Time-Reversible Born-Oppenheimer Molecular Dynamics

Anders M. N. Niklasson,[1,2,*] C. J. Tymczak,[1] and Matt Challacombe[1]

[1]*Theoretical Division, Los Alamos National Laboratory, Los Alamos, New Mexico 87545, USA*
[2]*Applied Materials Physics, Department of Materials Science and Engineering, Royal Institute of Technology,*
SE-100 44 Stockholm, Sweden
(Received 28 April 2006; published 18 September 2006)

We present a time-reversible Born-Oppenheimer molecular dynamics scheme, based on self-consistent Hartree-Fock or density functional theory, where both the nuclear and the electronic degrees of freedom are propagated in time. We show how a time-reversible adiabatic propagation of the electronic degrees of freedom is possible despite the nonlinearity and incompleteness of the self-consistent field procedure. With a time-reversible lossless propagation the simulated dynamics is stabilized with respect to a systematic long-term energy drift and the number of self-consistency cycles can be kept low thanks to a good initial guess given from the electronic propagation. The proposed molecular dynamics scheme therefore combines a low computational cost with a physically correct time-reversible representation, which preserves a detailed balance between propagation forwards and backwards in time.

PACS numbers: 31.15.Qg, 31.15.Ew, 34.10.+x, 71.15.Pd

Ab initio molecular dynamics based on Hartree-Fock or density functional theory [1–9] has become an important tool for simulations of an increasingly wider range of problems in geology, material science, chemistry, and biology. *Ab initio* molecular dynamics, where the atomic positions move along classical trajectories, can be categorized in two major groups: Lagrangian Car-Parrinello molecular dynamics and Born-Oppenheimer molecular dynamics [4–12]. The success of Car-Parinello molecular dynamics, invented two decades ago [4], is based on its efficient implementations and low computational cost. However, unless a Car-Parinello simulation is performed carefully, it may yield results different from Born-Oppenheimer molecular dynamics [5,6,13–17]. In Born-Oppenheimer molecular dynamics the atomic positions are propagated by forces that are calculated at the self-consistent electronic ground state for each instantaneous arrangement of the ions. Born-Oppenheimer molecular dynamics is computationally expensive compared to Car-Parinello dynamics because of the requirement to reach a self-consistent field (SCF) solution in each time step. However, the number of SCF cycles and thus the computational cost can be strongly reduced by using an initial guess for the electronic degrees of freedom $\rho(t_{n+1})$ (here represented by the electron density), which is given by an extrapolation from previous time steps [5,8,9,18–20]. The electronic extrapolation scheme, combined with the SCF procedure, can be seen as an adiabatic propagation of the electronic degrees of freedom, where

$$\rho(t_{n+1}) = \text{SCF}[\rho(t_n), \rho(t_{n-1}), \ldots]. \quad (1)$$

Unfortunately, this approach has a fundamental problem. Because of the nonlinear and irreversible SCF procedure, which in practice is never complete, the time-reversal symmetry of the electronic propagation is broken. This problem does not occur in Lagrangian Car-Parinello mo-

lecular dynamics [17], where both the nuclear and the electronic degrees of freedom can be propagated with time-reversible Verlet integrators [21]. The main purpose of this Letter is to show how an effective time-reversible lossless propagation of the electronic degrees of freedom is possible in Born-Oppenheimer molecular dynamics, despite an irreversible and approximate SCF procedure.

Computational schemes using time-reversible integrators give in general an energy meandering around a value that does not systematically drift with time. This property follows from the time-reversal symmetry, which excludes a steady increase or decrease of the energy for a periodic motion. Since the simulation time usually is very short compared to the Poincaré time, i.e., the period of the system, the total energy may still shift. Deviations in the energy can be bounded by using numerical integrators that exactly fulfill certain symmetries of Newtonian mechanics, such as the symplectic condition [22]. In this Letter we show how it is possible to drastically improve the accuracy and the stability of Born-Oppenheimer molecular dynamics by imposing the time-reversal symmetry on the propagation of the electronic degrees of freedom.

The irreversibility of the electronic propagation in Eq. (1), which also affects the integration of the nuclear degrees of freedom, leads to a small but systematic energy drift in the evolution of a microcanonical ensemble, with an accumulating phase-space error. The error can be systematically reduced by improving the SCF convergence or be removed completely by using an initial guess in the SCF procedure that is independent of previous time steps [8,9]. Both these remedies are computationally expensive and using an initial guess that is independent from previous time steps often require a significantly increased number of SCF cycles.

The basic principles for the time-reversible lossless integration of the electronic degrees of freedom is described

here in terms of the propagation of the electron density $\rho(t)$. However, time-reversible Born-Oppenheimer trajectories can be constructed by replacing the density by other parameters governing the electronic degrees of freedom, such as the effective single-particle potential, Hamiltonian, density matrix, or wave functions. Our approach is therefore general and applicable to a large number of electronic structure schemes based on self-consistent Hartree-Fock or density functional theory.

A reversible propagation of the electron density $\rho(t)$ in finite time steps of δt ($t_n = t_0 + n\delta t$) can be constructed from a lossless filter process analogous to, for example, the lossless wavelet transform used in data compression [23]. The principle of the process, which is the key result in this Letter, is shown in the upper part of Fig. 1. The scheme is divided into two channels; the upper channel, with an approximate auxiliary density (denoted by a tilde), $\tilde{\rho}_{n+1} \equiv \tilde{\rho}(t_{n+1})$, is used as an initial guess for the Born-Oppenheimer density, ρ_{n+1}, in the lower channel. The Born-Oppenheimer density is given through the nonlinear, in practice incomplete, and numerically lossy SCF procedure,

$$\rho_{n+1} = \text{SCF}[\tilde{\rho}_{n+1}]. \tag{2}$$

The function $U(\rho_n)$ is an update filter for the propagation of the auxiliary density,

$$\tilde{\rho}_{n+1} = U(\rho_n) - \tilde{\rho}_{n-1}. \tag{3}$$

It is easy to see that the filter process is perfectly lossless and reversible by running the process backwards in time with the $\oplus$ sign replaced by a $\ominus$ sign, as shown in the lower part of Fig. 1. The scheme is therefore a bijective map
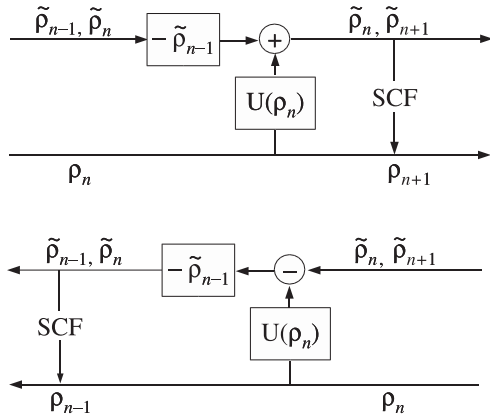

FIG. 1. The principle behind the time-reversible lossless filter process for propagation of the electronic density. In the forward filter process (upper part), $\tilde{\rho}_{n+1} = -\tilde{\rho}_{n-1} + U(\rho_n)$ and in the lossless backward reconstruction (lower part), $-\tilde{\rho}_{n-1} = \tilde{\rho}_{n+1} - U(\rho_n)$. The dual propagation with the auxiliary density $\tilde{\rho}_n$ allows a perfect reconstruction of the self-consistent Born-Oppenheimer density ρ_n, despite an irreversible, incomplete, and approximate SCF procedure.

which allows perfect reconstruction of ρ_n, despite the fact that the SCF procedure is an irreversible lossy transform.

The auxiliary density $\tilde{\rho}_{n+1}$ in Eq. (3) will be close to the self-consistent Born-Oppenheimer density ρ_{n+1} if the lossless filter process in Fig. 1 approximates the evolution of the density on the Born-Oppenheimer potential energy surface. This reduces the number of SCF cycles necessary to reach the new self-consistent density. One way to achieve this is to construct the update filter $U(\rho_n)$ in Eq. (3) from the time-reversible Verlet integrator [21] such that

$$\tilde{\rho}_{n+1} = [2\rho_n + \delta t^2 \ddot{\rho}_n] - \tilde{\rho}_{n-1}. \tag{4}$$

This integrator fulfills time-reversal symmetry since it remains the same if we switch the sign of δt and thus interchange $\tilde{\rho}_{n-1}$ with $\tilde{\rho}_{n+1}$. Note that perfect lossless reconstruction is a necessary, but not a sufficient, condition for time-reversal symmetry. The simplest time-reversible approximation of the acceleration $\ddot{\rho}_n = \partial^2 \rho_n / \partial t^2$ is to set it equal to zero, such that

$$\tilde{\rho}_{n+1} = 2\rho_n - \tilde{\rho}_{n-1}. \tag{5}$$

This surprisingly simple difference approximation fulfills time-reversal symmetry, allows a perfect reconstruction backwards in time, and avoids an unstable exponential error growth since the characteristic equation (with ρ_n replaced by $\tilde{\rho}_n$) has no roots outside the unit circle. If the auxiliary density $\tilde{\rho}_{n-1}$ in Eq. (5) is replaced by the self-consistent Born-Oppenheimer density ρ_{n-1} the propagation scheme is identical to a linear interpolation. Below we demonstrate how such a modification drastically affects phase-space conservation and the energy drift.

A quite general scheme for constructing update filters $U(\rho_n)$ for efficient time-reversible integrators is given by a least square fit of the ansatz

$$\tilde{\rho}(t) = \sum_{m=0}^{M} a_m t^{2m} - \tilde{\rho}(-t), \tag{6}$$

to Born-Oppenheimer densities $\rho(t_n)$ at N successive time steps. A similar least square approximation, but without the constraint of time reversibility, was recently proposed by Pulay and Fogarasi for the extrapolation of the single-particle Hamiltonian in a highly efficient Fock matrix dynamics (FMD) method [8,9]. Often only 2–3 SCF cycles are necessary in their scheme, but because of irreversibility (discussed in terms of hysteresis by Pulay and Fogarasi), a small but systematic energy drift occurs.

By choosing different numbers of fitted values N we can calculate the expansion coefficients a_m in Eq. (6) and express them in terms of previous Born-Oppenheimer densities. For example, a least square fit using the ansatz in Eq. (6), with $M = 1$ for 6 previous densities, leads to the time-reversible approximation

$$\tilde{\rho}_{n+1} = \frac{1}{13}[30(\rho_n + \rho_{n-4}) - 3(\rho_{n-1} + \rho_{n-3}) - 28\rho_{n-2}]$$
$$- \tilde{\rho}_{n-5}. \qquad (7)$$

Several alternative schemes for constructing time-reversible update filters $U(\rho_n)$ are possible. For example, we can correct the first order propagation of $\tilde{\rho}_n$ in Eq. (5) in its deviation from the exact self-consistent ground state using an energy gradient or a constrained functional gradient [24]. In such schemes the step length along the gradient can be associated with the inverse fictitious electron mass in Lagrangian Car-Parrinello dynamics.

To demonstrate the time-reversible lossless Born-Oppenheimer molecular dynamics we have implemented the scheme in MONDOSCF [25], a suite of programs using Gaussian basis sets for electronic structure calculations based on self-consistent Hartree-Fock or density functional theory. The density in the time-reversible propagation in Eqs. (5) and (6) has been replaced by the effective single-particle Hamiltonian, i.e., in this case the Fockian. For the nuclear coordinates a Verlet integrator was used [21]. The number of SCF cycles is measured in the number of constructions of Hamiltonians, which is the most time-consuming step.

Figure 2 shows the total energy as a function of time for an F_2 molecule. The modification of the time-reversible linear propagation in Eq. (5) to a linear interpolation leads to a systematic drift in the energy. Using more SCF cycles per time step reduces the energy drift, but it never really disappears. As a comparison we also show the (4−2) Fock matrix dynamics (FMD) scheme by Pulay and Fogarasi [8], based on a second order polynomial least square fit using four previous data points. This scheme gives, in principle, a more accurate extrapolation, which is noticed in a smaller
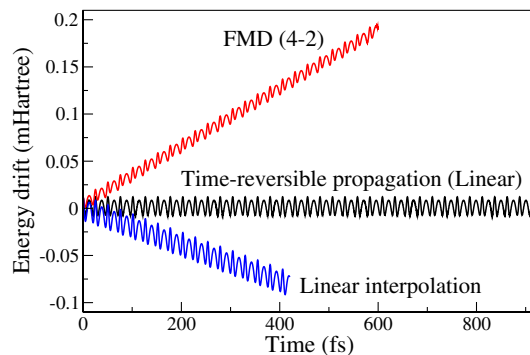
amplitude of the energy oscillations. However, there is a systematic drift in the energy. For the time-reversible linear integrator in Eq. (5), using only 2 SCF iterations per time step, any energy drift was less than 10^{-8} Hartree/ps. The phase space is also conserved with the time-reversible integration, as illustrated in Fig. 3. Thus, the simulated dynamics is physically accurate even with an incomplete SCF convergence.

The time-reversible propagators we have found so far have characteristic equations with all their roots on the unit circle. Roots inside the unit circle would improve stability but also lead to an exponential loss of memory. Because of the perfect lossless reconstruction, any error that occurs in the calculations will propagate throughout the simulation. This leads to a random noise that increases with time. Because of this noise the auxiliary density $\tilde{\rho}_n$ may slowly move away from the self-consistent solution. This means that it is not possible to combine long time steps with few SCF cycles. An increased number of SCF cycles (or shorter time steps) is in general necessary to reach a sufficiently accurate Born-Oppenheimer density for longer simulation times. Because of the dual filter process a conventional error analysis does not necessarily apply and the additional integration over the nuclear degrees of freedom further complicates the picture. Figure 4 shows the fluctuations in the total energy for a C_2F_4 molecule during 1 ps of simulation time at a temperature $T \approx 500$ K. The extrapolation scheme in Eq. (7) was used and 3 SCF cycles per time step were applied using Pulay's direct inversion in the iterative subspace (DIIS) algorithm to accelerate the convergence [26,27].

The ability of a perfect reconstruction of the dynamical data backwards in time and the local time reversibility is kept also using approximate arithmetic, thanks to the lossless bijective filter process illustrated in Fig. 1. The lossless property is therefore kept also when small elements below
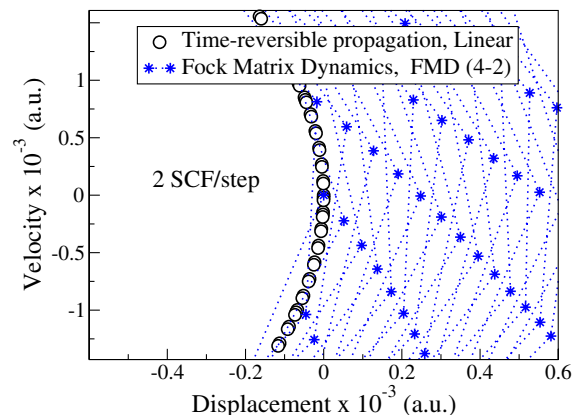

FIG. 2 (color online). The fluctuations in the total energy as a function of time for a F_2 molecule using Hartree-Fock theory with a Gaussian basis set (RHF/3-21G). The time-reversible propagation based on Eq. (5), with the density replaced by the effective single-particle Hamiltonian, is compared to the energy drift using the corresponding linear interpolation from previous time steps. The (4−2) Fock matrix dynamics (FDM) scheme by Pulay and Fogarasi is shown as a comparison. The time step was chosen to $\delta t = 0.25$ fs with 2 SCF/step.

FIG. 3 (color online). Detail of the phase space during 500 fs for one of the atoms in the F_2 molecule (Hartree-Fock theory with a Gaussian basis set RHF/3-21G, 2 SCF/step). The time-reversible propagation preserves the phase space whereas the non-time-reversible dynamics has a small but noticeable drift.
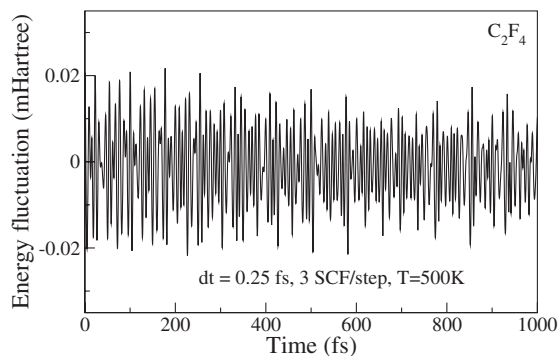
FIG. 4. The energy fluctuations around the average energy as a function of time for a C_2F_4 molecule at a temperature $T \approx$ 500 K (Hartree-Fock theory with a Gaussian basis set RHF/3-21G). Three SCF cycles per time step of length $\delta t = 0.25$ fs were used.

some drop tolerance are set to zero in the SCF optimization of ρ_n and in the update function $U(\rho_n)$. This is potentially important in the study of very large systems using, for example, linear scaling electronic structure methods [28] that in general require a lower numerical accuracy compared to conventional schemes.

We have shown that an effective time-reversible propagation of the electronic degrees of freedom is possible in self-consistent Born-Oppenheimer molecular dynamics, despite the irreversible, nonlinear, and in practice always approximate SCF procedure. The framework for lossless time-reversible Born-Oppenheimer molecular dynamics may open the door to a number of different methods, e.g., higher order symplectic integrators [22,29–32].

In summary, we have presented a scheme for time-reversible Born-Oppenheimer molecular dynamics that combines a low computational cost with a physically correct time-reversible propagation, which improves accuracy and stability of the dynamics. The principle is based on a lossless filter integration of the electronic degrees of freedom, where a dual propagation with an auxiliary density allows a perfect reconstruction backwards in time of the self-consistent Born-Oppenheimer density.

*Corresponding author.
 Email address: amn@lanl.gov

[1] C. C. J. Roothaan, Rev. Mod. Phys. **23**, 69 (1951).
[2] P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964).
[3] W. Kohn and L. J. Sham, Phys. Rev. **140**, A1133 (1965).
[4] R. Car and M. Parrinello, Phys. Rev. Lett. **55**, 2471 (1985).
[5] M. C. Payne, M. P. Teter, D. C. Allan, T. A. Arias, and J. D. Joannopoulos, Rev. Mod. Phys. **64**, 1045 (1992).
[6] D. Marx and J. Hutter, *Modern Methods and Algorithms of Quantum Chemistry*, edited by J. Grotendorst (John von Neumann Institute for Computing, Jülich, Germany, 2000), 2nd ed.
[7] H. B. Schlegel, J. M. Millam, S. S. Iyengar, G. A. Voth, A. D. Daniels, G. Scusseria, and M. J. Frisch, J. Chem. Phys. **114**, 9758 (2001).
[8] P. Pulay and G. Fogarasi, Chem. Phys. Lett. **386**, 272 (2004).
[9] J. Herbert and M. Head-Gordon, Phys. Chem. Chem. Phys. **7**, 3269 (2005).
[10] C. Leforestier, J. Chem. Phys. **68**, 4406 (1978).
[11] R. N. Barnett, U. Landman, A. Nitzan, and G. Rajagopal, J. Chem. Phys. **94**, 608 (1991).
[12] R. M. Wentzcovitch and J. L. Martins, Solid State Commun. **78**, 831 (1991).
[13] D. A. Gibson, I. V. Ionova, and E. Carter, Chem. Phys. Lett. **240**, 261 (1995).
[14] P. Tangney and S. Scandolo, J. Chem. Phys. **116**, 14 (2002).
[15] J. C. Grossman, E. Schwegler, E. W. Draeger, F. Gygi, and G. Galli, J. Chem. Phys. **120**, 300 (2004).
[16] E. Schwegler, J. C. Grossman, F. Gygi, and G. Galli, J. Chem. Phys. **121**, 5400 (2004).
[17] P. Tangney, J. Chem. Phys. **124**, 044111 (2006).
[18] T. Arias, M. Payne, and J. Joannopoulos, Phys. Rev. Lett. **69**, 1077 (1992).
[19] J. Millan, V. Bakken, W. Chen, L. Hase, and H. B. Schlegel, J. Chem. Phys. **111**, 3800 (1999).
[20] C. Raynaud, L. Maron, J.-P. Daudey, and F. Jolibois, Phys. Chem. Chem. Phys. **6**, 4226 (2004).
[21] L. Verlet, Phys. Rev. **159**, 98 (1967).
[22] R. D. Ruth, IEEE Trans. Nucl. Sci. **30**, 2669 (1983).
[23] A. R. Calderbank, I. Daubechies, W. Sweldens, and B. L. Yeo, Appl. Comput. Harmon. Anal. **5**, 332 (1998).
[24] A. M. N. Niklasson *et al.* (to be published).
[25] M. Challacombe, E. Schwegler, C. J. Tymczak, C. K. Gan, K. Nemeth, V. Weber, A. M. N. Niklasson, and G. Henkelman, computer code MONDOSCF $v1.0\alpha9$, 2001, http://www.t12.lanl.gov/home/mchalla/.
[26] P. Pulay, Chem. Phys. Lett. **73**, 393 (1980).
[27] P. Pulay, J. Comput. Chem. **3**, 556 (1982).
[28] S. Goedecker, Rev. Mod. Phys. **71**, 1085 (1999).
[29] E. Forest and R. Ruth, Physica (Amsterdam) **D43**, 105 (1990).
[30] B. J. Leimkuhler and R. D. Skeel, J. Comput. Phys. **112**, 117 (1994).
[31] D. I. Okunbor, J. Comput. Phys. **120**, 375 (1995).
[32] G. J. Martyna and M. Tuckerman, J. Chem. Phys. **102**, 8071 (1995).