Efficient formalism for large scale ab initio molecular dynamics based on time-dependent density functional theory

J. L. Alonso,1,2 X. Andrade,3 P. Echenique,1,2 F. Falceto,1,2 D. Prada-Gracia,4,2 and A. Rubio3

1Departamento de Física Teórica, Universidad de Zaragoza, Pedro Cerbuna 12, E-50009 Zaragoza, Spain.
2Instituto de Biocomputación y Física de Sistemas Complejos (BIFI).
3European Theoretical Spectroscopy Facility, Departamento de Física de Materiales, Universidad del País Vasco, Centro Mixto CSIC-UPV, and DIPC, Edificio Korta, Av. Tolosa 72, E-20018 San Sebastián, Spain.
4Departamento de Física de la Materia Condensada, Universidad de Zaragoza.

(Dated: August 29, 2008)

A new “on the fly” method to perform Born-Oppenheimer ab initial molecular dynamics (AIMD) is presented. Inspired by Ehrenfest dynamics in time-dependent density functional theory, the electronic orbitals are evolved by a Schrödinger-like equation, where the orbital time derivative is multiplied by a parameter. This parameter controls the time scale of the fictitious electronic motion and speeds up the calculations with respect to standard Ehrenfest dynamics. In contrast to other methods, wave function orthogonality needs not be imposed as it is automatically preserved, which is of paramount relevance for large scale AIMD simulations.

PACS numbers: 71.10.-w, 71.15.Pd, 31.15.Ew

Ab initio molecular dynamics (AIMD) on the ground state Born-Oppenheimer (gsBOMD) potential energy surface for the nuclei has become a standard tool for simulating the conformational behaviour of molecules, bio- and nano-structures and condensed matter systems from first principles [1]. However, gsBOMD (in the DFT [2] picture) requires that the Kohn-Sham (KS) energy functional be minimized for each value of the nuclei positions. As this minimization can be very demanding, Car and Parrinello (CP) [3] proposed an elegant and efficient “on the fly” scheme in which the KS orbitals are propagating with a fictitious dynamics that mimics gsBOMD. The CP method has had a tremendous impact in many scientific areas [4, 5]. Nevertheless, the numerical cost of AIMD hinders the application of the method to large scale simulations, such as those of interest in biochemistry or material science. Recently, new methods that allow larger systems and longer simulation times to be studied have been reported [6], but the cost associated with the wave function orthogonalization is still a potential bottleneck for both gsBOMD and CP.

Time-dependent density functional theory (TDDFT) [7, 8] has been for a long time recognized as an orthogonalization-free alternative for both ground state [4] and excited state AIMD. In its simplest implementation, Ehrenfest TDDFT, the ions are treated classically following electronic Hellmann-Feynman forces. For systems where the gap between the ground and the first excited state is large, Ehrenfest tends to gsBOMD and can mimic adiabatic dynamics [9]. However, the rapid movement of the electrons in TDDFT requires the use of a very small time step, which, in many occasions, renders its numerical application non-practical [10].

In this letter, we borrow some of the ideas of CP and introduce a new TDDFT Ehrenfest dynamics that reduces the cost of AIMD simulations while keeping the accuracy of the results in tolerable levels, similar to CP. The whole scheme can be obtained from the following Lagrangian (atomic units are used throughout this paper):

\[ \mathcal{L} = i \frac{\mu}{2} \sum_{j=1}^{N} \int \left( \phi_j^* \dot{\phi}_j - \dot{\phi}_j^* \phi_j \right) \text{d}r + K_I - E[\phi, R], \quad (1) \]

where \( K_I = \frac{1}{2} \sum_I M_I \dot{R}_I \cdot \dot{R}_I \) is the kinetic energy of the nuclei, \( M_I \) their masses and \( E \) the KS energy. Note that the major modification with respect to TDDFT is the scaling of the electronic velocities by a parameter \( \mu \) (TDDFT is recovered when \( \mu = 1 \)). We show in what follows that, in the \( \mu \rightarrow 0 \) limit, the trajectories of the system approach gsBOMD, and practical calculations can be done for values of \( \mu \gg 1 \), thus allowing for more efficient implementations than TDDFT while retaining its advantageous properties: the conservation of the total energy and of the orthogonality of the orbitals. Also, from the computational point of view, the new scheme is simple and can be easily incorporated into existing codes.

The equations of motion obtained from (1) for the electronic (\( \phi_j \)) and nuclear (\( \mathbf{R}_I \)) degrees of freedom are:

\[ i \mu \dot{\phi}_j = \frac{\delta E[\phi, R]}{\delta \phi_j^*} = -\frac{1}{2} \nabla^2 \phi_j + v_{\text{eff}}(r, t) \phi_j, \quad (2a) \]

\[ M_I \ddot{\mathbf{R}}_I = -\nabla_I E[\phi, R], \quad (2b) \]

where \( v_{\text{eff}} \) is the time-dependent KS effective potential.

In contrast to CP, the new dynamics conserves the physical energy \( E_{\text{phys}} := K_I + E[\phi, R] \) as well as the scalar product among the orbitals \( \phi_j \). The first is a direct consequence of \( \mathcal{L} \) being linear in the velocities \( \dot{\phi}_j \) and \( \dot{\phi}_j^* \), and not depending explicitly on \( t \). The conservation of the scalar product requires more attention due to...
the nonlinear character of the term $\delta E / \delta \phi^*_j$. To prove it, note that $E[\phi, R]$ is invariant under any unitary transformation mixing the orbitals $\phi \rightarrow U \phi$, with $U = e^{-i(\epsilon / \mu)A}$, being $A$ an $N \times N$ Hermitian matrix. From this invariance and eq. (2), we have

$$\frac{d}{dt} \int A_{jk} \phi_j^* \phi_k \, d\mathbf{r} = \int \left( A_{jk} \dot{\phi}_j \phi_k + A_{jk} \phi_j^* \dot{\phi}_k \right) \, d\mathbf{r} = -\frac{i}{\mu} \int \left( A_{jk} \frac{\delta E}{\delta \phi_j} \phi_k - A_{jk} \phi_j^* \frac{\delta E}{\delta \phi_k} \right) \, d\mathbf{r} = \frac{d}{d\varepsilon} \left[ e^{-i\varepsilon A} \phi, R \right] \bigg|_{\varepsilon=0} = 0.$$

(3)

Now, since $\int A_{jk} \phi_j^* \phi_k \, d\mathbf{r}$ is constant for all $A = A^\dagger$, the scalar product of any pair $\phi_j, \phi_k$ is a constant as well. Hence, if we start from an orthonormal set, we will not have to reorthonormalize the orbitals during the MD simulation. Numerically, this means that the formal scaling of the new scheme is quadratic with the number of atoms, while for CP and gsBOMD it is cubic due to the orthogonalization. In addition, the time propagation is naturally parallelizable by distributing the orbitals among different processors, as the evolution of each orbital is almost independent from the others’.

An important question is whether the new method reproduces gsBOMD. We show that the $\mu \rightarrow 0$ limit accounts for this solution. To do so, we recall that the BO Lagrangian reads

$$\mathcal{L}_{BO} = K_I - E[\phi, R] + \sum_{jk} \Lambda^\text{BO}_{jk} \left( \int \phi_j^* \phi_k \, d\mathbf{r} - \delta_{jk} \right),$$

(4)

where $\Lambda^\text{BO}_{jk}$ are the Lagrange multipliers which ensure the orthonormality of the orbitals. Clearly, as the orthonormality is automatically satisfied by the propagator in our approach, the limit $\mu \rightarrow 0$ gives the BO Lagrangian without the $\Lambda^\text{BO}$ term. Note, however, that one could have started from a different Lagrangian $\mathcal{L}' = \mathcal{L} + \sum_{jk} \Lambda^\text{BO}_{jk} \left( \int \phi_j^* \phi_k \, d\mathbf{r} - \delta_{jk} \right)$ for which the $\mu \rightarrow 0$ limit is $\mathcal{L}_{BO}$, and then, using (for $\mu \neq 0$) the gauge symmetry of $\mathcal{L}'$ ($\phi' = e^{iA_\phi} \phi$ and $\Lambda' = e^{iA} \Lambda e^{-iA} - i\mu e^{iA} \frac{\delta E}{\delta \phi} e^{-iA}$), where $A$ is a time-dependent Hermitian matrix, one can send $\Lambda'$ to zero recovering the dynamics of $\mathcal{L}$. This limit has a simple physical interpretation. The effect of $\mu$ is to scale the TDDFT excitation energies by a $1/\mu$ factor.

So for $\mu > 1$ the gap of the artificial system is decreased, increasing the non-adiabatic coupling, while for small values $\mu$ the excited states are pushed up in energy forcing the system to stay in the adiabatic regime.

Next, to provide an estimation of the performance improvements of our method in comparison with Ehrenfest dynamics, we write the left hand side of (2a) as

$$\mu (d\phi / dt) = d\phi / dt_e.$$  

With this transformation, (2a) can be seen as a standard TDDFT propagation, and we find that the maximum time step for our method in terms of $\mu$ is $\Delta t = \Delta t_e$, where $\Delta t_e$ is the maximum electronic time step, determined by the system and the propagation scheme. In the case of CP, on the other hand, $\Delta t \propto \sqrt{\mu CP}$. Note, however, that this difference does not imply anything about the relative performance of both methods, since the two parameters are not directly comparable (e.g., they have different dimensions).

The dependence of the accuracy on $\mu$ must also be taken into account, as we show later. Additionally, the ionic motion imposes a constraint in the maximum value of $\Delta t$, but usually this limit is much higher.

Now, although our method approaches the reference gsBOMD as $\mu \rightarrow 0$, this limit is not practical from a numerical point of view because it implies a time step $\Delta t \rightarrow 0$. But, as $\mu = 1$ is already close to gsBOMD for large gap systems, we shall mainly focus on how close we can stay to this limit for $\mu \gg 1$. In this regime, numerical simulations are in principle $\mu$ times faster than standard TDDFT, so we first made a detailed study of how large $\mu$ can be in CP and Ehrenfest for a simple 2-band model (see supplementary material), to conclude that the new scheme shows a performance similar to CP.

To further investigate in real systems the efficiency of this new approach, we implemented it, together with CP, in the first principle Octopus code [11]. For Ehrenfest dynamics, the Approximated Enforced Time Reversal Symmetry method [12] is used to propagate the electronic wave functions. In the case of CP the electronic part is integrated by a RATTLE/Velocity Verlet algorithm described in Ref. [13]. In both cases, velocity Verlet algorithm is used for the ionic equations of motion. The ions are represented using norm-conserving pseudo-potentials and the exchange correlation term is approximated by the Adiabatic LDA functional.

With respect to implementation there are some differences to remark. For Ehrenfest, the propagation must be performed using complex wave functions while for CP it can be performed using real wave functions for finite systems or for gamma point super-cell calculations in periodic systems. Also, due to the second order dynamics of CP, two sets of wave functions must be propagated while only one is needed for Ehrenfest. Finally, in the velocity Verlet algorithm, a temporary third set of wave functions is required to store the previous time step.

In parallel architectures, CP methods are known to scale very well based on domain decomposition [14]. This also applies to Ehrenfest dynamics and, on top of that, we can add a new level of parallelization by distributing groups of different states among processors. As the evolution of each state is independent, this is a very effective approach where communication is only required to calculate quantities that involve sums over all states, like the density or the forces. As these operations are performed only once per time step, it can scale efficiently even over slow interconnections. In the case of CP, due to orthogonalization between states, this parallelization
scheme is more complex to implement \cite{15} and requires more much communication.

The first real system we simulated was the Nitrogen molecule (see supplementary material \cite{19}). We observed that, for $\mu = 20$, the simulation remains steadily close to the BO potential energy surface, and there is only a 3.4\% deviation of the vibrational frequency. For $\mu = 30$ the system starts to strongly separate from the gsBO surface by mixing with higher BO surfaces.

Next, we applied the method to the benzene molecule. We set-up the atoms in the equilibrium geometry with a random Maxwell-Boltzmann distribution for 300$^\circ$K. Each run was propagated for a period of time of $\sim 400$ [fs] with a time step of $\mu \times 0.001$ [fs] (that provides a reasonable convergence in the spectra). Vibrational frequencies were obtained from the Fourier transform of the velocity auto-correlation function. In table 1, we show some low, medium and high frequencies of benzene as a function of $\mu$. The general trend is a red-shift of the frequencies with a maximum deviation of 7\% for $\mu = 15$. Still, to make a direct comparison with experiment, we computed the infrared spectra as the Fourier transform of the electronic dipole operator. In Fig. 1 we show how the spectra changes with $\mu$. For large $\mu$, besides the red-shift, spurious peaks appear above the higher vibrational frequency (not shown). We performed equivalent CP calculations for different values of $\mu_{CP}$, and found that, as shown in Fig. 2, it is possible to compare the physical error induced in both methods and establish and a relation between $\mu$ and $\mu_{CP}$.

Having established this link, we address the numerical performance of our new method compared to CP in terms of system size. To do this, we simulate several benzene molecules in a cell. For the new scheme, a value of $\mu = 15$ is used while for CP $\mu_{CP} = 750$, (values that yield a similar deviation from the BO surface, according to Fig. 1). The time steps used are 3.15 [a.u.] and 7.26 [a.u.] respectively. The computational cost is measured as the simulation time required to propagate one atomic unit of time. We performed the comparison both for serial and parallel calculations; the results are shown in Fig. 2. In the serial case, CP is 3.5 times faster for small systems, but the difference reduces to only 1.7 times faster for the larger ones. Extrapolating the results, we predict that the new dynamics will become less demanding than CP for around 1100 atoms. In the parallel case, the difference is reduced, CP being only 2 times faster than our method for small systems, and with a crossing point below 750 atoms. This is due to the better scalability of the Ehrenfest approach, as seen on Fig. 2. Moreover, memory requirements for our approach are lower than for CP: in the case of 480 atoms the ground state calculation requires a maximum of 3.5 GB, whereas in the MD, Ehrenfest requires 5.6 GB and CP 10.5 GB.

To close the computational assessment of the new formalism, we illustrate our method (using $\mu = 5$) for the calculation of the infrared spectrum of a prototype molecule, C_60. The calculated IR spectra is in very good agreement with the experiment (see Fig. 3) for low and high energy peaks (which are the most sensitive to $\mu$ as seen in Fig. 1). The result is robust and independent of the initial condition of the simulation. The low energy splitting of IR spectrum starts to be resolved for simulations longer than 2 [ps].

In conclusion, we have presented a new approach to AIMD based on a generalization of TDDFT Ehrenfest dynamics. Our approach introduces a parameter $\mu$ that controls the trade-off between the closeness of the simulation to the gsBO surface and the numerical cost of the calculation, analogously to the role of the fictitious electronic mass in CP. We have made direct comparisons of the numerical performance with CP, and, while quantitatively our results are system- and implementation-dependent, they prove that our method can outperform CP in some relevant cases, namely, for large scale systems that are of interest in several research areas and that can only be studied from first principles MD in massively parallel computers. To increase its applicability, it would also be important to study if the improvements developed to optimize CP can be combined with our approach \cite{6}, in particular, techniques to treat small-gap or metallic systems \cite{18}.

Finally, note that the introduction of the parameter $\mu$ comes at a cost, as we change the time scale of the movements of the electrons with respect to the Ehrenfest

| $\mu$ | 398 961 1209 1623 3058 |
|-------|------------------------|
| $\mu = 1$ | 396 958 1204 1620 3040 |
| $\mu = 5$ | 391 928 1185 1611 2969 |
| $\mu = 10$ | 381 938 1181 1597 2862 |
| $\mu = 15$ | 372 948 1178 1594 2762 |

**TABLE I:** Selected vibrational frequencies (in cm$^{-1}$) for the benzene molecule, obtained using different values of $\mu$. 

**FIG. 1:** Calculated infrared spectrum for benzene for different values of $\mu$, compared to CP dynamics and to experiment \cite{16}. 

![Infrared spectrum for benzene](image-url)
FIG. 2: Computational performance comparisons of our method and CP for an array of benzene molecules with finite boundary conditions and a spacing of 0.6 [a.u.]. Performance is measured as the computational time required to propagate one atomic unit of time. a) Single processor computational cost for different system sizes. (inset) Polynomial extrapolation for larger systems. Performed in one core of an Intel Xeon E5435 processor. b) Parallel computational cost for different system sizes. Performed in 32 × Intel Itanium 2 (1.66 GHz) processor cores of a SGI Altix. c) Parallel scaling with respect to the number of processor for a system of 480 atoms in a SGI Altix system. In both cases a mixed states-domain parallelization is used to maximize the performance.

case, which implies a shift in the electronic excitation energies. This must be taken into account when we extend the applicability of our method for non-adiabatic MD and MD under electromagnetic fields, in particular for the case of Raman spectroscopy, general resonant vibrational spectroscopy, and laser induced molecular bond rearrangement (work in progress).

We acknowledge support from MEC: FIS2006-12781-C02-01, FPA2006-02315, FIS2007-65702-C02-01, BFM2002-00113 and postdoc grant for P. E.; EC: Nanoquanta (NMP4-CT-2004-500198), SANES (NMP4-CT-2006-017310) and e-I3 (INFRA-211956); UPV/EHU (SGiker Arina) and Barcelona Supercomputing Center. We thank F. Mauri for helpful discussions and the refer-

FIG. 3: Infrared spectrum of C60. The (blue) dashed line corresponds to the calculated one (µ = 5 and 2 [ps] of time) while the black bars are the experimental values from Ref. [17].

[1] D. Marx and J. Hutter, in Modern Methods and Algorithms of Quantum Chemistry, edited by J. Grotendorst (John von Neumann Institute for Computing, Jülich, 2000), vol. 3, pp. 329–477.
[2] W. Kohn and L. J. Sham, Phys. Rev. 140, A1133 (1965).
[3] R. Car and M. Parrinello, Phys. Rev. Lett. 55, 2471 (1985).
[4] W. Andreoni, D. Marx, and M. Sprik, Chem. Phys. Chem. 6, 1671 (2005).
[5] M. E. Tuckerman, J. Phys.: Condens. Matter 14, R1297 (2002).
[6] T. D. Kühl, M. Krack, F. R. Mohamed, and M. Parrinello, Phys. Rev. Lett. 98, 066401 (2007).
[7] M. A. L. Marques, C. A. Ullrich, F. Nogueira, A. Rubio, K. Burke, and E. K. U. Gross, eds., Time Dependent Density Functional Theory (TDDFT), vol. 706 of Lecture Notes in Physics (Springer Verlag, 2006).
[8] E. Runge and E. K. U. Gross, Phys. Rev. Lett. 52, 997 (1984).
[9] A. Selloni, P. Carnevali, R. Car, and M. Parrinello, Phys. Rev. Lett. 59, 823 (1987).
[10] J. Theilhaber, Phys. Rev. B 46, 12990 (1992).
[11] A. Castro, M. A. L. Marques, H. Appel, M. Oliveira, C. A. Rozzi, X. Andrade, F. Lorenzen, E. K. U. Gross, and A. Rubio, Phys. Status Solidi B 243, 2465 (2006).
[12] A. Castro, M. A. L. Marques, and A. Rubio, J. Chem. Phys. 121, 3425 (2004).
[13] M. E. Tuckerman, and M. Parrinello, J. Chem. Phys. 101, 1316 (1994).
[14] J. Hutter, and A. Curioni, Parallel Comput. 31, 1 (2005).
[15] F. Lorenzen, Massively-parallel eigensolver for the octopus code in Science and Supercomputing in Europe 2007 (to appear).
[16] NIST Mass Spec Data Center, S. E. Stein, director, in NIST Chemistry WebBook, NIST Standard Reference Database Number 69, edited by P. J. Linstrom and W. G. Mallard (National Institute of Standards and Technology, Gaithersburg MD, 20899, 2005).
[17] T. Cabioc’h, A. Kharbach, A. Le Roy, and J. P. Riviè re, Chem. Phys. Lett. 285, 216 (1998).
[18] N. Marzari, D. Vanderbilt, and M. C. Payne, Phys. Rev. Lett. 79, 1337 (1997).

[19] Available online at http://neptuno.unizar.es/files/public/supp_materials/A102008PRLsupp.pdf

[20] Assuming no linear scaling techniques are used. Note that these techniques could also be combined with our method, where the orthogonalization-free propagation would be reflected in a smaller scaling prefactor.

This sets the maximum value of $\mu$ as the electronic gap divided by the highest vibration energy in the system.