Time-reversible quantum molecular dynamics

Anders M. N. Niklasson *

Theoretical Division, Los Alamos National Laboratory, Los Alamos, New Mexico 87545 and
Applied Materials Physics, Department of Materials Science and Engineering,
Royal Institute of Technology, SE-100 44 Stockholm, Sweden

C. J. Tymczak
Department of Physics, Texas Southern University, Houston, Texas 77004 and
Theoretical Division, Los Alamos National Laboratory, Los Alamos, New Mexico 87545

Matt Challacombe
Theoretical Division, Los Alamos National Laboratory, Los Alamos, New Mexico 87545
(Dated: December 7, 2006)

Time-reversible ab initio molecular dynamics based on a lossless multi-channel decomposition for the integration of the electronic degrees of freedom [Phys. Rev. Lett. 97, 123001 (2006)] is explored. We present a lossless time-reversible density matrix molecular dynamics scheme. This approach often allows for stable Hartree-Fock simulations using only one single self-consistent field cycle per time step. An analysis for the necessary condition for stability is presented. We also present a generalization, introducing an additional “forcing” term, that in a special case includes a hybrid Lagrangian, i. e. Car-Parrinello-like, method, which can systematically be constrained to the Born-Oppenheimer potential energy surface by using an increasing number of self-consistency cycles in the nuclear force calculations. Furthermore, in analogue to the reversible and symplectic leapfrog or velocity Verlet schemes, we propose a density Verlet formalism for time-reversible Born-Oppenheimer molecular dynamics.

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

I. INTRODUCTION

In classical molecular dynamics it is well known that computational schemes that are not reversible in time often lead to numerical instabilities and difficulties with the conservation of the total energy [1, 2]. Irreversible schemes do not conserve the phase space volume and they do not satisfy the detailed balance required for hybrid Monte Carlo methods [3]. By integrating Newton’s equations of motion with a numerical method that preserves the time-reversal symmetry these problems can often be avoided. An example is the velocity Verlet integrator, which is both time-reversible and symplectic [4–9]. As a result the method is very good for long time simulations with time-reversible trajectories that conserve the total energy well. However, in ab initio Born-Oppenheimer molecular dynamics, based on self-consistent Hartree-Fock or density functional theory [10–16], the application of the Verlet algorithm for the integration of the nuclear motion is problematic [17, 18]. In Born-Oppenheimer molecular dynamics the electronic degrees of freedom, e. g. the density $\rho(t)$, is propagated by an extrapolation from previous time steps, $\rho(t - n\delta t)$. This is necessary to provide an accurate initial guess for the iterative self-consistent field (SCF) optimization, which constrains the solution to the Born-Oppenheimer potential energy surface. A good initial guess often reduces the computational cost by an order of magnitude [15, 17–21]. The electron extrapolation (predictor) combined with the SCF optimization (corrector) constitutes an adiabatic propagation of the electronic degrees of freedom,

$$\rho(t + \delta t) = \text{SCF}[\rho(t), \rho(t - \delta t), \ldots]$$

Unfortunately, a straightforward propagation of the electron density in Eq. (1) is inconsistent with a time-reversible dynamics. Because of the nonlinearity and irreversibility of the SCF procedure in Eq. (1) the forces acting on the atomic cores are not calculated under the constraint of a time-reversible adiabatic evolution of the electron density. The symmetry of the nuclear Verlet integrator is therefore broken. This typically leads to a hysteresis effect in the dynamics and a systematic long term energy drift with an error accumulation in the phase space [17, 18, 22]. The remedy for this problem was previously to force the SCF optimization to a very high degree of accuracy, which is computationally expensive, though less so if the initial guess is very good [17, 18]. At “exact” SCF convergence the optimized density is independent of the input density. In practice this means that only the nuclear degrees of freedom are propagated, without breaking the time-reversal symmetry in the Verlet integration. However, since the SCF optimization never is perfect, this remedy only reduces the problem with the systematic error accumulation. The problem does not disappear. A more efficient approach would be to re-

*Corresponding author: Email amn@lanl.gov
store the time-reversal symmetry in the electron propagation. This was achieved with the introduction of time-reversible Born-Oppenheimer molecular dynamics [22].

In time-reversible Born-Oppenheimer molecular dynamics the time-reversal symmetry in the propagation of the electronic degrees of freedom is restored, for example, by using a dual filter process, where an approximate auxiliary electron density \( \tilde{\rho}(t) \) is integrated in parallel to the self-consistent Born-Oppenheimer density \( \rho(t) \). This propagation can be made both lossless and time-reversible [22]. More recently Kühne et al. [23] suggested a similar approach to time-reversible Born-Oppenheimer molecular dynamics, but instead of using a dual filter process their electron propagation is based on the predictor-corrector scheme by Kolafa [24], where the time-reversibility is constrained to some chosen high order in each time-step.

In this paper we extend the original ideas behind time-reversible Born-Oppenheimer molecular dynamics [22] based on a dual lossless filter-integration. First we briefly review the lossless time-reversible integration. Thereafter we present a density matrix scheme for the time-reversible linear propagation of the electronic degrees of freedom. The density matrix approach often allows a hybrid SCF cycle per time step. A stability analysis is given in the appendix. We also present a generalization of the density matrix scheme by introducing an additional “forcing” term, that in a special case includes a hybrid Lagrangian, i. e. Car-Parrinello-like, method [25–28]. This hybrid scheme can systematically be constrained to the Born-Oppenheimer potential energy surface by including an increasing number of self-consistency cycles. Furthermore, in analogue to the reversible and symplectic leapfrog or velocity Verlet schemes, we propose a density Verlet-like formalism for time-reversible Born-Oppenheimer molecular dynamics, where the density velocity or the density acceleration are propagated through a lossless multi-channel integration.

The key advantage with time-reversible Born-Oppenheimer molecular dynamics is that, thanks to time-reversibility, the dynamics is stabilized with respect to energy fluctuations and phase-space errors. In fact, though unproven, the full symplectic symmetry of the velocity Verlet algorithm for the integration of the nuclear degrees of freedom may be restored. We are therefore no longer required to calculate forces with a precision necessary to achieve sufficient global accuracy over the total simulation length, including maybe thousands of time steps. Instead it is only necessary to determine forces with a precision appropriate for local time intervals. The computational cost can therefore be substantially reduced.

FIG. 1: The principle of lossless dual-channel filter integration. We use the short hand notation \( \rho_n = \rho(t_0 + n\delta t) \). The propagation is driven by the update filter \( U \), where \( \tilde{\rho}_{n+1} = U[\rho_n] - \tilde{\rho}_{n-1} \). The filter scheme is perfectly reversible, either by going backwards and changing the \( \oplus \) sign to a \( \ominus \) sign, or equivalently, by changing \( \delta t \) to \(-\delta t\).

II. LOSSLESS TIME-REVERSIBLE DUAL CHANNEL INTEGRATION

In the present study the nuclear degrees of freedom, represented by the coordinates and velocities of the atomic cores, are propagated with the time-reversible and symplectic velocity Verlet algorithm [4–8]. Thanks to the symplectic symmetry, areas of the phase space are preserved in the propagation. The dynamics therefore remain stable also when small numerical errors are introduced. However, if the forces are not calculated under the constraint of a time-reversible (or a symplectic) electron propagation the corresponding symmetry properties of the nuclear integrator are no longer guaranteed. Thus, we need a time-reversible propagation of the electronic degrees of freedom to recover the stabilizing properties of the velocity Verlet scheme for the nuclear propagation. This is achieved by using a lossless time-reversible dual channel integration. The more recently proposed solution to time-reversible Born-Oppenheimer molecular dynamics by Kühne et al. [23], where the time-reversibility is constrained to some chosen order in each time-step, is an interesting alternative, but it will not be considered in this article.

The principle of lossless time-reversible dual channel integration of the electronic degrees of freedom is illustrated in Fig. 1. The idea is to replace the irreversible adiabatic propagation of the electron density in Eq. (1) by a two-step procedure using an auxiliary density denoted by a tilde:

\[
\tilde{\rho}(t + \delta t) = U[\rho(t)] - \tilde{\rho}(t - \delta t),
\]

\[
\rho(t + \delta t) = \text{SCF}[\tilde{\rho}(t + \delta t)].
\]

Thanks to the approximate auxiliary density \( \tilde{\rho} \) the
scheme is perfectly reversible, despite the irreversibility and non-linearity of the iterative SCF procedure, which in practice always is approximate and never complete. The reverse propagation is given simply by replacing $\delta t$ with $-\delta t$ in Eqs. (2) and (3).

In the initial two time steps we set $\tilde{\rho}(t) = \rho(t)$. Thereafter, if we choose an appropriate update filter $U[\rho(t)]$, $\tilde{\rho}(t)$ will stay close to the self-consistent Born-Oppenheimer density $\rho(t)$, so that only few SCF cycles are necessary to find the self consistent ground state solution [22]. This strongly reduces the computational cost and, thanks to the time-reversal symmetry, the dynamics is stabilized, in particular with respect to a long-term energy drift.

III. TIME-REVERSIBLE AB INITIO DENSITY MATRIX MOLECULAR DYNAMICS

Instead of using the electron density $\rho(t)$ we can use other quantities representing the electronic degrees of freedom. Previously we used the effective single-particle Hamiltonian or Fockian, though the theory is valid for other representations of the electronic degrees of freedom [22]. In this article we will use the single-particle density matrix $P_n$, where

$$P_n = P(t_0 + n\delta t).$$

As for the auxiliary density we denote auxiliary density matrices by a tilde, $\tilde{P}_n$. We will assume that the density matrix is in an orthogonal representation. A non-orthogonal representation can always be transformed into an orthogonal form by a congruence transform based on, for example, the inverse Cholesky or the Löwdin factorization of the overlap matrix [25, 29–33], and a generalization to a non-orthogonal representation is straightforward, e.g. see Refs. [34–37].

The time-reversible density matrix molecular dynamics scheme is given by

$$\tilde{P}_{n+1} = 2P_n - \tilde{P}_{n-1},$$

$$P_{n+1} = \text{SCF} \left[ \theta(\mu I - \tilde{P}_{n+1}) \right].$$

Equation (6) includes a projection of the auxiliary density matrix $\tilde{P}_n$, using the Heaviside step function $\theta$ with the step formed at the chemical potential $\mu$, separating the $N_e$ occupied states from the unoccupied. This projection can be performed using McWeeny purification [26, 38, 39] or more efficiently with trace correcting purification [37, 40], which guarantees both idempotency, i.e. $P = P^2$, and the correct occupation, $Tr(P) = N_e$, at convergence. In trace correcting purification

$$X_0 = \tilde{P}_{n+1},$$

$$X_{k+1} = X_k + \sigma_k(X_k - X_k^2); \quad k = 0, 1, 2, \ldots,$$

where

$$\sigma_k = \arg \min_{\sigma = \pm 1} |Tr[X_{k+1}(\sigma)] - N_e|,$$

i.e., the sign factor $\sigma_k = \pm 1$ is determined from the criteria of minimizing the deviation of the trace $Tr[X_{k+1}]$ from the correct occupation $N_e$. At convergence

$$\theta(\mu I - \tilde{P}_{n+1}) = \lim_{k \to \infty} X_k.$$

The step is automatically formed in the gap separating the occupied from the unoccupied states. This is achieved without prior knowledge of the chemical potential $\mu$. Since the auxiliary density matrix $\tilde{P}$ usually is approximately idempotent and because of the rapid convergence of purification, only a couple of iterations are necessary in each time step.

1. Energy stabilization

The time-reversible linear integrator for the density matrix in Eq. (5) would be identical to a linear interpolation if the auxiliary density matrix $\tilde{P}_{n-1}$ was replaced by the self-consistent Born-Oppenheimer density matrix $P_{n-1}$. In this case there is a propagation only of the Born-Oppenheimer density matrix, which is given by

$$P_{n+1} = \text{SCF} \left[ \theta(\mu I - 2P_n + P_{n-1}) \right],$$

which is irreversible and lossy.

Figure 2 illustrates the effect of replacing the lossy linear propagation in Eq. (11) with the time-reversible lossless linear integration in Eqs. (5) and (6). The number of SCF cycles are measured in the number of constructions of effective single-particle Hamiltonians, i.e. Fock builds, since this is the most time consuming part. To accelerate
SCF convergence we used Pulay’s direct inversion in the iterative subspace (DIIS) algorithm [41, 42]. With 3 SCF cycles per time step the irreversible extrapolation scheme in Eq. (11) has a systematic drift in the total energy that is avoided, even when only 1 SCF cycle per time step is used with the lossless time-reversible linear integration. The irreversible linear propagation scheme in Eq. (11) was unstable and diverged if we applied only 1 or 2 SCF cycles per time step. This clearly demonstrates the advantage of imposing time-reversibility on the numerical integration.

2. Approximate dynamics using only 1 (or few) SCF cycles per step

We have found that the density matrix formulation of the time-reversible linear propagation of the electronic degrees of freedom often allows for stable Hartree-Fock simulations using only 1 SCF cycle per time step. A more detailed discussion regarding the necessary conditions for stability is given in the appendix. Figure 3 shows the fluctuations in the total energy for a water cluster using only 1 SCF cycle per iteration with a time step of 0.5 fs. The energy is fluctuating upward and downward, but it does not show any sign of a systematic long term energy drift. The same behavior can be seen in Fig. 4, which shows the energy drift for a C2F4 molecule using only 1 SCF update per time step.

Figure 5 shows the difference in energy fluctuations for the C2F4 molecule using 1 or 4 SCF cycles per time step. The difference has only a minor effect on the vibrational properties of the molecule. In fact, the time evolution of the C-C bond length in the simulation with only 1 SCF per time step has almost the identical behavior as the C-C bond length for a simulation with half the time step and with 4 SCF cycles per time step as shown in Fig. 6. The approximate scheme using only 1 (or few) SCF cycles can thus provide an accurate simulation of the dynamics.

By using a time-reversible integration of the electronic degrees of freedom, the error measured by the global root mean square (RMS) error of the energy,

$$\text{Error}_{RMS} = \sqrt{\frac{1}{N} \sum_{n=1}^{N} (\langle E \rangle - E_n)^2},$$  

will be dominated by the accuracy of the (symplectic [47]) velocity Verlet integration of the nuclear coordinates. Figure 7 shows the RMS error averaged over the first 1000 time steps as a function of the step length δt, i.e. for $N = 1000$ in Eq. (12). We find that using only 1 SCF cycle per time step essentially gives the same order of accuracy as when 4 SCF cycles per step are used (in combination with DIIS). This would not necessarily be the case if we had used an irreversible integrator of the electronic degrees of freedom, since the energy drift and stability in that case strongly depends on the time step and the number of SCF cycles.

In conclusion, since the energy fluctuations are stabilized thanks to time-reversibility we are no longer required to calculate forces with a precision necessary to achieve sufficient global accuracy over the total simulation length, including maybe thousands of time steps. Instead it is only necessary to determine forces with a precision appropriate for local time intervals. The computational cost can therefore be substantially reduced.

IV. GENERALIZED TIME-REVERSIBLE LINEAR INTEGRATION

The time-reversible dual filter integration in Fig. 1 can be generalized by including a contribution also from the auxiliary density $\tilde{\rho}(t)$ in the update filter. Thus, instead of using the update filter $U = U[\rho(t)]$ we can use the
slightly more general filter $U = U[\tilde{\rho}(t), \rho(t)]$. The generalization is illustrated in Fig. 8. Using the generalized filter update the time-reversible density matrix dynamics in Eq. (5) can be extended with an additional “forcing” term $G(P_n, \tilde{P}_n)$ such that

$$\tilde{P}_{n+1} = 2P_n - \alpha G(P_n, \tilde{P}_n) - \tilde{P}_{n-1},$$

(13)

$$P_{n+1} = \text{SCF} \left[ \theta(\mu I - \tilde{P}_{n+1}) \right],$$

(14)

with two choices for $\tilde{P}_n$:

$$\tilde{P}_n = \begin{cases} P_n, & \alpha = 0, \\ \tilde{P}_n, & \alpha \neq 0. \end{cases}$$

(15)

Other combinations of $\alpha$ and $\tilde{P}_n$ in Eq. (15) can be used, but are not considered here. The $\alpha = 0$ case in Eq. (13) corresponds to the time-reversible linear propagation in Eq. (5), whereas the choice $\alpha \neq 0$ includes the forcing term $G$. The additional forcing term can be chosen, for example, to keep $P_n$ and $\tilde{P}_n$ close together or closer to the exact self-consistent ground state, which can be achieved with $G$ given from the derivative of some measure of the distance between $P_n$ and $\tilde{P}_n$ or their distance from the self-consistent ground state. While there certainly are other choices of the forcing term $G$ that may work better, one particular choice has the interesting feature that it generalizes to Lagrangian Car-Parrinello molecular dynamics [25–28],

$$G(P_n, \tilde{P}_n) = \nabla_{\tilde{P}_n} F_{\text{tot}}[\tilde{P}_n, H(P_n)]|_{P_n=P_2} = \{\tilde{P}_n, [\tilde{P}_n, H(P_n)]\}.$$  

(16)
density matrix $P_n$ from Eq. (14). If $\tilde{P}_n$ would be at the self-consistent ground state, $\tilde{P}_n$ commutes with $H(P_n)$ and the $G$ term is zero. The energy gradient in Eq. (16) is taken under the constraint of idempotency, i.e. $\tilde{P}_n = \tilde{P}_n^2$, which corresponds to the criteria of wavefunction orthogonality in the original Car-Parrinello scheme. Thus, the forcing term $G$ in Eq. (16) can be seen as an electronic Lagrangian force. Including this Lagrangian force does not necessarily give a better result, but it offers an interesting comparison and generalization of extended Lagrangian Car-Parrinello dynamics. By comparing the generalized density matrix propagation in Eq. (13), using the forcing term given by Eq. (16) to Lagrangian Car-Parrinello molecular dynamics [25–28], in the case when no SCF updates are applied, we find that

$$\alpha = \frac{\delta t^2}{2\mu_e} .$$

Here $\mu_e$ is the fictitious electron mass parameter used in Car-Parrinello molecular dynamics. The scalar prefactor $\alpha$ of the forcing term $G$ in Eq. (13) is thus proportional to the inverse fictitious electron mass parameter in Car-Parrinello molecular dynamics. When multiple SCF updates of $P_n$ are used in the nuclear force calculation the corresponding Lagrangian is constrained closer to the Born-Oppenheimer potential energy surface. In this sense Eq. (13) including the $G$ term is a Lagrangian-like approach only for the electronic degrees of freedom since $\tilde{P}_n$ never is updated through the SCF procedure for $\alpha \neq 0$ in Eq. (15). By increasing the number of SCF cycles in the calculation of $P_{n+1}$ in Eq. (14) the nuclear motion may successively be brought closer to the exact Born-Oppenheimer potential energy surface. Because of the lossless dual filter process this does not affect the time-reversal symmetry, not even for incomplete SCF convergence.

3. Successive transition between Lagrangian and Born-Oppenheimer molecular dynamics

When the forcing term $G$ in Eq. (16) is included in Eq. (13), and when no SCF cycles are performed, the time-reversible integration corresponds to a Lagrangian density matrix scheme. In this case our scheme is quite similar to the approach proposed by Schlegel et al. [26, 28], but in contrast to their algorithm we do not break the time-reversal symmetry in the purification step. In this sense our scheme is closer to the extended Lagrangian scheme developed by Herbert and Head-Gordon [25]. They use a unitary (and thus time reversible) exponential parametrization in the integration of the electronic degrees of freedom, which avoids the purification step. However, thanks to the time-reversal filter integration our scheme allows for repeated SCF updates in the force calculations. Each additional SCF cycle systematically constrains the nuclear degrees of freedom to the Born-Oppenheimer potential energy surface. This is possible thanks to the time-reversible dual filter integration. Figure 9 shows a comparison between energy fluctuations for $\alpha = 0.05$ and $\delta t = 0.1$ fs using 1, 2, or 3 SCF cycles. For each additional SCF cycle the Lagrangian-like scheme has smaller energy oscillations with a dynamics that becomes successively closer to the Born-Oppenheimer potential energy surface. As a comparison we show the energy fluctuations for the time-reversible linear propagator in Eq. (5), i.e. the time-reversible linear integrator in Eq. (5), is close to the “exact” solution using multiple SCF cycles. The time step $\delta t = 0.1$ fs.

FIG. 9: The total energy fluctuations with the Lagrangian-like integration in Eq. (13) of the electronic degrees of freedom for $\alpha = 0.05$ using 1, 2, and 3 SCF cycles per time step. The $\alpha = 0$ case, which corresponds to the time-reversible linear integrator in Eq. (5), using only 1 SCF cycle is virtually identical to the “exact” solution using multiple time-steps.

FIG. 10: The total energy fluctuations with the Lagrangian-like integration in Eq. (13) of the electronic degrees of freedom for $\alpha = 0.03$ using 2 and 3 SCF cycles per time step. Using only 1 SCF/step was not stable. The $\alpha = 0$ case using only 1 SCF cycle, i.e. the time-reversible linear integrator in Eq. (5), is close to the “exact” solution using multiple SCF cycles. The time step $\delta t = 0.1$ fs.
V. DENSITY VERLET FORMALISM

In analogue to the symplectic leapfrog or velocity Verlet schemes for the nuclear degrees of freedom [4-7] we may divide the integration of the electron density in a density \(\rho(t, \mathbf{r})\), a density velocity \(\dot{\rho}(t, \mathbf{r})\), and a density acceleration component \(\ddot{\rho}(t, \mathbf{r})\):

\[
\rho(t + \delta t) = \rho(t) + \delta t \dot{\rho}(t) + \frac{1}{2} (\delta t)^2 \ddot{\rho}(t), \tag{18}
\]

\[
\dot{\rho}(t + \delta t) = \dot{\rho}(t) + \frac{1}{2} \delta t [\ddot{\rho}(t) + \dddot{\rho}(t + \delta t)]. \tag{19}
\]

However, we need to find a formulation for the density Verlet scheme above that can be used in time-reversible Born-Oppenheimer molecular dynamics. Using the idea behind the lossless multi-channel integration we can find several solutions. Though there are many different versions, of which several in principle are equivalent except for trivial differences in the initial conditions, we present four examples: A) density acceleration (leapfrog) Verlet,

\[
\tilde{P}_{n-1} = 2(P_{n-1} - 2P_{n-2} + P_{n-3}) - \tilde{P}_{n-3}, \tag{20}
\]

\[
\tilde{P}_{n-3/2} = \tilde{P}_{n-2} + \frac{1}{2} \tilde{P}_{n-2}, \tag{21}
\]

\[
\tilde{P}_n = P_{n-1} + \tilde{P}_{n-3/2} + \frac{1}{2} \tilde{P}_{n-1}, \tag{22}
\]

\[
\ddot{P}_{n-1} = \tilde{P}_{n-3/2} + \frac{1}{2} \ddot{P}_{n-1}, \tag{23}
\]

\[
P_n = SCF[\theta(\mu I - \tilde{P}_n)], \; n = 5, 6, \ldots, \tag{24}
\]

with the initial conditions

\[
\dot{\tilde{P}}_2 = (P_3 - P_1)/2, \quad \ddot{\tilde{P}}_2 = (P_3 - 2P_2 + P_1), \quad \dddot{\tilde{P}}_3 = (P_4 - 2P_3 + P_2), \quad \tilde{P}_k = P_k, \; k = 1, 2, 3, 4. \tag{25}
\]

B) density velocity Verlet I,

\[
\dot{\tilde{P}}_{n-1} = P_{n-1} - P_{n-3} - \tilde{P}_{n-3}, \tag{26}
\]

\[
\tilde{P}_n = P_{n-1} + \dot{\tilde{P}}_{n-1}, \tag{27}
\]

\[
P_n = SCF[\theta(\mu I - \tilde{P}_n)], \; n = 5, 6, \ldots, \tag{28}
\]

with the initial conditions

\[
\dot{\tilde{P}}_2 = (P_3 - P_1)/2, \quad \dot{\tilde{P}}_3 = (P_4 - P_2)/2, \quad \tilde{P}_k = P_k, \; k = 1, 2, 3, 4. \tag{29}
\]

C) density velocity (leapfrog) Verlet II,

\[
\dot{\tilde{P}}_{n-1/2} = 2(P_{n-1} - P_{n-2}) - \tilde{P}_{n-5/2}, \tag{30}
\]

\[
\tilde{P}_n = P_{n-1} + \dot{\tilde{P}}_{n-1/2}, \tag{31}
\]

\[
P_n = SCF[\theta(\mu I - \tilde{P}_n)], \; n = 4, 5, \ldots, \tag{32}
\]

with the initial conditions

\[
\dot{\tilde{P}}_{3/2} = (P_2 - P_1), \quad \dot{\tilde{P}}_{5/2} = (P_3 - P_2), \quad \tilde{P}_k = P_k, \; k = 1, 2, 3. \tag{33}
\]

and finally D) the density velocity Verlet III scheme,

\[
\dot{\tilde{P}}_{n-1} = \dot{\tilde{P}}_{n-2} + \dot{\tilde{P}}_{n-3/2} - 2\tilde{P}_{n-2} + \tilde{P}_{n-3}, \tag{34}
\]

\[
\tilde{P}_n = P_{n-1} + \dot{\tilde{P}}_{n-1}, \tag{35}
\]

\[
P_n = SCF[\theta(\mu I - \tilde{P}_n)], \; n = 4, 5, \ldots, \tag{36}
\]

with the initialization

\[
\dot{\tilde{P}}_2 = (P_3 - P_2)/2, \quad \dot{\tilde{P}}_k = P_k, \; k = 1, 2, 3. \tag{37}
\]

The density Verlet schemes presented above are reversible and, depending on the initial boundary condition that can be made more accurate, the schemes have a performance that is similar or equivalent to the time-reversible linear propagator in Eq. (5). The differences in the four schemes above are essentially determined by the choice of auxiliary densities used to preserve reversibility through the SCF optimization. Despite not offering any obvious improvements, the density Verlet-like schemes above potentially provide an approach to construct higher order symplectic integrators for the electronic degrees of freedom. Combinations using various additional G-terms, \(G(P_n, \dot{\tilde{P}}_n, \tilde{P}_n)\), as discussed previously, could also provide interesting opportunities.

VI. DISCUSSION AND CONCLUSIONS

We have proposed time-reversible ab initio density matrix schemes for quantum molecular dynamics simulations based on lossless multi-channel integration. The first scheme in Eq. (5) is a time-reversible linear density matrix molecular dynamics algorithm that we have found particularly simple and useful. It often allows for a stable Hartree-Fock molecular dynamics using only one single SCF update per time step. An analysis regarding the stability conditions is outlined in the appendix. The second
scheme in Eq. (13) is a generalization that forms a quite general framework for the propagation of the electronic degrees of freedom. In the case when the forcing term $G$ in Eq. (16) is used the scheme corresponds to a hybrid Lagrangian, i.e. Car-Parrinello-like, approach to Born-Oppenheimer molecular dynamics that allows for multiple SCF updates in the force calculation. By increasing the number of SCF cycles the nuclear degrees of freedom can be constrained to the Born-Oppenheimer potential energy surface. However, the hybrid Lagrangian-like approach to Born-Oppenheimer molecular dynamics in Eq. (13) with $G$ given by Eq. (16) has several disadvantages compared to the time-reversible linear propagator without the forcing term. In the case when we use only 1 SCF cycle per time step we find that the time-reversible linear propagation yields a dynamics closer to the exact self-consistent Born-Oppenheimer potential energy surface compared to the Lagrangian-like dynamics. In addition, the time-reversible linear integrator can be used with longer time steps and it has no dependency on a fictitious electron mass parameter.

Besides the approach to time-reversible Born-Oppenheimer molecular dynamics presented in Eqs. (5) and (13) we also proposed some density Verlet-like schemes that propagate not only the density, but also the density velocity or the density acceleration. This alternative approach may open for the construction of higher order symplectic integrators [6, 8, 43-45].

The propagation of the electronic degrees of freedom is coupled to the integration of the nuclear coordinates. By restoring the time-reversibility (or other symmetries) in the propagation of the electronic degrees of freedom, the full symplectic and/or time-reversal symmetry of the nuclear velocity Verlet integrator can be restored. Therefore, as a consequence, time-reversible Born-Oppenheimer molecular dynamics may already in its present form make it possible to apply higher-order symplectic methods for the integration of the nuclear degrees of freedom that previously were limited only to classical molecular dynamics.

When multiple SCF cycles are applied in combination with some SCF convergence acceleration technique, for example the DIIS scheme, different mixing coefficients are used in each individual time step. While the local time-reversal symmetry as well as the global reversibility is still preserved, the global time-reversal symmetry is broken. This may potentially have an effect on the energy stability. However, the global time-reversal symmetry can be restored by using the idea behind the time-reversal propagation of the electronic degrees of freedom also for the propagation of the (DIIS) mixing coefficients $c_k$, for example, with

$$\tilde{c}_k(t + \delta t) = 2c_k(t) - \tilde{c}_k(t - \delta t), \quad (38)$$

where the mixing coefficients without a tilde $c_k$ (for SCF cycle $k$) are constructed from DIIS or some other SCF acceleration algorithm. Higher order propagations can be used as well [22]. Modifying the number of SCF cycles has typically a larger effect on the energy conservation than the variation of the DIIS coefficients. We therefore always chose to keep the number of SCF cycles constant. Another possible solution, as discussed in the appendix, is to use a linear mixing scheme with a constant mixing parameter.

An interesting aspect of time-reversible Born-Oppenheimer molecular dynamics, which was briefly mentioned in the introduction, is for applications in hybrid Monte Carlo methods [3]. A time-reversible propagation satisfies the requirement of a detailed balance between propagation forwards and backwards in time, which may otherwise not be fulfilled.

Of importance in time-reversible Born-Oppenheimer molecular dynamics is the stability of simulations of very large systems using linear scaling electronic structure technologies based on approximate sparse matrix algebra. The property of perfect reconstruction as well as the local time-reversal symmetry and the stability conditions should still be preserved using approximate arithmetics.

The major advantage with time-reversible Born-Oppenheimer molecular dynamics is that, thanks to time-reversibility, the dynamics is stabilized with respect to energy fluctuations and phase-space errors. We are therefore no longer required to calculate forces with a precision necessary to achieve sufficient global accuracy over the total simulation length, including maybe thousands of time steps. Instead it is only necessary to determine forces with a precision appropriate for local time intervals. The computational cost can therefore be substantially reduced. In fact, the stability analysis in the appendix shows that by choosing an appropriate SCF optimization scheme only 1 single SCF update is necessary for stability in time-reversible Born-Oppenheimer molecular dynamics. Though, not based on the propagation of the density but on the density matrix, this was also demonstrated. The analysis shows that stability should be possible to preserve also when approximate arithmetics in linear scaling theory is used, since it should have only minor effects on the eigenvalues of the self-consistency repsonse kernel in Eq. (44).

We hope that the discussion and analysis of time-reversible Born-Oppenheimer molecular dynamics given in this article may inspire further development and exploration.

VII. APPENDIX: STABILITY ANALYSIS USING ONLY A SINGLE SCF UPDATE

Assume that the auxiliary density is given by some time-reversible guess from $m$ previous time steps

$$\tilde{\rho}_{n+1}(r) = \sum_{i=n}^{n-m+1} a_i \rho_i(r) + b_i \tilde{\rho}_i(r). \quad (39)$$
The updated Born-Oppenheimer density $\rho_{n+1}(r)$ is then given by

$$\rho_{n+1}(r) = \text{SCF}[\tilde{\rho}_{n+1}(r)],$$  \hspace{1cm} (40)

where the SCF optimization procedure is determined by some scheme, for example, linear mixing, the DIIS scheme by Pulay or some other functional minimization method. In a single SCF iteration we have that

$$\rho_{n+1}^{\text{new}}(r) = D[\rho_{n+1}^{\text{input}}(r)],$$  \hspace{1cm} (41)

where $D[\rho_{n+1}^{\text{input}}(r)]$ is the new density given by the input density. For example, this operation is given by

$$D[\tilde{\rho}_{n+1}(r)] = \theta[uI - H(\tilde{\rho}_{n+1}(r))],$$  \hspace{1cm} (42)

where $H(\tilde{\rho}_{n+1}(r))$ is the effective single-particle Hamiltonian constructed from the auxiliary density $\tilde{\rho}_{n+1}(r)$ and $\theta$ is the Heaviside step function with the step formed at the chemical potential $\mu$. At the exact self-consistent solution $\rho_{n+1}^*(r)$ we have that

$$\rho_{n+1}^*(r) = D[\rho_{n+1}^*(r)].$$  \hspace{1cm} (43)

Performing a first order expansion around this solution we find that

$$D[\tilde{\rho}_{n+1}(r)] \approx \rho_{n+1}^*(r) + \int f(r, r') \delta \rho_{n+1}(r') dr'.$$  \hspace{1cm} (44)

where $\delta \rho_{n+1}(r') = \tilde{\rho}_{n+1}(r) - \rho_{n+1}^*(r)$ and $f(r, r')$ is a SCF response kernel, e.g., see Ref. [46] for details. In some given basis-set representation we can write this first order expansion in matrix-vector representations as

$$D[\tilde{\rho}_{n+1}] \approx \rho_{n+1}^* + \Lambda \delta \rho_{n+1}.$$  \hspace{1cm} (45)

Using only a single SCF cycle and a linear mixing scheme,

$$\rho_{n+1} = \gamma D[\tilde{\rho}_{n+1}] + (1 - \gamma)\rho_{n+1},$$

$$\approx \gamma (\rho_{n+1}^* + \Lambda \delta \rho_{n+1}) + (1 - \gamma)(\rho_{n+1}^* + \delta \rho_{n+1})$$

$$\approx \rho_{n+1}^* + [\gamma \Lambda + (1 - \gamma)] \delta \rho_{n+1},$$  \hspace{1cm} (46)

we find that the condition for stability is given when

$$\|\gamma \Lambda + (1 - \gamma)\| < 1.$$  \hspace{1cm} (47)

If we assume that the linear mixing parameter $\gamma$ is a constant scalar, its optimal choice can be determined by the eigenvalues of the SCF response kernel $\Lambda$, which are related to the static dielectric constant [46]. Thus, by choosing an appropriate linear mixing parameter $\gamma$, it is possible to reach stability using only a single SCF iteration.

[47] Notice, that even if we have restored the time-reversal symmetry and removed the systematic energy drift, we have not formally proven that the combined integration of the electronic and nuclear degrees of freedom fulfills the symplectic condition [9].