Curvy-steps approach to constraint-free extended-Lagrangian *ab initio* molecular dynamics, using atom-centered basis functions: Convergence toward Born–Oppenheimer trajectories

John M. Herbert[a)] and Martin Head-Gordon[b)]
Department of Chemistry, University of California, Berkeley, California 94720

A dynamical extension of the ''curvy-steps'' approach to linear-scaling self-consistent field calculations is presented, which yields an extended-Lagrangian formulation of *ab initio* molecular dynamics. An exponential parametrization of the one-electron density matrix, expressed in terms of atom-centered Gaussian basis functions, facilitates propagation along the manifold of density matrices in a geometrically correct fashion that automatically enforces idempotency constraints. The extended Lagrangian itself is constraint free, thus neither density matrix purification nor expensive, iterative solution for Lagrange multipliers is required. Propagation is highly efficient, and time steps compare favorably to those used in Car–Parrinello molecular dynamics simulations. The behavior of the method, especially with regard to the maintenance of adiabatic decoupling of nuclei and electrons, is examined for a sequence of diatomic molecules, and comparison is made to trajectories propagated on the converged Born–Oppenheimer surface. Certain claims to the contrary notwithstanding, our results demonstrate that vibrational frequencies may depend on the value of the fictitious mass parameter, even in an atom-centered basis. Light-atom stretching frequencies can be significantly redshifted, even when the nuclear and electronic energy scales are well separated. With a sufficiently small fictitious mass and a short time step, accurate frequencies can be obtained; we characterize appropriate values of these parameters for a wide range of vibrational frequencies. © *2004 American Institute of Physics.* [DOI: 10.1063/1.1814934]

I. INTRODUCTION

Over the past two decades, Car–Parrinello molecular dynamics (CPMD) (Refs. 1–6) has become the *de facto* tool for simulation of reactive chemical systems and myriad other environments where analytic potentials are either unavailable or unreliable. While the electronic basis functions in traditional CPMD are plane waves, atom-centered basis functions are in many ways more suitable for nonperiodic molecular systems. Gaussian atomic orbitals are intrinsically localized, which not only reduces basis-set demands, relative to plane waves (by facilitating compact expansions of both core and valence molecular orbitals), but also greatly expedites reduced- and linear-scaling algorithms for computing the electronic energy and its derivatives. The presence of overlap derivatives in an atom-centered basis does introduce a small (but much-maligned) amount of additional overhead; however, computation of such quantities has been a standard feature of geometry optimization since the early days of quantum chemistry.[7]

One of the most expensive steps in the CPMD algorithm is the iterative solution, at every time step, for a matrix of Lagrange multipliers.[8,9] These define a ''constraint force'' in the equations of motion that is necessary in order to maintain orthogonality of the molecular orbitals. In this paper, we introduce an alternative to the Car–Parrinello formulation of extended-Lagrangian molecular dynamics (ELMD), in which the Lagrangian is free of constraints. This simplification is achieved by formulating the self-consistent field (SCF) problem in terms of the one-electron density matrix, rather than the molecular orbitals (MOs); by an exponential parametrization of a unitary density-matrix update; and finally, by transforming to generalized electronic coordinates in which idempotency of the density matrix is maintained automatically. These generalized coordinates propagate without constraints. Our ELMD formalism can be combined with either a plane-wave or an atom-centered basis, and here we choose the latter, in order to exploit the advantages extolled above.

This approach to ELMD represents a dynamical extension of the curvy-steps approach to linear-scaling SCF calculations that our group has pioneered.[10,11] We call this method ''curvy steps'' in order to emphasize that the exponentially transformed density matrix parametrizes geodesics (shortest-distance paths) along the Grassmann manifold of admissible density matrices.[12] This manifold is curved as a consequence of the idempotency (or orthogonality) constraints on the SCF problem. Because we incorporate these constraints automatically into a geometrically correct energy gradient, our curvy-steps ELMD technique is highly efficient, requiring at each time step only a few matrix multiplications in addition to the usual Fock build and energy gradient computations.

Recently, Schlegel and co-workers[13–16] have introduced an alternative, Gaussian-orbital-based ELMD method that they call ''adiabatic density matrix propagation'' (ADMP).

a)Electronic mail: herbert@bastille.cchem.berkeley.edu
b)Electronic mail: mhg@bastille.cchem.berkeley.edu

As the name implies, the density matrix elements are propagated directly, without the transformation to generalized coordinates, and this necessitates iterative solution of Lagrange multiplier equations, as well as iterative purification[17,18] of a nonidempotent (and therefore nonphysical) density matrix at each time step in the propagation. From a geometrical point of view, ADMP propagates the density matrix along straight-line trajectories, as opposed to curved ones, and such linear updates cannot remain on the Grassmann manifold.[11]

This paper presents the details of our algorithm (Sec. II) and its application to simulate vibrational spectra for a sequence of diatomic molecules (Sec. III). Our main goal at present is to characterize curvy-steps ELMD in terms of time step and fictitious mass parameters, and to make detailed comparison to Born–Oppenheimer molecular dynamics (BOMD), in which the SCF calculation is converged at each time step. However, these simple systems allow us to examine carefully the important question of whether—and to what extent—ELMD vibrational frequencies depend upon the fictitious mass parameter. While such a dependence has been documented in CPMD,[19–21] the developers of ADMP claim,[15] based upon examination of the NaCl molecule, that the use of atom-centered basis functions eliminates this undesirable feature of ELMD.

Although we concur with this result for NaCl, it unfortunately lacks generality: we demonstrate cases—especially, light-atom stretching vibrations—for which the vibrational frequency decreases significantly as the fictitious mass is increased, even while the nuclear and electronic energy scales remain separated by several times the energy of the vibrational fundamental. Other standard metrics used to judge the quality of ELMD simulations also belie this fictitious mass dependence. The only reliable indications of the accuracy of ELMD frequencies are the maximum value of the fictitious kinetic energy, along with a quantitative estimate of the slowest electronic time scale.

II. THEORY

A. Curvy-steps extended Lagrangian

At the heart of the curvy-steps SCF method[10,11] is a unitary update of a reference density matrix $\mathbf{P}(0)$,

$$\mathbf{P}(\lambda) = e^{\lambda\mathbf{\Delta}}\mathbf{P}(0)e^{-\lambda\mathbf{\Delta}}, \tag{1}$$

parametrized in terms of a skew-symmetric matrix $\mathbf{\Delta} = -\mathbf{\Delta}^\dagger$ of curvy-steps variables. In this work, $\mathbf{P}$ is expressed in an orthogonal basis, as a matter of convenience rather than necessity. The curvy-steps update Eq. (1) preserves idempotency and normalization of the density matrix $\mathbf{P}(\lambda)$, provided that $\mathbf{P}(0)$ satisfies these constraints. Although the unitary transformation that connects $\mathbf{P}(\lambda)$ to $\mathbf{P}(0)$ can be parametrized in several ways,[22–26] the exponential parametrization alone has the *geodesic property*: Eq. (1) represents the shortest path between the points $\mathbf{P}(\lambda)$ and $\mathbf{P}(0)$, along the curved manifold of one-electron density matrices.[11,12] A similar parametrization of the MOs has been exploited for constraint-free CPMD,[27] but to our knowledge this technique has not been applied to dynamics in the density matrix representa-

tion. As discussed in the following section, the density-matrix approach is formally simpler and exhibits better scaling than the MO-based method.

A curvy-steps SCF calculation involves minimizing the energy, for a fixed Fock matrix, along the geodesic $\mathbf{P}(\lambda)$, but here we take the transformation variables Δ_{ij} ($i<j$) as generalized coordinates representing the electronic degrees of freedom. This suggests an extended Lagrangian of the form

$$\mathcal{L} = \frac{1}{2}\sum_I M_I \dot{R}_I^2 + T_F - E_{\text{SCF}} - V_{\text{nuc}}, \tag{2}$$

in which E_{SCF} is the SCF electronic energy, V_{nuc} is the nuclear Coulomb repulsion energy, and

$$T_F = \frac{\mu}{2}\sum_{i<j}\dot{\Delta}_{ij}^2 = -\frac{\mu}{4}\text{Tr}(\dot{\mathbf{\Delta}}^2) \tag{3}$$

is the fictitious kinetic energy, with "mass" parameter μ. (μ actually has dimensions of a moment of inertia.) For future reference, we express the electronic energy

$$E_{\text{SCF}} = \text{Tr}[(\mathbf{h} + \tfrac{1}{2}\mathbf{H})\mathbf{P}] \tag{4}$$

in terms of a core Hamiltonian $\mathbf{h}$ that is independent of $\mathbf{P}$, along with a $\mathbf{P}$-dependent exchange and/or correlation contribution $\mathbf{H}$. (In Hartree-Fock theory, for example, $\mathbf{H}$ is equal to $\mathbf{P}$ contracted with the two-electron integrals.[28,29]) The Fock matrix $\mathbf{F} = \mathbf{h} + \mathbf{H}$ is a function of both the nuclear Cartesian coordinates $\{R_I\}$ and the curvy-steps electronic variables $\{\Delta_{ij}|i<j\}$, while $\mathbf{P}$ in our formulation is viewed as a function of $\mathbf{\Delta}$ only.

The Lagrangian in Eq. (2) leads, via the stationary-action principle,[4] to Euler-Lagrange equations of motion

$$M_I\ddot{R}_I = -\frac{\partial(E_{\text{SCF}} + V_{\text{nuc}})}{\partial R_I} \tag{5a}$$

for the nuclear Cartesian coordinates and

$$\mu\ddot{\Delta}_{ij} = -\frac{\partial E_{\text{SCF}}}{\partial\Delta_{ij}} \tag{5b}$$

for the curvy-steps variables.

The standard CPMD Lagrangian[1,4,30] differs from Eq. (2) in that the electronic variables in CPMD are the occupied Kohn-Sham MOs ψ_i (with velocities $\dot{\psi}_i$), while in ADMP (Refs. 13–16) the density matrix elements P_{ij} serve as electronic coordinates. Relative to ADMP we have transformed the Lagrangian into generalized coordinates in which the constraints disappear, since $\mathbf{P}^2(t) = \mathbf{P}(t)$ is maintained at all times t, and therefore the generalized coordinates can be varied freely. Otherwise, the Lagrangian must contain an additional term representing the idempotency violation of $\mathbf{P}$ or the orthogonality violation of the MOs. Our formulation thus obviates one of the most expensive steps[8,9] in traditional ELMD, namely, iterative solution for the Lagrange multipliers that define the constraint force. In place of this force, we have a gradient $\partial E_{\text{SCF}}/\partial\mathbf{\Delta}$ and an electronic trajectory $\mathbf{P}(\lambda)$ that are geometrically correct for the Grassmann manifold on which the SCF problem resides.[11,12,24,31]

B. Curvy-steps ELMD algorithm

Our *ab initio* molecular dynamics scheme consists in integrating the equations of motion, Eqs. (5), using the velocity Verlet algorithm,[32]

$$r(t+\delta t)=r(t)+v(t)\delta t+\tfrac{1}{2}(\delta t)^2 a(t), \quad (6a)$$

$$v(t+\delta t)=v(t)+\tfrac{1}{2}\delta t[a(t)+a(t+\delta t)], \quad (6b)$$

with r, v, and a replaced by the appropriate nuclear or electronic coordinates, velocities, and accelerations, respectively. In particular, for the electronic degrees of freedom we propagate $r\equiv\Delta_{ij}$, $v\equiv\dot\Delta_{ij}$, and $a\equiv-\mu^{-1}(\partial E_{SCF}/\partial\Delta_{ij})$, for each $i<j$. Propagation of the nuclear variables is straightforward but propagation of the electronic degrees of freedom warrants some comment.

First of all, based upon its satisfactory performance in large-molecule SCF calculations,[10] we have opted to use a basis of Cholesky-orthogonalized atomic orbitals (AOs). The Cholesky basis is defined by the decomposition

$$\mathbf{S}_{AO}=\mathbf{LL}^\dagger=\mathbf{Z}^{-1}(\mathbf{Z}^{-1})^\dagger \quad (7)$$

of the AO overlap matrix $\mathbf{S}_{AO}$, with $\mathbf{L}=\mathbf{Z}^{-1}$ a lower triangular matrix. The Fock and density matrices in this basis are

$$\mathbf{F}=\mathbf{Z}\mathbf{F}_{AO}\mathbf{Z}^\dagger \quad (8)$$

and

$$\mathbf{P}=\mathbf{L}^\dagger\mathbf{P}_{AO}\mathbf{L}. \quad (9)$$

Formally, inversion of $\mathbf{L}$ is inconsistent with linear scaling and ought to be replaced by an algorithm that forms the sparse matrix $\mathbf{Z}$ directly, via biorthogonalization of $\mathbf{S}_{AO}$,[33–35] and furthermore eliminates the need for $\mathbf{L}$ by working in the AO basis where necessary. For systems amenable to modern hardware and algorithms, however, formation of both $\mathbf{Z}$ and $\mathbf{L}$ constitutes a very small fraction of the total computational effort,[10] so in our implementation both are computed using standard dense-matrix LAPACK routines.

The electronic forces contain a derivative of $\mathbf{P}$ with respect to Δ, by virtue of Eqs. (1) and (4). This derivative can be expressed in closed form,[27] but only when $\Delta=0$ does one obtain an expression that can be evaluated exactly without diagonalizing Δ:

$$\mathbf{G}\equiv\frac{\partial E_{SCF}}{\partial\Delta}\bigg|_{\Delta=0}=\mathbf{FP}-\mathbf{PF}. \quad (10)$$

(As Δ and the nuclear coordinates $\vec{R}$ are complementary variables, it is implied that the latter are fixed in derivatives with respect to the former.) Diagonalization of Δ is avoided by resetting Δ to the zero matrix at the end of every time step, which amounts to a translation of the electronic coordinate frame. Commensurate translation ("parallel transport") of the electronic velocities is nontrivial, however, since the Grassmann manifold is non-Euclidean, but in practice we find this to be numerically unimportant and thus we do not alter $\dot\Delta$ when Δ is reset to zero. Further exposition concerning parallel transport and other geometric aspects of curvy-steps ELMD is provided in Sec. II C.

First let us explicate our updating scheme for the electronic variables. Following the nth time step, a new matrix Δ_{n+1} of curvy-steps variables is generated according to

$$\Delta_{n+1}=\delta t\dot\Delta_n-[(\delta t)^2/2\mu]\mathbf{G}_n, \quad (11)$$

which is merely the velocity Verlet update of Eq. (6a), starting from $\Delta=0$. The nuclear coordinates $\vec{R}_n$ are updated at the same time, following which the nuclear and electronic half-step velocities,[32]

$$v(t+\delta t/2)=v(t)+\tfrac{1}{2}a(t)\delta t, \quad (12)$$

are computed, since these require only the forces saved from the previous step. For the electronic variables, this corresponds to a velocity update

$$\dot\Delta_n\rightarrow\dot\Delta_n-(\delta t/2\mu)\mathbf{G}_n. \quad (13)$$

Having computed nuclear and electronic half-step velocities, we update the density matrix according to

$$\mathbf{P}_{n+1}=e^{\Delta_{n+1}}\mathbf{P}_n e^{-\Delta_{n+1}}. \quad (14)$$

Exact evaluation of $\exp(\Delta_{n+1})$ requires diagonalization of the $N_{AO}\times N_{AO}$ matrix Δ_{n+1}. We circumvent this expensive scenario by means of a truncated Baker-Campbell-Hausdorff (BCH) nested commutator expansion of Eq. (14), which is extremely efficient given the small size of a typical ELMD time step. In an orthogonal basis, only one matrix multiplication per BCH order is required,[10] and even with highly conservative thresholds we find that four to six terms are sufficient to converge the expansion. Idempotency of the density matrix is maintained below the BCH truncation threshold.

In view of this rapid convergence, we speculate that series expansion of $\exp(\Delta)$ might improve the efficiency of MO-based exponential transformations as well. In these methods, the MO coefficient matrix $\mathbf{C}$ is updated according to $\mathbf{C}_{n+1}=\mathbf{C}_n\exp(\Delta_{n+1})$, which forms the basis of a constraint-free CPMD technique introduced by Hutter, Parrinello, and Vogel[27] and has also seen widespread use in static SCF calculations (consult Refs. 22–25 for a bibliography). Because the energy depends only on the N_{occ} occupied columns of $\mathbf{C}$, this update can be accomplished by diagonalizing an $N_{occ}\times N_{occ}$ matrix rather than an $N_{AO}\times N_{AO}$ matrix,[24,27] but for large systems this is liable to remain more expensive than a truncated expansion of $\exp(\Delta)$, especially in dynamics methods, where the density matrix changes little in a single step.

Returning to our own algorithm, following the density matrix update we next generate new transformation matrices $\mathbf{Z}_{n+1}$ and $\mathbf{L}_{n+1}$ to reflect the new molecular geometry and corresponding overlap matrix. Using $\mathbf{Z}_{n+1}$, $\mathbf{P}_{n+1}$ is then transformed into the new AO basis for use in the subsequent Fock build, following which the new nuclear and electronic forces are computed. (Derivation of the nuclear gradient is

the topic of Sec. II D.) These forces correspond to $a(t+\delta t)$ in the velocity Verlet update of Eq. (6b), which is finally completed, e.g.,

$$\dot{\boldsymbol{\Delta}}_{n+1}=\dot{\boldsymbol{\Delta}}_n-(\delta t/2\mu)\mathbf{G}_{n+1}. \qquad (15)$$

Note that $\boldsymbol{\Delta}$ and $\dot{\boldsymbol{\Delta}}$ are never transformed with $\mathbf{Z}$ or $\mathbf{L}$; rather, these quantities propagate in the orthogonal basis.

Importantly, the electronic gradient $\mathbf{G}$ is free of occupied-occupied and virtual-virtual contributions. Introducing $\mathbf{Q}=\mathbf{I}-\mathbf{P}$, the projector onto the virtual space, any matrix $\mathbf{M}$ may be expressed identically as

$$\mathbf{M}\equiv\mathbf{PMP}+\mathbf{QMQ}+\mathbf{QMP}+\mathbf{PMQ}. \qquad (16)$$

The terms on the right define, respectively, the OO, VV, VO, and OV components of $\mathbf{M}$ (O=occupied, V=virtual). Decomposing $\mathbf{G}$ in this manner, one finds that

$$\mathbf{G}=\mathbf{QGP}+\mathbf{PGQ}, \qquad (17)$$

which demonstrates that the quantities $\mathbf{P\Delta P}$, $\mathbf{Q\Delta Q}$, $\mathbf{P\dot{\Delta}P}$, and $\mathbf{Q\dot{\Delta}Q}$ are constants of the motion.

Equation (17) reflects the fact that only OV and VO rotations alter the **SCF** energy, and the SCF problem is converged when $\mathbf{QFP}=\mathbf{0}$. That $\mathbf{P\dot{\Delta}P}$ and $\mathbf{Q\dot{\Delta}Q}$ are constant implies that the fictitious kinetic energy T_F, just like E_{SCF}, is invariant to OO and VV rotations. However, standard mass-preconditioning schemes for ELMD destroy this invariance in T_F. Specifically, in order to achieve longer time steps, it is standard practice to give larger fictitious masses to either the high-energy plane waves (in CPMD[8,36]) or the high-energy density matrix elements (in ADMP[14]). In the ADMP version of this technique,[14] for example, the scalar μ is replaced by a diagonal matrix $\boldsymbol{\mu}$, and the fictitious kinetic energy is expressed as $T_F=\text{Tr}[(\boldsymbol{\mu}^{1/2}\dot{\mathbf{P}})^2]/2$, which is equivalent to associating a mass $(\mu_i\mu_j)^{1/2}$ with the density matrix element P_{ij}. A mass-preconditioned version of curvy-steps ELMD could be formulated similarly, with $(\mu_i\mu_j)^{1/2}$ the mass for coordinate Δ_{ij}.

An unfortunate consequence of all these mass-preconditioning methods is the introduction of nonvanishing OO and VV contributions to the electronic gradient. This is readily verified by replacing μ with $(\mu_i\mu_j)^{1/2}$ in electronic equation of motion, Eq. (5b), whereupon $\mathbf{P\Delta P}$ and $\mathbf{Q\Delta Q}$ vanish only in the case that $\mu_i=\mu_j$ for each i and j. Nonvanishing OO and VV accelerations ultimately manifest as OO and VV rotations of the density matrix. In CPMD[8,36] and ADMP,[14] it is evidently the case that increased control over fluctuations in T_F, which facilitates a larger time step, compensates for the inherent decrease in efficiency arising from the introduction of electronic coordinate variations that do not alter the energy. In curvy-steps ELMD the electronic coordinate steps are optimally small, and it remains to be tested whether mass-preconditioning is beneficial. In the interest of minimizing the number of adjustable parameters, however, we shall at present employ only a scalar fictitious mass. We achieve time steps comparable to ADMP sans preconditioning, which are somewhat larger than nonpreconditioned CPMD time steps.[13]

C. Geometric aspects of curvy-steps ELMD

As a postscript to the procedure outlined in the preceding section, we wish to discuss certain geometric aspects of this algorithm that we consider important, even though in its present implementation they have no practical bearing on curvy-steps ELMD. As pointed out in the seminal work of Edelman, Arias, and Smith[12] (recapitulated in our own work on curvy steps[10,11,24]), the Grassmann manifold that characterizes the SCF problem is not flat, and consequently vectors that are tangent to this manifold at a certain point (i.e., for a certain density matrix or set of MOs) generally do not lie in the tangent space at other points on the manifold. When tangent vectors associated with a previous time step are required, these vectors must be *parallel transported*[12,37] along the geodesic defined by the curvy steps variables, in order to bring them into the tangent space at the current time step.

Our focus on the matrix $\boldsymbol{\Delta}$ of curvy-steps variables potentially obfuscates the fact that points on the Grassmann manifold are represented by density matrices $\mathbf{P}$; antisymmetric matrices $\boldsymbol{\Delta}$ are not points on the manifold, but rather for a given point $\mathbf{P}(0)$, these matrices parametrize directions of transit (i.e., geodesics) originating from that point. $\mathbf{P}(\lambda)$ in Eq. (1) represents the geodesic whose initial tangent vector (at $\lambda=0$) is

$$\dot{\mathbf{P}}(0)=\dot{\boldsymbol{\Delta}}\mathbf{P}(0)-\mathbf{P}(0)\dot{\boldsymbol{\Delta}}. \qquad (18)$$

This expression can be used to compute the instantaneous rate of change in the density matrix elements. In fact, Eq. (18) represents the general form of a tangent vector, in a projection-operator representation of the Grassmann manifold,[38] and can be derived from geometric considerations alone. Alternatively, this expression follows from the chain rule:

$$\dot{\mathbf{P}}=\sum_{i<j}\frac{\partial\mathbf{P}}{\partial\Delta_{ij}}\frac{d\Delta_{ij}}{dt}, \qquad (19)$$

with $\partial\mathbf{P}/\partial\boldsymbol{\Delta}$ obtained from Eq. (1).

The antisymmetric matrix $\dot{\boldsymbol{\Delta}}$ thus parametrizes a tangent vector whose origin is $\mathbf{P}(0)$. Parallel transport of tangent vectors is realized by transformation of their skew-symmetric direction matrices,[12] and it follows that $\dot{\boldsymbol{\Delta}}$ ought to be transformed, following each velocity Verlet update, according to the prescription

$$\dot{\boldsymbol{\Delta}}\rightarrow e^{\boldsymbol{\Delta}}\dot{\boldsymbol{\Delta}}, \qquad (20)$$

which engenders parallel transport of $\dot{\mathbf{P}}$ along the geodesic defined by $\boldsymbol{\Delta}$. The update of $\dot{\boldsymbol{\Delta}}$ in Eq. (20) is fundamentally different from the exponential transformation of the density matrix in Eq. (1), but this is no cause for alarm, given the different rôles occupied by these two quantities.

As a practical matter, because the electronic velocities are kept small (else ELMD breaks down, for reasons unrelated to differential geometry), we find in practice that a series expansion of Eq. (20) consistently converges at $O(\boldsymbol{\Delta}^2)$, even with highly conservative thresholds, and furthermore parallel transport of $\dot{\boldsymbol{\Delta}}$ has an undetectable effect on the dynamics. This may be attributable to the velocity Verlet

updating scheme that—when implemented in the two-stage version described in Sec. II B—never combines tangent vectors that differ in time by more than $\delta t/2$. Specifically, the update $\dot{\boldsymbol{\Delta}}(t) \rightarrow \dot{\boldsymbol{\Delta}}(t+\delta t/2)$, Eq. (13), utilizes the gradient $\mathbf{G}(t)$ and is therefore geometrically correct, whereas the second stage of the velocity update, $\dot{\boldsymbol{\Delta}}(t+\delta t/2) \rightarrow \dot{\boldsymbol{\Delta}}(t)$ [Eq. (15)] uses the gradient $\mathbf{G}(t+\delta t)$, which is not quite geometrically consistent because the update occurs at time $t+\delta t/2$, not $t+\delta t$. We plan to explore this issue further in future work.

Formally speaking, our practice of resetting $\boldsymbol{\Delta}$ to zero following each time step may lead to an algorithm that is not strictly time reversible—indeed, it cannot be rigorously reversible so long as $\dot{\boldsymbol{\Delta}}$ is not properly transported—but numerical experiments indicate that this is not important. Specifically, for the systems discussed in Sec. III, we have performed experiments in which the signs of all nuclear and electronic velocities are reversed at the end of a multiple-picosecond trajectory. In each case, and to very high precision, the algorithm successfully propagates back to the original initial conditions.

D. Nuclear gradients

Evaluation of the nuclear forces requires derivatives $\partial E_{\mathrm{SCF}}/\partial R_I$ in the case that the SCF calculation is not converged, whereas the usual derivation of the SCF gradient[28] makes explicit use of the convergence condition, $\mathbf{FP}=\mathbf{PF}$. Our presentation of the more general formulation shall be terse, as the gory details can be found, for the ADMP method, in the work of Schlegel et al.[13] Apart from the absence of a constraint force in our formulation, the derivation is the same.

Briefly,

$$\frac{\partial E_{\mathrm{SCF}}}{\partial R_I}\bigg|_{\boldsymbol{\Delta}} = \frac{\partial E_{\mathrm{SCF}}}{\partial R_I}\bigg|_{\mathbf{P}} = g_I^{\mathrm{HF}} + g_I^{P} \tag{21}$$

consists of an integral-derivative (Hellmann-Feynman) contribution g_I^{HF} and an overlap-derivative (Pulay[7]) term g_I^{P}. The former,

$$g_I^{\mathrm{HF}} = \mathrm{Tr}\left[\left(\frac{\partial \mathbf{h}_{\mathrm{AO}}}{\partial R_I} + \frac{1}{2}\frac{\partial \mathbf{H}_{\mathrm{AO}}}{\partial R_I}\bigg|_{\mathbf{P}_{\mathrm{AO}}}\right)\mathbf{P}_{\mathrm{AO}}\right], \tag{22}$$

has the same form irrespective of whether or not the density matrix is converged, which allows us to recycle much existing SCF gradient code. For nonconverged density matrices, the new twist is the form of g_I^{P}:

$$g_I^{P} = \mathrm{Tr}\left(\mathbf{F}_{\mathrm{AO}}\frac{\partial \mathbf{P}_{\mathrm{AO}}}{\partial R_I}\bigg|_{\mathbf{P}}\right) = 2\,\mathrm{Tr}(\mathbf{FPD}_I), \tag{23}$$

where

$$\mathbf{D}_I = \frac{\partial \mathbf{Z}}{\partial R_I}\mathbf{L} = -\mathbf{Z}\frac{\partial \mathbf{L}}{\partial R_I} \tag{24}$$

represents the contribution from the transformation to the Cholesky-orthogonalized basis. At SCF convergence, Eq.

(23) reduces to the more familiar expression $g_I^{P} = \mathrm{Tr}[\mathbf{W}(\partial \mathbf{S}_{\mathrm{AO}}/\partial R_I)]$, where $\mathbf{W}=\mathbf{P}_{\mathrm{AO}}\mathbf{F}_{\mathrm{AO}}\mathbf{P}_{\mathrm{AO}}$ is the energy-weighted density matrix.[28,29]

Following Schlegel et al.,[13] a simple expression for $\mathbf{D}_I$ is obtained by differentiating the identity $\mathbf{Z}^{\dagger}\mathbf{S}_{\mathrm{AO}}\mathbf{Z}=\mathbf{I}$ and exploiting the fact that $\mathbf{D}_I$ is lower triangular. The result is

$$(\mathbf{D}_I)_{ij} = \begin{cases} -[\mathbf{Z}(\partial \mathbf{S}_{\mathrm{AO}}/\partial R_I)\mathbf{Z}^{\dagger}]_{ij} & \text{if } i<j \\ -[\mathbf{Z}(\partial \mathbf{S}_{\mathrm{AO}}/\partial R_I)\mathbf{Z}^{\dagger}]_{ii}/2 & \text{if } i=j. \\ 0 & \text{if } i>j \end{cases} \tag{25}$$

Combining this expression with Eqs. (22) and (23) affords an expression for the gradient that requires only matrix multiplications, not diagonalization, and is therefore consistent with linear scaling. This is one advantage of Cholesky decomposition over, say, symmetric (Löwdin) orthogonalization.

A subtle consequence of the presence of overlap derivatives is that the nuclear derivatives $\partial E_{\mathrm{SCF}}/\partial R_I$ depend upon the choice of representation for the electronic variables, that is, they depend upon how the AOs are orthogonalized. Specifically, while the first equality in Eq. (23) provides a general definition of g_I^{P} for any orthogonal basis, the second equality is particular to the Cholesky orthogonalized basis employed here. In the Appendix, we provide a general proof that the nuclear forces are representation dependent whenever atom-centered basis functions are employed, and furthermore illustrate how this dependence disappears at SCF convergence, yielding BOMD nuclear forces that are representation independent. (Similar arguments can be found in Ref. 16.) From the expressions derived in the Appendix, it follows that the representation dependence of the nuclear forces will be minor, provided that the electronic gradient $\mathbf{G}$ is small and oscillatory (always a necessary condition if ELMD is to reproduce long-time BOMD averages), and furthermore provided that the transformation matrix from the AO to the orthogonalized AO basis is differentiable. That $\mathbf{Z}(t)$ and $\mathbf{L}(t)$ are smooth functions of time has been verified, empirically, in our calculations. In addition, ADMP calculations using Cholesky-versus Löwdin-orthogonalized AOs suggests that the differences are small and do not alter pertinent nuclear observables.[13,16]

III. NUMERICAL RESULTS

Curvy-steps ELMD has been implemented in a developmental version of the Q-CHEM electronic structure software package;[39] we have also implemented BOMD in order to make comparison. Our goal in this initial investigation is to characterize acceptable values of μ and commensurate time steps, which will depend upon the fastest nuclear motion of the system in question, and we have chosen to investigate a sequence of diatomic molecules as an easy means to generate a variety of nuclear time scales. Our molecular test set consists of NaCl, F_2, HCl, DCl, HF, and DF, whose vibrational frequencies span the range from 300 to 4000 cm^{-1}. Actually, much of our analysis will focus on HF, as this is the fastest vibration and therefore the most difficult case, and because

J. Chem. Phys., Vol. 121, No. 23, 15 December 2004

Lagrangian molecular dynamics   11547

TABLE I. Parameters for curvy-steps ELMD simulations, in atomic units: 1 a.u. (fictitious mass) ≈ 1820 amu bohr^2; 1 a.u. (time) ≈ 0.0242 fs.

Simulation	μ (a.u.)	δt	
		a.u.	fs
i	45	5	0.12
ii	90	7	0.17
iii	180	10	0.24
iv	360	14	0.34

high-frequency H-atom stretching vibrations are relevant to the extensive body of ELMD simulations of liquid water and water clusters.[4–6]

A. Simulation parameters

All of our simulations utilize the B3LYP density functional.[49] (Note that this and other functionals containing Hartree-Fock exchange greatly increase the cost of plane-wave CPMD,[4,40] but no such penalty is incurred in a Gaussian basis set.) Some all-electron results will be described below, but most of our calculations utilize the Stevens–Basch–Krauss–Josien–Cundari (SBKJC) pseudopotential[41] and corresponding valence basis set, consisting of $(3s) \rightarrow [2s]$ contracted basis functions for hydrogen and $(4sp) \rightarrow [2sp]$ functions for second-row elements. All simulations were initiated from a converged SCF density matrix at the minimum-energy bond length, with $\dot{\Delta} = 0$. (This last choice is convenient but may somewhat exaggerate energy fluctuations.[2]) Initial nuclear velocities correspond to harmonic zero-point energy in the vibrational mode, as determined by an SCF frequency calculation.

Four ELMD simulations were performed for each molecule, using the time step and fictitious mass parameters listed in Table I; we shall refer to these simulations by the Roman numerals given in that table. In the interest of obtaining well-resolved vibrational spectra, each simulation was propagated for 20 000 time steps, corresponding to 2.4–6.8 ps of simulated time. Since the maximum possible time step (for a given level of energy conservation, say) ought to scale as $\mu^{1/2}$ [cf. Eq. (11)], as should the lowest-energy electronic oscillation, the simulation parameters vary in such a way that the ratio $\delta t/\mu^{1/2}$ is approximately constant, up to our decision to use integer values (in atomic units) for both parameters. We have not bothered to tune μ by less than factors of 2, but modulo this coarse-grained tuning, the value of δt for each μ is as large as it can be made without inducing unacceptably large fluctuations ($\gg 10^{-3}E_h$) in the energy. Notably, the largest permissable time step is the same for all of the systems studied here, and thus represents the limit of velocity Verlet integration of the fast electronic motion.

Table II provides a summary of fluxional quantities obtained from each ELMD simulation. The quantity $\mathcal{H}$ introduced in that table is the value of the classical Hamiltonian corresponding to the extended Lagrangian $\mathcal{L}$, and fluctuations in $\mathcal{H}$ characterize finite time step integration error. In contrast to $\mathcal{H}$, the "real" energy

$$E(t) = \mathcal{H}(t) - T_F(t) \qquad (26)$$

is not a conserved quantity, but its fluctuations also characterize the quality of the extended-Lagrangian dynamics. Also listed in Table II is the maximum value of the time derivative of T_F in each simulation, another fluxional quantity that has been suggested[14,16] as a measure of the extent of adiabatic decoupling of nuclei and electrons. For curvy-steps ELMD, this derivative may be calculated according to

$$\left. \frac{dT_F}{dt} \right|_{\Delta=0} = \sum_{i<j} \frac{d\dot{\Delta}_{ij}}{dt} \left. \frac{\partial T_F}{\partial \dot{\Delta}_{ij}} \right|_{\Delta=0} = \frac{1}{2}\text{Tr}(\dot{\Delta}\mathbf{G}). \qquad (27)$$

However, our experience suggests that this is not a useful metric by which to characterize ELMD simulations. This will be discussed in Sec. III B.

Perusing Table II, one finds that fluctuations in both $\mathcal{H}$ and E are maintained at acceptable, submillihartree levels in all cases except simulation (iii) for HF and simulation (iv) for HCl, DF, and HF. At all times and in all simulations, T_F is maintained at less than $8 \times 10^{-4}E_h$, and because our simulations are quite long, this effectively demonstrates that T_F is free of systematic drift. (Additional evidence that the electrons do not "heat up" is presented in the following section, where we discuss the behavior of dT_F/dt.) Note that T_F, rather than μ, is the quantity to compare across various ELMD methods, because the fictitious "mass"—really a moment of inertia—depends upon the choice of electronic coordinates.

While it is not our goal in this report to benchmark the efficiency of curvy-steps ELMD, at the risk of digressing we cannot help but call attention to the tabulation, in Table II, of the average order $\bar{n}_{\text{BCH}}$ necessary to converge the BCH expansion to an exceedingly conservative threshold of 10^{-14} in every commutator matrix element. (In comparison, the drop tolerance for AO integrals is $10^{-11}E_h$.) In all cases $\bar{n}_{\text{BCH}} \leq 7$, which means that seven or fewer matrix multiplications are typically sufficient to update the density matrix. In a multiple time-scale implementation of ELMD, such as the one described in Ref. 42, these multiplications would constitute essentially the entire cost of each short electronic time step, while comparatively expensive Fock builds and integral derivative calculations would be required only at the conclusion of each longer, nuclear time step.

B. Adiabaticity

We next wish to examine the extent to which curvy-steps ELMD is a faithful representation of classical dynamics on the Born–Oppenheimer surface. First, however, we must demonstrate that our BOMD simulations indeed represent these dynamics, as such simulations are known to suffer a systematic drift in the energy (and presumably other properties) unless the SCF calculation is converged rather tightly.[2,4,43] In our implementation of BOMD, the SCF convergence threshold is set to $10^{-8}E_h$ in the OV elements of the Fock matrix, which is Q-CHEM's default threshold whenever energy gradients are required. In Table III we list the fluctuations in $E(t) - E(0)$ and also $|E(t) - E(0)|/|E(0)|$, for long BOMD simulations of each test molecule. The fluctuations are extremely small, reflecting our conservative choice of time step. More importantly, the mean of the

TABLE II. Summary of fluxional quantities (in atomic units), averaged over 20 000 time steps, for curvy-steps ELMD simulations on the B3LYP/SBKJC potential surface. Consult the text for definitions.

Simulation		$\bar{n}_{BCH}$	Energy fluctuations/10^{-4} E_h				T_F max	$\|\mathbf{G}\|$ max /10^{-4}	dT_F/dt max /10^{-6}
			$\mathcal{H}(t)-\mathcal{H}(0)$		$E(t)-E(0)$				
			rms	max	rms	max			
NaCl	(i)	4.0	0.001	0.004	0.001	0.005	0.004	0.002	0.012
	(ii)	4.2	0.002	0.009	0.003	0.012	0.008	0.005	0.017
	(iii)	4.5	0.003	0.017	0.005	0.023	0.017	0.009	0.024
	(iv)	4.8	0.006	0.034	0.010	0.045	0.033	0.019	0.037
F_2	(i)	5.0	0.026	0.113	0.027	0.114	0.035	0.129	0.251
	(ii)	5.0	0.048	0.223	0.049	0.227	0.072	0.267	0.363
	(iii)	5.1	0.102	0.458	0.104	0.476	0.141	0.549	0.502
	(iv)	5.4	0.201	0.941	0.205	0.874	0.296	1.111	0.723
DCl	(i)	5.2	0.110	0.564	0.132	0.683	0.149	0.221	0.729
	(ii)	5.5	0.228	1.086	0.273	1.376	0.312	0.447	1.109
	(iii)	5.7	0.465	1.507	0.552	2.552	0.636	1.010	1.531
	(iv)	6.0	1.278	5.297	1.407	6.422	1.308	2.404	2.268
HCl	(i)	5.4	0.199	0.968	0.247	1.196	0.319	0.467	1.557
	(ii)	5.8	0.488	2.181	0.570	2.595	0.595	1.581	2.468
	(iii)	6.0	1.348	5.824	1.462	6.659	1.254	2.538	3.164
	(iv)	6.4	2.029	9.588	2.388	11.638	2.662	5.686	4.893
DF	(i)	5.4	0.396	1.851	0.448	2.081	0.263	0.500	1.462
	(ii)	5.7	0.742	3.411	0.827	3.829	0.507	1.622	2.142
	(iii)	6.0	1.874	8.416	2.001	9.173	1.123	3.418	3.268
	(iv)	6.3	3.942	16.969	4.337	19.076	2.275	6.650	4.498
HF	(i)	5.6	0.768	3.299	0.867	3.756	0.498	1.171	2.678
	(ii)	6.0	1.700	8.342	1.874	9.304	1.114	2.672	4.041
	(iii)	6.3	3.688	16.496	4.071	18.572	2.284	7.174	6.368
	(iv)	7.0	9.271	45.702	10.055	66.191	6.619	40.820	18.401

signed energy fluctuations $E(t)-E(0)$, which places a bound on any systematic drift in the BOMD energy, is less than $4\times10^{-5}E_h$ for each simulation. This indicates that our BOMD simulations are free of drift over the simulation lengths used in this work. At the risk of belaboring the point, however, we show in Fig. 1 the normalized energy fluctuations $|E(t)-E(0)|/|E(0)|$ for the initial and final 2000 steps of a 20 000-step BOMD simulation of HF, with $\delta t = 10$ a.u. By the end of the simulation the fluctuations have grown slightly larger, but there is no evidence of drift. (Some systematic drift is observed if the convergence criterion is relaxed to $10^{-5}E_h$, Q-CHEM's default for single-point energy calculations.)

Having demonstrated that our BOMD forces are fully converged, let us compare them to the ELMD ones. Actually we find it more convenient to deal with the acceleration $\ddot{R}$ along the bond axis, and in Fig. 2 we plot the ELMD and BOMD values of $\ddot{R}(t)$ for F_2 simulation (iv) and HF simulation (i). Consistent with the conclusions of previous studies,[2,19,44] the ELMD force oscillates rapidly around the BOMD force, at least for appropriate values of μ. Ten vibrational periods are plotted for each of the examples in Fig. 2, and on this time scale the ELMD force, averaged over the fast oscillations, is faithful to the BOMD force, although for HF a small but systematic drift has begun to appear by the

TABLE III. Fluxional parameters from BOMD simulations with $\delta t = 10$ a.u.; numbers in parenthesis denote powers of ten.

	$|E(t)-E(0)|/|E(0)|$		$(E(t)-E(0))/E_h$	
	rms	Max	Mean	rms
NaCl[a]	8.4(−10)	3.2(−9)	5.1(−9)	2.8(−8)
F_2[b]	8.6(−9)	2.6(−8)	5.8(−7)	4.2(−7)
DCl[a]	2.6(−7)	8.5(−7)	5.7(−6)	4.1(−6)
HCl[a]	5.2(−7)	1.7(−6)	1.1(−5)	8.3(−5)
DF[a]	5.0(−7)	1.7(−6)	1.8(−6)	1.3(−5)
HF[b]	1.0(−6)	3.4(−6)	3.6(−5)	2.6(−5)

[a]5 000 time step simulation.
[b]20 000 time step simulation.

J. Chem. Phys., Vol. 121, No. 23, 15 December 2004

Lagrangian molecular dynamics   11549
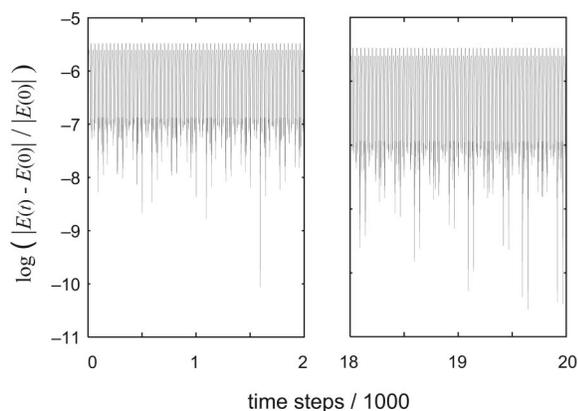
FIG. 1. Normalized energy fluctuations for a BOMD simulation of HF with $\delta t = 10$ a.u.

tenth vibrational period. The origin of this feature is discussed below.

It is quite intentional that in Fig. 2 we have compared simulation (i) for HF—which employs the smallest value of μ studied here—to simulation (iv) for F_2, which uses our largest value of μ. This is a first example of a general principle that will figure prominently in our discussion of ELMD: the accuracy of the simulation (iv) results, for molecules not containing hydrogen, is quite acceptable over many vibrational periods, and is comparable to or better than the accuracy of the simulation (i) results for hydrogen-containing molecules.

The first indication of possible difficulties with

FIG. 3. Acceleration $\ddot{R}$ along the HF bond, from ELMD simulations (ii)–(iv). The broken lines are BOMD results.

hydrogen-containing species comes from examining $\ddot{R}$ for HF as a function of μ, Fig. 3. Like simulation (i), the average ELMD force in simulation (ii) oscillates tightly around the BOMD force for at least ten vibrational periods, albeit with a slightly faster drift than was observed in (i). By (iii), however, the fluctuations about the BOMD force are noticeably larger in amplitude and the drift is unmistakably faster. Finally, for simulation (iv) the ELMD force is no longer a reasonable facsimile of the BOMD force beyond two or three vibrational periods.

The slow drift in the ELMD nuclear forces away from the BOMD ones, evident in all of our simulations but most severe for hydrogen-containing molecules, arises because these forces cannot respond instantaneously when the BOMD forces turn over. These lag times add coherently, so that the ELMD forces grow farther out of phase each time a classical turning point is reached. Increasing μ exacerbates this problem in two ways: first, by slowing down the electronic oscillations, so that the electrons respond less quickly when the nuclei reverse direction; and second, by increasing the amplitude of the fluctuations in the ELMD nuclear force about the BOMD force, so that they must affect an even larger change when the BOMD force turns over. This latter effect is evident from the increasingly jagged features in the

FIG. 2. Acceleration $\ddot{R}$ along the diatom bond for (a) F_2, from simulation (iv); and (b) HF, from simulation (i). Circles represent the BOMD result.

ELMD force at large μ (Fig. 3). These asperous features can be mollified by coupling thermostats to the electronic velocities, a technique that has been shown to correct nuclear velocity autocorrelation functions when large values of μ are employed.[30] In our experience, however, the application of such techniques to hydrogen-containing molecules is nontrivial, and is beyond the scope of this report.

The drift in the ELMD nuclear force must eventually manifest as a drift in the ELMD trajectory away from the BOMD result. This is illustrated by the phase-space trajectories for HF depicted in Fig. 4. Note especially the distinctly nonsinusoidal nature of the nuclear velocity in simulation (iv). Compared to $R(t)$, the velocity $\dot{R}(t)$ offers a better indication of the failure of this simulation. In Sec. III C we will demonstrate that the phase lag in the ELMD forces and trajectories is manifested as a redshift in the nuclear vibrational frequency.

The ELMD forces for HF simulation (iv) might be sufficiently different from sinusoidal as to sound klaxons of alarm in one's mind, but the results of simulation (iii), say, are more ambiguous, and the situation is apt to become even more convoluted in polyatomic molecules, where the nuclear forces do not vary sinusoidally at all. Insofar as the *raison d'être* of ELMD is to circumvent the high cost of BOMD, it is imperative to possess criteria, not requiring an explicit BOMD calculation, which indicate whether the ELMD forces are likely to be oscillating about the BOMD ones. The quantity dT_F/dt, introduced in Eq. (27), has been proffered as one such "adiabaticity index" for ELMD simulations:[14,16] based upon the observation that dT_F/dt oscillates rapidly about zero without drift, Iyengar *et al.*[16] concluded that certain ADMP simulations were indeed faithful reproductions of BOMD nuclear dynamics.

With this in mind, in Fig. 5 we plot dT_F/dt for the first 100 fs of HF simulations (i) and (iv). Fluctuations in this quantity are qualitatively similar over the remaining several picoseconds of these simulations, and the mean value of dT_F/dt, averaged over all 20 000 time steps, is 2.8×10^{-11} a.u. for simulation (i) and -8.9×10^{-10} a.u. for simulation (iv)—that is to say, zero, on the scale that dT_F/dt fluctuates. Certainly, these fluctuations grow larger in amplitude as μ increases, but as an adiabaticity index dT_F/dt falls flat, utterly failing to detect the impending disaster in the ELMD forces for $\mu = 360$ a.u. That dT_F/dt oscillates about zero appears to be a weak criterion that is necessary but not sufficient to obtain reliable ELMD results.

In CPMD, the validity of the adiabatic decoupling hypothesis is sometimes interrogated by computing the spectrum of fictitious electronic oscillations,[44,45] which can be compared to the (nuclear) vibrational spectrum to determine whether, in fact, a separation of energy scales exists. Both spectra are computed in the same way, as the Fourier transform of a velocity autocorrelation function. For curvy-steps ELMD, the electronic velocity autocorrelation function is defined as

$$C_{\text{el}}(t) = \sum_{i<j} \langle \dot{\Delta}_{ij}(t) \Delta_{ij}(0) \rangle, \qquad (28)$$

in which the average $\langle \cdots \rangle$ runs over time origins but not over
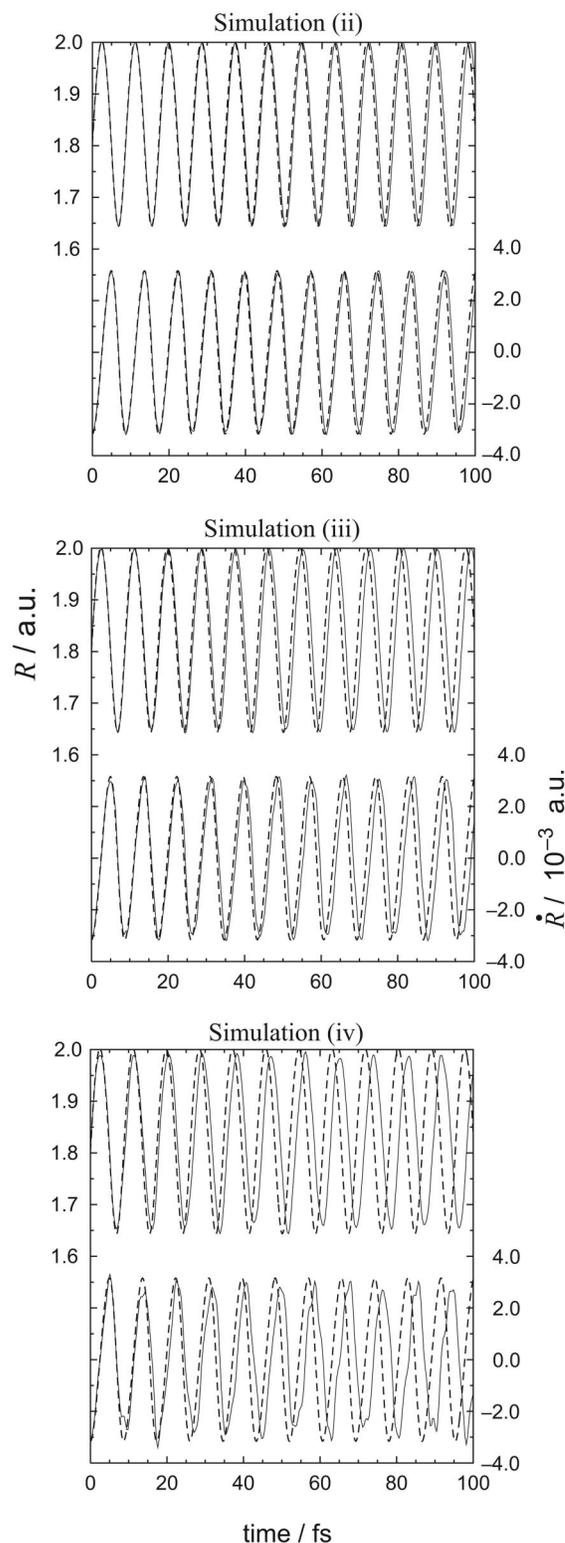
FIG. 4. Comparison of ELMD trajectories (solid curves) and BOMD trajectories (broken curves) for HF, starting from the same initial conditions. In each panel, the upper curves depict the bond length $R(t)$, and the lower curves illustrate the nuclear velocity along this coordinate $\dot{R}(t)$.

any molecular ensemble, as we consider only isolated molecules. Otherwise, $C_{\text{el}}(t)$ is computed just like any other classical time correlation function,[32] and its cosine transform [since $C_{\text{el}}(t) = C_{\text{el}}(-t)$] affords the spectrum of bound elec-
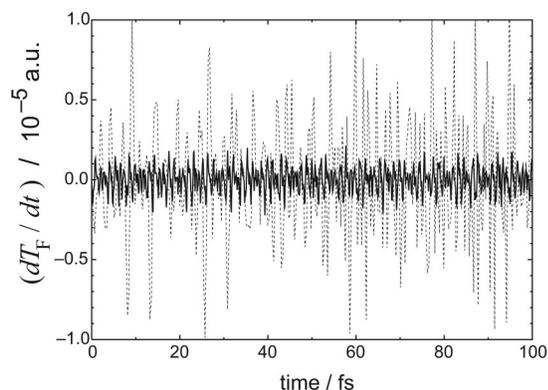
FIG. 5. Plots of dT_F/dt for HF simulations (i) (darker, solid line) and (iv) (lighter, broken line).

tronic vibrations. These we term fictons, as they do not describe real physics, and we reserve the word *vibrations* to mean nuclear oscillations.

Unfortunately, the use of ambulating (atom-centered) basis functions complicates the ficton spectrum considerably, because the motion of the basis functions necessarily introduces correlations into $C_{el}(t)$ that decay on a nuclear time scale. This is illustrated in Fig. 6, which depicts sections of both the nuclear and electronic velocity autocorrelation functions for HF. The latter exhibits an unmistakable, fast electronic motion superimposed upon the slow oscillation of the basis functions. An approximate, *a priori* estimate of the fast oscillation time scale is afforded by the simple formula[4,44]

$$\tau_{el} \lesssim 2\pi \sqrt{\mu/(\varepsilon_{LUMO} - \varepsilon_{HOMO})}. \tag{29}$$

At the B3LYP/SBKJC level, the highest occupied molecular orbital-lowest unoccupied molecular orbital (HOMO-LUMO) gap for HF is $0.442E_h$, whence $\tau_{el} \sim 1.5$ fs for $\mu = 45$ a.u. This coincides nicely with a cursory examination of the electronic correlation function in Fig. 6: the time between successive local minima in $C_{el}(t)$ is roughly 2 fs.

Whenever atom-centered basis sets are used, the ficton spectrum will contain Fourier components at all of the fundamental vibrational frequencies and, resolution permitting, at overtones and combinations of these frequencies as well. This makes for a rather complicated ficton spectrum, as vibrational overtones may lie among the ficton lines, even if
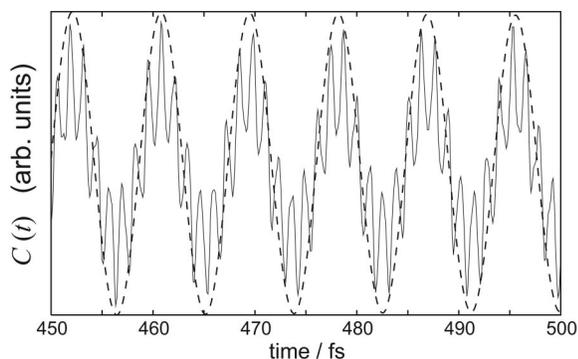
FIG. 6. Section of the electronic (solid line) and nuclear (broken line) velocity autocorrelation functions for HF, from simulation (i).
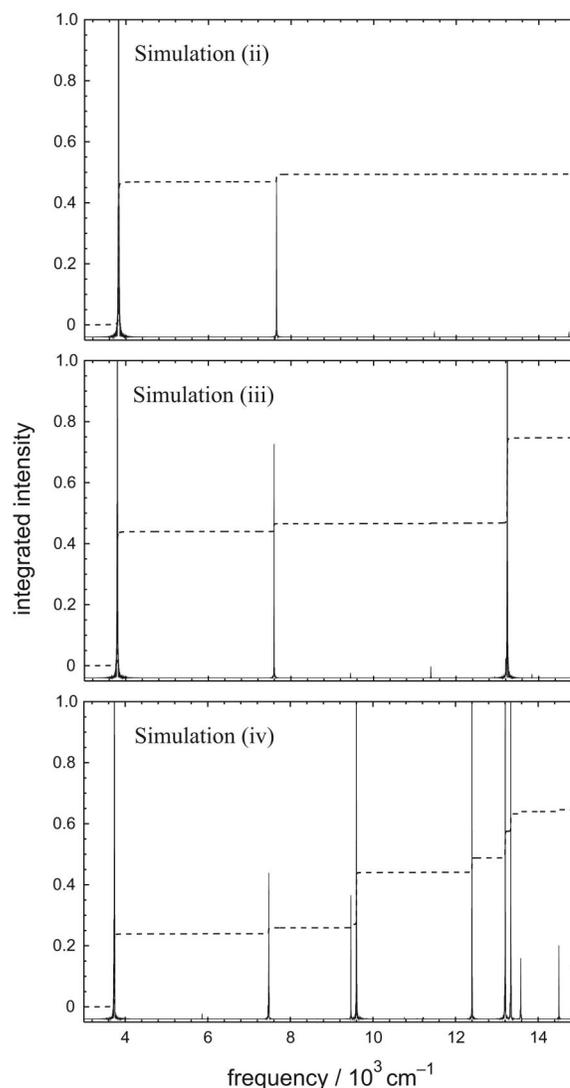
FIG. 7. Low-energy ficton (electronic oscillation) spectra for HF; the total integrated intensities are shown as a broken line. The most intense peaks have been truncated in order to accentuate the weaker lines in the spectrum.

the fundamental is well separated in energy, and our decision to focus on diatomic molecules in this initial investigation is partly intended to simplify the spectra as much as possible. This being said, it merits noting that in polyatomic molecules—and especially in short simulations—there is little chance that vibrational overtones or the corresponding ficton lines will be resolved.

The low-energy portions of the ficton spectra for HF simulations (ii)–(iv) are shown in Fig. 7. (The full spectra exhibit peaks up to about 75 000 cm^{-1}.) The lowest-energy and most intense peak in each spectrum corresponds exactly—in both line position and width—to the vibrational fundamental, while the next peak in each spectrum (~ 7500 cm^{-1}) represents the first vibrational overtone. The second overtone, at 11 510 cm^{-1}, is just barely visible in the spectrum from simulation (iii), though it is resolved in all three spectra and would be visible upon sufficient magnification. Despite these complicating features, there is still a clear separation between these vibrationally induced ficton lines and the much higher-energy features arising from re-

TABLE IV. Spectral data from ELMD simulations: ω is the fundamental vibrational frequency and $\delta\omega$ $=\omega_{BOMD}-\omega$ is its (red) shift relative to the BOMD result. τ_{el} and τ_{nuc} are the electronic and nuclear time scales, as defined in the text, with the former either observed from the spectrum or calculated from Eq. (29). Here "gap" refers to the adiabatic gap defined in the text.

Simulation		Frequency/cm^{-1}			T_F/cm^{-1}		τ_{el}/fs		
		ω	$\delta\omega$	Gap	rms	Max	Obs.	Calc.	τ_{nuc}/τ_{el}
NaCl	(i)	317	0	14498	0.02	0.09	2.3	2.6	45.7
	(ii)	317	0	10158	0.04	0.18	3.2	3.7	32.9
	(iii)	317	0	7103	0.07	0.36	4.5	5.3	23.4
	(iv)	317	0	4920	0.14	0.73	6.4	7.4	16.4
F$_2$	(i)	989	0	26231	0.12	0.78	1.2	2.1	28.1
	(ii)	989	0	23596	0.25	1.57	1.4	2.9	24.1
	(iii)	989	0	16411	0.50	2.90	1.9	4.1	17.7
	(iv)	988	1	11325	1.03	6.22	2.7	5.9	12.5
DCl	(i)	1902	0	22313	0.63	3.26	1.4	1.8	12.5
	(ii)	1900	2	16580	1.29	6.86	1.8	2.5	9.8
	(iii)	1896	6	11190	2.58	13.96	2.5	3.5	7.0
	(iv)	1888	14	7400	5.03	28.70	3.6	4.9	4.9
HCl	(i)	2648	3	20827	1.30	6.99	1.4	1.8	9.0
	(ii)	2642	9	15858	2.65	13.05	1.8	2.5	7.0
	(iii)	2633	18	10372	4.92	27.52	2.6	3.5	4.9
	(iv)	2614	37	6726	10.71	58.42	3.6	4.9	3.6
DF	(i)	2788	7	20502	1.28	5.78	1.4	1.5	8.6
	(ii)	2783	12	15692	2.26	11.16	1.8	2.2	6.7
	(iii)	2773	22	10358	4.87	24.65	2.5	3.1	4.8
	(iv)	2751	44	6671	10.79	49.93	3.5	4.3	3.5
HF	(i)	3836	17	22294	2.47	10.93	1.3	1.5	6.7
	(ii)	3824	29	14726	5.20	24.06	1.8	2.2	4.8
	(iii)	3796	57	9444	10.42	50.12	2.5	3.1	3.5
	(iv)	3738	115	5727	25.12	145.27	3.5	4.3	2.5

coupling of the orthogonalized AOs due to changes in the electronic structure. Peaks of the latter variety are absent in the portion of the spectrum shown for simulation (ii), but these peaks shift into view as the fictitious mass is increased.

In Table IV we compile a list of fundamental vibrational frequencies and adiabatic gaps, as calculated from our ELMD ficton spectra. The *adiabatic gap* is defined here as the separation between the vibrational fundamental and the lowest ficton frequency not coincident with any line in the vibrational spectrum. Also listed for each simulation is the electronic time scale τ_{el} corresponding to the lowest ficton frequency ω_{el}. As expected, the measured values of τ_{el} scale as $\mu^{1/2}$, and moreover the estimate of τ_{el} from Eq. (29) proves, in each instance, to be an upper bound to the observed electronic time scale. This is important, since Eq. (29) can be used to determine an appropriate value of μ using only a known or estimated value for the HOMO-LUMO gap.

The adiabatic gap is a highly seductive measure because it, too, can be calculated from the ELMD simulation alone, without the need for BOMD. Based upon CPMD results for liquid water, Grossman *et al.*[45] concluded that the onset of deviations from BOMD is a direct consequence of overlapping ficton and vibrational spectra. It is rather unnerving, therefore, to note that even for HF simulation (iv), in which the ELMD forces exhibit large deviations from BOMD, the adiabatic gap is still 5727 cm^{-1}, and furthermore the lowest ficton peak is $\sim$1000 cm^{-1} higher in energy than the first

vibrational overtone (see Fig. 7). A better harbinger of the problems with this simulation is the observation that the fundamental vibrational period τ_{nuc} exceeds τ_{el} only by a factor of 2.5, which evidently is not a sufficient separation of time scales. Of course, the corresponding frequencies ω_{nuc} and ω_{el} contain the same information as the time scales τ_{nuc} and τ_{el}, but this content is packaged differently. Whereas a frequency difference $\omega_{nuc}-\omega_{el}\sim5700$ cm^{-1}, in conjunction with the CPMD results of Grossman *et al.*,[45] might tempt one to conclude that the simulation is adiabatically decoupled, a time-scale ratio $\tau_{nuc}/\tau_{el}\sim2.5$ seems less convincing, as the vacillating ELMD force then has time for only 1.25 oscillations between nuclear turning points. This rapidly gives rise to the phase differences seen in Fig. 3.

C. μ dependence of vibrational frequencies

A recurring issue in the CPMD literature is the dependence of vibrational frequencies on the fictitious mass parameter. Such a dependence has been reported both in small molecules[20,21] and bulk materials,[19] in bulk MgO, over a range of μ otherwise thought to provide sufficient adiabatic decoupling. Having already pointed out in this work the slow drift of ELMD trajectories away from BOMD ones, it comes as no surprise that the computed frequencies differ, though it is important to quantify the extent of the discrepancy. The $\delta\omega$ values listed in Table IV are vibrational frequency differ-

J. Chem. Phys., Vol. 121, No. 23, 15 December 2004

Lagrangian molecular dynamics   11553

ences relative to the correct values obtained from BOMD. For NaCl and F_2, ELMD simulations reproduce the BOMD vibrational frequency across the entire range of μ employed here. The agreement for NaCl is especially significant in light of a previous ADMP study[15] that also found the NaCl vibrational frequency to be independent of μ over a wide range of the latter. These authors went on to claim that the use of atom-centered Gaussian orbitals eliminates the μ dependence of ELMD vibrational frequencies in general. Our results reveal this conclusion to be erroneous.

In fact, NaCl is the most well-behaved molecule studied here, according to every single metric that we have introduced, while vibrational frequencies of other molecules (especially those containing hydrogen atoms) exhibit a pronounced redshift as μ is increased. HF is the worst case, where even simulations (i) and (ii) afford shifts of 17 and 29 cm^{-1}, respectively. For this molecule we have clearly been too greedy in selecting a fictitious mass; however, our main point is to emphasize that this fact is not obvious, absent BOMD data, and is concealed by several of the usual measures for assessing ELMD simulation results. In particular, adiabatic gaps, real energy fluctuations, and dT_F/dt are each deficient metrics in this respect. Separation of time scales is a much more incisive measure: in all cases considered here, a time-scale separation $\tau_{nuc}/\tau_{el} \sim 10$ is sufficient to obtain a vibrational frequency that is within a few cm^{-1} of the BOMD frequency.

Figure 8 presents a graphical illustration of the shift in the HF vibrational frequency as a function of μ. In the lower panel of that figure, we plot the same spectra, shifted to the blue by an amount equal to the maximum value of T_F in each simulation. In the case of simulation (iv), where μ is much too large and the ELMD forces are rather erratic, this adjustment overcompensates for the vibrational frequency shift, but for the other three simulations this correction affords remarkable agreement with the BOMD vibrational spectrum. For polyatomic molecules, the maximum value of T_F will increase (all else being equal), since T_F is an extensive quantity, but we expect some correlation between the maximum T_F and the vibrational frequency shift per mode. However, the distribution is unlikely to be statistical, hence it is doubtful that a similar *a posteriori* correction is possible in polyatomic molecules. On the other hand, in CPMD simulations of isolated H_2O and CO_2 molecules,[20] all three vibrational frequencies were shown to shift linearly with μ, opening the possibility of extrapolating the vibrational spectrum to the $\mu = 0$ limit *in toto*. This remains a topic for future investigation.

In order to verify that these frequency shifts are not an artifact of the treatment of the electronic structure, we have performed all-electron B3LYP/6-31G** ELMD simulations for HF, using parameters ($\mu = 45$ a.u., $\delta t = 2.5$ a.u.) in one simulation and ($\mu = 180$ a.u., $\delta t = 5.0$ a.u.) in a second simulation. The corresponding BOMD simulation used $\delta t = 10$ a.u., and all three simulations were propagated for 5000 time steps. For the first of the ELMD simulations, the maximum value of T_F was 18.4 cm^{-1} and fluctuations in E and $\mathcal{H}$ were maintained below $5.0 \times 10^{-4} E_h$. For the second ELMD simulation, $T_F \lesssim 72.9$ cm^{-1} and fluctuations in E and
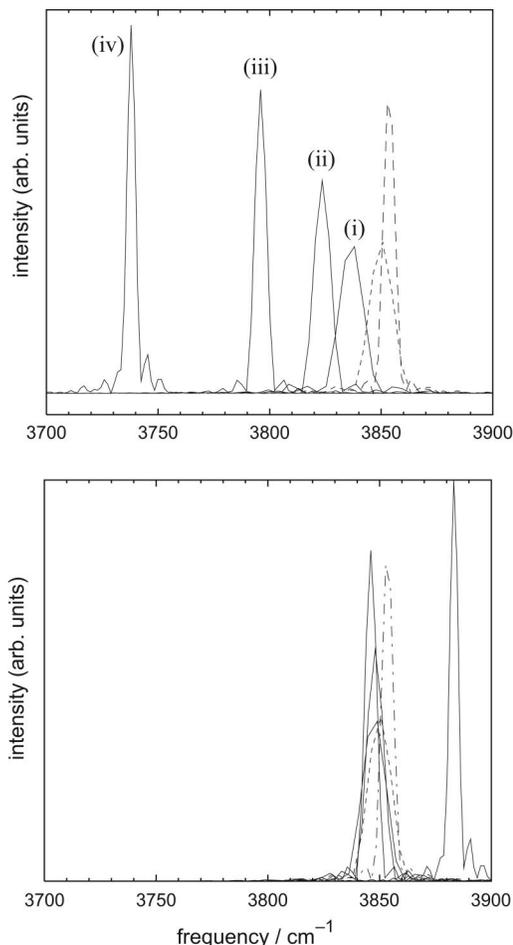
FIG. 8. Upper panel: calculated vibrational spectra for HF. Solid lines are the ELMD results and broken lines are BOMD spectra obtained from a 2.4 ps trajectory (broader spectrum) and a 4.8 ps trajectory (narrower spectrum). Lower panel: result of shifting the ELMD spectra by an amount equal to the maximum value of T_F for each simulation.

$\mathcal{H}$ were less than $1.8 \times 10^{-3} E_h$. As depicted in Fig. 9, the larger fictitious mass results in a vibrational frequency that, within the limited resolution afforded by this short-time spectrum, appears to be redshifted by precisely the maximum value of T_F. For the other all-electron ELMD simulation, the maximum fictitious kinetic energy is substantially smaller
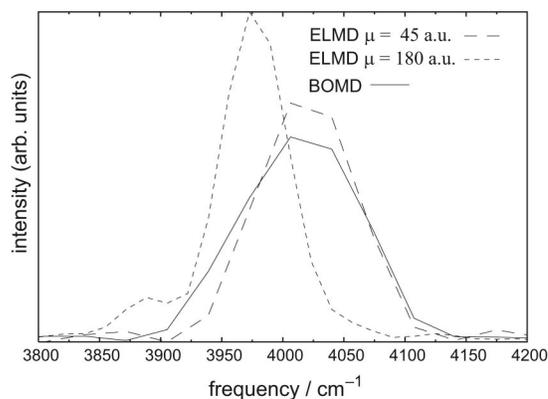
FIG. 9. Vibrational spectra for HF from all-electron B3LYP/6-31G** simulations.
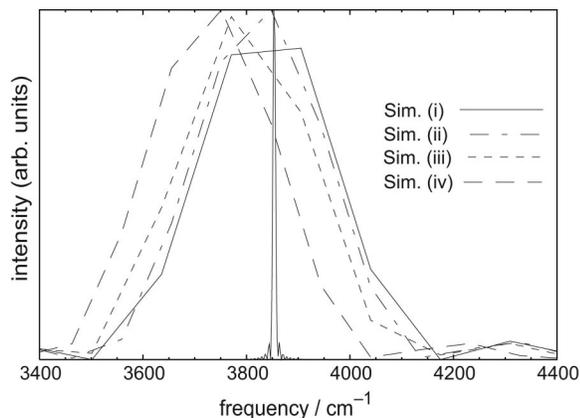
FIG. 10. Vibrational spectra for HF, obtained from 100 fs ELMD trajectories. The narrow spectrum is calculated from a 4.8 ps BOMD trajectory with the same initial conditions.

than the width of the spectrum, and the calculated vibrational frequency is in good agreement with BOMD.

These results demonstrate convincingly that the error in ELMD vibrational frequencies increases as the maximum value of T_F increases. The simplest way to suppress fluctuations in T_F is to reduce μ; for H-atom stretching motions above $3000\,\text{cm}^{-1}$ ($\tau_{\text{nuc}} < 12$ fs) and a target time-scale separation $\tau_{\text{nuc}}/\tau_{\text{el}} \gtrsim 10$, this requires at best our smallest value of μ. The fictitious mass, in conjunction with the HOMO-LUMO gap, sets the electronic time scale, and the finite time step integration error in turn determines the maximum permissable time step, which is limited in our B3LYP/SBKJC calculations to $\delta t \approx 0.12$ fs for $\mu = 45$ a.u. For comparison, in the absence of mass preconditioning, thermostats, or other techniques designed to suppress the most rapid electronic fluctuations, it appears that $\delta t \approx 0.07$ fs is appropriate for CPMD simulations of liquid water,[46] increasing to perhaps 0.1 fs when electronic thermostats are employed.[47]

On the other hand, for the HF simulations discussed here, we have demonstrated that ELMD phase-space trajectories are an accurate approximation to BOMD trajectories (with the same initial conditions) only for ~100 fs, while our ELMD simulations are propagated for 2.4–6.8 ps. These long-time results represent the true, converged vibrational frequency in the ELMD universe, though one might argue that for large systems, where propagation of picosecond trajectories is exorbitantly expensive, one will often be limited to simulating only the broad envelope of the vibrational spectrum, which can be obtained from a relatively short simulation. Thus we have computed B3LYP/SBKJC vibrational spectra for HF based on 100 fs of total propagation time, using the same μ and δt parameters as in the longer pseudopotential simulations (i)–(iv). These spectra are shown in Fig. 10, set against the narrow spectrum from a 4.8 ps BOMD trajectory. For the short versions of simulations (i) and (ii), the broad spectra are centered at the BOMD frequency, in contrast to the long-time results, even though the maximum and r.m.s. values of T_F are essentially the same as in the longer simulations. On the other hand, short versions of simulations (iii) and (iv) broaden the spectrum with respect to the long-time result but do not significantly alter the

position of the line. This is not surprising because for these simulations, 100 fs of propagation time is sufficient to observe substantial deviations from the BOMD trajectory (see Fig. 4).

IV. SUMMARY

We have described an approach to *ab initio* molecular dynamics—"curvy steps" ELMD—based on a transformation into generalized electronic coordinates in which the extended Lagrangian is free of constraints. Although this part of the formalism works equally well in a plane-wave basis, we have chosen to implement the method using atom-centered Gaussian orbitals. Propagation is highly efficient, requiring only a few matrix multiplications (in order to update the electronic density matrix) in between Fock build and integral derivative calculations. Neither purification nor solution for Lagrange multipliers is required. Time steps comparable to those typical of CPMD are possible, without the need to resort to thermostats, and using only a single fictitious mass parameter. We expect that the use of multiple time-scale integration techniques[8,42,48] will substantially accelerate this method by obviating a significant fraction of the integral and integral derivative calculations, and in the future we shall present a detailed analysis of the efficiency of the method.

In this preliminary report, we have focused on a microscopic examination and characterization of the dynamics, as a function of the fictitious mass parameter. Our results denigrate the value of most fluxional quantities as criteria to assess the validity of ELMD results. Especially for hydrogen-atom stretching vibrations above $3000\,\text{cm}^{-1}$, we have exhibited cases where both the real and the total energy fluctuations are kept to $\lesssim 10^{-4}$ a.u.; where the time derivative dT_F/dt of the fictitious kinetic energy oscillates around zero with a mean of 10^{-10} a.u. and no systematic drift; and where the fictitious electronic oscillations are separated from the nuclear vibrations by several times the energy of the latter; yet even still the ELMD forces are not an accurate representation of the BOMD forces, even when averaged over fast electronic oscillations. In such cases, ELMD vibrational frequencies may exhibit a pronounced redshift relative to the correct BOMD values (though the gross features of the vibrational spectrum, obtained from a short simulation, can still be correct). These results contradict an earlier claim[15] that the use of atom-centered basis functions eliminates the μ dependence of vibrational frequencies, and we suggest that the developers of the ADMP method[13–16] should reexamine this dependence, using more difficult test cases.

Two criteria prove to be useful characterizations of the accuracy of ELMD: the ratio $\tau_{\text{nuc}}/\tau_{\text{el}}$ of the fastest nuclear time scale to the slowest electronic time scale, and the maximum value of T_F, the fictitious kinetic energy. For ELMD simulations that are near, but not within, the adiabatic decoupling régime, the latter offers a reliable estimate of the vibrational redshift, at least for diatomic molecules. More importantly, the ratio $\tau_{\text{nuc}}/\tau_{\text{el}}$, where τ_{el} may be estimated accurately using the fictitious mass and the HOMO-LUMO gap, is a reliable indication of whether the ELMD simulation is within the adiabatic decoupling régime. A time-scale sepa-

ration $\tau_{nuc}/\tau_{el} \sim 10$ appears sufficient to obtain hydrogen-atom stretching frequencies that are correct to within a few cm^{-1}. Our results indicate that these difficult cases may require a time step on the order of 0.1 fs or less (consistent with the CPMD liquid water results of Schwegler *et al.*[46]), although mass preconditioning may increase this value somewhat.

APPENDIX: GRADIENTS IN AN ARBITRARY ORTHOGONAL BASIS

In Sec. II D we gave an expression for $\partial E_{SCF}/\partial R_I$ that was particular to the Cholesky-orthogonalized AO basis, our representation of choice for the calculations in Sec. III. In this appendix we examine the form of the nuclear gradient for an arbitrary (bi)orthogonalization of the AO density matrix, i.e., for transformation matrices $\mathbf{A}$ and $\mathbf{B}$ such that $\mathbf{S}_{AO} = \mathbf{BA}$, where

$$\mathbf{P} = \mathbf{A}\mathbf{P}_{AO}\mathbf{B} \tag{A1}$$

is the density matrix in the biorthogonal basis. The Fock matrix transforms contravariantly with respect to $\mathbf{P}$, so

$$\mathbf{F} = \mathbf{B}^{-1}\mathbf{F}_{AO}\mathbf{A}^{-1}. \tag{A2}$$

For brevity, we introduce the notation $\mathbf{M}_x = \partial\mathbf{M}/\partial x$ and write the nuclear gradient as $g_x^{HF} + g_x^P$. The Hellmann-Feynman derivative g_x^{HF} in our arbitrary biorthogonal basis is the same as that given in Eq. (22), which is obvious since this expression involves only AO matrices and derivatives at fixed $\mathbf{P}_{AO}$ (rather than fixed $\mathbf{P}$). For the Pulay term g_x^P, on the other hand, the general result is

$$g_x^P = -\operatorname{Tr}(\mathbf{A}_x\mathbf{A}^{-1}\mathbf{PF} + \mathbf{B}^{-1}\mathbf{B}_x\mathbf{FP}), \tag{A3}$$

which reduces to Eq. (23) for $\mathbf{B} = \mathbf{L} = \mathbf{A}^\dagger$.

Now consider an arbitrary unitary transformation of the Fock and density matrices,

$$\tilde{\mathbf{F}} = \mathbf{U}\mathbf{F}\mathbf{U}^\dagger = \tilde{\mathbf{B}}^{-1}\mathbf{F}_{AO}\tilde{\mathbf{A}}^{-1} \tag{A4a}$$

$$\tilde{\mathbf{P}} = \mathbf{U}\mathbf{P}\mathbf{U}^\dagger = \tilde{\mathbf{A}}\mathbf{P}_{AO}\tilde{\mathbf{B}}, \tag{A4b}$$

where $\tilde{\mathbf{A}} = \mathbf{UA}$ and $\tilde{\mathbf{B}} = \mathbf{BU}^\dagger$. Let $\widetilde{g_x^P}$ denote the Pulay derivative in this new basis, defined as in Eq. (A3) but with tildes on all of the matrices. It can then be shown that

$$\widetilde{g_x^P} - g_x^P = 2\operatorname{Tr}[\mathbf{U}_x\mathbf{U}^\dagger(\mathbf{PF} - \mathbf{FP})], \tag{A5}$$

which renders manifest the fact that $\widetilde{g_x^P} - g_x^P = 0$ at SCF convergence, hence the BOMD nuclear forces are independent of the choice of representation. Away from SCF convergence, $\partial E_{SCF}/\partial R_I$ depends upon the choice of orthonormal basis through the transformation derivative term $\mathbf{U}_x\mathbf{U}^\dagger$. (The same conclusion was also reached in Ref. 16.) So long as the derivative $\mathbf{U}_x$ is well-behaved—assuming, for example, that the transformation from AOs to orthogonalized AOs does not

change discontinuously as a function of the nuclear geometry—Eq. (A5) suggests that the smallness of the electronic gradient $\mathbf{G} = \mathbf{FP} - \mathbf{PF}$ is important for achieving near independence of representation. This bodes well for the scalar-mass curvy-steps ELMD scheme, in which the occupied-occupied and virtual-virtual components of $\mathbf{G}$ are identically zero. By this logic, it is plausible that mass preconditioning may exacerbate the representation dependence of the nuclear forces, though this has not been investigated. We have, however, confirmed that for the diatomic molecules studied in Sec. III, $\mathbf{Z}(t)$ and $\mathbf{L}(t)$ are smooth functions of time.

[1] R. Car and M. Parrinello, Phys. Rev. Lett. **55**, 2471 (1985).
[2] D. K. Remler and P. A. Madden, Mol. Phys. **70**, 921 (1990).
[3] M. C. Payne, M. P. Teter, D. C. Allan, T. A. Arias, and J. D. Joannopoulos, Rev. Mod. Phys. **64**, 1045 (1992).
[4] D. Marx and J. Hutter, in *Modern Methods and Algorithms of Quantum Chemistry*, NIC Series vol. 3, 2nd ed., edited by J. Grotendorst (John von Neumann Institute for Computing, Jülich, Germany, 2000) p. 329.
[5] U. Rothlisberger, Comput. Chem. (Oxford) **6**, 33 (2001).
[6] M. E. Tuckerman, J. Phys.: Condens. Matter **14**, R1297 (2002).
[7] P. Pulay, Mol. Phys. **17**, 197 (1969).
[8] M. E. Tuckerman and M. Parrinello, J. Chem. Phys. **101**, 1316 (1994).
[9] J. Hutter, M. E. Tuckerman, and M. Parrinello, J. Chem. Phys. **102**, 859 (1995).
[10] Y. Shao, C. Saravanan, M. Head-Gordon, and C. A. White, J. Chem. Phys. **118**, 6144 (2003).
[11] M. Head-Gordon, Y. Shao, C. Saravanan, and C. A. White, Mol. Phys. **101**, 37 (2003).
[12] A. Edelman, T. A. Arias, and S. Smith, SIAM J. Matrix Anal. Appl. **20**, 303 (1998).
[13] 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 (2002).
[14] S. S. Iyengar, H. B. Schlegel, J. M. Millam, G. A. Voth, G. E. Scuseria, and M. J. Frisch, J. Chem. Phys. **115**, 10291 (2002).
[15] H. B. Schlegel, S. S. Iyengar, X. Li, J. M. Millam, G. A. Voth, G. E. Scuseria, and M. J. Frisch, J. Chem. Phys. **117**, 8694 (2002).
[16] S. S. Iyengar, H. B. Schlegel, G. A. Voth, J. M. Millam, G. E. Scuseria, and M. J. Frisch, Isr. J. Chem. **42**, 191 (2002).
[17] R. McWeeny, Rev. Mod. Phys. **32**, 335 (1960).
[18] R. Pino and G. E. Scuseria, Chem. Phys. Lett. **360**, 117 (2002).
[19] P. Tangney and S. Scandolo, J. Chem. Phys. **116**, 14 (2002).
[20] V. Wathelet, B. Champagne, D. H. Mosley, J.-M. André, and S. Massidda, Chem. Phys. Lett. **275**, 506 (1997).
[21] V. Wathelet, B. Champagne, D. H. Mosley, É. A. Perpète, and J.-M. André, J. Mol. Struct.: THEOCHEM **425**, 95 (1998).
[22] G. Chaban, M. W. Schmidt, and M. S. Gordon, Theor. Chem. Acc. **97**, 88 (1997).
[23] H. Larsen, J. Olsen, P. Jørgensen, and T. Helgaker, J. Chem. Phys. **115**, 9685 (2001).
[24] T. van Voorhis and M. Head-Gordon, Mol. Phys. **100**, 1713 (2002).
[25] J. M. Herbert and J. E. Harriman, J. Chem. Phys. **118**, 10835 (2003).
[26] W. Liang and M. Head-Gordon, J. Chem. Phys. **120**, 10379 (2004).
[27] J. Hutter, M. Parrinello, and S. Vogel, J. Chem. Phys. **101**, 3862 (1994).
[28] M. Frisch, M. Head-Gordon, and J. Pople, Chem. Phys. **141**, 189 (1990).
[29] C. Ochsenfeld and M. Head-Gordon, Chem. Phys. Lett. **270**, 399 (1997).
[30] M. E. Tuckerman and M. Parrinello, J. Chem. Phys. **101**, 1302 (1994).
[31] D. Raczkowski, C. Y. Fong, P. A. Schultz, R. A. Lippert, and E. B. Stechel, Phys. Rev. B **64**, 155203 (2001).
[32] M. P. Allen and D. J. Tildesley, *Computer Simulation of Liquids* (Oxford University Press, New York, 1987).
[33] M. Challacombe, J. Chem. Phys. **110**, 2332 (1999).
[34] M. Benzi, J. K. Cullum, and M. Tůma, SIAM J. Sci. Comput. (USA) **22**, 1318 (2001).
[35] M. Benzi, R. Kouhia, and M. Tůma, Comput. Methods Appl. Mech. Eng. **190**, 6533 (2001).
[36] F. Tassone, F. Mauri, and R. Car, Phys. Rev. B **50**, 10561 (1994).
[37] T. Frenkel, *The Geometry of Physics* (Cambridge University Press, Cambridge, UK, 1997).
[38] In Ref. 12, a quotient representation of the Grassmann manifold is em-

ployed, which affords a different form for tangent vectors. The quotient representation is germane when molecular orbitals are used in the electronic structure calculation, whereas a projection-operator representation is appropriate for density matrix-based SCF calculations.

[39] J. Kong, C. A. White, A. I. Krylov *et al.*, J. Comput. Chem. **21**, 1532 (2000).

[40] S. Chawla and G. A. Voth, J. Chem. Phys. **108**, 4697 (1998).

[41] W. J. Stevens, M. Krauss, and H. Basch, J. Chem. Phys. **81**, 6026 (1984).

[42] D. A. Gibson and E. A. Carter, J. Phys. Chem. **97**, 13429 (1993).

[43] P. Pulay and G. Fogarasi, Chem. Phys. Lett. **386**, 272 (2004).

[44] G. Pastore, E. Smargiassi, and F. Buda, Phys. Rev. A **44**, 6334 (1991).

[45] J. C. Grossman, E. Schwegler, E. W. Draeger, F. Gygi, and G. Galli, J. Chem. Phys. **120**, 300 (2004).

[46] E. Schwegler, J. C. Grossman, F. Gygi, and G. Galli, J. Chem. Phys. **121**, 5400 (2004).

[47] I.-F. W. Kuo, C. J. Mundy, M. J. McGrath *et al.*, J. Phys. Chem. B **108**, 12990 (2004).

[48] M. Tuckerman, B. J. Berne, and G. J. Martyna, J. Chem. Phys. **97**, 1990 (1992).

[49] A. D. Becker, J. Chem. Phys. **98**, 5648 (1993); P. J. Stephens, J. F. Devlin, C. F. Chabalowski, and M. J. Frisch, J. Phys. Chem. **98**, 11623 (1994).