Higher-order Extended Lagrangian Born-Oppenheimer Molecular Dynamics for Classical Polarizable Models

Alex Albaugh[1,4], Teresa Head-Gordon[1-5*], Anders M. N Niklasson[6*]
[1]Department of Chem. and Biomolecular Engineering, [2]Department of Chem., [3]Department of Bioengineering, [4]Pitzer Center for Theoretical Chem.
[5]Chem. Sciences Division, Lawrence Berkeley National Laboratory
University of California, Berkeley, CA 94720
[6]Theoretical Division, Los Alamos National Laboratory, Los Alamos, New Mexico 87545

Generalized extended Lagrangian Born-Oppenheimer molecular dynamics (XLBOMD) methods provide a framework for fast iteration-free simulations of models that normally require expensive electronic ground state optimizations prior to the force evaluations at every time step. XLBOMD uses dynamically driven auxiliary degrees of freedom that fluctuate about a variationally optimized ground state of an approximate 'shadow' potential which approximates the true reference potential. While the requirements for such shadow potentials are well understood, constructing such potentials in practice has previously been *ad hoc*, and in this work we present a systematic development of XLBOMD shadow potentials that match the reference potential to any order. We also introduce a framework for combining friction-like dissipation for the auxiliary degrees of freedom with general-order integration, a combination that was not previously possible. These developments are demonstrated with a simple fluctuating charge model and point induced dipole polarization models.

Corresponding authors:

*thg@berkeley.edu
*amn@lanl.gov

INTRODUCTION

Under the Born-Oppenheimer approximation[1] the time scale separation between the motions of the nuclear and electronic degrees of freedom allows for the latter to be solved iteratively to self-consistency at each time step for a fixed nuclear configuration during a molecular dynamics trajectory. However, the undesirable added computational cost of iterative optimization, combined with numerical errors that destroy energy conservation if the self consistent field (SCF) is not tightly converged, has led to treating the electronic degrees of freedom as additional dynamical variables in an extended system in which they are integrated forward in time with no SCF iterations. The extended Lagrangian (EL) scheme for *ab initio* molecular dynamics by Car and Parrinello,[2] where the electronic degrees of freedom are included as classical dynamical variables, has been used with success in condensed phase simulations and in applications with classical polarization solutions for inducible point dipoles,[3] Drude particles,[4] and fluctuating charges.[5-7] However, the time steps are typically much smaller than those of SCF converged Born-Oppenheimer simulations in order to maintain numerical stability and energy conservation, thus negating the benefit of eliminating the SCF calculation during the MD simulation.

Over the last decade Niklasson and colleagues[8-18] have introduced and developed a more generalized extended Lagrangian Born-Oppenheimer molecular dynamics approach (XLBOMD), in which an auxiliary set of electronic degrees of freedom are used as either a time reversible initial guess for a self-consistent solver or as part of a well designed approximate potential, which we term a "shadow potential" (which conceptually is closely related to a shadow Hamiltonian[19]). The shadow potential is derived as a variationally *fully minimized* approximate functional for which exact forces can be calculated at low cost. With the right functional, one can obtain truly iteration-free dynamics by exactly integrating an approximate potential, as opposed to approximately integrating the exact potential. This idea of a backward error analysis is frequently used in applied mathematics and it is a key concept behind the construction of geometric integration schemes in classical dynamics.[20] In XLBOMD the approach was initially applied to non-linear self-consistent field theory such as Hartree-Fock and density functional theory (DFT).[16-18] While the XLBOMD methods can be hindered by resonance or accumulating numerical errors in the integration of the auxiliary equations of motion, the inclusion of dissipation into the integration of the extended electronic degrees of freedom in the form of a Langevin-like friction has proven very effective for small systems at short time-scales that are more typical in *ab initio* molecular dynamics (AIMD).[11]

Recently, Albaugh and co-workers extended the XLBOMD ideas to classical induced dipole[21-22] and Drude[23] polarizable force fields, as well as linear scaling DFT[24], by controlling

numerical errors and resonances not through dissipation, but by thermostating the auxiliary degrees of freedom. The primary benefit of the thermostating approach over dissipation is that time-reversibility is satisfied, however condensed phase system sizes must be large enough such that a thermodynamic temperature is well-defined.[21] Using thermostated auxiliary induced dipoles as an initial guess to an iterative SCF solver, one can reduce the number of SCF iterations required for tight convergence of a polarizable potential or electron density calculation by half or more in the iEL/SCF method[21, 24]. Furthermore, we have shown that by formulating an approximate polarization potential one can completely do away with the need for an iterative solver altogether in the so-called iEL/0-SCF method, which proved effective for both the induced dipole[22] and Drude[23] polarization models where significant increases in time step are now possible. The iEL/0-SCF formulation thus has the benefits of an EL approach, i.e. no iterations are performed at each time step, while eliminating the drawbacks of defining fictitious masses or dual temperature schemes, or the necessity of reducing the time step to maintain energy conservation and numerical stability.

In this work we show that the iEL/0-SCF approach is part of a general class of XLBOMD approaches whose success rests on the proper formulation of a shadow potential functional, i.e. a potential energy functional that is at its variationally optimized solution, and is a close approximation to the reference potential, for which exact forces can be evaluated. Although the general theoretical requirements in the construction of the shadow potential, in principle, are fairly well understood[16-17] previous approaches to building such potentials tended to be *ad hoc*[17, 22] as per the iEL/0-SCF method. However, in contrast to previous theoretical formulations, the electronic degrees of freedom will not be used as auxiliary dynamical variables and instead an approximation to the interaction operator acting on the electronic or polarizable degrees of freedom is used as an extended dynamical degree of freedom. Furthermore, we combine our shadow potential formalism with Langevin-like friction[11-12, 25] introduced through higher-order integration schemes that increase the accuracy and stability of the molecular dynamics and better minimize the energy drift. While this has previously been found to be difficult,[13] in this work we develop a general structure for systematically increasing the order of integration and dissipation to better satisfy time reversibility and energy conservation. We show that by combining the systematic improvements in the shadow potential formulation with increasing order in the intrinsic integration error with commensurate order of dissipation, there are synergistic effects on accuracy and stability in the XLBOMD formalism, as illustrated on a simple fluctuating charge model, but trivially extensible to point dipole formalism for polarization. Finally, we show how the iEL/0-SCF method fits within the general XLBOMD framework, whose success we show is an interplay of an optimal shadow

potential and a time-reversible auxiliary temperature control on the auxiliary equation of motion that mimics higher order dissipation.

THEORY

Here we present a systematic way of constructing the 'shadow potential energy functional', that satisfies both an adiabatic separation of the time scales between the electronic and the nuclear degrees of freedom, while also delivering an accurate approximation of the underlying exact potential that is given through an iteration-free exact optimization. This is coupled to the formulation of the equations of motion that drive the dynamics of the nuclear (particle) and electronic degrees of freedom using higher order integration schemes with comparable order in friction-like dissipation. The combination of all aspects then yields a general framework for classical polarization whereby we can choose or construct shadow potentials, dissipation, and integration, each to any order, allowing us to control accuracy and energy drift to an arbitrary degree. We present the theory in terms of a fluctuating charge model, for which we have formulated numerical experiments in the Results section, as well as the small variations needed to formulate the same theory for an inducible dipole model described in Appendix A.

Higher Order Shadow Potentials. The exact electrostatic potential energy surface for a fluctuating charge model is given from the constrained charge optimization given in Eq. (1)

$$U^{el}(\boldsymbol{r}^N, \boldsymbol{q}) = \boldsymbol{\chi}^{\boldsymbol{T}}\boldsymbol{q} + \frac{1}{2}\boldsymbol{q}^{\boldsymbol{T}}\boldsymbol{C}\boldsymbol{q} \tag{1}$$

where $\boldsymbol{\chi}$ is a vector of all the electronegativities of the N atoms in the system ($\boldsymbol{\chi} = [\chi_1 \dots \chi_N]^T$), $\boldsymbol{q}$ is a vector of the fluctuating partial charges of the N atoms ($\boldsymbol{q} = [q_1 \dots q_N]^T$), $\boldsymbol{r}^N$ denotes that the potential energy surface, U^{el}, is a function of the positions of the N atoms, and $\boldsymbol{C} = \boldsymbol{C}(\boldsymbol{r}^N)$ is a symmetric interaction tensor whose off-diagonal terms describe Coulombic interactions between charges q_i and q_j ($C_{ij} = 1/r_{ij}$ for $i \neq j$ in atomic units) and whose diagonal terms are twice the electronegative hardness of the i^{th} atom ($C_{ii} = 2H_i$).[13] Ignoring enforcement of charge neutrality for simplicity, the set of charges that minimizes the potential surface in Eq. (1) is thus the ground state Born-Oppenheimer solution for the system as given by Eq. (2).

$$\frac{\partial U^{el}}{\partial \boldsymbol{q}} = 0 = \boldsymbol{\chi}^{\boldsymbol{T}} + \boldsymbol{C}\boldsymbol{q} \quad \Longrightarrow \quad \boldsymbol{q}_{min} = -\boldsymbol{C}^{-\boldsymbol{1}}\boldsymbol{\chi} \tag{2}$$

The potential energy surface is then given as the variational, equilibrated charge minimum, i.e. $U^{el}(\boldsymbol{r}^N) = U^{el}(\boldsymbol{r}^N, \boldsymbol{q}_{min})$ from Eq. (2). Note that solving for $\boldsymbol{C}^{-\boldsymbol{1}}$ through matrix inversion or via an SCF procedure is the costly step that needs to be avoided.

We now introduce a dynamical auxiliary variable $\boldsymbol{X}$, which we design to be a good approximation to $\boldsymbol{C^{-1}}$. We then build a 'shadow' potential functional around $\boldsymbol{X}$ using some function $f(\boldsymbol{X})$ that we require to give a better approximation of $\boldsymbol{C^{-1}}$ than $\boldsymbol{X}$ itself, and for which we can control the degree to which this shadow potential matches the reference potential of Eq. (1):

$$U_{shadow}^{el}(\boldsymbol{r}^N, \boldsymbol{X}, \boldsymbol{q}) = \boldsymbol{\chi}^T\boldsymbol{q} + \frac{1}{2}\boldsymbol{q}^T[f(\boldsymbol{X})]^{-1}\boldsymbol{q} \tag{3}$$

Eq. (4) is a method for approximating the inverse of the $\boldsymbol{C}$ matrix such that $f(\boldsymbol{X}) \rightarrow \boldsymbol{C^{-1}}$ in the limit that $m \rightarrow \infty$,

$$f(\boldsymbol{X}) = \boldsymbol{C^{-1}}[\boldsymbol{I} - (\boldsymbol{I} - \boldsymbol{CX})^m] \tag{4}$$

whose general form in Eq. (4) was introduced by Niklasson[26] as a generalization of the Schulz method originally derived for the case of $m = 2$.[27]

We can determine the set of charges, $\boldsymbol{q}$, that minimize this new form of the shadow potential functional, given by substituting Eq. (4) into Eq. (3).

$$\frac{\partial U_{shadow}^{el}(\boldsymbol{r}^N, \boldsymbol{X}, \boldsymbol{q})}{\partial \boldsymbol{q}} = 0 = \boldsymbol{\chi}^T + [f(\boldsymbol{X})]^{-1}\boldsymbol{q} \tag{5a}$$

$$\Longrightarrow \boldsymbol{q}_{min} = -f(\boldsymbol{X})\boldsymbol{\chi} = -\boldsymbol{C^{-1}}[\boldsymbol{I} - (\boldsymbol{I} - \boldsymbol{CX})^m]\boldsymbol{\chi}$$

$$U_{shadow}^{el}(\boldsymbol{r}^N, \boldsymbol{X}) = \frac{1}{2}\boldsymbol{\chi}^T\boldsymbol{q}_{min} = -\frac{1}{2}\boldsymbol{\chi}^T\boldsymbol{C^{-1}}[\boldsymbol{I} - (\boldsymbol{I} - \boldsymbol{CX})^m]\boldsymbol{\chi} = -\frac{1}{2}\boldsymbol{\chi}^T f(\boldsymbol{X})\boldsymbol{\chi} \tag{5b}$$

From Eq. (5a) we can see that the set of charges that minimize the shadow potential functional are given by the negative action of any choice for $f(\boldsymbol{X})$ acting on the electronegativities, a straightforward operation, guaranteeing the variationally optimized minimum condition. Eq. (5b), then, gives various equivalent forms of this shadow potential energy surface at the variationally optimized minimum described by Eq. (5a) for different values of m. Note that while $\boldsymbol{C^{-1}}$ does appear explicitly in Eq. (5) it is annihilated when multiplied into the expansion and does not appear when writing $f(\boldsymbol{X})$ for some finite m (Table 1).

Following the previous work of Niklasson and colleagues[16-17] one can now define an extended Lagrangian for a system with the shadow potential of Eq. (5b). Instead of introducing auxiliary degrees of freedom that represent the electronic degrees of freedom themselves, (i.e. auxiliaries representing $\boldsymbol{q}$), we build the extended Lagrangian with the auxiliary $\boldsymbol{X}$ matrix, which should dynamically follow the behavior of the $\boldsymbol{C^{-1}}$ operator matrix, as a dynamical degree of freedom with a fictitious mass, m_X. This extended Lagrangian is given by

$$\mathcal{L}\left(\boldsymbol{r}^N, \dot{\boldsymbol{r}}^N, \boldsymbol{X}, \dot{\boldsymbol{X}}\right) = \frac{1}{2}\sum_{i=1}^{N} m_i\dot{\boldsymbol{r}}_i^2 + \frac{1}{2}m_X\mathrm{tr}\left[\dot{\boldsymbol{X}}^2\right] - U(\boldsymbol{r}^N) + \tag{6}$$

$$\frac{1}{2}\boldsymbol{\chi}^{\boldsymbol{T}} f(\boldsymbol{X})\boldsymbol{\chi} - \frac{1}{2} m_X \omega^2 \mathrm{tr}\left[\left(\widetilde{\boldsymbol{C}}^{-1} - \boldsymbol{X}\right)^2\right]$$

where $U(\boldsymbol{r}^N)$ are other non-electrostatic potential energy functions we may have in the system such as bonds, angles, van der Waals interactions, etc. that are independent of the fluctuating charges. The final term in Eq. (6) is a harmonic oscillator with frequency ω that seeks to keep the auxiliary $\boldsymbol{X}$ close to the ground state solution. This oscillator fluctuates about $\widetilde{\boldsymbol{C}}^{-1}$, which is given as a general placeholder for $\boldsymbol{C}^{-1}$. $\widetilde{\boldsymbol{C}}^{-1}$ can explicitly be $\boldsymbol{C}^{-1}$, which is not practical as inverting the $\boldsymbol{C}$ matrix is what we want to avoid, so $\widetilde{\boldsymbol{C}}^{-1}$ can be any approximation to $\boldsymbol{C}^{-1}$, as long as $\widetilde{\boldsymbol{C}}^{-1}$ is a better approximation to $\boldsymbol{C}^{-1}$ than $\boldsymbol{X}$ itself.

Table 1: *The m^{th}-order $f(\boldsymbol{X})$, and the corresponding variationally minimized set of atomic partial charges, $\boldsymbol{q}$, for that order along with the shadow potential, U^{el}_{shadow}, and equations of motion for the nuclear degrees of freedom and the auxiliary matrix, $m_i\ddot{\boldsymbol{r}}_i$ and $m_X\ddot{\boldsymbol{X}}$, respectively.*

m	**2**	**3**	**4**
$f(\boldsymbol{X})$	$2\boldsymbol{X} - \boldsymbol{XCX}$	$3\boldsymbol{X} - 3\boldsymbol{XCX} + \boldsymbol{XCXCX}$	$4\boldsymbol{X} - 6\boldsymbol{XCX} + 4\boldsymbol{XCXCX} - \boldsymbol{XCXCXCX}$
$\boldsymbol{q}$	$-(2\boldsymbol{X} - \boldsymbol{XCX})\boldsymbol{\chi}$	$-(3\boldsymbol{X} - 3\boldsymbol{XCX} + \boldsymbol{XCXCX})\boldsymbol{\chi}$	$-(4\boldsymbol{X} - 6\boldsymbol{XCX} + 4\boldsymbol{XCXCX} - \boldsymbol{XCXCXCX})\boldsymbol{\chi}$
U^{el}_{shadow}	$-\frac{1}{2}\boldsymbol{\chi}^T(2\boldsymbol{X} - \boldsymbol{XCX})\boldsymbol{\chi}$	$-\frac{1}{2}\boldsymbol{\chi}^T(3\boldsymbol{X} - 3\boldsymbol{XCX} + \boldsymbol{XCXCX})\boldsymbol{\chi}$	$-\frac{1}{2}\boldsymbol{\chi}^T(4\boldsymbol{X} - 6\boldsymbol{XCX} + 4\boldsymbol{XCXCX} - \boldsymbol{XCXCXCX})\boldsymbol{\chi}$
$m_i\ddot{\boldsymbol{r}}_i = -\frac{\partial U(\boldsymbol{r}^N)}{\partial \boldsymbol{r}_i} + \cdots$	$-\frac{1}{2}\boldsymbol{\chi}^T\boldsymbol{X}^T\frac{\partial \boldsymbol{C}}{\partial \boldsymbol{r}_i}\boldsymbol{X}\boldsymbol{\chi}$	$-\frac{1}{2}\boldsymbol{\chi}^T\boldsymbol{X}^T\left(3\frac{\partial \boldsymbol{C}}{\partial \boldsymbol{r}_i} - \frac{\partial \boldsymbol{C}}{\partial \boldsymbol{r}_i}\boldsymbol{XC} - \boldsymbol{CX}\frac{\partial \boldsymbol{C}}{\partial \boldsymbol{r}_i}\right)\boldsymbol{X}\boldsymbol{\chi}$	$-\frac{1}{2}\boldsymbol{\chi}^T\boldsymbol{X}^T\left(6\frac{\partial \boldsymbol{C}}{\partial \boldsymbol{r}_i} - 4\frac{\partial \boldsymbol{C}}{\partial \boldsymbol{r}_i}\boldsymbol{XC} + \frac{\partial \boldsymbol{C}}{\partial \boldsymbol{r}_i}\boldsymbol{XCXC} - 4\boldsymbol{CX}\frac{\partial \boldsymbol{C}}{\partial \boldsymbol{r}_i} + \boldsymbol{CXCX}\frac{\partial \boldsymbol{C}}{\partial \boldsymbol{r}_i} + \boldsymbol{CX}\frac{\partial \boldsymbol{C}}{\partial \boldsymbol{r}_i}\boldsymbol{XC}\right)\boldsymbol{X}\boldsymbol{\chi}$
$\ddot{\boldsymbol{X}} =$	$\omega^2(\boldsymbol{X} - \boldsymbol{XCX})$	$\omega^2(2\boldsymbol{X} - 3\boldsymbol{XCX} + \boldsymbol{XCXCX})$	$\omega^2(3\boldsymbol{X} - 6\boldsymbol{XCX} + 4\boldsymbol{XCXCX} - \boldsymbol{XCXCXCX})$

Applying the Euler-Lagrange equation to the Lagrangian of Eq. (6) we can derive the equations of motion for our system where Eq. (7a) is the equation of motion for the i^{th} real particle and Eq. (7b) for the auxiliary matrix, $\boldsymbol{X}$.

$$m_i\ddot{\boldsymbol{r}}_i = -\frac{\partial U(\boldsymbol{r}^N)}{\partial \boldsymbol{r}_i} + \frac{1}{2}\boldsymbol{\chi}^T\frac{\partial f(\boldsymbol{X})}{\partial \boldsymbol{r}_i}\boldsymbol{\chi} - m_X\omega^2(\widetilde{\boldsymbol{C}}^{-1} - \boldsymbol{X})\frac{\partial \widetilde{\boldsymbol{C}}^{-1}}{\partial \boldsymbol{r}_i} \tag{7a}$$

$$m_X\ddot{\boldsymbol{X}} = \frac{1}{2}\frac{\partial}{\partial \boldsymbol{X}}[\boldsymbol{\chi}^T f(\boldsymbol{X})\boldsymbol{\chi}] + m_X\omega^2\left(\widetilde{\boldsymbol{C}}^{-1} - \boldsymbol{X}\right) - m_X\omega^2(\widetilde{\boldsymbol{C}}^{-1} - \boldsymbol{X})\frac{\partial \widetilde{\boldsymbol{C}}^{-1}}{\partial \boldsymbol{X}} \tag{7b}$$

From the equations of motion in Eq. (7) we assume a classical adiabatic separation between the time scales of the particle and auxiliary motion where $\frac{\partial}{\partial \boldsymbol{X}}[\boldsymbol{\chi}^T f(\boldsymbol{X})\boldsymbol{\chi}]$ decays as ω^{-2} or faster. With this assumption and taking the limit $m_X \to 0$ and $\omega \to \infty$, the resulting equations of motion are given by Eq. (8), where we have explicitly evaluated $\frac{\partial f(\boldsymbol{X})}{\partial \boldsymbol{r}_i}$ and used Eq. (4) for $f(\boldsymbol{X})$.

$$m_i\ddot{\boldsymbol{r}}_i = -\frac{\partial U(\boldsymbol{r}^N)}{\partial \boldsymbol{r}_i} - \frac{1}{2}\boldsymbol{\chi}^T\boldsymbol{C}^{-1}\left[\frac{\partial \boldsymbol{C}}{\partial \boldsymbol{r}_i}\boldsymbol{C}^{-1}(\boldsymbol{I}-(\boldsymbol{I}-\boldsymbol{C}\boldsymbol{X})^m) - \sum_{j=0}^{m-1}(\boldsymbol{I}-\boldsymbol{C}\boldsymbol{X})^j\frac{\partial \boldsymbol{C}}{\partial \boldsymbol{r}_i}\boldsymbol{X}(\boldsymbol{I}-\boldsymbol{C}\boldsymbol{X})^{m-j-1}\right]\boldsymbol{\chi} \quad (8a)$$

$$\ddot{\boldsymbol{X}} = \omega^2\left(\widetilde{\boldsymbol{C}}^{-1}-\boldsymbol{X}\right) = \omega^2(\boldsymbol{C}^{-1}[\boldsymbol{I}-(\boldsymbol{I}-\boldsymbol{C}\boldsymbol{X})^m]-\boldsymbol{X}) \quad (8b)$$

Eq. (8a) gives the particle equation of motion, whose electrostatic force (second term on the right hand side) is now explicitly dependent on the auxiliary variable $\boldsymbol{X}$. Eq. (8b) defines an equation of motion that will propagate the auxiliary $\boldsymbol{X}$, and by choosing $\widetilde{\boldsymbol{C}}^{-1} = f(\boldsymbol{X})$ guarantees that the energy is variationally minimized by construction. The shadow potential then better matches the exact reference potential from Eq. (1) to an increasing degree of accuracy as m increases. Notice that for both Eq. (8a) and (8b), the equations of motion that drive this system no longer depend on an auxiliary mass parameter, m_X. One may note, again, that $\boldsymbol{C}^{-1}$ appears explicitly in Eq. (8), but this inverse is annihilated in the expansions of these equations for finite m, as seen in Table 1. Table 1 also gives explicit expressions for the equations of motion (Eq. (8)), $f(\boldsymbol{X})$, U^{el}_{shadow}, and $\boldsymbol{q}$ for several values of m. Finally, at the beginning of a simulation $\boldsymbol{X}$ is initialized to $\boldsymbol{C}^{-1}$ (or a close approximation), but this expensive operation only needs to be calculated once.

Using Eq. (8) we can now build XLBOMD based schemes that are time reversible, have exact analytical agreement between forces and energies, yield ground state atomic partial charges that minimize the potential energy at every time step, using a potential energy function that can be made arbitrarily close to the reference energy function by choosing higher values of m. One can see that this approach is similar to that of Brooks and colleagues[28-29] who build the polarization energy as a perturbation of the electrostatics truncated at a certain order, then statistically extrapolating to an infinite order to, in principle, recover the true energy minimized mutual polarization response. By contrast, we obtain an exact minimization of an approximate potential energy, which matches the true potential energy to a degree of our choosing.

Higher Order Dissipation. In practice the equation of motion of the auxiliary degree of freedom (Eq. (8b)) can suffer from resonance effects or instabilities caused by its coupling to the real degrees of freedom or due to numerical noise in the integration[11, 21]. These numerical artifacts need to be dissipated away lest the auxiliary matrix $\boldsymbol{X}$ drift too far from $\boldsymbol{C}^{-1}$ and cause the simulation to become unstable or energy to drift. To combat this problem Niklasson and

colleagues[11-12, 25] introduced a modified Verlet integration scheme for the auxiliary equation of motion, which has an additional dissipative term similar to the friction term used in Langevin dynamics. In order to introduce some amount of dissipation into the Verlet integration of the auxiliary equation of motion, Eq. (8b), we will introduce a friction-like term that is dependent on the history of the auxiliary matrix $\boldsymbol{X}$, as given in Eq. (9).

$$\boldsymbol{X}(t+\Delta t) = 2\boldsymbol{X}(t) - \boldsymbol{X}(t-\Delta t) + \kappa\,[\boldsymbol{C^{-1}}\,[\boldsymbol{I} - [\boldsymbol{I} - \boldsymbol{C}\boldsymbol{X}(t)]^m] - \boldsymbol{X}(t)] + \alpha \sum_{k=0}^{K} c_k \boldsymbol{X}(t-k\Delta t) \tag{9}$$

where Δt is the time step of the simulation, α controls the strength of the dissipation, and K is the order of the dissipation. We have also introduced $\kappa = \omega^2 \Delta t^2$, a dimensionless parameter that now controls the frequency of the auxiliary harmonic oscillation. While Eq. (9) does break time reversibility, the coefficients of the friction term, c_k, are designed to only break time reversibility up to $O(\Delta t^{2K-3})$ by removing odd-order terms in Δt from an expansion of the equation of motion[11]. These coefficients are reproduced for several orders governed by K in Table 2.

Table 2: *Coefficients for friction-like dissipation[11].*

K	c_0	c_1	c_2	c_3	c_4	c_5	c_6	c_7	c_8	c_9
3	-2	3	0	-1						
4	-3	6	-2	-2	1					
5	-6	14	-8	-3	4	-1				
6	-14	36	-27	-2	12	-6	1			
7	-36	99	-88	11	32	-25	8	-1		
8	-99	286	-286	78	78	-90	42	-10	1	
9	-286	858	-936	364	168	-300	184	-63	12	-1

Higher Order Integrators. The integration of the equation of motion given in Eq. (8), and combined with dissipation, is typically done by Verlet integration, shown in Eq. (9). While Verlet integration is generally robust, being symplectic, time-reversible, and energy conserving, one may want to use higher-order geometric integration schemes in order to realize a higher degree of accuracy or to use larger time steps. Such higher-order integrators have a general multi-step form ($l = 1,2,\dots,L$) described by Eq. (10).

$$\dot{\boldsymbol{X}}\left(t + \sum_{j=1}^{l} b_j \Delta t\right) = \dot{\boldsymbol{X}}\left(t + \sum_{j=1}^{l-1} b_j \Delta t\right) + b_l \Delta t \ddot{\boldsymbol{X}}\left(t + \sum_{j=1}^{l-1} a_j \Delta t\right), \quad l = 1,2,\dots,L \tag{10a}$$

$$\boldsymbol{X}\left(t + \sum_{j=1}^{l} a_j \Delta t\right) = \boldsymbol{X}\left(t + \sum_{j=1}^{l-1} a_j \Delta t\right) + a_l \Delta t \dot{\boldsymbol{X}}\left(t + \sum_{j=1}^{l} b_j \Delta t\right), \quad l = 1,2,\dots,L \tag{10b}$$

For an integrator with L intermediate integration steps, Eq. (10) represents the l^{th} update of the velocity and position where the overall time step Δt is divided into L segments. The coefficients a_l and b_l are specific to the integrator and are subject to the condition $\sum_{l=1}^{L} a_l = \sum_{l=1}^{L} b_l = 1$. Previously the use of friction-like dissipation when integrating the auxiliary equation of motion, Eq. (8b), was not well-defined in the context of higher-order integrators of the form of Eq. (10) due to the use of velocities as intermediates for higher order integrators.[13]

Instead we define a generalization of the friction-like dissipation terms of the modified Verlet integration scheme, Eq. (9), to the higher order integrator schemes of Eq. (10). As a general integration scheme this combination of friction-like dissipation and general order integration takes the form of Eq. (11). Eq. (11) introduces dissipation as a friction-like force term as the last term in Eq. (11a), thus only appearing in the velocity updates. The key insight is that L sets of K previous positions are stored, and during the calculations of friction terms of the l^{th} velocity update only the positions at integer multiples, k, of the full integration time step, Δt, from previous l^{th} integration steps are considered. Position updates stay the same and do no require any special consideration.

$$\begin{aligned}\dot{\boldsymbol{X}}\left(t+\sum_{j=1}^{l} b_j\Delta t\right) = \dot{\boldsymbol{X}}\left(t+\sum_{j=1}^{l-1} b_j\Delta t\right) + \ddot{\boldsymbol{X}}\left(t+\sum_{j=1}^{l-1} a_j\Delta t\right) \\ + \frac{b_l\alpha}{\Delta t}\sum_{k=0}^{K} c_k \boldsymbol{X}\left(t-k\Delta t+\sum_{j=1}^{l-1} a_j\Delta t\right)\end{aligned} \tag{11a}$$

$$\boldsymbol{X}\left(t+\sum_{j=1}^{l} a_j\Delta t\right) = \boldsymbol{X}\left(t+\sum_{j=1}^{l-1} a_j\Delta t\right) + a_l\Delta t\dot{\boldsymbol{X}}\left(t+\sum_{j=1}^{l} b_j\Delta t\right) \tag{11b}$$

Applying the general integration of Eq. (11) to the auxiliary equation of motion, Eq. (8b), then takes the following form:

$$\begin{aligned}\dot{\boldsymbol{X}}\left(t+\sum_{j=1}^{l} b_j\Delta t\right) = \dot{\boldsymbol{X}}\left(t+\sum_{j=1}^{l-1} b_j\Delta t\right) + \\ \frac{b_l\kappa}{\Delta t}\left[\boldsymbol{C}^{-1}\left[\boldsymbol{I}-\left[\boldsymbol{I}-\boldsymbol{C}\boldsymbol{X}\left(t+\sum_{j=1}^{l-1} a_j\Delta t\right)\right]^{m}\right]-\boldsymbol{X}\left(t+\sum_{j=1}^{l-1} a_j\Delta t\right)\right] \\ + \frac{b_l\alpha}{\Delta t}\sum_{k=0}^{K} c_k \boldsymbol{X}\left(t-k\Delta t+\sum_{j=1}^{l-1} a_j\Delta t\right)\end{aligned} \tag{12a}$$

$$\boldsymbol{X}\left(t+\sum_{j=1}^{l} a_j \Delta t\right) = \boldsymbol{X}\left(t+\sum_{j=1}^{l-1} a_j \Delta t\right) + a_l \Delta t \dot{\boldsymbol{X}}\left(t+\sum_{j=1}^{l} b_j \Delta t\right) \tag{12b}$$

Table 3 gives the optimal integration parameters a_l and b_l to minimize the error of the integration for a velocity Verlet method[30] and also several optimal higher-order methods as described by McLachlan and Atela[31].

Table 3: *Integration schemes of various order and their parameters*[31]. Note that the 5th-order optimal method requires a time step to be broken down into six velocity and position updates, $L = 6$. In nomenclature for the remainder of this paper $L = 2$, $L = 3$, $L = 4$, and $L = 6$ will refer to the 2nd-order, 3rd-order, 4th-order, and 5th-order optimal methods, respectively.

Name	$\boldsymbol{L}$	$\boldsymbol{a_l}$	$\boldsymbol{b_l}$
velocity Verlet	2	$a_1 = 1$ $a_2 = 0$	$b_1 = \frac{1}{2}$ $b_2 = \frac{1}{2}$
2nd –order optimal	2	$a_1 = \frac{1}{\sqrt{2}}$ $a_2 = 1 - a_1$	$b_1 = a_2$ $b_2 = a_1$
3rd –order optimal	3	$a_1 = 0.919661523017399857$ $a_2 = \frac{1}{4a_1} - \frac{a_1}{2}$ $a_3 = 1 - a_1 - a_2$	$b_1 = a_3$ $b_2 = a_2$ $b_3 = a_1$
4th –order optimal	4	$a_1 = 0.5153528374311229364$ $a_2 = -0.085782019412973646$ $a_3 = 0.4415830236164665242$ $a_4 = 0.1288461583653841854$	$b_1 = 0.1344961992774310892$ $b_2 = -0.2248190307942080 58$ $b_3 = 0.7563200005156682911$ $b_4 = 0.3340036032863214255$
5th –order optimal	6	$a_1 = 0.339839625839110000$ $a_2 = -0.088601336903027329$ $a_3 = 0.5858564768259621188$ $a_4 = -0.603039356536491888$ $a_5 = 0.3235807965546976394$ $a_6 = 0.4423637942197494587$	$b_1 = 0.1193900292875672758$ $b_2 = 0.6989273703824752308$ $b_3 = -0.1713123582716007754$ $b_4 = 0.4012695022513534480$ $b_5 = 0.0107050818482359840$ $b_6 = -0.0589796254980311632$

Combining Shadow Potentials, Dissipation, and Integration of Varying Order. Using the methods proposed in previous sections one can now construct a shadow potential to match the reference potential for any order, selecting from Table 1 and driving the equations of motion in Eq. (8). We can also select a dissipative scheme that will only break time reversibility up to some chosen order by selecting from Table 2. Finally one can choose an integration scheme that is correct to some order in Δt by integrating with Eq. (12) and choosing a method from Table 3. This sets forth a general framework for extended Lagrangian Born-Oppenheimer dynamics for classical polarization. While we have illustrated this formalism with a classical fluctuating charge model we also present the formalism adapted for an induced dipole model in Appendix A.

Table 4: *Optimal κ and α values for a combination of dissipative and integration orders, K and L, respectively.* The value of the dissipation for each combination, $[\rho(\mathbf{T})]_{\min}$, is also given.

Integrator Name	L	K	κ	α	$[\rho(T)]_{min}$
velocity Verlet	2	3	1.776	0.112	0.5785
	2	4	1.738	0.0655	0.8278
	2	5	1.752	0.0248	0.9084
	2	6	1.769	0.00825	0.9487
	2	7	1.790	0.00250	0.9708
	2	8	1.802	0.000750	0.9833
	2	9	1.818	0.000212	0.9906
2nd-order optimal	2	3	2.183	0.190	0.7315
	2	4	2.279	0.0712	0.7487
	2	5	2.271	0.0292	0.8597
	2	6	2.281	0.0101	0.9150
	2	7	2.295	0.00320	0.9469
	2	8	2.311	0.000958	0.9665
	2	9	2.327	0.000276	0.9787
3rd-order optimal	3	3	3.856	0.403	0.9178
	3	4	4.025	0.155	0.8727
	3	5	4.172	0.0475	0.8290
	3	6	4.312	0.0125	0.7792
	3	7	4.376	0.00349	0.7644
	3	8	4.350	0.00117	0.8326
	3	9	4.323	0.000382	0.8735
4th-order optimal	4	3	3.891	0.363	0.9154
	4	4	4.061	0.139	0.8704
	4	5	4.187	0.0430	0.8259
	4	6	4.292	0.0116	0.7731
	4	7	4.332	0.00340	0.7646
	4	8	4.298	0.00121	0.8296
	4	9	4.288	0.000384	0.8703
5th-order optimal	6	3	3.973	0.358	0.9206
	6	4	4.133	0.139	0.8781
	6	5	4.255	0.0434	0.8366
	6	6	4.361	0.0117	0.7895
	6	7	4.424	0.00316	0.7425
	6	8	4.385	0.00115	0.8167
	6	9	4.371	0.000371	0.8602

Optimal Parameters. The integration described by Eq. (11) is dependent on two key parameters not yet discussed, κ and α, which are given in Table 4. The parameter κ describes the frequency of the auxiliary harmonic and should be as high as possible to drive $\boldsymbol{X}$ to the ground state solution and to enforce an adiabatic decoupling to the nuclear motion. We want α to be as high as possible as well, to give the maximum possible amount of dissipation to stay close to the ground

state solution and to dampen resonance and numerical noise. With these conditions in mind, Appendix B examines what the optimal sets of κ and α parameters are (summarized in Table 4), which rounds out the necessary information to build a general XLBOMD method.

RESULTS

To give a clear illustration of our general approach let us choose the lowest order of shadow potential, dissipation, and integration by selecting $m = 2$ from Table 1, $K = 3$ from Table 2, and $L = 2$ (second-order optimal) from Table 3. The equations of motion become Eq. (13a) for the particles and Eq. (13b) for the auxiliary, $\boldsymbol{X}$.

$$m_i\ddot{\boldsymbol{r}}_i = -\frac{\partial U(\boldsymbol{r}^N)}{\partial \boldsymbol{r}_i} - \frac{1}{2}\boldsymbol{\chi}^T\boldsymbol{X}^T\frac{\partial \boldsymbol{C}}{\partial \boldsymbol{r}_i}\boldsymbol{X}\boldsymbol{\chi} \tag{13a}$$

$$\ddot{\boldsymbol{X}} = \omega^2(\boldsymbol{X} - \boldsymbol{X}\boldsymbol{C}\boldsymbol{X}) \tag{13b}$$

If we now integrate Eq. (13) with 2nd-order optimal integration and 3rd–order dissipation, a single full integration time step would look like

$$\begin{aligned}\dot{\boldsymbol{X}}(t + b_1\Delta t) = \dot{\boldsymbol{X}}(t) &+ \frac{b_1\kappa}{\Delta t}[\boldsymbol{X}(t) - \boldsymbol{X}(t)\boldsymbol{C}\boldsymbol{X}(t)] \\ &+ \frac{b_1\alpha}{\Delta t}[c_0\boldsymbol{X}(t) + c_1\boldsymbol{X}(t - \Delta t) + c_2\boldsymbol{X}(t - 2\Delta t) + c_3\boldsymbol{X}(t - 3\Delta t)]\end{aligned} \tag{14a}$$

$$\boldsymbol{X}(t + a_1\Delta t) = \boldsymbol{X}(t) + a_1\Delta t\dot{\boldsymbol{X}}(t + b_1\Delta t) \tag{14b}$$

$$\begin{aligned}\dot{\boldsymbol{X}}(t + \Delta t) = \dot{\boldsymbol{X}}(t + b_1\Delta t) &+ \frac{b_2\kappa}{\Delta t}[\boldsymbol{X}(t + a_1\Delta t) - \boldsymbol{X}(t + a_1\Delta t)\boldsymbol{C}\boldsymbol{X}(t + a_1\Delta t)] \\ &+ \frac{b_2\alpha}{\Delta t}[c_0\boldsymbol{X}(t + a_1\Delta t) + c_1\boldsymbol{X}(t + a_1\Delta t - \Delta t) + c_2\boldsymbol{X}(t + a_1\Delta t - 2\Delta t) \\ &+ c_3\boldsymbol{X}(t + a_1\Delta t - 3\Delta t)]\end{aligned} \tag{14c}$$

$$\boldsymbol{X}(t + \Delta t) = \boldsymbol{X}(t + a_1\Delta t) + a_2\Delta t\dot{\boldsymbol{X}}(t + \Delta t) \tag{14d}$$

For brevity the time dependence of the $\boldsymbol{C}$ matrices has been dropped, but these need to be updated with the positions, $\boldsymbol{r}^N$. The particle equation of motion should be integrated using the corresponding 2nd-order optimal integration scheme as the auxiliary $\boldsymbol{X}$, but without the dissipative force term. In this case the integration of the particles would therefore be given by Eq. (15), which would be interleaved with Eq. (14) at equal time intervals.

$$\dot{\boldsymbol{r}}_i(t + b_1\Delta t) = \dot{\boldsymbol{r}}_i(t) + \frac{b_1}{m_i\Delta t}\left[-\frac{\partial U\big(\boldsymbol{r}^N(t)\big)}{\partial \boldsymbol{r}_i(t)} - \frac{1}{2}\boldsymbol{\chi}^T\boldsymbol{X}^T(t)\frac{\partial \boldsymbol{C}}{\partial \boldsymbol{r}_i}\boldsymbol{X}(t)\boldsymbol{\chi}\right] \tag{15a}$$

$$\boldsymbol{r}_i(t + a_1\Delta t) = \boldsymbol{r}_i(t) + a_1\Delta t\dot{\boldsymbol{r}}_i(t + b_1\Delta t) \tag{15b}$$

$$\dot{\boldsymbol{r}}_i(t+\Delta t) = \dot{\boldsymbol{r}}_i(t+a_1\Delta t) + \frac{b_2}{m_i \Delta t}\left[-\frac{\partial U\left(\boldsymbol{r}^N(t+a_1\Delta t)\right)}{\partial \boldsymbol{r}_i(t+a_1\Delta t)} - \frac{1}{2}\boldsymbol{\chi}^T \boldsymbol{X}^T(t+a_1\Delta t)\frac{\partial \boldsymbol{C}}{\partial \boldsymbol{r}_i}\boldsymbol{X}(t+a_1\Delta t)\boldsymbol{\chi}\right] \quad (15c)$$

$$\boldsymbol{r}_i(t+\Delta t) = \boldsymbol{r}_i(t+a_1\Delta t) + a_2 \Delta t \dot{\boldsymbol{r}}_i(t+\Delta t) \quad (15d)$$

We next test the theory presented above using a dimensionless charge equilibration model, involving three particles with electrostatic interactions described by Eq. (1) along with a harmonic restraining potential for each particle, which constitutes the entirety of the system. For further simplicity the system was confined to a single dimension. This basic system allows for efficient testing of the potentials, integrators, and dissipation schemes. Furthermore, the simple charge equilibration model does not exhibit a well-defined statistical temperature, and if successful, the higher-order dissipative integration schemes should therefore be applicable also to first-principles Born-Oppenheimer molecular dynamics[24].

Dissipation. To illustrate why the XLBOMD schemes require dissipation, we introduce a perturbation after 1000 time steps into the simulation, by swapping the auxiliary matrix, $\boldsymbol{X}$, with its value with the value from the three previous time steps. This modeled perturbation simulates a spike in numerical noise or a resonance instability one may encounter in a more complex system. Figure 1 shows the scaled energy deviation from the initial system energy over the course of a simulation; for good energy conservation we expect this quantity to stay close to 0 with little drift. We can see that this perturbation is quickly corrected when we use a dissipative scheme (red curve), however, when no dissipation is used the auxiliary variables have no way to remove the momentary disturbance and their equation of motion quickly becomes unstable.

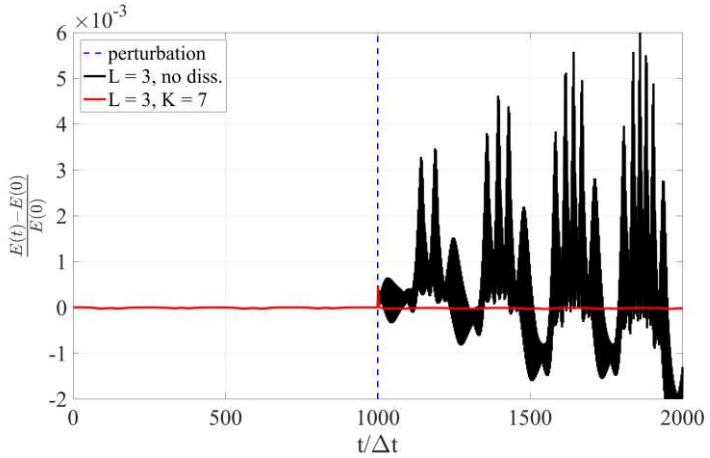

Figure 1: *The response of the scaled energy deviation to a perturbation for a dissipative (red) and non-dissipative (black) integration scheme.* Both schemes used a 3rd-order optimal integrator (L = 3) and a 7th-order dissipative scheme (K = 7). The introduced perturbation (dotted blue line) is a swap of the **X** matrix at the 1000th time step with its previous value at the 997th time step.

One downside to the friction-like dissipative schemes described is that they break time reversibility, which can lead to energy drift. So while dissipation is necessary to account for numerical noise or resonance effects if we use a dissipative scheme as described in this text we can achieve low energy drift rates by using higher order dissipation, where the dissipative c_k ($k = 0, 1, \ldots, K$) coefficients are designed to only break time reversibility up to some order K in the time step[13]. In Figure 2a we give an example of the energy drift over the course of a trajectory for a specific integration order and shadow potential order; we can see that the energy drift decreases as the dissipative order increases. This trend is then replicated for many combinations of integration order and shadow potential order in Figure 2b, where we report the fitted energy drift rates as a function of dissipative order.

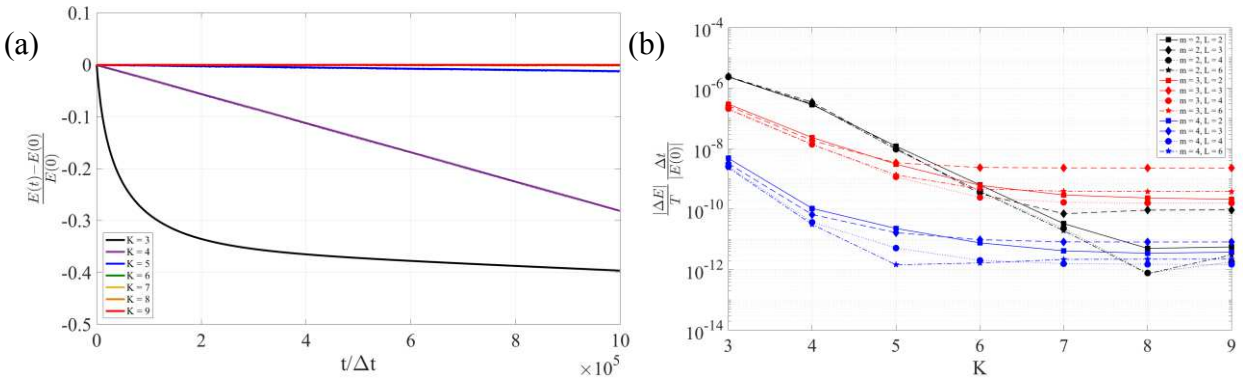

Figure 2: *Energy conservation properties as a function of dissipation.* (a) The scaled energy deviation from the total initial energy over trajectories using the second order optimal integration method ($\mathrm{L} = 2$) and various dissipative orders, K. (b) The fitted energy drifts as a function of dissipative order, K, for various combinations of integration order, L, and shadow potential order, m. In (b) line color denotes shadow potential order and line and symbol shape denote integration order. The time step has been increased by a factor of 5 to clearly demonstrate the effect of dissipative order on energy drift. The energy drift is calculated by taking a linear fit of the energy over total simulation time, $\Delta \mathrm{E}/\mathrm{T}$, and non-dimensionalized by the time step and initial energy.

We note that the dissipation schemes presented here are not the only way to introduce dissipation into the auxiliary integration. One can also define an auxiliary temperature and use a thermostat to control that temperature, essentially ensuring that the inertia of the auxiliaries does not diverge[21-23]. While this method works well for large systems where the temperature better manifests as a macroscopic quantity, we have previously shown that it does not perform well if system sizes are too small[24], where the temperature becomes quite variable and hard to control, such as in our test system. However, for larger systems the amount of information that needs to be stored increases as $O(N^2)$, as $\boldsymbol{X} \sim N^2$, and the highest order dissipative methods required for excellent energy conservation can become very memory intensive. Hence for larger systems, auxiliary temperature

control may be the more attractive option, since it does not require storing any history and relies only on information in the current time step.

Integration. Using the formulation we developed in Eq. (11) we can now use general order dissipative methods with general order integrators. While a higher order integration scheme requires more force evaluations per time step (which can be expensive), the benefit that is derived is a more stable simulation and more accurately calculated properties, such as energy shown in Figure 3 for a second order ($m = 2$) shadow potential. In Figure 3a and 3b we see that for the same time step, and more force evaluations, the higher order integrators give a lower deviation in the energy along a trajectory. If instead we maintain a fixed cost (i.e. higher order integrators use larger time steps so that the number of force evaluations per total simulated time is fixed) then the higher the order the integrator, the greater the deviation that is seen for a second order shadow potential.

Figure 3c shows the interplay of integration order with the order of the shadow potential. For this particular system the benefits of going to fourth ($L = 4$) and fifth ($L = 6$) order optimal integration are marginal when using just a second order shadow potential, but we can achieve ever greater accuracy by increasing the integration order with higher order shadow potentials. Whereas for a second order shadow potential the energy deviation became worse with increasing integration order for fixed cost, by using third and fourth order shadow potentials now a fixed cost simulation can yield a greater accuracy by using a higher integrator order. This demonstrates that the more accurate integration is able to reveal the adequacy of the approximation of the shadow potential to the true solution.

As Figures 3d and 3e show, the difference between the potential energies and forces with respect to an exactly converged reference potential scale as Δt^{2m} and $\Delta t^{2(m-1)}$, respectively. This means that higher order shadow potentials will scale more rapidly in their deviation from the reference potential as the time step increases and will therefore require more accurate (higher order) integration to realize this greater agreement with the reference potential.

Shadow Potentials. Since the potential is variationally optimized at each time step by construction with Eq. (5b), with no explicit matrix inversion, we are at the ground state of the shadow potential at every time step. Similar to integrator order, in the limit of high shadow potential order we obtain the true self-consistent solution to the reference potential, which is the highest possible level of accuracy one could obtain in terms of potential energy convergence.

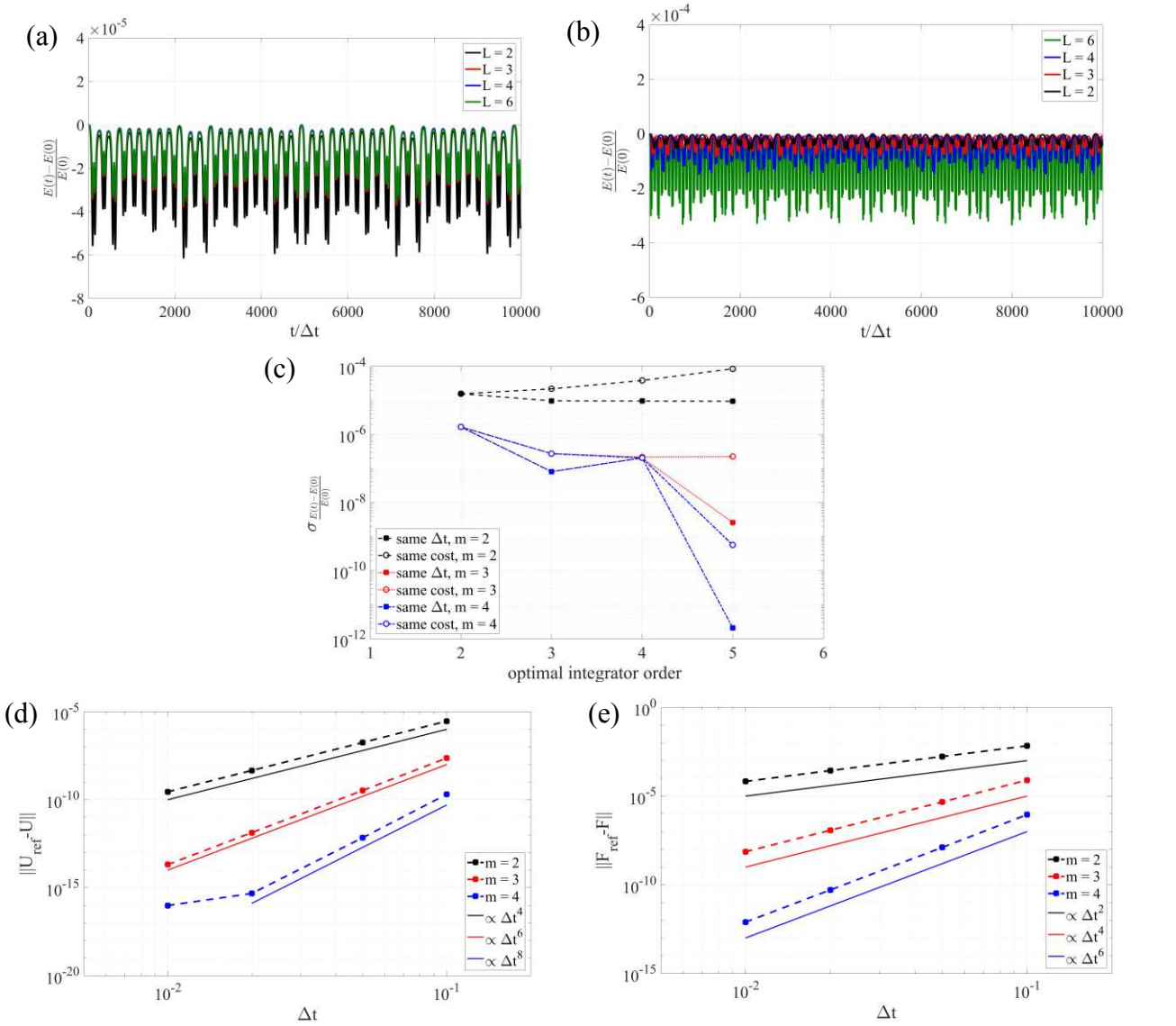

Figure 3: *Scaled energy deviation and standard deviation over the course of a trajectory using fifth order dissipation ($K = 5$) with a second order shadow potential ($m = 2$) for various integration orders (L) and scaling of errors in potential energy and forces with time step*. (a) For each order of the integrator the same time step is used, but differ in the number of force evaluations used. (b) The time steps are adjusted so that each order of integrator uses the same number of force evaluations per time so that the higher order the integrator the larger the time step. (c) Standard deviation using fixed time step or fixed number of force evaluations (fixed cost) over the course of a trajectory for different orders of the shadow potential for various optimal integration orders. The color and line style denotes the shadow potential order (black dashed is second order, m = 2; red dotted is third order, m = 3; blue red/dotted is fourth order, m = 4) and symbol denotes same time step (solid square) or same cost (hollow circle). The difference (L2 norm) between the potential energy (d) and forces (e) compared to the exactly converged reference potential as a function of time step and for various orders of the shadow potential. Fitted dependencies on time step, Δt, are plotted as solid lines for comparison. Note that the limit of machine precision here is $\sim 10^{-16}$, which is why the $m = 4$ curve in (e) does not scale to $\Delta t = 0.01$.

Figure 4a gives an example of this, where we see the energy trajectory over the course of a simulation for a specific value of integration and dissipation order, but with varying orders of the shadow potential. Figure 4b shows that we can reach complete convergence of the simple fluctuating charge model with just a third order shadow potential for integration schemes up to fourth order ($L = 4$). For a fifth order integration ($L = 6$) we need to use a fourth order shadow potential ($m = 4$) to achieve a level of convergence akin to the fully converged reference potential. Therefore, we can choose exactly how closely we need to match the reference potential for a given application. In practice, the necessary order of the shadow potential could be dictated by time scale of decay of the true ground state solution. If the electronic degrees of freedom decay on a time scale much longer than the time step, as we have shown for classical polarization models[21], then the shadow potential may not need to be of a high order to accurately follow the dynamics.

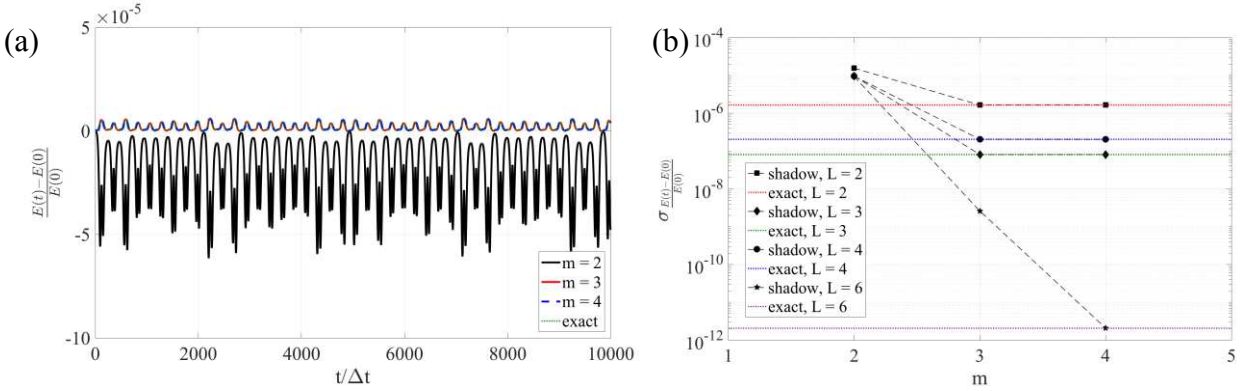

Figure 4: (a) *Scaled energy deviation and (b) standard deviation over the course of a trajectory using fifth order dissipation ($K = 5$) with a second order optimal integration ($L = 2$) for various shadow potential orders (m).* In both plots the 'exact' dynamics are given as a point of comparison where the **C** matrix is inverted directly and used with the reference potential, Eq. (1) to drive the dynamics for a given integrator order.

Figure 5 gives a succinct summary of the interplay between the shadow potential order, order of the integration scheme, and order of dissipation in terms of the energy deviation from the initial energy over the course of each type of simulation. For the simulation that uses a low dissipative order, low integration order, and low shadow potential order, the energy exhibits large fluctuations and energy drift. Upon increasing the dissipative order, the energy drift has been corrected since the time irreversibility of the dissipation is pushed out to a higher order in the time step. By increasing the integration order the underlying shadow potential is more accurately integrated and the energy deviations decrease. Finally, increasing the order of the shadow potential gives a better approximation to the underlying fully converged reference potential and the deviations decrease even further.

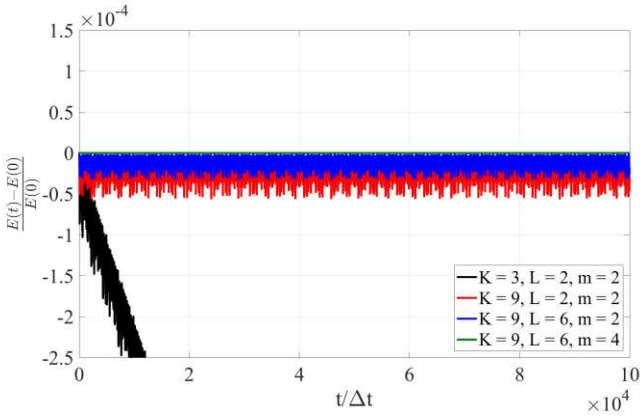

Figure 5: *Scaled energy deviations for selected combinations of dissipation order, K; integration order, L; and shadow potential order, m.* The combinations are low dissipation, low integration, low shadow potential (black); high dissipation, low integration, low shadow potential (red); high dissipation, high integration, low potential (blue); and high dissipation, high integration, and high shadow potential (green).

DISCUSSION

The methods discussed here present a generalized starting point from which to work with the XLBOMD methods. Here we show that the iEL/0-SCF method[22-23] formally fits within the general framework, and furthermore that it leads to a reduction in cost by replacing the matrix-matrix with a matrix-vector calculation.

Taking the case of a 2nd-order potential ($m = 2$) and ignoring dissipation for a moment, the equations of motion are given in Eq. (16)

$$m_i \ddot{\boldsymbol{r}}_i = -\frac{\partial U(\boldsymbol{r}^N)}{\partial \boldsymbol{r}_i} - \frac{1}{2}\boldsymbol{\chi}^T \boldsymbol{X}^T \frac{\partial \boldsymbol{C}}{\partial \boldsymbol{r}_i}\boldsymbol{X}\boldsymbol{\chi} \tag{16a}$$

$$\ddot{\boldsymbol{X}} = \omega^2(2\boldsymbol{X} - \boldsymbol{X}\boldsymbol{C}\boldsymbol{X} - \boldsymbol{X}) \tag{16b}$$

From a computational efficiency perspective there is no great expense in the particle equation of motion, Eq. (16a), as we only ever have to perform matrix-vector multiplications which are $O(N)$. This can easily be seen as $\boldsymbol{X}\boldsymbol{\chi}$ and $\boldsymbol{\chi}^T\boldsymbol{X}^T$ are matrix-vector operations resulting in vectors and $[\boldsymbol{\chi}^T\boldsymbol{X}^T]\frac{\partial \boldsymbol{C}}{\partial \boldsymbol{r}_i}[\boldsymbol{X}\boldsymbol{\chi}]$ would become a matrix-vector operation, as well. The auxiliary equation of motion, Eq. (16b), however, has a matrix-matrix-matrix multiplication, $\boldsymbol{X}\boldsymbol{C}\boldsymbol{X}$, which is $O(N^3)$ and would scale poorly with system size; for larger or condensed phase systems, $\widetilde{\boldsymbol{C}}^{-1}$ could instead be approximated by using real space cutoffs and leveraging sparse linear algebra techniques for the multiplication in order to improve its scaling.

One practical remedy that is used in the iEL/0-SCF method is to perform a change of variables and dynamically integrate $\boldsymbol{a} = -\boldsymbol{X}\boldsymbol{\chi}$ instead of $\boldsymbol{X}$. In this case, the dynamical variables

correspond to the matrix operator acting on a vector, as opposed to the matrix operator itself. This substitution is also convenient for 'bulk' simulations where periodic boundary conditions or Ewald summations are used and explicitly calculating the interaction matrix operator itself is difficult. Making this substitution the equations of motion now become

$$m_i \ddot{\boldsymbol{r}}_i = -\frac{\partial U(\boldsymbol{r}^N)}{\partial \boldsymbol{r}_i} - \frac{1}{2}\boldsymbol{a}^T \frac{\partial \boldsymbol{C}}{\partial \boldsymbol{r}_i}\boldsymbol{a} \tag{17a}$$

$$\ddot{\boldsymbol{a}} = \omega^2(\boldsymbol{q}[\boldsymbol{a}] - \boldsymbol{a}) \tag{17b}$$

In Eq. (17b) we have made use of the identity in Eq. (5b), $\boldsymbol{q} = -f(\boldsymbol{X})\boldsymbol{\chi}$. The new auxiliary, $\boldsymbol{a}$, can be thought of as an auxiliary variable that will stay close to the ground state fluctuating charges, $\boldsymbol{q}$, and $\boldsymbol{q}[\boldsymbol{a}]$ then becomes an approximation of the ground state charges. Furthermore, $\boldsymbol{a}$ is a vector and we avoid any costly matrix-matrix multiplication. However this substitution trick is only applicable to the special case of m=2, since for $m > 2$, $\boldsymbol{X}$ is not always multiplied into $\boldsymbol{\chi}$ like it is in particle equation of motion, Eq. (16a).

This result is very similar to previously described methods[16-17, 22-23], which we now see as a specific type of approximation to the general formalism outlined here. For example, to recover the previously described iEL/0-SCF method[22] the specific form of $\boldsymbol{q}[\boldsymbol{a}]$ used is described by Eq. (18a), which is a single iteration using $\boldsymbol{a}$ as an initial guess. A further refinement for the auxiliary equation of motion, Eq. (18b) is introduced as a corrector-like step with a tunable correction parameter, γ, which can range from 0 to 1.

$$\boldsymbol{q}[\boldsymbol{a}] = -\boldsymbol{D}^{-1}\boldsymbol{\chi} - \boldsymbol{D}^{-1}(\boldsymbol{L} + \boldsymbol{U})\boldsymbol{a} \tag{18a}$$

$$\boldsymbol{q}_{SCF}[\boldsymbol{a}] \approx \gamma \boldsymbol{q}[\boldsymbol{a}] - (1-\gamma)\boldsymbol{a} \tag{18b}$$

In Eq. (18) $\boldsymbol{D}$, $\boldsymbol{L}$, and $\boldsymbol{U}$ are the diagonal, lower triangular, and upper triangular components of $\boldsymbol{C}$, respectively. Since the diagonal components, $\boldsymbol{D}$, are simply fixed parameters (the atomic hardness) the inverse $\boldsymbol{D}^{-1}$ is straightforward and need only be calculated once at the beginning of a simulation. One further refinement of the iEL/0-SCF method is that $\boldsymbol{q}[\boldsymbol{a}]$ is used in the particle equation of motion, as opposed to $\boldsymbol{a}$. Since $\boldsymbol{q}[\boldsymbol{a}]$ is now dependent on particle positions, whereas $\boldsymbol{a}$ is not as it is an independent dynamic degree of freedom, we must account for this additional position dependence in the particle equation of motion. Combining all of these refinements we recover the iEL/0-SCF equations of motion, Eq. (19).

$$m_i \ddot{\boldsymbol{r}}_i = -\frac{\partial U(\boldsymbol{r}^N)}{\partial \boldsymbol{r}_i} - \frac{1}{2}\boldsymbol{q}^T[\boldsymbol{a}]\frac{\partial \boldsymbol{C}}{\partial \boldsymbol{r}_i}\boldsymbol{q}[\boldsymbol{a}] + (\boldsymbol{\chi}^T + \boldsymbol{q}^T[\boldsymbol{a}]\boldsymbol{C})\boldsymbol{D}^{-1}\frac{\partial \boldsymbol{C}}{\partial \boldsymbol{r}_i}\boldsymbol{a} \tag{19a}$$

$$\ddot{\boldsymbol{a}} = \gamma\omega^2(\boldsymbol{q}[\boldsymbol{a}] - \boldsymbol{a}) \tag{19b}$$

The resulting iEL/0-SCF method is a specific example of how one might use this general XLBOMD formalism. The formal method establishes the relevant criteria, and although the formal methods can be computationally expensive, they can be built upon to introduce approximations to reduce cost. While the general method of Eq. (16) involves expensive, poorly scaling matrix-matrix multiplication, subsequent approximations that led to Eq. (19) gives the iEL/0-SCF method, which has already been proven effective for large, condensed phase systems using point dipoles and Drude polarization.[22-23] It is interesting to note that the dissipation approach described here was not used for the iEL/0-SCF method, since even for the higher order dissipation method it does not perform as well as time-reversible auxiliary temperature control to the auxiliary equation of motion.[21] While the auxiliary thermostating scheme is arguably the best choice for long time scale simulations of systems with many degrees of freedom, for small test systems or small *ab initio* systems the dissipative schemes may be suitable.

CONCLUSIONS

We have presented a general and flexible framework on which to build XLBOMD methods for treating models that require self-consistent optimization at each time step. This general framework combines increasing orders of shadow potentials that are designed to systematically improve agreement with the reference potential, increasing orders of dissipation to correct for unwanted numerical noise or resonance effects, and higher order integrators to provide greater accuracy in the simulation of properties of a given potential, such as energy conservation. When used together we obtain equations of motion for an auxiliary matrix, $\boldsymbol{X}$, a dynamically driven approximation to the inverse interaction operator that would normally solve the true system exactly. Within this general framework, then, one can choose a combination of dissipation, integration, and shadow potential suitable for a given application and in terms of what is acceptable for energy drift and accuracy. In summary, greater dissipation order will lead to less energy drift and greater integration and shadow potential order will lead to better accuracy. We illustrated our results with a small, simple charge equilibration system.

In addition, we now have a better understanding as to why the previously proposed iEL/0-SCF method works as well as it does for larger condensed phase systems by casting it within this general framework. First is that by proposing an alternative auxiliary integration variable, only applicable to low orders of the shadow potential ($m = 2$), we can avoid the matrix-matrix multiplication by integrating a vector auxiliary quantity instead of a matrix, which ultimately recovered the iEL/0-SCF equations of motion. Furthermore, while the cost of the dissipation

schemes presented here may also prove to be too expensive for larger systems due to the necessity of storing matrices that scale as $O(N^2)$ from previous time steps, the use of auxiliary thermostats[21] is largely analogous to implementation of a very high order dissipation scheme.

One final outstanding problem is how to leverage the method of shadow potentials for *ab initio* methods. For the classical models presented here, the fluctuating charge in the Theory section and induced dipole in Appendix A, one needs to solve a linear equation to obtain the ground state electronic degrees of freedom (Eqs. (2) and (A2)). For an *ab initio* method one needs to solve a non-linear optimization. Determining how to leverage a general order shadow potential for such a model requiring non-linear optimization is a point of ongoing work[17], however, using it successfully for a classical model in this work is an excellent proof of concept.

ACKNOWLEDGMENTS. AA and THG thank the National Science Foundation Grant No. CHE-1665315 for support of this work. A.M.N.N. acknowledges support from the Department of Energy Offices of Basic Energy Sciences (Grant No. LANL2014E8AN).

Appendix A: Shadow Potentials for Induced Dipole Polarization

While the dissipation and integrations schemes laid out in the theory section are generally applicable, the shadow potentials were illustrated with a classical fluctuating charge model. One could, in principle, adapt the discussion of higher order shadow potentials to a range of models and simulation techniques. Here we present the development of higher order shadow potentials for a classical dipole polarization model. For the most part the details mirror those already laid out in the theory section so we will focus on brevity here.

Dipole Polarization. The potential energy surface for an inducible dipole model is given by the constrained minimization of Eq. (A1).

$$U^{pol} = \frac{1}{2}\boldsymbol{\mu}^{\boldsymbol{T}}\boldsymbol{C}\boldsymbol{\mu} - \boldsymbol{E}^{T}\boldsymbol{\mu} \tag{A1}$$

In Eq. (A1), $\boldsymbol{\mu}$ represents a set of N inducible dipoles on the N atoms of the system. $\boldsymbol{E}$ is the permanent electrostatic field created by any fixed electric multipoles in the system, and $\boldsymbol{C} = \boldsymbol{C}(\boldsymbol{r}^N) = \boldsymbol{\alpha}^{-1} - \boldsymbol{T}$ where $\boldsymbol{\alpha}$ is a diagonal matrix with the values of atomic polarizability for each atom, α_i, on the diagonal and $\boldsymbol{T} = \boldsymbol{T}(\boldsymbol{r}^N)$ is the dipole-dipole interaction matrix. The ground state (self-consistent field) solution of Eq. (A1) is the set of induced dipoles that minimizes the polarization energy, which is given in Eq. (A2).

$$\frac{\partial U^{pol}}{\partial \boldsymbol{\mu}} = 0 = -\boldsymbol{E} + \boldsymbol{C\mu} \quad \Longrightarrow \quad \boldsymbol{\mu}_{min} = -\boldsymbol{C^{-1}E} \tag{A2}$$

As before we can now introduce a dynamically driven auxiliary matrix, $\boldsymbol{X}$, that should follow the behavior of $\boldsymbol{C^{-1}}$ and build a shadow potential functional around this auxiliary matrix as shown in Eq. (A3).

$$U^{pol}_{shadow}(\boldsymbol{r}^N, \boldsymbol{X}, \boldsymbol{\mu}) = \frac{1}{2}\boldsymbol{\mu^T}[f(\boldsymbol{X})]^{-1}\boldsymbol{\mu} - \boldsymbol{E}^T\boldsymbol{\mu} \tag{A3}$$

We make the same choice of $f(\boldsymbol{X})$ as before, and reproduced in Eq. (A4). This form, again, will produce a better estimate of $\boldsymbol{C^{-1}}$ from $\boldsymbol{X}$.

$$f(\boldsymbol{X}) = \boldsymbol{C^{-1}}[\boldsymbol{I} - (\boldsymbol{I} - \boldsymbol{CX})^m] \tag{A4}$$

We can also show that the shadow potential functional we have constructed in Eq. (A3) is necessarily minimized for the set of induced dipoles given in Eq. (A5), which is simply the function of Eq. (A4) dotted into the permanent electrostatic field.

$$\begin{aligned} \frac{\partial U^{pol}_{shadow}(\boldsymbol{r}^N, \boldsymbol{X}, \boldsymbol{\mu})}{\partial \boldsymbol{\mu}} = 0 = \left[\boldsymbol{C^{-1}}[\boldsymbol{I} - (\boldsymbol{I} - \boldsymbol{CX})^m]\right]^{-1}\boldsymbol{\mu} - \boldsymbol{E} \quad \Longrightarrow \\ \boldsymbol{\mu}_{min} = \boldsymbol{C^{-1}}[\boldsymbol{I} - (\boldsymbol{I} - \boldsymbol{CX})^m]\boldsymbol{E} = f(\boldsymbol{X})\boldsymbol{E} \end{aligned} \tag{A5}$$

Making the substitution of Eq. (A5) into Eq. (A3) we can now build a Lagrangian with this shadow potential and $\boldsymbol{X}$ as an additional dynamical degree of freedom, shown in Eq. (A6).

$$\begin{aligned} &\mathcal{L}\left(\boldsymbol{r}^N, \dot{\boldsymbol{r}}^N, \boldsymbol{X}, \dot{\boldsymbol{X}}\right) \\ &\quad = \frac{1}{2}\sum_{i=1}^{N} m_i \dot{\boldsymbol{r}}_i^2 + \frac{1}{2}m_X \mathrm{tr}\left[\dot{\boldsymbol{X}}^2\right] - U(\boldsymbol{r}^N) - \frac{1}{2}\boldsymbol{E^T} f(\boldsymbol{X})\boldsymbol{E} \\ &\quad - \frac{1}{2}m_X\omega^2 \mathrm{tr}\left[\left(\widetilde{\boldsymbol{C}}^{-1} - \boldsymbol{X}\right)^2\right] \end{aligned} \tag{A6}$$

In Eq. (A6), $U(\boldsymbol{r}^N)$ gives components of the potential that are independent of the induced dipoles, $\boldsymbol{r}$ and $\dot{\boldsymbol{r}}$ are the positions and velocities of the atoms, and m_X is introduced as a fictitious mass for the $\boldsymbol{X}$ degree of freedom. $\widetilde{\boldsymbol{C}}^{-1}$ represents some approximation to $\boldsymbol{C}^{-1}$ that is valid as long as $\widetilde{\boldsymbol{C}}^{-1}$ is closer to $\boldsymbol{C}^{-1}$ than $\boldsymbol{X}$. Applying the Euler-Lagrange equation to Eq. (A6) we obtain Eq. (A7).

$$m_i\ddot{\boldsymbol{r}}_i = -\frac{\partial U(\boldsymbol{r}^N)}{\partial \boldsymbol{r}_i} + \frac{1}{2}\boldsymbol{E}^T\frac{\partial f(\boldsymbol{X})}{\partial \boldsymbol{r}_i}\boldsymbol{E} - f(\boldsymbol{X})\frac{\partial \boldsymbol{E}}{\partial \boldsymbol{r}_i} - m_X\omega^2(\widetilde{\boldsymbol{C}}^{-1} - \boldsymbol{X})\frac{\partial \widetilde{\boldsymbol{C}}^{-1}}{\partial \boldsymbol{r}_i}$$	(A7a)
$$m_X\ddot{\boldsymbol{X}} = -\frac{1}{2}\frac{\partial}{\partial \boldsymbol{X}}[\boldsymbol{E}^T f(\boldsymbol{X})\boldsymbol{E}] + m_X\omega^2\left(\widetilde{\boldsymbol{C}}^{-1} - \boldsymbol{X}\right) - m_X\omega^2(\widetilde{\boldsymbol{C}}^{-1} - \boldsymbol{X})\frac{\partial \widetilde{\boldsymbol{C}}^{-1}}{\partial \boldsymbol{X}}$$	(A7b)

From Eq. (7) we assume a classical adiabatic separation between particle and auxiliary motion where $\frac{\partial}{\partial \boldsymbol{X}}[\boldsymbol{E}^T f(\boldsymbol{X})\boldsymbol{E}]$ decays as ω^{-2} or faster. Then taking the limit that $m_X \to 0$ and $\omega \to \infty$ the resulting equations of motion are given by Eq. (A8), where we have also evaluated the derivate $\frac{\partial f(\boldsymbol{X})}{\partial \boldsymbol{r}_i}$ and substituted for $f(\boldsymbol{X})$.

$$m_i\ddot{\boldsymbol{r}}_i = -\frac{\partial U(\boldsymbol{r}^N)}{\partial \boldsymbol{r}_i} + \frac{1}{2}\boldsymbol{E}^T\boldsymbol{C}^{-1}\left[\frac{\partial \boldsymbol{C}}{\partial \boldsymbol{r}_i}\boldsymbol{C}^{-1}(\boldsymbol{I}-(\boldsymbol{I}-\boldsymbol{C}\boldsymbol{X})^m) + \sum_{j=0}^{m-1}(\boldsymbol{I}-\boldsymbol{C}\boldsymbol{X})^j\frac{\partial \boldsymbol{C}}{\partial \boldsymbol{r}_i}(\boldsymbol{I}-\boldsymbol{C}\boldsymbol{X})^{m-j-1}\right]\boldsymbol{E}$$	(A8a)
$$\ddot{\boldsymbol{X}} = \omega^2\left(\widetilde{\boldsymbol{C}}^{-1}-\boldsymbol{X}\right) = \omega^2(\boldsymbol{C}^{-1}[\boldsymbol{I}-(\boldsymbol{I}-\boldsymbol{C}\boldsymbol{X})^m]-\boldsymbol{X})$$	(A8b)

We now choose $\widetilde{\boldsymbol{C}}^{-1} = f(\boldsymbol{X})$, as shown in the second right hand side of Eq. (A8b), which guarantees the auxiliary $\boldsymbol{X}$ oscillates about an approximation to $\boldsymbol{C}^{-1}$ that is better than itself. The equations of motion in Eq. (A8) can be integrated and dissipation can be introduced as laid out in the subsequent Theory sections, as before. One may be concerned that $\boldsymbol{C}^{-1}$ appears explicitly in Eq. (A7), however, for some finite m the expansion in m and subsequent algebraic simplification removes any $\boldsymbol{C}^{-1}$ terms. We now have a shadow potential for dipole polarization that can match the base potential to an arbitrary degree, yet is guaranteed to be minimized via Eq. (A5) at any point with no iteration.

Appendix B. Optimization of κ and α parameters

Under the requirement that $\widetilde{\boldsymbol{C}}^{-1}$ gives a better approximation to $\boldsymbol{C}^{-1}$ than $\boldsymbol{X}$ and that $\boldsymbol{X}$ is close to $\boldsymbol{C}^{-1}$ we can describe $\widetilde{\boldsymbol{C}}^{-1}$ to be given through an approximate linearized optimization process $\boldsymbol{\Gamma}$ acting on $\boldsymbol{X}$, $\widetilde{\boldsymbol{C}}^{-1} = \boldsymbol{\Gamma}\boldsymbol{X}$[11-12]. For a convergent optimization we can expect the eigenvalues of $\boldsymbol{\Gamma}$ to be $|\gamma| < 1$. By replacing $\boldsymbol{\Gamma}$ by its maximum eigenvalue, γ, we can examine the stability of the integration under incomplete convergence to the ground state reference potential. The equation of motion in Eq. (8b) of the text then becomes Eq. (B1)

$$\ddot{\boldsymbol{X}}(t) = \omega^2(\gamma-1)\boldsymbol{X}(t) \tag{B1}$$

And we can examine its behavior for a range of convergences ($\gamma \in [-1,1]$ with $\gamma = 0$ corresponding to complete convergence, $\boldsymbol{C}^{-1} = \boldsymbol{X}$) and for different integration methods.

We can now integrate the equation of motion, Eq. (B1), using our combined higher order dissipation and integration schemes, given by Eq. (11) in the text. This integration can be described as a mapping of $\boldsymbol{X}$ and its velocity $\dot{\boldsymbol{X}}$ at one time step to the next. This mapping is given by Eq. (B2)

$$\begin{bmatrix} \dot{\boldsymbol{X}}(t+\Delta t) \\ \boldsymbol{X}(t+\Delta t) \\ \boldsymbol{X}(t) \\ \vdots \\ \boldsymbol{X}(t-(K+1)\Delta t) \end{bmatrix} = \boldsymbol{T} \begin{bmatrix} \dot{\boldsymbol{X}}(t) \\ \boldsymbol{X}(t) \\ \boldsymbol{X}(t-\Delta t) \\ \vdots \\ \boldsymbol{X}(t-K\Delta t) \end{bmatrix} \tag{B2}$$

where $\boldsymbol{T}$ is a matrix that describes the mapping (integration) process and is a function of a_l, b_l, κ, α, and γ (see Supplementary Information for more specifics). In Eq. (B2) we make the implicit variable substitution $\dot{\boldsymbol{X}} \rightarrow \Delta t\dot{\boldsymbol{X}}$ so that all elements of the $\boldsymbol{T}$ matrix are dimensionless. If the integration is to be stable then the maximum absolute eigenvalue of $\boldsymbol{T}$ (its spectral radius) must be no greater than 1.0, otherwise the mapping corresponds to an exponential increase that diverges into instability. Spectral radii less than 1.0, on the other hand, represent dissipation in the integration. With this metric we can determine the optimal values of κ and α for a given choice of L and K that will maintain integration stability while maximizing dissipation.

Figure B1 shows the analysis of the $\boldsymbol{T}$ matrix for the specific case of $K = 3$ and $L = 2$ (2^{nd}-order optimal). Discrete values of α are chosen and tested over a range of κ values. For each set of values of α and κ the range of γ from -1 to 1 is tested and the maximum spectral radius of $\boldsymbol{T}$ from that range of γ is recorded, which is shown in Figure B1a. For each value of α we can determine the maximum possible κ value as the value at which the spectral radius becomes greater than 1.0. These values, then, define the curve of α_{max} as a function of κ given in Figure B1c. Using this set of κ and corresponding α_{max} values we also look at their specific spectral radius behavior as a function of γ, shown in Figure B1b. The lower the spectral radius is the more dissipation that particular combination of α_{max} and κ achieves. We define the maximum dissipation as the spectral radius at $\gamma = 0$, which we call $[\rho(\boldsymbol{T})]_{min}$ which is plotted as a function of κ in Figure B1c, as well.

While we do want to maximize κ we also want to maximize the dissipation, $[\rho(\boldsymbol{T})]_{min}$ (lower values of $[\rho(\boldsymbol{T})]_{min}$ correspond to more dissipation). Figure B1c shows that the maximum possible κ occurs when there is no dissipation ($\alpha_{max} = 0$, $[\rho(\boldsymbol{T})]_{min} = 1$), which is not optimal. Fortunately, $[\rho(\boldsymbol{T})]_{min}$ has a local minima near the maximum possible value of κ. We choose this point to be the optimal combination of large κ and maximum dissipation. In this specific case we find that for $K = 3$ and $L = 2$ (2^{nd}-order optimal) the optimal values for κ and α are 2.183 and 0.190, respectively, with $[\rho(\boldsymbol{T})]_{min} = 0.7315$. An analysis like that described above and shown in

Figure B1 is repeated for a wide combination of K and L values and the optimal κ and α for each is given in Table 4.

(a)

(b)

(c)

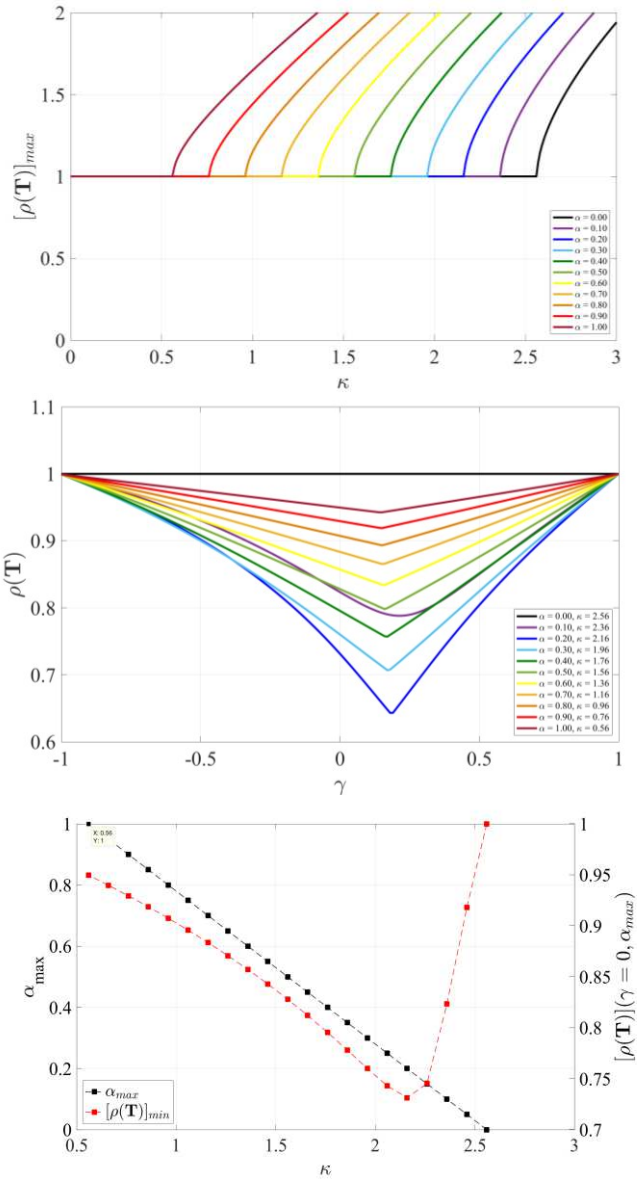

Figure B1: *Stability analysis of the integration schemes.* Analysis of the $\boldsymbol{T}$ matrix (Eq. 17) for the specific case of $K = 3$ and $L = 2$. (a) The maximum spectral radius of $\mathbf{T}$, $[\rho(\mathbf{T})]_{\max}$, as a function of κ for a range of α values. (b) The point at which $[\rho(\mathbf{T})]_{\max}$ becomes greater than 1.0 defines a pair of κ and $\alpha_{\max}$ and for those sets of points we plot the spectral radius, $\rho(\mathbf{T})$, as a function of γ.(c) We define the maximum dissipation as the spectral radius at $\gamma = 0$ and call this $[\rho(\mathbf{T})]_{\min}$, which is a function of κ and $\alpha_{\max}$. Both $\alpha_{\max}$ (black) and $[\rho(\mathbf{T})]_{\min}$ (red) are given as function of κ.

REFERENCES

1. Marx, D.; Hutter, J., *Ab initio molecular dynamics*. Cambridge University Press: New York, 2009.
2. Car, R.; Parrinello, M., Unified approach for molecular dynamics and density-functional theory. *Phys. Rev. Lett.* **1985,** *55*, 2471-2474.

3. Van Belle, D.; Froeyen, M.; Lippens, G.; Wodak, S. J., Molecular dynamics simulation of polarizable water by an extended lagrangian method. *Mol. Phys.* **1992,** *77*, 239-255.
4. Lamoureux, G.; Roux, B. t., Modeling induced polarization with classical drude oscillators: Theory and molecular dynamics simulation algorithm. *J. Chem. Phys.* **2003,** *119*, 3025-3039.
5. Sprik, M., Computer simulation of the dynamics of induced polarization fluctuations in water. *J. Phys. Chem.* **1990,** *95*, 2283-2291.
6. Rick, S. W.; Stuart, S. J.; Berne, B. J., Dynamical fluctuating charge force fields: Application to liquid water. *J. Chem. Phys.* **1994,** *101*, 6141-6156.
7. Kaminski, G. A.; Stern, H. A.; Berne, B. J.; Friesner, R. A., Development of an accurate and robust polarizable molecular mechanics force field from ab initio quantum Chem. *J. Phys. Chem. A* **2004,** *108*, 621-627.
8. Niklasson, A. M.; Tymczak, C.; Challacombe, M., Time-reversible born-oppenheimer molecular dynamics. *Phys. Rev. Lett.* **2006,** *97*, 123001.
9. Niklasson, A. M.; Tymczak, C.; Challacombe, M., Time-reversible ab initio molecular dynamics. *J. Chem. Phys.* **2007,** *126*, 144103.
10. Niklasson, A. M., Extended born-oppenheimer molecular dynamics. *Phys. Rev. Lett.* **2008,** *100*, 123004.
11. Niklasson, A. M.; Steneteg, P.; Odell, A.; Bock, N.; Challacombe, M.; Tymczak, C.; Holmström, E.; Zheng, G.; Weber, V., Extended lagrangian born–oppenheimer molecular dynamics with dissipation. *J. Chem. Phys.* **2009,** *130*, 214109.
12. Steneteg, P.; Abrikosov, I. A.; Weber, V.; Niklasson, A. M. N., Wave function extended lagrangian born-oppenheimer molecular dynamics. *Phys. Rev. B* **2010,** *82*, 075110.
13. Odell, A.; Delin, A.; Johansson, B.; Cawkwell, M. J.; Niklasson, A. M. N., Geometric integration in born-oppenheimer molecular dynamics. *J. Chem. Phys.* **2011,** *135*, 224105.
14. Niklasson, A. M. N.; Cawkwell, M. J., Fast method for quantum mechanical molecular dynamics. *Phys. Rev. B* **2012,** *86*, 174308.
15. Souvatzis, P.; Niklasson, A. M. N., Extended lagrangian born-oppenheimer molecular dynamics in the limit of vanishing self-consistent field optimization. *J. Chem. Phys.* **2013,** *139*, 214102.
16. Niklasson, A. M.; Cawkwell, M. J., Generalized extended lagrangian born-oppenheimer molecular dynamics. *J. Chem. Phys.* **2014,** *141*, 164123.
17. Niklasson, A. M. N., Next generation extended lagrangian first principles molecular dynamics. *arXiv* **2017**.
18. Souvatzis, P.; Niklasson, A. M. N., First principles molecular dynamics without self-consistent field optimization. *J. Chem. Phys.* **2014,** *140*, 044117.
19. Gans, J.; Shalloway, D., Shadow mass and the relationship between velocity and momentum in symplectic numerical integration. *Phys. Rev. E* **2000,** *61*, 4587-4592.
20. Hairer, E.; Lubich, C.; Wanner, G., Geometric numerical integration illustrated by the störmer–verlet method. *Acta Numer.* **2003,** *12*, 399-450.
21. Albaugh, A.; Demerdash, O.; Head-Gordon, T., An efficient and stable hybrid extended lagrangian/self-consistent field scheme for solving classical mutual induction. *J. Chem. Phys.* **2015,** *143*, 174104.
22. Albaugh, A.; Niklasson, A. M. N.; Head-Gordon, T., Accurate classical polarization solution with no self-consistent field iterations. *J. Phys. Chem. Lett.* **2017,** *8*, 1714-1723.
23. Alex Albaugh, T. H.-G., A new method for treating drude polarization in classical molecular simulation. *J. Chem. Theo. Comp.* **2017**.

24. Vitale, V.; Dziedzic, J.; Albaugh, A.; Niklasson, A. M. N.; Head-Gordon, T.; Skylaris, C.-K., Performance of extended lagrangian schemes for molecular dynamics simulations with classical polarizable force fields and density functional theory. *J. Chem. Phys.* **2017,** *146*, 124115.
25. Zheng, G.; Niklasson, A. M. N.; Karplus, M., Lagrangian formulation with dissipation of born-oppenheimer molecular dynamics using the density-functional tight-binding method. *J. Chem. Phys.* **2011,** *135*, 044122.
26. Niklasson, A. M. N., Iterative refinement method for the approximate factorization of a matrix inverse. *Phys. Rev. B* **2004,** *70*, 193102.
27. Schulz, G., Iterative berechung der reziproken matrix. *ZAMM - J. App. Math. Mech.* **1933,** *13*, 57-59.
28. Simmonett, A. C.; Pickard, F. C.; Shao, Y.; Cheatham, T. E.; Brooks, B. R., Efficient treatment of induced dipoles. *J. Chem. Phys.* **2015,** *143*, 074115.
29. Simmonett, A. C.; Pickard, F. C.; Ponder, J. W.; Brooks, B. R., An empirical extrapolation scheme for efficient treatment of induced dipoles. *J. Chem. Phys.* **2016,** *145*, 164101.
30. Swope, W. C.; Andersen, H. C.; Berens, P. H.; Wilson, K. R., A computer simulation method for the calculation of equilibrium constants for the formation of Phys. clusters of molecules: Application to small water clusters. *J. Chem. Phys.* **1982,** *76*, 637-649.
31. McLachlan, R. I.; Atela, P., The accuracy of symplectic integrators. *Nonlin.* **1992,** *5*, 541.

TABLE OF CONTENTS GRAPHIC