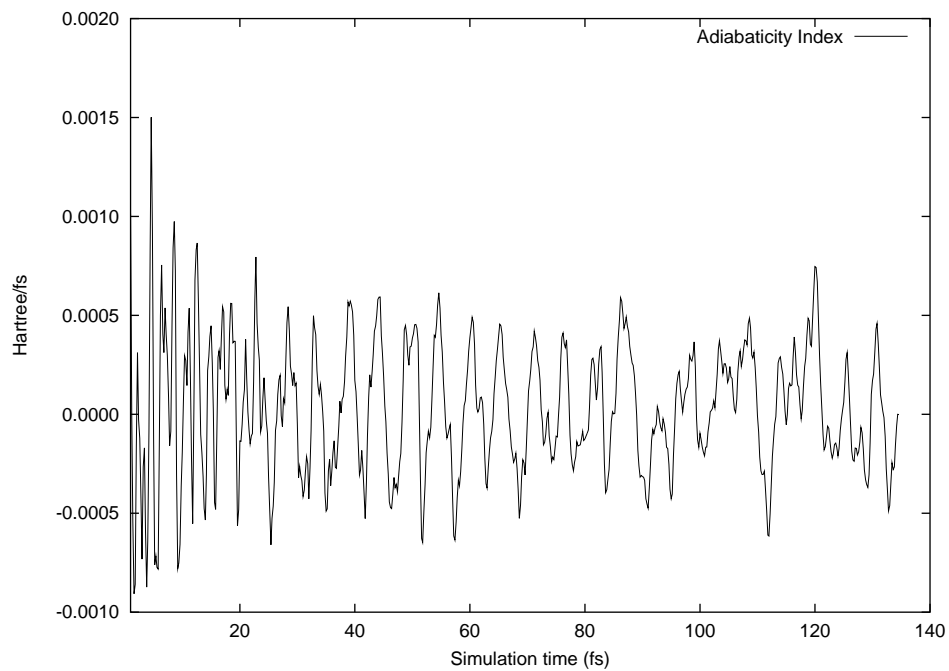


FIG. 2: The oscillatory nature of the adiabaticity index in Eq. (17) is well depicted here for the $\text{Cl}^-(\text{H}_2\text{O})_{25}$ cluster. This illustrates the conditions of adiabaticity derived in Ref.⁹.