Title
Improved Methods for Polarizable Classical Molecular Dynamics Simulations

Author
Albaugh, Alex

Publication Date
2018

Peer reviewed|Thesis/dissertation

Improved Methods for Polarizable Classical Molecular Dynamics Simulations

By

Alex Todd Albaugh

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Chemical Engineering

and the Designated Emphasis

in

Computational and Data Science and Engineering

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Teresa Head-Gordon, Chair
Professor Martin Head-Gordon
Professor Ali Mesbah

Spring 2018

Abstract

Improved Methods for Polarizable Classical Molecular Dynamics Simulations

by

Alex Todd Albaugh

Doctor of Philosophy in Chemical Engineering

Designated Emphasis in Computational and Data Science and Engineering

University of California, Berkeley

Professor Teresa Head-Gordon, Chair

Polarization is the ability of a molecule's electron density to respond to and influence its environment and is the leading order many-body interaction for advanced electrostatics used in classical molecular simulation. It has proven to be an important interaction that is necessary to accurately simulate certain molecular systems. Polarization helps to capture intermolecular interactions of ligand-macromolecule complexes, heterogeneity at interfaces, electric field environments of heterogeneous systems such as proteins, and structure and dynamics of peptide-water solutions. In general, systems that can benefit most from the inclusion of polarization effects are heterogeneous, non-bulk systems that give rise to asymmetric environments. Additionally, polarization has been shown to be more transferable across the phase diagram beyond regions where the force field was initially parameterized.

The main drawback of including polarization in molecular simulation, however, is the computational expense of calculating explicit polarization interactions. The most common approach is to approximate the polarization solution using an iterative self-consistent field (SCF) method, which accounts for about half the cost of a polarizable simulation. Another approach is that of extended Lagrangians (EL), which treat polarization degrees of freedom dynamically and do not require iterations. EL methods, however, suffer from instability and require prohibitively small simulation time steps.

The focus of this dissertation is the reduction of the computational cost of polarizable classical molecular simulations while maintaining the high level of accuracy associated with these simulations. I present several new methods that combine the stability of SCF methods with the iteration-free dynamics of EL methods into a hybrid EL/SCF framework. The key to these EL/SCF methods is the introduction of auxiliary polarization degrees of freedom, which can be dynamically integrated and drive the real polarization degrees of freedom. The first approach is a relatively simple method for polarization that reduces the number of iterative cycles required for an SCF solution. This method also introduces thermostat control of auxiliary variables and is called iEL/SCF. A more sophisticated approach that eliminates the need for SCF iteration altogether, iEL/0-SCF, is also presented. This method is developed for both induced dipole and Drude polarization models. I also present a generalized and complete theory

for classical iteration-free polarizable EL/SCF dynamics and explore combining iteration-free dynamics with other advanced high efficiency methods such as RESPA multi-time stepping and stochastic-isokinetic integration, which work complementarily with EL/SCF to further increase computational efficiency.

In summary, the developments presented in this dissertation are methods and theories that significantly reduce the cost of classical polarizable molecular dynamics without sacrificing accuracy. This work represents an important step in moving the scientific community toward the broader adoption of advanced potential energy surfaces embodied by polarizable force fields.

Dedicated to the memory of my grandparents
Lloyd “Todd” and Marion Albaugh
who saw me embark on this journey
but were not able to see my return.

Contents

Acknowledgments

I cannot take sole credit for this dissertation, not by a long, long way. The work presented here is the result of many collaborators, mentors, educators, and supporters.

Before all others I would like to thank my parents, Cindy and Todd Albaugh. From an early age they indulged my curiosities and ambitions and I would be nowhere without them. They are the greatest supporters a person could ask for. My parents are my heroes and I think of what they have done for me every day. I also need to thank my best friend, Michael Albaugh. I can always count on Michael for a laugh, a good time, or a meaningful conversation. While we have been physically separated for many years our brotherly bond endures and has been a great comfort to me.

My partner in life, Ellyn Gendler, has played a central role through this graduate work. She once told me after she graduated from law school that it seemed like we had both graduated from law school, having shared all of the highs and lows together so closely. I think that sentiment is now reciprocated many fold on my part. Ellyn's love, kindness, and companionship sustain me and I am excited about what the future holds for us. I'm also thankful for the support and kindness of Lisa, Dan, and Karl Gendler.

This work would not have been possible without my advisor Professor Teresa Head-Gordon. At several points in graduate school I stumbled and her unwavering support allowed me to continue, something I will always remember. Most of the work seen in the following pages is the result of many lengthy, detailed, and sometimes confused conversations between the two of us. Research never quite goes the way you expect, but I think we had a good time rolling with those punches. I am grateful to my collaborator Dr. Anders Niklasson whose brilliant theories form the basis of this work and who I had the absolute honor of working with. As he told me, "Theory is a matter of will." I can only dream of matching his will. I would also like to thank Professor Mark Tuckerman, who I also had the pleasure of working with and whose insights have been most helpful. My thanks to Professor Chris-Kriton Skylaris, Dr. Jacek Dziedzic, and Valerio Vitale, as well, with whom I had a very fruitful collaboration, as well. I also thank Professors Martin Head-Gordon and Ali Mesbah for serving on my dissertation committee and Professors Susan Muller, David Graves, and Berend Smit for serving on my qualifying committee.

Getting me to this point is the result of the combined efforts of many talented educators. Starting at Palmyra Area High School in my hometown of Palmyra, PA I must thank Mrs. Judi Iskowitz who put me on the path of chemistry, Mrs. Kim Grissinger who continued to develop this passion, and Mr. Tim Refi and Mrs. Katherine Baum who provided me with an incredible foundation in mathematics that serves me well to this day. Mr. Anatoly Krashunsky, my clarinet and saxophone instructor, taught me the value of persistence and improvisation. At the University of Michigan I received the world's best education in large thanks to the efforts of many like Professors Ronald Larson and Indranil Saha Dalal who took me under their wings and allowed me to do research with them leading to my interest in computational science. I would also like to acknowledge Dr. Susan Montgomery, the excellent advisor who encouraged me to go to graduate

school, Dr. Andrew Tadd who graciously wrote recommendation letters for me, and the late Dr. Jason Daida who taught me how to program a computer.

At UC Berkeley I must acknowledge Dr. Lisa Felberg, my labmate, roommate, colleague, friend, confidant, and intellectual superior. Lisa is the sister I never had and whose walks and talks were always appreciated and constantly missed now. I owe a great deal to Dr. Omar Demerdash, as well, whose insight and encyclopedic knowledge were invaluable for my research progress. I also appreciate all the teaching experiences I had, both good and bad, with my officemate and good friend Dr. Sukanya Sasmal. I will always cherish the friendships that were formed in the Head-Gordon lab with Dr. Saurabh Belsare, Dr. Asmit Bhowmick, David Brookes, Jamie Lincoff, Dr. Luis Ruiz Pestaña, Tara Armand, Dr. Efrem Braun, Michelle Liu, Matt Witman, and Eugene Yedvabny. Also at UC Berkeley I must thank Professor Kranthi Mandadapu whose conversations I always enjoy and learn from and Dr. Henrik Wallman who I enjoyed teaching with.

I am grateful to have the support and love of many close friends. I was fortunate to meet an incredible group of people in my formative years and still hold them all close to my heart. Sarah Ditzler, Natalie DeSantis, Laura Fenstermacher, Amy Mastrangelo, Nick Kauffman, Shayne Bement, and Blake Davis are always there when I need them to be and I am incredibly fortunate for that. As a member of the Michigan Marching Band in college I met my second family. Jon and Lauren Gentile, Kyle Mooney, Ethan Hain, Catherine and Josh Huffman, Catelyn VanDerKolk, Matt Schottler, Stuart and Amber Marshall, Paul Gorbsky, Gina Angileri-Lange, and Jessie Fletcher are the truest friends I could hope to have. I am also deeply thankful for those friendships that were forged in the fires of my graduate school experience- Dr. Alex Wang, my funny and trusted better looking twin, the kind-hearted Sarah Wang, and the always-thoughtful Mike Goldstein.

I am also fortunate to have the support of a large extended family that I do not get to see enough, but am profoundly thankful for. Thank you so much for your love and support Pam, Keith, Roxie, Gary, Nikki, Josh, Sammy, Jess, Dan, Lydia, Maddie, Lisa, Kyle, Drew, Ben, Tracy, Tom, Annie, Sophie, and Claire. We all owe a debt of gratitude to my late grandparents Lloyd “Todd” and Marion Albaugh, to whom I have dedicated this work. Grandma and grandpa worked hard to raise this great family and they serve as an inspiration to us all.

To all of these people and many more I will never be able to fully express my thanks, but hopefully this is some small start.

1.
Introduction

Polarization is the ability of a molecule's electron density to respond to and influence its environment and is the leading order many-body interaction for advanced electrostatics used in classical molecular simulation[1-4]. It has proven to be an important interaction that is necessary to accurately simulate certain molecular systems. Polarization helps to capture intermolecular interactions of ligand-macromolecule complexes [5-6], heterogeneity at interfaces[7-8], electric field environments of heterogeneous systems such as proteins[9-13], and structure and dynamics of peptide-water solutions[14-15]. In general, systems that can benefit most from the inclusion of polarization effects are heterogeneous, non-bulk systems that give rise to asymmetric environments[4, 16-18]. Additionally, polarization has been shown to be more transferable across the phase diagram of liquid water beyond the regions where the force field was initially parameterized[19].

The main drawback from including polarization in molecular simulation, however, is the computational expense of calculating explicit polarization interactions. In principle this calculation can be done exactly by matrix inversion or Cholesky factorization, but this becomes prohibitively expensive for systems larger than tens of atoms[4]. The most common approach is to approximate the solution using an iterative self-consistent field (SCF) method[20-21], which accounts for about half the cost of a polarizable simulation[22]. The focus of this dissertation is the reduction of the computational cost of polarizable classical molecular simulations while maintaining the high level of accuracy associated with these simulations. Chapter 2 presents an initial, relatively simple method for polarization that reduces the number of iterative cycles required for an SCF solution[23]. Chapter 3 presents a more sophisticated approach that eliminates the need for SCF iteration altogether for an induced dipole polarization model[22] and Chapter 4 extends this approach to a Drude polarization model[24]. Chapter 5 presents a generalized and complete theory for classical iteration-free polarizable dynamics[25]. Finally, Chapter 6 explores combining iteration-free dynamics with other advanced high efficiency methods such as RESPA multi-time stepping[26] and stochastic-isokinetic integration[27]. The following sections of this introduction give an introduction to polarizable models (Section 1.1), polarizable force fields used in this study (Section 1.2), and current polarizable methods including a brief introduction to the methods presented in this work (Section 1.3).

In summary, the developments presented in this dissertation are methods and theories that significantly reduce the cost of classical polarizable molecular dynamics without sacrificing accuracy, all at the cost of only the chosen level of the permanent electrostatic model. This work represents an important step toward moving the scientific community toward the broader adoption of advanced potential energy surfaces embodied in polarizable potentials[2].

1.1 Polarizable Models

Polarization can be described in classical electrostatics through three models: induced dipoles[17, 19, 28-39], Drude oscillators[40-46], and fluctuating charges[47-53]. Each of these models will be introduced below along with a brief discussion of its benefits and drawbacks. This dissertation makes use of all three models at various points, but the main model used is the induced dipole model because of its strong theoretical grounding as the first-order response with respect to an applied external electric field[54] and also its efficient implementation when combined with multipolar electrostatics[4].

1.1.1 *Induced Dipole Model*

The induced dipole model has a history that goes back to Warshel and Levitt's Nobel Prize-winning work studying enzyme-substrate dynamics via simulation in 1976[28]. More recently induced dipoles have found a place in the AMOEBA polarizable force field[19, 31, 33-34], a force field used extensively in this work and discussed in Section 1.2.1, and other state of the art potentials, as well[2].

A point induced dipole is a dipole created by the electric field experienced by an atom in order to minimize potential energy. Specifically, an induced dipole is proportional to the total electric at an atomic center with the constant of proportionality called the polarizability, as shown in Eq. (1.1).

$$\boldsymbol{\mu_i} = \alpha_i \boldsymbol{E}_i^{tot} = \alpha_i \boldsymbol{E}_i + \alpha_i \sum_{j=1}^{N} \boldsymbol{T}'_{ij} \boldsymbol{\mu}_j \tag{1.1}$$

In Eq. (1.1) $\boldsymbol{\mu}_\mathrm{i}$ is the induced dipole at the *i*-th atomic center, $\boldsymbol{E}_i^{tot}$ is the total electric field contribution at the *i*-th atomic center, and α_i is the atomic polarizability of the *i*-th atom. The total electric field can be broken down into contributions due to other permanent (non-polarizable) electrostatic moments (like fixed charges, fixed dipoles, fixed quadrupoles, etc.) and the electric field created by other induced dipoles in the system, as the second right-hand side of Eq. (1.1) shows. Here $\boldsymbol{E}_i$ is the electric field due to permanent electrostatic moments in the system and $\boldsymbol{T}'_{ij}$ is the dipole interaction matrix between atomic centers *i* and *j* (its elements consist of derivatives of $1/r_{ij}$ Coulomb interaction terms). From Eq. (1.1) we can begin to see why a polarization model is expensive. The induced dipole terms appear on both the right- and left-hand sides and need to be self-consistent. This equation can be written in super matrix form, where induced dipoles for every atomic center in the system can be represented with a single equation, given in Eq. (1.2).

$$\boldsymbol{C}'\boldsymbol{\mu} = \boldsymbol{E} \tag{1.2}$$

In Eq. (1.2) $\boldsymbol{C}'$ is a *3N* by *3N* matrix whose blocks are defined by $\boldsymbol{C}'_{ij} = (\alpha_j^{-1}\delta_{ij} - \boldsymbol{T}'_{ij})$ between atomic centers *i* and *j* for *N* total atomic centers in the system. The $\boldsymbol{\mu}$ and $\boldsymbol{E}$ matrices are then *3N* by *3* where the *i*-th row is given by the transpose of the *i*-th induced dipole or permanent electric field, respectively. The self-consistent solution for the induced dipoles is then given by Eq. (1.3).

$$\boldsymbol{\mu}_{SCF} = \boldsymbol{C'}^{-1}\boldsymbol{E} \tag{1.3}$$

Inverting the $\boldsymbol{C'}$ matrix is computationally expensive and scales as $O(N^3)$ so direct inversion is not practical for condensed phase simulation where systems of tens of thousands of atoms are common, but it does mathematically define the true ground state electronic solution. With these self-consistent dipoles the potential energy due to polarization is defined by Eq. (1.4).

$$U^{polar} = -\frac{1}{2}\boldsymbol{\mu}_{\boldsymbol{SCF}}^{\boldsymbol{T}}\boldsymbol{E} \tag{1.4}$$

The induced dipole model is arguably the most robust polarizable model because its foundation arises from an expansion of the electric potential about an applied electric field and therefore it can be systematically generalized to higher-order moments like induced quadrupoles, induced octupoles, etc.[54]

1.1.2 *Drude Model*

The Drude model was first introduced by Paul Drude in 1900 to study the optical properties of materials[55-56] and actually predates the advent of quantum mechanics. The model was first introduced into classical molecular simulation in 1974[40] and is sometimes known as the shell model[40-41].

In the Drude model an additional particle, known as the Drude particle, is bound to each atomic center with a harmonic spring whose equilibrium position is its parent atom's position with a force constant of k_D. Each Drude particle also carries a charge so polarization is included via an approximate induced dipole created by charge separation between the Drude particle and atomic center. The position of the *i*-th atom is then $\boldsymbol{r_i}$, the position of the corresponding *i*-th Drude particle is $\boldsymbol{r_{D,i}}$, and their relative displacement is $\boldsymbol{d_i} = \boldsymbol{r_{D,i}} - \boldsymbol{r_i}$. If the *i*-th atomic center usually carries a charge of q_i^{net} then in the Drude model the parent atom will now carry a charge of q_i and the Drude particle will carry a charge of $q_{D,i}$ such that $q_i^{net} = q_i + q_{D,i}$. It can be shown that the relationship between the charge partitioned to the Drude particle, $q_{D,i}$, the Drude spring constant, k_D, and the atomic polarizability, α_i, is given by Eq. (1.5) [43]. If we know the atomic polarizability of an atom then we only need to introduce one new parameter as the Drude charge and force constant are not independent.

$$\alpha = \frac{q_D^2}{k_D} \tag{1.5}$$

The Drude model then requires two modification of the potential energy. The first of which is the inclusion of a potential energy term due to the harmonic bonding of the Drude particle to its parent atom, shown in Eq. (1.6).

$$U^{Drude} = \frac{1}{2}k_D\sum_i \left|\boldsymbol{r_i} - \boldsymbol{r_{D,i}}\right|^2 \tag{1.6}$$

The second modification of the potential energy is that the electrostatic energy must include interactions between not just atomic centers and other atomic centers, but also atomic centers and Drude particles and Drude particles and other Drude particles, as given by Coulombic interactions in in Eq. (1.7).

$$U^{electrostatic} = \sum_i \sum_{j>i} \frac{q_i q_j}{|\boldsymbol{r_i} - \boldsymbol{r_j}|} + \sum_i \sum_j \frac{q_i q_{D,j}}{|\boldsymbol{r_i} - \boldsymbol{r_{D,j}}|} + \sum_i \sum_{j>i} \frac{q_i q_{D,j}}{|\boldsymbol{r_{D,i}} - \boldsymbol{r_{D,j}}|} \tag{1.7}$$

As with the induced dipole model we can define a ground state, self-consistent solution for the positions of Drude particles by $k_D \boldsymbol{d}_i = q_{D,i} \boldsymbol{E}_{D,i}$, where $\boldsymbol{E}_{D,i}$ is the electric field at the *i*-th Drude particle. This is also a difficult to solve equation as $\boldsymbol{E}_{D,i}$ is dependent on the other displacements, $\boldsymbol{d}_j$, and can be thought of as analogous to Eq. (1.1).

The main advantage of the Drude model is that it is relatively straightforward to implement. Most molecular dynamics codes and force fields incorporate Coulombic (charge-charge) interactions already so adding a Drude model amounts to simply doubling the number of particles in the system and implementing the harmonic springs of Eq. (1.6). One disadvantage of the Drude model is that Drude particles are typically treated with extended Lagrangian schemes (described in Section 1.3.2) where the Drude particles carry part of the parent atom's mass. This mass repartitioning between parent and Drude can cause numerical instability as the Drude mass must be small to properly follow electronic dynamics, leading to either instability or the need for intractably small time steps.

1.1.3 *Fluctuating Charge Model*

The fluctuating charge model introduces polarization by allowing the charges on each atom in the system to change value in order to minimize the potential energy of the system, subject to the constraint that the net charge in the system remains unchanged. Eq. (1.8) gives the combined electrostatic/polarization potential energy for a fluctuating charge model[47].

$$U^{electrostatic} = \boldsymbol{\chi}^{\boldsymbol{T}} \boldsymbol{q} + \frac{1}{2} \boldsymbol{q}^{\boldsymbol{T}} \boldsymbol{C} \boldsymbol{q} \tag{1.8}$$

In Eq. (1.8) $\boldsymbol{\chi}$ is a vector of all the electronegativities of the *N* atoms in the system, χ_i; $\boldsymbol{q}$ is a vector of the fluctuating partial charges of the *N* atoms, q_i; and $\boldsymbol{C}$ 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 ($C_{ii} = 2H_i$). The self-consistent fluctuating charges for a given atomic configuration are the set of charges that minimize this potential energy and are given by Eq. (1.9).

$$\boldsymbol{q}_{SCF} = -\boldsymbol{C}^{-1} \boldsymbol{\chi} \tag{1.9}$$

Similar to the case of induced dipoles in Eq. (1.3) the self-consistent ground state fluctuating charges are given by a matrix inversion that becomes expensive for all but the

smallest systems. In practice this equation is solved with iterative methods or other methods described in Section 1.3. In a sense the fluctuating charge model can be thought of as a reduced dimensional induced dipole model. In such an analogy fluctuating charge electronegativities, χ_i, parallel induced dipole permanent electric field, $\boldsymbol{E}_i$; the $\boldsymbol{C}$ matrix of the fluctuating charge model is a charge-charge interaction matrix and the $\boldsymbol{C}'$ matrix of the induced dipole model is a dipole-dipole interaction matrix; and finally, the fluctuating charges, q_i, are analogous to the induced dipoles, $\boldsymbol{\mu}_i$. This dimensionality reduction, however, is one of the drawbacks of the fluctuating charge model. Induced dipoles and Drude particles can account for out-of-plane polarization because of the directional nature of induced dipole vectors and Drude particle displacements. The polarization response of fluctuating charges, however, is constrained to the lines or planes along which charges lie in molecular geometry. For example, for a carbon dioxide molecule a fluctuating charge model can only polarize along the oxygen-carbon-oxygen axis, whereas induced dipoles or Drude displacements can point away from this axis. Because of this the fluctuating charge model is a more restrictive polarization model than induced dipoles or Drude particles.

1.2 Polarizable Force Fields

A force field, in the molecular mechanics sense, is a specific set of functional forms and parameterizations for those functional forms that define the potential energy for a configuration of atoms. In short, a force field is a function that turns atomic coordinates into potential energy, $U(\boldsymbol{r})$. Polarizable force fields include a polarizable model from Section 1.1 and a corresponding parameterization for that model, as well as terms to describe bonded valence interactions and other non-bonded interactions like permanent electrostatics and van der Waals interactions. The bulk of this dissertation is performed with the AMOEBA (atomic multipole optimized energetics for biomoleular simulation) force field [19, 31, 33-34] and work with the Drude model was performed with the PSPC (polarizable simple point charge) force field for water[42-43].

1.2.1 *AMOEBA Force Field*

The AMOEBA force field is a general-purpose biomolecular force field with parameters for water, ions, small molecules, proteins, and DNA. The AMOEBA classical functional form is broken into several components, given in Eq. (1.10).

$$U^{AMOEBA} = U^{bond} + U^{angle} + U^{bond-angle} + U^{out-of-plane} + U^{torsional} + U^{vdW} + U^{electrostatic} + U^{polarization} \tag{1.10}$$

The first five terms of Eq. (1.10) describe bonded interactions that keep the internal geometry of a molecule close to its equilibrium conformation. Respectively, these are bond stretching, angle bending, bond-angle stretch/bend cross terms, out-of-plane bending, and torsional rotation. The first four of these are described by high-order anharmonic springs with equilibrium positions corresponding to the equilibrium bond lengths and angles of a molecule. The torsional potential is given by cosine terms that describe energy barriers associated with rotating atoms separated by three bonds about a common axis. The final three terms of Eq. (1.10) describe through-space non-bonded

interactions. The first are the van der Waals interactions, which include long-range dispersion and short-range repulsion. For the AMOEBA model these are described by a buffered 14-7 functional form[57]. The electrostatic contribution comes from fixed (non-inducible) multipole moments[58] in the system. The AMOEBA force field includes fixed electrical monopoles (charges), dipoles, and quadrupoles for each atom in the system. Finally, the polarization is implemented using an induced dipole model described in Section 1.1.1. One notable side effect of any classical polarizable model is that if two atoms come close together they may mutually polarize without bound, a problem known as the polarization catastrophe. This is avoided in the AMOEBA force field by using short-range polarization damping[30].

1.2.2 *PSPC Force Field*

The PSPC force field has a relatively simple functional form given by Eq. (1.11).

$$U^{PSPC} = U^{vdW} + U^{electrostatic} + U^{Drude} \tag{1.11}$$

Unlike the AMOEBA force field, the PSPC force field has no bonded terms. For the PSPC model the intramolecular geometry is constrained to be in its equilibrium configuration and does not use spring-like potentials that allow this geometry to deform, like the AMOEBA force field. The rigid model constraint enforcement is satisfied with a SHAKE/RATTLE algorithm[59]. Non-bonded forces then define the potential energy and the van der Waals interactions are described by simple a Lennard-Jones potential. The polarization is implemented using a Drude model so the electrostatics include sums over the Drude positions and charges, Eq. (1.7), and we must account for the harmonic binding of Drude particles to parent atoms, Eq. (1.6). The fixed electrostatics of the PSPC force field only include fixed charges and not higher-order multipoles like AMOEBA.

1.3 Polarization Solutions

There are two standard approaches to treating polarization, which are self-consistent field iterative solvers (SCF) and extended Lagrangian (EL) dynamics. This dissertation lays out a third new hybrid approach by combining favorable aspects of EL and SCF, which I term EL/SCF. This section will give background on SCF and EL approaches as well as a very brief introduction to the EL/SCF that are the subject of this dissertation. Other novel alternative approaches to treating polarization will also be briefly reviewed. For simplicity most approaches will be discussed in the context of an induced dipole polarization model (Section 1.1.1).

1.3.1 *Self-Consistent Field Iteration*

As mentioned previously, the solution to a polarizable system given in Eq. (1.2) is almost always solved through a self-consistent solution with an iterative optimizer. One simple example of such an SCF iterative approach is successive over-relaxation (SOR) [60], given by Eq. (1.12).

$$\boldsymbol{\mu}_i^{n+1} = (1-\omega)\boldsymbol{\mu}_i^n + \omega\left[\boldsymbol{\mu}_i^n + \alpha_i \sum_{j=1}^{N} \boldsymbol{T}'_{ij}\boldsymbol{\mu}_j\right] \tag{1.12}$$

In Eq. (1.12) the n-th iteration of the induced dipole $\boldsymbol{\mu}_i$ is updated to the n+1-th iteration using a linear mixing of the n-th iteration and a combination of the n-th iteration and the mutual response of the n-th iteration. Here the linear mixing is controlled by the relaxation parameter, ω. After enough iteration, the dipoles will eventually converge to a ground state solution. The convergence threshold is given by the root-mean-square change between the induced dipoles from one iteration to the next and once enough iterations are performed to reach this threshold the iterative procedure is stopped and the dipoles of the final iteration are declared close enough to the true solution, $\boldsymbol{\mu}_i^{SCF}$. Convergence thresholds are typically around 10^{-6} RMS Debye (D).

The SOR method is illustrative, but not practically efficient. Typically more sophisticated SCF solvers are employed such as the preconditioned conjugated gradient (PCG) [20] or direct inversion of the iterative subspace (DIIS) [21, 61]. Another recent development along these lines is the truncated conjugate gradient (TCG) method, which minimizes the number of matrix-vector multiplications and is amenable to scaling on modern high-performance computing platform[62]. To give an idea of scale, simulating bulk water with the AMOEBA force field typically requires around 10-12 iterations per time step with SOR and 5-8 iterations per time step with PCG. Each iteration comes at significant expense because the cost of calculating the $\sum_{j=1}^{N} \boldsymbol{T}'_{ij}\boldsymbol{\mu}_j$ term in Eq. (1.12) (or its equivalent) is non-negligible. The more efficient methods developed in this dissertation will compare against the PCG iterative solver as a baseline.

One common addition to an SCF solver is a predictor[63], which takes a linear combination of previous solutions of the induced dipoles during a simulation and uses the result as an initial guess for the SCF solver, Eq. (1.13).

$$\boldsymbol{\mu}_i^0(t) = \sum_{k=1}^{K} b_k \boldsymbol{\mu}_i^{SCF}(t - k\Delta t) \tag{1.13}$$

In Eq. (1.13) the initial guess for the SCF solver at the current time t, $\boldsymbol{\mu}_i^0(t)$, is a linear combination of the K previous iterative solutions at times $t - k\Delta t$, where Δt is the time step of the simulation. The constants b_k then give the exact form of this linear combination. The use of predictors can accelerate SCF solvers, but leads to overall degradation in energy conservation and a more detailed discussion of their use is presented in Chapter 2.

In general, self-consistent iterative solution parallels the Born-Oppenheimer approximation[64] in *ab initio* (quantum) molecular simulation. Under the Born-Oppenheimer approximation 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 atomic configuration during a molecular dynamics trajectory.

1.3.2 *Extended Lagrangian Dynamics*

Extended Lagrangian methods typically appear in molecular simulation in order to simulate in various thermodynamic ensembles. The natural ensemble for molecular simulation is constant particle, constant volume, and constant energy (NVE). By including additional thermostat degrees of freedom coupled to particle velocities[65-66] one can move to an NVT ensemble. By further treating the simulation box sides as dynamic degrees of freedom we can move to the NPT ensemble. The Lagrangian is defined as the kinetic energy minus the potential energy of a system. When additional degrees of freedom (thermostats, barostats, and as I will show, auxiliary polarization variables) are included in a simulation corresponding kinetic energies are added to and corresponding potential energies are subtracted from the Lagrangian, hence extended Lagrangian.

Electronic degrees of freedom, such as induced dipoles for polarization, can also be treated dynamically in an extended Lagrangian approach. For such a method we first define the extended Lagrangian, where we now include the induced dipoles as dynamic degrees of freedom as given in Eq. (1.14).

$$\mathcal{L} = \frac{1}{2}\sum_{i=1}^{N} m_i \dot{\boldsymbol{r}}_i^2 + \frac{1}{2}\sum_{i=1}^{N} m_i^{\mu} \dot{\boldsymbol{\mu}}_i^2 - U(\boldsymbol{r}) + \frac{1}{2}\boldsymbol{\mu}_{\boldsymbol{SCF}}^{\boldsymbol{T}}\boldsymbol{E} \tag{1.14}$$

In Eq. (1.14) the first term represents the kinetic energy of the particles, the second term represents kinetic energy associated with the now dynamic induced dipoles, which now have associated masses m_i^{μ}, the third term represents the non-polarization potential energy (the potential energy not dependent on the induced dipoles), and the final term is the potential energy associated with the induced dipoles from Eq. (1.4). We can derive equations of motion from a Lagrangian by using the Euler-Lagrange equation, Eq. (1.15), where q is a general coordinate.

$$\frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{q}} = \frac{\partial \mathcal{L}}{\partial q} \tag{1.15}$$

Applying Eq. (1.15) on Eq. (1.14) with respect to the induced dipoles and leveraging Eq. (1.1) gives an induced dipole equation of motion, Eq. (1.16).

$$m_i^{\mu}\ddot{\boldsymbol{\mu}}_i = \boldsymbol{E}_i^{tot} - \alpha_i^{-1}\boldsymbol{\mu}_i \tag{1.16}$$

Using Eq. (1.16) we can now numerically integrate this equation of motion for the induced dipoles, just as we numerically integrate the equations of motion of the atomic degrees of freedom. In this way we have replaced SCF iteration with a much more efficient dynamic integration in time. We should note that if the definition of the ground state induced dipoles in Eq. (1.1) is satisfied then the right-hand side of Eq. (1.16) is zero and there is no force on the induced dipoles. In this way the induced dipoles are driven towards and fluctuate about the ground state solution, but do not necessarily (and in practice rarely ever exactly) match the true solution.

One may think that extended Lagrangians are the ideal solution, however, to achieve realistic dynamics the dipole mass parameter m_i^{μ} must be small. Due to this

small mass the time step of the simulation must also be small to accurately follow the dynamics, so what is gained in not doing iteration is lost in the cost of a smaller time step. In practice EL methods are not popular due to their unstable nature except at prohibitively small time steps.

The use of extended Lagrangians to treat polarization is directly analogous to and a product of Car and Parrinello's seminal work in 1985[67] where they treated the orbitals of density functional theory (DFT) as dynamic degrees of freedom in an extended Lagrangian approach. Their work was an attempt to alleviate the computationally restrictive nature of Born-Oppenheimer dynamics. Since then such approaches have been applied to all models of polarization- induced dipoles (described above)[68], Drude particles[43], and fluctuating charges[47, 50, 69].

1.3.3 *Extended Lagrangian/Self-Consistent Field Methods*

The ideal method for treating polarization would combine the stability of SCF solutions to the Born-Oppenheimer condition with the iteration-free efficiency of extended Lagrangians used in methods such as Car-Parrinello. Recently Niklasson and colleagues have pursued just such an approach for *ab initio* method[70-81], which forms the inspiration and groundwork for this dissertation. By combining aspects of EL and SCF into a single, fundamental framework for classical polarization I have developed methods that can reduce the number of required SCF iterations using a relatively simple approach (Chapter 2) [23], eliminate the need for SCF iteration altogether while maintaining SCF stability and accuracy (Chapter 3 for the induced dipole model, Chapter 4 for the Drude model) [22, 24], described the theoretical framework that binds all of these approaches together (Chapter 5) [25], and paired these approaches with other advanced molecular dynamics acceleration approaches to achieve truly remarkable efficiency (Chapter 6). Using these methods now guarantees that polarization is now no more computationally expensive than the underlying permanent electrostatic model of the system. What will follow is a very brief outline of the approach, which is developed more extensively in subsequent chapters.

The main insight is to build an extended Lagrangian, but instead of using the electronic degrees of freedom directly in the Lagrangian (induced dipole in this example), we introduce an auxiliary set of electronic degrees of freedom, $\boldsymbol{a}$. Here $\boldsymbol{a}_i$ is an auxiliary induced dipole corresponding to the real induced dipole $\boldsymbol{\mu}_i$. The extended Lagrangian we build then takes the form of Eq. (1.17).

$$\mathcal{L} = \frac{1}{2}\sum_{i=1}^{N} m_i \dot{\boldsymbol{r}}_i^2 + \frac{1}{2}\sum_{i=1}^{N} m_i^a \dot{\boldsymbol{a}}_i^2 - U(\boldsymbol{r}) - U^{pol} - \frac{1}{2}\omega^2 \sum_{i=1}^{N} m_i^a \left(\boldsymbol{\mu}_i - \boldsymbol{a}_i\right)^2 \qquad (1.17)$$

In Eq. (1.17) U^{pol} is left in a general form because depending on the flavor of the method it may need to be modified. We have also introduced a separate harmonic term into the Lagrangian, the final term on the right-hand side of Eq. (1.17), to maintain each auxiliary close to its corresponding real dipole with frequency parameter ω. By applying the Euler-Lagrange equation along with the limit that the auxiliary mass parameter m^a goes to 0, we obtain equations of motion for the real and auxiliary degrees of freedom, given in Eq. (1.18).

$$m_i \ddot{\boldsymbol{r}}_i = -\frac{\partial (U(\boldsymbol{r}) + U^{pol})}{\partial \boldsymbol{r}_i}\Big|_{\boldsymbol{a}} \tag{1.18a}$$

$$\ddot{\boldsymbol{a}}_i = \omega^2 (\boldsymbol{\mu}_i - \boldsymbol{a}_i) \tag{1.18b}$$

If we want to simply reduce the number of iterations required for an SCF solver we could use the auxiliary dipole, $\boldsymbol{a}$, s at each time step as an initial guess for the SCF solver and use the standard form of U^{pol}. Using the auxiliaries as an initial guess is superior to a predictor because the auxiliaries are integrated time reversibly, as will be discussed in Chapter 2. To achieve iteration-free dynamics we need to modify the form of U^{pol} and estimate the true SCF solution using linear mixing, which modifies the auxiliary equation of motion, which is discussed in Chapters 3 and 4. In general, though, we are using an extended Lagrangian for an auxiliary set of polarization degrees of freedom while maintaining key aspects of SCF, leading to the family of EL/SCF methods.

1.3.4 *Alternative Polarizable Methods*

Other approaches whose aim is to reduce the cost of polarization have been developed, as well. Presented here is a brief overview of notable methods along with discussion of their benefits and drawbacks.

The iAMOEBA approach[19], which is based on a reparameterization of the AMOEBA force field, only accounts for direct polarization, which is the polarization response due to fixed electrostatics in the system, and ignores the mutual dipole response. Mathematically this equates to truncating the second right-hand side of Eq. (1.1) to just the first term. One can clearly see then that no iteration is necessary to find the polarization response, it is simply proportional to the fixed electrostatic field. While eliminating the need for iteration this approach requires force field reparameterization and does not account for the full polarization response, thus being less accurate in principle.

Extrapolated perturbation theory (ExPT) is a method that expands the polarization energy as a series of perturbations, truncates the series at low order, and then statistically extrapolates to infinite order to recover an approximation of the true, fully converged SCF solution[82-83]. The cost of this method is on the order of 2-3 iteration cycles, which is excellent, but suffers from inaccuracy due to the nature of statistically extrapolating a polarization response that must be fitted.

Polarization, being inherently a many-body or N-body effect, can be written as an infinite expansion of one-body, two-body, three-body, four-body, etc. terms. A method taking advantage of this fact explores the possibility of treating polarization by truncating this expansion to the three-body term, thus evaluating polarization as the sum of exact one-, two-, and three-body interactions[84-85]. This is, of course, an approximation to the true polarization response, but was found to be accurate for some systems and realized computational gains, albeit with major code restructuring to efficiently calculate three-body interactions.

While solving for the polarization response using direct inversion, Eq. (1.3), is computationally intractable for large systems, it may be done efficiently for very small systems. Taking advantage of this is the coupled polarization-matrix inversion and

iteration (CPII) scheme[86]. This approach does direct inversion to find the intramolecular polarization response, which is feasible for small molecules, and then uses that result to solve for the overall intermolecular response iteratively. This approach can reduce the number of iterations required, but not completely.

2.
iEL/SCF:
Reducing the Cost of Self-Consistent Iteration for Polarization

2.1 Introduction

Polarizable empirical force fields offer a clear and systematic improvement over current generation fixed charge force fields by including many-body effects that allow for a molecular response to evolving heterogeneous environments[2, 4, 6, 11, 16-17, 42]. However, the primary computational expense of a classical polarization model for large systems resides in the solution of a linear system of equations for the induced dipoles. One choice is to implement a self-consistent field (SCF) iterative scheme such as a preconditioned conjugate gradient approach (PCG)[20] or direct inversion in the iterative subspace (DIIS)[21, 87] along with a predictor to accelerate the convergence of the SCF problem. However, predictors use information from previous steps and are time irreversible, leading to an inevitable degradation in energy conservation. Another approach is to replace the SCF step with an extended Lagrangian (EL) formulation to avoid any iterative SCF costs[16-17, 67-68]. However, this approach can be plagued with problems of accuracy since EL formulations allow the induced dipoles to fluctuate around an average orientation that does not strictly conform to the true electric field vector. Moreover, this approach can suffer from problems of stability and energy conservation in the context of a MD trajectory that forces the time step to be unacceptably short.

Recently, Niklasson and co-workers[70-81] have introduced a hybrid extended Lagrangian self-consistent field (EL/SCF) scheme in the context of Born-Oppenheimer molecular dynamics (BOMD) wherein an extended set of auxiliary electronic degrees of freedom serve as an initial guess for the SCF solver. This allows less strict convergence of the ground state electron density due to the benefits of a time reversible Verlet algorithm that realizes excellent energy conservation. In this chapter I present an adapted EL/SCF approach for the problem of classical polarization. In this context the initial guess for the induced dipole SCF calculation is given by auxiliary induced dipole variables evolved via a time-reversible velocity Verlet scheme[88]. The benefit to such an approach is that one can now integrate reversible equations of motion of the auxiliary dipole initial guess at the same large time step as the atomic positions, while maintaining superior energy conservation since the polarization response stays near the true solution (Born-Oppenheimer surface) at looser convergence levels relative to standard SCF solvers.

However, like the original hybrid EL/SCF approach used in BOMD for the electron density matrix, I find that the set of auxiliary induced dipoles also exhibit a similar problem in numerical stability. This is manifested as a continued (and likely boundless) increase in the number of SCF cycles to meet even loose (10^{-1} D) convergence tolerances over the course of a 1.0 ns trajectory of the AMOEBA14 water model[34]. I

have diagnosed the problem in the hybrid EL/SCF scheme applied to classical polarization as arising from resonances in the equations of motion[27, 89-90] that manifests as a buildup of inertia for the auxiliary dipoles. Although in principle I could address the resonance problem with a smaller integration time step, instead I have formulated a new restrained inertia EL/SCF (iEL/SCF) method that for all intents and purposes controls for the resonance problem, analogous to other isokinetic approaches[27], such that the equations of motion of the auxiliary dipoles remain stable and time reversible. This inertial restraining is implemented with a thermostatting scheme, illustrated using Berendsen weak coupling[91] and Nosé-Hoover chain thermostats[65-66], applied to the auxiliary dipole velocities.

This iEL/SCF approach is a clear improvement over standard SCF approaches to classical mutual induction calculations. The results on the AMOEBA polarization model show that the iEL/SCF method exhibits excellent energy conservation and thermodynamic and dynamic properties in both the NVE and NVT ensembles, but at greatly relaxed real dipole convergence tolerances, which reduce the number of SCF cycles relative to standard SCF solvers used for the classical polarization solution. As such, the iEL/SCF scheme clearly offers a better choice for classical mutual induction calculations compared to many EL and SCF and hybrid alternatives. iEL/SCF is also straightforward to implement, as it requires no modification of potential energy functions or previously implemented SCF solvers. Additionally this iEL/SCF approach has been proven to be useful for *ab initio* molecular dynamics, as well[92].

2.2 Methods

2.2.1 *Polarizable Model*

In this work we develop our approach on the classical polarizable force field AMOEBA (Atomic Multipole Optimized Energetics for Biomolecular Applications)[31-34]. In addition to fixed multipole electrostatics, the AMOEBA model provides a consistent treatment of intramolecular and intermolecular polarization, and uses a physically motivated Thole damping scheme for local polarization effects to avoid the well-known polarization catastrophe that results when mutually inducible sites polarize each other to infinity at short inter-site separation[30, 93]. AMOEBA's many-body polarization energy, U^{polar}, is given by Eq. (2.1a) with the SCF (ground state) dipoles given by Eq. (2.1b).

$$U^{polar} = -\frac{1}{2}\sum_{i=1}^{N}\left(\boldsymbol{\mu}_{SCF,i}\right)^{T}\boldsymbol{E}_{i} = -\frac{1}{2}\sum_{i=1}^{N}\left(\boldsymbol{\mu}_{SCF,i}\right)^{T}\sum_{j\neq i}\boldsymbol{T}_{ij}\boldsymbol{M}_{j} \tag{2.1a}$$

$$\boldsymbol{\mu}_{SCF,i} = \alpha_{i}\left(\sum_{j\neq i}\boldsymbol{T}_{ij}\boldsymbol{M}_{j} + \sum_{j\neq i}\boldsymbol{T}'_{ij}\boldsymbol{\mu}_{SCF,j}\right) \tag{2.1b}$$

In Eq. (2.1) $\boldsymbol{\mu}_{SCF,i}$ is the inducible dipole at atom site i, α_i is the isotropic polarizability of atom i, $\boldsymbol{T}_{ij}$ is the rank-two interaction tensor between atoms i and j containing derivatives of $1/r_{ij}$ according to the permanent multipole expansion, $\boldsymbol{T}'_{ij}$ is the

corresponding interaction tensor for dipole-dipole interactions, and $\boldsymbol{M}_j$ are the permanent multipole moments. The $\boldsymbol{T}$ and (rank-one) $\boldsymbol{M}$ tensors encompass the 13 permanent multipole moments for the AMOEBA potential (q, μ_x, μ_y, μ_z, Q_{xx}, Q_{xy}, Q_{xz}, Q_{yx}, Q_{yy}, Q_{yz}, Q_{zx}, Q_{zy}, Q_{zz}). The AMOEBA force field uses special scaling factors used for intramolecular electrostatic interactions[33] and in practice this involves treating two sets of induced dipoles for an SCF or EL calculation and an additional set of auxiliary dipoles for EL/SCF methods. Details on this complexity are discussed in Appendix C.

2.2.2 *Self-Consistent Field Method*

The SCF method for AMOEBA implemented in TINKER software package is a preconditioned conjugate gradient SCF method (PCG) using a predictor[20], while other AMOEBA implementations have used direct inversion of the iterative subspace[21, 61, 87, 94]. These more advanced SCF solvers are more efficient iterative methods that converge in fewer steps compared to more naïve methods like successive over-relaxation (SOR)[60]. In my comparisons to the various mutual induction calculation approaches such as EL[68] and hybrid EL/SCF schemes, I used the default TINKER PCG with a predictor.

2.2.3 *Extended Lagrangian Method*

I have implemented an extended Lagrangian formalism[68] (directly analogous to the Car-Parrinello approach[67]) for AMOEBA in the TINKER package that treats the polarization degrees of freedom as additional dynamic variables in the system, allowing us to integrate them on the same footing as the atomic positions and avoid using self-consistent iteration to obtain polarization near the Born-Oppenheimer surface. The extended Lagrangian for point dipoles is given by Eq. (2.2).

$$\mathcal{L}^{dipole} = \frac{1}{2}\sum_{i=1}^{N} m_i \dot{\boldsymbol{r}}_i^2 + \frac{1}{2}\sum_{i=1}^{N} m_i^{\mu} \dot{\boldsymbol{\mu}}_i^2 - U(\boldsymbol{r}, \boldsymbol{\mu}) \tag{2.2}$$

In Eq. (2.2) m_i is the mass of atom i, $\boldsymbol{r}_i$ is the position of atomic center i, $U(\boldsymbol{r}, \boldsymbol{\mu})$is the potential energy from the AMOEBA force field with the only difference being that the dipoles are now dynamically integrated and not iteratively converged. As a result there is a kinetic energy contribution from the dipoles given as the second term on the right hand side of Eq. (2), where m_i^{μ} is a fictitious dipole mass given in units of $ps^2/Å^3$. In addition I also thermostat these extended system degrees of freedom to a very low temperature (~1 K)[16, 42] to maintain the polarization close to the Born-Oppenheimer surface using Nosé-Hoover (NH) thermostats[65-66]. The complete extended system equations of motion for Nosé-Hoover temperature control on both the atomic centers and induced dipoles using a single Nosé-Hoover thermostat are given by Eq. (2.3).

$$m_i \ddot{\boldsymbol{r}}_i = \boldsymbol{F}_{\boldsymbol{r},i} - m_i \dot{\boldsymbol{r}}_i \dot{\eta} \tag{2.3a}$$

$$m_i^{\mu} \ddot{\boldsymbol{\mu}}_i = \boldsymbol{F}_{\boldsymbol{\mu},i} - m_i^{\mu} \dot{\boldsymbol{\mu}}_i \dot{\eta}_* \tag{2.3b}$$

$$Q\ddot{\eta} = \sum_{i=1}^{N} m_i \dot{\boldsymbol{r}}_i^2 - N_f k_B T \tag{2.3c}$$

$$Q_* \ddot{\eta}_* = \sum_{i=1}^{N} m_i^{\mu} \dot{\boldsymbol{\mu}}_i^2 - N_{f*} k_B T_* \tag{2.3d}$$

In Eq. (2.3) the thermostat "position", η, couples to the physical system to control the temperature and the thermostat mass, Q, is related to a characteristic time parameter, τ, by $Q = N_f k_B T \tau^2$. In these equations "*" denotes quantities associated with polarization degrees of freedom so N_f and T are the number of degrees of freedom and temperature of the atomic coordinates, respectively, and N_{f*} and T_* are the degrees of freedom and the temperature of polarization degrees of freedom, respectively, associated with a low T_*. It should be noted that $\boldsymbol{F}_{\boldsymbol{\mu},i} = \boldsymbol{E}_i - \alpha_i^{-1} \boldsymbol{\mu}_i$, which goes to 0 in the limit of a self-consistent solution of the dipoles, $\boldsymbol{\mu}_i = \alpha_i \boldsymbol{E}_i$.

2.2.4 *Simulation Details*

All results reported in this chapter are for pure water systems of 512 molecules, although all methods described should be generalizable to any molecular system. I used the water parameters of the AMOEBA14 water model[34]. All simulations started from a pre-equilibrated box and long-range electrostatics were treated with particle-mesh Ewald[95] with a real-space cutoff of 9 Å. The equations of motion of the atomic degrees of freedom were integrated using the velocity Verlet method[88] and the Nosé-Hoover formalism with a fourth-order chain was used for temperature control[66] with a τ of 0.1 ps. I used a time step of 1.0 fs for base AMOEBA simulations. For the extended Lagrangian simulations, I used a T_* of 1.0 K and a τ_* of 0.1 ps with 1500 iterations of the thermostat per step and the "mass" associated with inducible dipoles was set to 3.6e-9 ps^2/Å^3. I explored time steps from 0.25 to 1.0 fs, depending on the method used, to solve mutual induction.

2.3 Theory

I have adapted the approach developed by Niklasson and colleagues[70-81], originally formulated for BOMD, to classical mutual induction calculations. In particular, this original hybrid EL/SCF method introduced an initial guess for SCF calculations by propagating a set of auxiliary electronic degrees of freedom in a time-reversible manner. In its original form these auxiliary variables corresponded to the electronic ground state density matrix and here I adapt these to the case of classical polarization formulated as induced dipoles.

In the spirit of Niklasson *et al*.[79] one can define an extended Lagrangian given by Eq. (2.4) for the induced dipoles.

$$\mathcal{L}_{hybrid}^{dipole} = \frac{1}{2}\sum_{i=1}^{N} m_i \dot{\boldsymbol{r}}_i^2 + \frac{1}{2}\sum_{i=1}^{N} m_i^a \dot{\boldsymbol{a}}_i^2 - U(\boldsymbol{r}, \boldsymbol{\mu}_{SCF}) - \frac{1}{2}\omega^2 \sum_{i=1}^{N} m_i^a \left(\boldsymbol{\mu}_{SCF,i} - \boldsymbol{a}_i\right)^2 \tag{2.4}$$

In Eq. (2.4) $\boldsymbol{\mu}_{SCF}$ represents the set of all converged real induced dipoles, and we introduce another set of auxiliary induced dipoles, $\boldsymbol{a}$, which are the initial guesses to the iterative solution of $\boldsymbol{\mu}_{SCF}$. This auxiliary set of induced dipoles is restrained to stay near

the true self-consistent values via the final term in Eq. (2.4) using a harmonic function where m_i^a and ω are the fictitious mass and a universal frequency that determines the curvature of the harmonic well, respectively. Applying the Lagrangian equation of motion (Euler-Lagrange equation) to Eq. (2.4) in the limit of $m_i^a \longrightarrow 0$, yields the equations of motion for atomic centers and induced dipoles, given by Eq. (2.5).

$$m_i \ddot{\boldsymbol{r}}_i = -\left.\frac{\partial U(\boldsymbol{r}, \boldsymbol{\mu}_{SCF})}{\partial \boldsymbol{r}_i}\right|_{\boldsymbol{a}} \tag{2.5a}$$

$$\ddot{\boldsymbol{\mu}}_i = \omega^2\left(\boldsymbol{\mu}_{SCF,i} - \boldsymbol{a}_i\right) \tag{2.5b}$$

Eq. (2.5a) shows that equations of motion for the atomic centers are propagated in the usual way, except that the iterative solution uses an initial guess that is propagated by the auxiliary electronic degrees of freedom in Eq. (2.5b). I now integrate both equations of motion using time-reversible velocity Verlet integration[88]. To determine $\boldsymbol{\mu}_{SCF}$ we still use PCG, but now the time reversible auxiliary dipoles serve as an initial guess to the real dipoles. I chose ω to be $\sqrt{2}/\Delta t$, where Δt is the time step[72], which I set to 1.0 fs. Overall this method is simple to implement in code, simply include an additional integration for the auxiliaries, which are then fed into the SCF solver as initial guesses.

The one drawback of the original hybrid EL/SCF schemeis that over longer trajectories the propagated auxiliary dipoles systematically degrade as a reasonable initial guess for the subsequent SCF steps using reasonable time steps. Figure 2.1a shows that just beyond the 100 fs time scale the original hybrid EL/SCF[70-72] method requires an increasing number of SCF cycles, eventually reaching up to 5 SCF iterations to meet even a loose criteria of 10^{-1} D after 1.0 ns. To put that in perspective, the standard PCG scheme with a predictor requires 5 SCF cycles to reach a convergence of 10^{-6} D at all timescales. In fact the hybrid EL/SCF scheme starts at 6 SCF steps at this corresponding 10^{-6} D convergence level and increases to 8 SCF cycles over the 1.0 ns trajectory. In an attempt to alleviate this problem I considered the use of higher-order symplectic integrators, which use multiple force evaluations per time step. The dotted line in Figure 2.1a gives the results of a 4th-order integrator[96] applied to this hybrid EL/SCF scheme at the lowest level of convergence. These results show that while higher-order integration stretches the timescale of the increase of SCF cycles it does not eliminate the problem. The number of SCF cycles would be expected to continue to increase based on the data shown in Figure 2.1a.

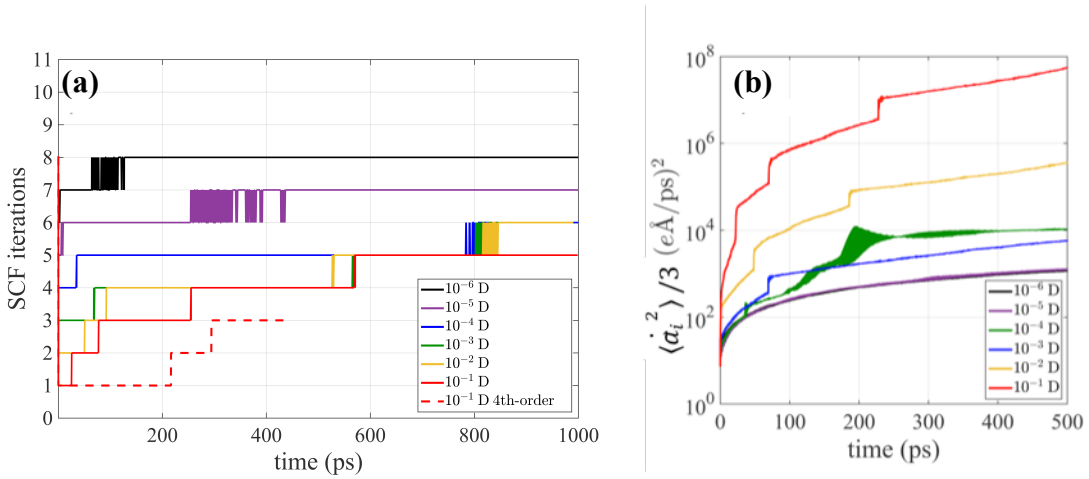


Figure 2.1: *The standard hybrid EL/SCF scheme without dissipation. (a) The number of SCF cycles needed by the EL/SCF hybrid scheme increases without limit over longer timescales since the initial guesses for the auxiliary dipoles are degrading over time. Also included at the lowest convergence are the results from a 4th-order integrator[96]. (b) The origin of the numerical instability in the standard hybrid EL/SCF scheme is that the mean squared auxiliary induced dipole velocities increase without bound, especially under loose convergence. Different curves represent different levels of convergence for the induced dipoles, given in terms of the root mean square change in the induced dipoles from one iterative step to another.*

The reason for this instability is presented in Figure 2.1b which shows that over the course of the simulation the ensemble average of the auxiliary dipole velocities $\langle \dot{\boldsymbol{a}}_i{}^2 \rangle /3$, a "pseudo temperature", increases continuously throughout the simulation, and this inertia eventually swamps the harmonic restoring force that aims to keep the auxiliary dipoles close to the real, converged dipoles, ultimately leading to instability in the equations of motion (Eq. 2.5b). This seems to be a problem with resonances[89] existing in the auxiliary dipole equations of motion that are on a faster time scale than that experienced by the real induced dipoles through the subsequent SCF solver. As Figure 2.2 shows, the auxiliary dipoles show high frequency behavior compared to their real counterparts owing to the optimal choice for their characteristic frequency, $\sqrt{2}/\Delta t$, and their direct coupling in the auxiliary potential (Eq. (2.4)) leads to corruption of the dynamics.

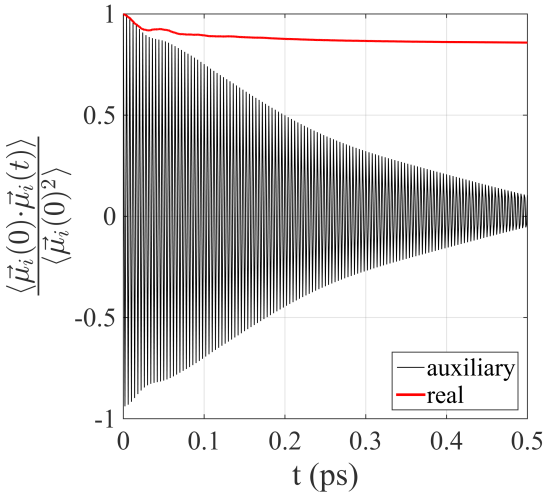


Figure 2.2: *Oxygen auxiliary dipole time correlation function. The optimal frequency, $\omega = \frac{\sqrt{2}}{\Delta t}$, of the auxiliary dipoles (black) causes the dipoles to move on a much shorter timescale than the real dipoles (red).*

Niklasson and co-workers attributed this behavior to accumulation of numerical errors throughout a simulation and sought to mitigate it with a Langevin-like scheme[73] that introduces an explicit dissipative force on the motion of the electronic degrees of freedom. The introduction of a dissipative force will inevitably lead to some time irreversibility, and hence an optimization scheme was introduced in that study to maximize stability and minimize the undesired time irreversibility that will degrade energy conservation. I have implemented the 9th-order version of this scheme and its results are given in Figure 2.3. Figure 2.3a shows that the dissipative scheme corrects for the increasing number of SCF cycles. However, Figure 2.3b and Table 2.1 show that the benefits of small numbers of SCF iterations comes at the cost of unacceptable energy drift at loose levels of convergence due to its time irreversibility. This leads to the conclusion that dissipation of pseudo kinetic energy is important in achieving a stable number of SCF iterations, but that dissipation schemes with acceptable energy conservation do not significantly reduce the number of SCF iterations relative to standard SCF solvers.

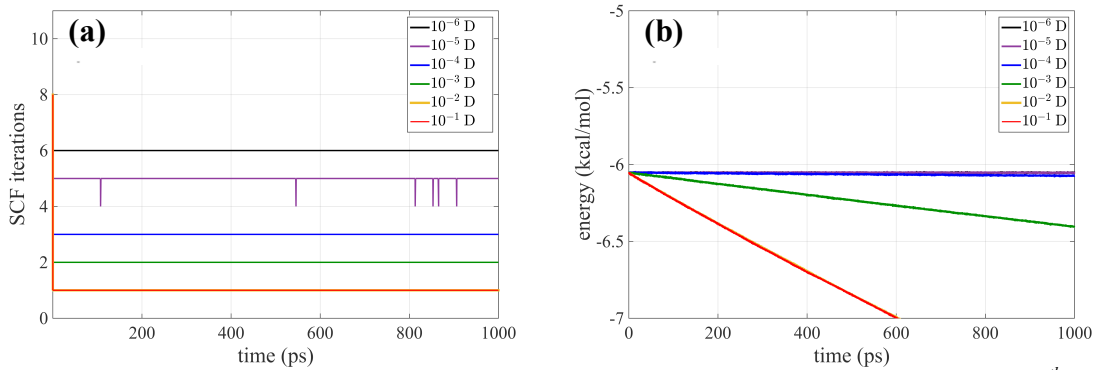


Figure 2.3: *Total energy conservation and stability using the adapted EL/SCF method with a 9th-order Langevin-like dissipative scheme from Niklasson et al[73]. (a) The required number of SCF iterations and (b) the total energy in the NVE ensemble. Different curves represent different levels of convergence for the*

induced dipoles, given in terms of the root mean square change in the induced dipoles from one iterative step to another.

Table 2.1. *Total energy drift rates for CG-SCF with predictor, the standard hybrid EL/SCF scheme with no dissipation, the hybrid EL/SCF scheme with Langevin-like dissipation described by Niklasson et al. [73], and a hybrid EL/SCF method which thermostats the auxiliary dipole velocities. Drift rates are fit to the simulated data given in Figures 2.3-2.5.*

Convergence (RMS Debye)	Energy Drift (kcal/mol/ps)				
	Standard SCF	**Hybrid EL/SCF**	**Hybrid EL/SCF with Dissipation**	**Hybrid EL/SCF with Berendsen**	**Hybrid EL/SCF with Nosé-Hoover**
10^{-6}	+4.63e-6	-6.09e-8	-4.86e-7	+1.21e-7	+6.20e-8
10^{-5}	-2.50e-5	+1.87e-7	-6.99e-7	+2.37e-7	-7.96e-8
10^{-4}	+2.10e-3	-6.62e-7	-2.08e-5	+2.48e-7	+7.96e-8
10^{-3}	+6.52e-4	-1.05e-6	-3.52e-4	+9.07e-8	-4.13e-7
10^{-2}	-1.24e-1	+5.16e-6	-1.49e-3	+2.32e-6	+2.75e-6
10^{-1}	-1.17e-1	+2.83e-4	-1.50e-3	+2.96e-5	+2.76e-5

Here I present a different solution to this instability problem by "thermostatting" the auxiliary dipoles through modification of their velocities in the time reversible velocity Verlet integration. It is known that corruption of the dynamics due to resonances can be controlled using the isokinetic ensemble[27], although one cannot formally implement an isokinetic scheme due to the $m_i^a \longrightarrow 0$ limit that yields Eqs. (2.5). Since there are no longer any contributions to the total energy from the auxiliary dipoles, and thus their kinetic energy or temperature cannot be formally defined, I nonetheless show that by rescaling the auxiliary velocities I can execute control on the mean squared velocity (or pseudo temperature), $\langle \dot{\boldsymbol{a}}_i{}^2 \rangle / 3$, that controls the buildup of this inertial pseudo kinetic energy quantity.

For iEL/SCF I have implemented both a weak coupling Berendsen velocity rescaling scheme[91] and a fourth-order time-reversible Nosé-Hoover chain (NHC) [66] to control for the inertial accumulation. In this context the distinction between Berendsen and NHC is likely unimportant since the auxiliary dipoles are serving as an initial guess to the SCF equations. Hence although Berendsen velocity rescaling does not lead to the correct limiting canonical ensemble as does the NHC scheme, I note that the primary property needed is a damping scheme for the auxiliary variables that can be formally proven to be numerically stable, close to the exact solution regardless of convergence level, minimally perturbs time reversibility, and keeps the pseudo kinetic energy of the auxiliary dipoles constant. More sophisticated schemes for isokinetic integrators would perhaps be more desirable[27, 90], but I show that the simple velocity rescaling approach is sufficient in this case, since I am thermostatting the auxiliary variables whose only role is to provide an initial guess to the SCF solver of the real inducible dipoles.

Here I define the Berendsen rescaling factor, α, which scales the velocities propagated by a reversible velocity Verlet integration at each time step in the weak coupling regime in Eq. (2.6).

$$\alpha = \sqrt{1 + \frac{\Delta t}{\tau}\left(\frac{T_{aux}}{\langle \dot{\boldsymbol{\mu}}_i{}^2 \rangle / 3} - 1\right)} \tag{2.6}$$

In Eq. (2.6) τ is a rescaling timescale parameter and T_{aux} is the set pseudo temperature of the auxiliary induced dipoles that corresponds to the desired value of $\langle \dot{\boldsymbol{a}}_i{}^2 \rangle$ and has units of $e^2\text{Å}^2/ps^2$. The pseudo temperature chosen for the auxiliary dipoles is chosen to approximately conform to equipartition of energy consistent with a classical harmonic oscillator given the form of the auxiliary dipole potential given in Eq. (2.7).

$$\frac{1}{2}N\omega^2 \langle \left(\boldsymbol{\mu}_{SCF,i} - \boldsymbol{a}_i\right)^2 \rangle + \frac{1}{2}N\langle \dot{\boldsymbol{a}}_i{}^2 \rangle = NT_{aux} \tag{2.7}$$

The maximum auxiliary dipole velocity can be estimated by using the square of the maximum displacement of the real induced dipole distribution to approximate $\langle \left(\boldsymbol{\mu}_{SCF,i} - \boldsymbol{a}_i\right)^2 \rangle$. Using $T_{aux} = \langle \dot{\boldsymbol{a}}_i{}^2 \rangle / 3$ and $\omega^2 = \frac{2}{\Delta t^2}$ with $\langle \left(\boldsymbol{\mu}_{SCF,i} - \boldsymbol{a}_i\right)^2 \rangle \sim$ (0.2 eÅ)^2, gives a pseudo temperature of ~10^5 $e^2\text{Å}^2/ps^2$ which is what I use here. A discussion of the numerical stability of the iEL/SCF method is given in Appendix A.

2.4 Results

In what follows I characterize the relative performance of the PCG SCF solver with predictor, the standard EL method, the hybrid EL/SCF scheme with no dissipation, and my new iEL/SCF scheme over a 1.0 ns simulation of the AMOEBA14 polarizable force field for water in the NVE and NVT ensembles. In all cases I use the PCG solver with predictor in which dipoles are converged to 10^{-6} D as the gold standard, with the understanding that over longer timescales there will be a noticeable energy drift with the PCG approach even at this relatively tight dipole convergence tolerance.

Figure 2.4 shows energy conservation in the NVE ensemble for a standard PCG scheme with predictor, a standard EL scheme, and the original EL/SCF scheme with no dissipation. On the nanosecond time scale examined here, one can see a significant energy drift for CG-SCF at a convergence level of 10^{-4} D and a severe lack of energy conservation with even looser convergence criteria (Figure 2.4a). Although the energy conservation of the SCF scheme looks stable for convergence levels of 10^{-5} D or tighter over the 1.0 ns trajectory, in fact the energy will eventually drift over longer timescales due to its one-sided convergence and use of a predictor. The standard EL scheme shows poor energy conservation, even for small time steps of 0.25 fs and 0.5 fs for which the EL method is stable (Figure 2.4b) and time steps larger than 0.5 fs are numerically unstable.

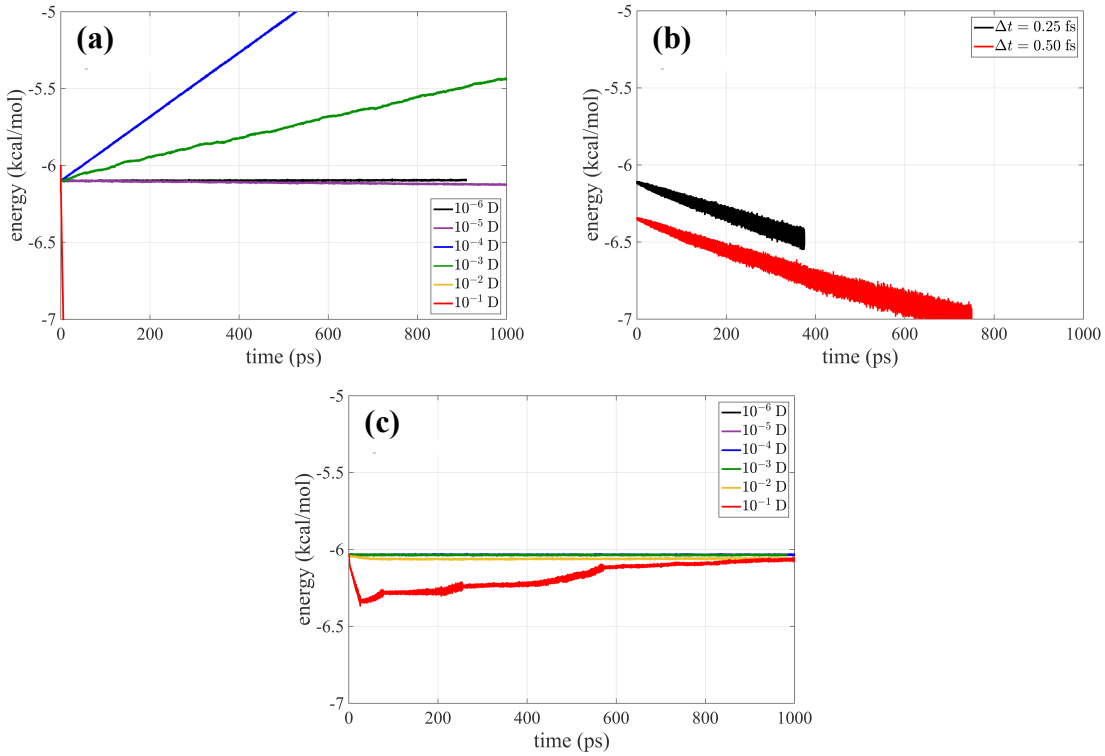


Figure 2.4: *Total energy conservation in the NVE ensemble (a) using a PCG solver with predictor, (b) using a basic EL method, and (c) using the standard EL/SCF method with no dissipation. Different curves represent different levels of convergence for the induced dipoles, given in terms of the root mean square change in the induced dipoles from one iterative step to another.*

For the hybrid EL/SCF without dissipation (Figure 2.4c), there is very good energy conservation up to a convergence level of 10^{-3} D, but for looser levels of convergence the system energy experiences continual adjustment due to the poor quality of the SCF guess generated by the auxiliary dipoles. These deviations do not appear to be divergent and/or large in magnitude compared to the standard SCF and EL methods, and eventually settle down at the end of the 1.0 ns trajectory. However, although there is some improvement in energy conservation for hybrid EL/SCF without dissipation compared to the standard PCG with predictor and EL schemes at a given level of convergence of the mutual induction, it comes at the cost of increasing numbers of SCF cycles due to the accumulated inertia over time (Figure 2.1a). In fact, the rebounding energy profile at the lowest level of induced dipole convergence for the no dissipation scheme corresponds exactly to a systematic bump up in the number of SCF cycles as the accumulated pseudo kinetic energy in the initial SCF guess gets worse over time (Figure 2.1a). Even so, the hybrid EL/SCF without dissipation outperforms the standard EL and SCF approaches when we fit the energy conservation profile to derive total energy drift rates (Table 2.1).

Figure 2.5 shows energy conservation for iEL/SCF using weak coupling Berendsen velocity rescaling and Nosé-Hoover chains (NHC) for controlling $\langle \dot{\boldsymbol{a}}_i{}^2 \rangle/3$, for which mutual induction tolerances less than 10^{-2} D using either thermostatting method show superior energy conservation compared to PCG at tight tolerances greater than 10^{-5} D, with the significant additional benefit that fewer SCF steps are required. Using the

iEL/SCF method only 4 SCF steps are needed for 10^{-2} D and 3 SCF steps for 10^{-1} D, compared to the 5 SCF steps needed by PCG for a convergence level of 10^{-6} D. Both velocity attenuation methods show superior energy conservation compared to all methods over the full range of induced dipole convergence (Table 2.1), and corrects for the "rebounding" energy behavior seen in the non-dissipative EL/SCF method at low convergence levels (Figure 2.4c). I note that both thermostatting methods begin to show systematic energy drift at 10^{-1} D levels of convergence. Since the NHC thermostat is formally formulated as a time reversible scheme, I conclude that an inherent limit to how low the convergence tolerance in the SCF solutions can be set, although the energy drift is at least one, if not many, orders of magnitude better than any other possible scheme at 3 SCF steps for 10^{-1} D.

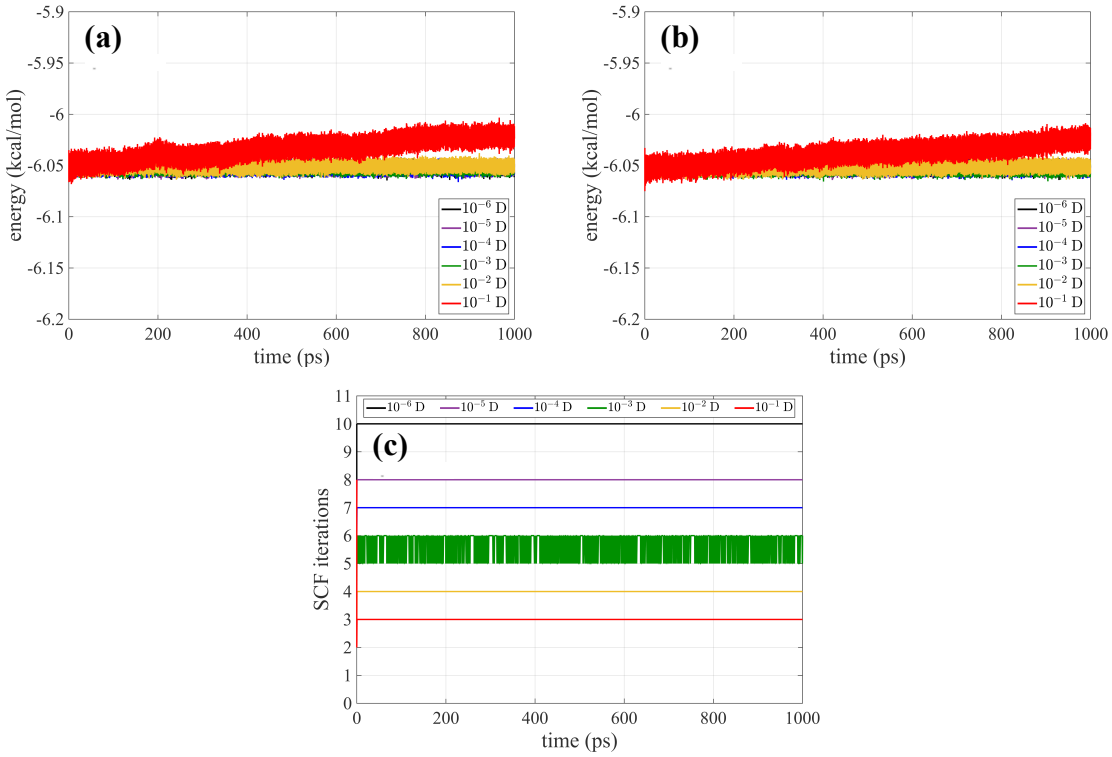
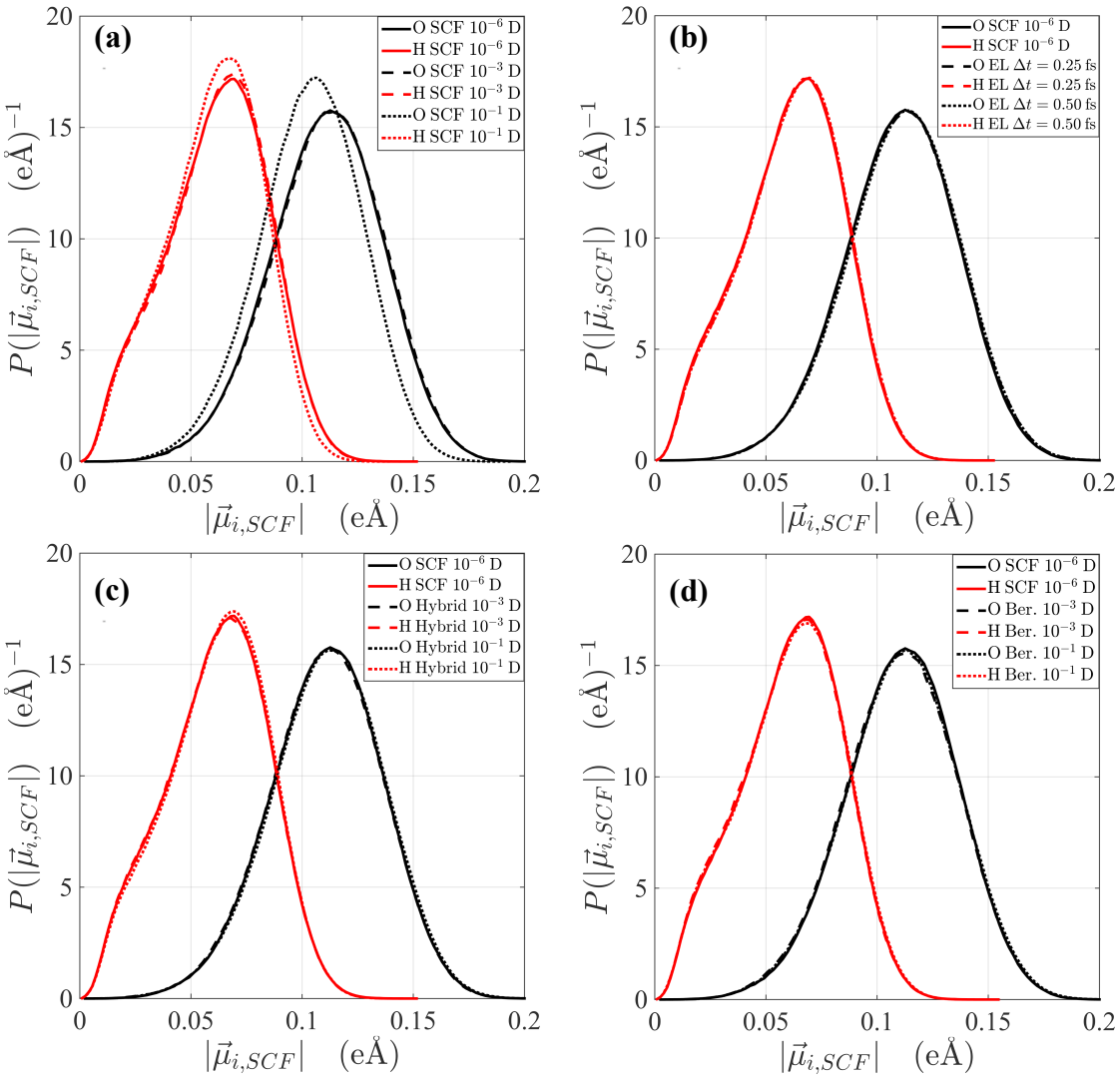


Figure 2.5: *Total energy conservation and stability of the adapted EL/SCF method with pseudo temperature control of the auxiliary dipoles. (a) Total energy in the NVE ensemble using Berendsen control, (b) total energy conservation in the NVE ensemble using Nosé-Hoover control, and (c) the number of SCF cycles needed by the EL/SCF hybrid scheme for a given convergence level. Different curves represent different levels of convergence for the induced dipoles, given in terms of the root mean square change in the induced dipoles from one iterative step to another. It is evident that the effective thermostats make even the 10^{-1}D level of convergence using the hybrid EL/SCF scheme superior to any known method for the classical mutual induction calculation, requiring only 3 SCF cycles to reach reasonable energy conservation.*

Next I consider in detail how the induced dipole polarization is accounted for in each method within an NVT simulation at 298.0 K. Figure 2.6 shows the probability density function for the induced dipole magnitude and Figure 2.7 shows the normalized dipole autocorrelation function at short times (< 1 ps) for the AMOEBA14 water model. Appendix B provides the probability density of the in-plane and out-of-plane induced

dipole angle as well. One can see that PCG with a predictor shows significant deviations from the gold standard (10^{-6} D) starting at convergence levels of the induced dipoles at 10^{-3} D. Whereas the standard EL method shows good agreement with the SCF converged result for time steps of 0.25 fs and 0.5 fs, it fails completely at larger time steps due to numerical instability. By contrast, the hybrid EL/SCF scheme without dissipation reproduces correct polarization probability distributions and the dipole autocorrelation well, even at 10^{-1} D, although, again, this comes at the cost of increasing numbers of SCF cycles. The iEL/SCF approach developed here also reproduces all polarization properties well at loose convergence and uses the fewest number of SCF cycles to do so. All of the methods show some degradation in the range of polarization properties at a given level of convergence when utilized in the NVE ensemble (Figures B.3, B.4, B.5, B.6, and Table B.1 in Appendix B), which is not unexpected since thermostatting of the real system variables can mask underlying numerical problems in integrators and/or poorly converged energy and forces. However, the hybrid EL/SCF schemes are superior to the standard EL and SCF at any dipole convergence tolerance.



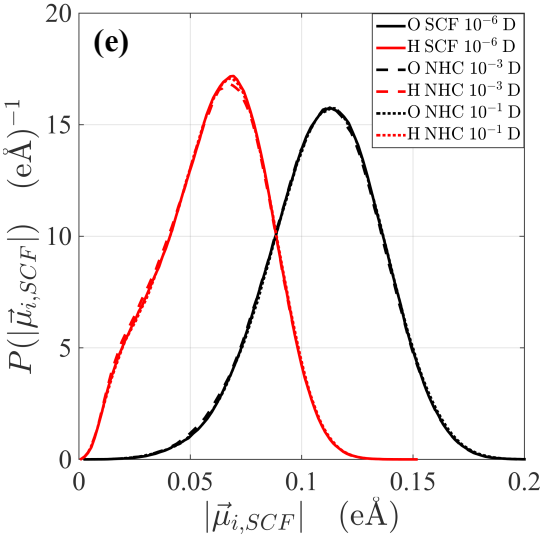
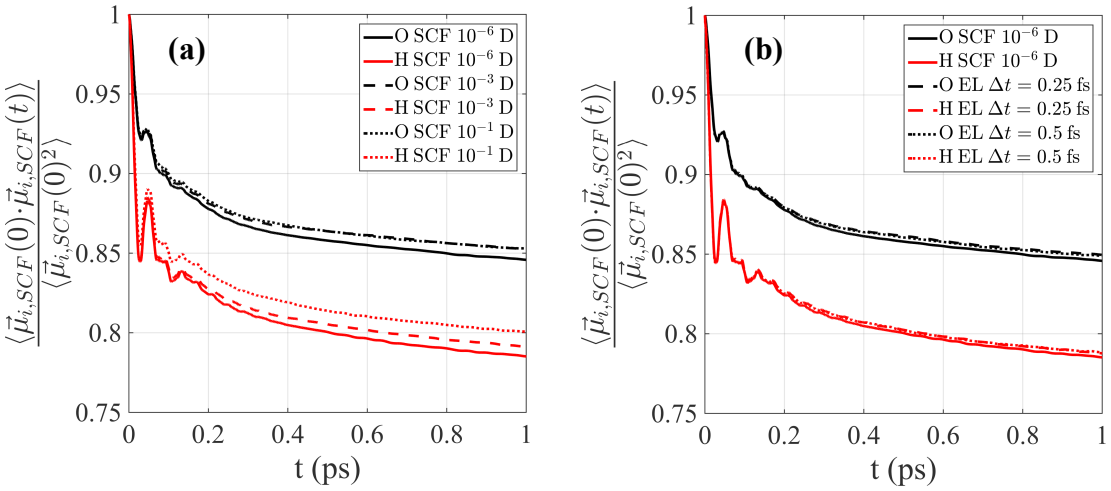


Figure 2.6: *Comparison of the ensemble averaged probability distributions for induced dipole magnitude. Using (a) standard PCG solver with predictor (SCF), (b) basic EL method, (c) the adapted EL/SCF method with no dissipation (hybrid), (d) the adapted EL/SCF method with pseudo temperature control of the auxiliary dipoles using Berendsen (Ber.) rescaling, and (e) the adapted EL/SCF method with pseudo temperature control of the auxiliary dipoles using a Nosé-Hoover thermostat (NHC). The continuous curves give the data obtained using base AMOEBA at a level of 10^{-6} RMS Debye convergence.*



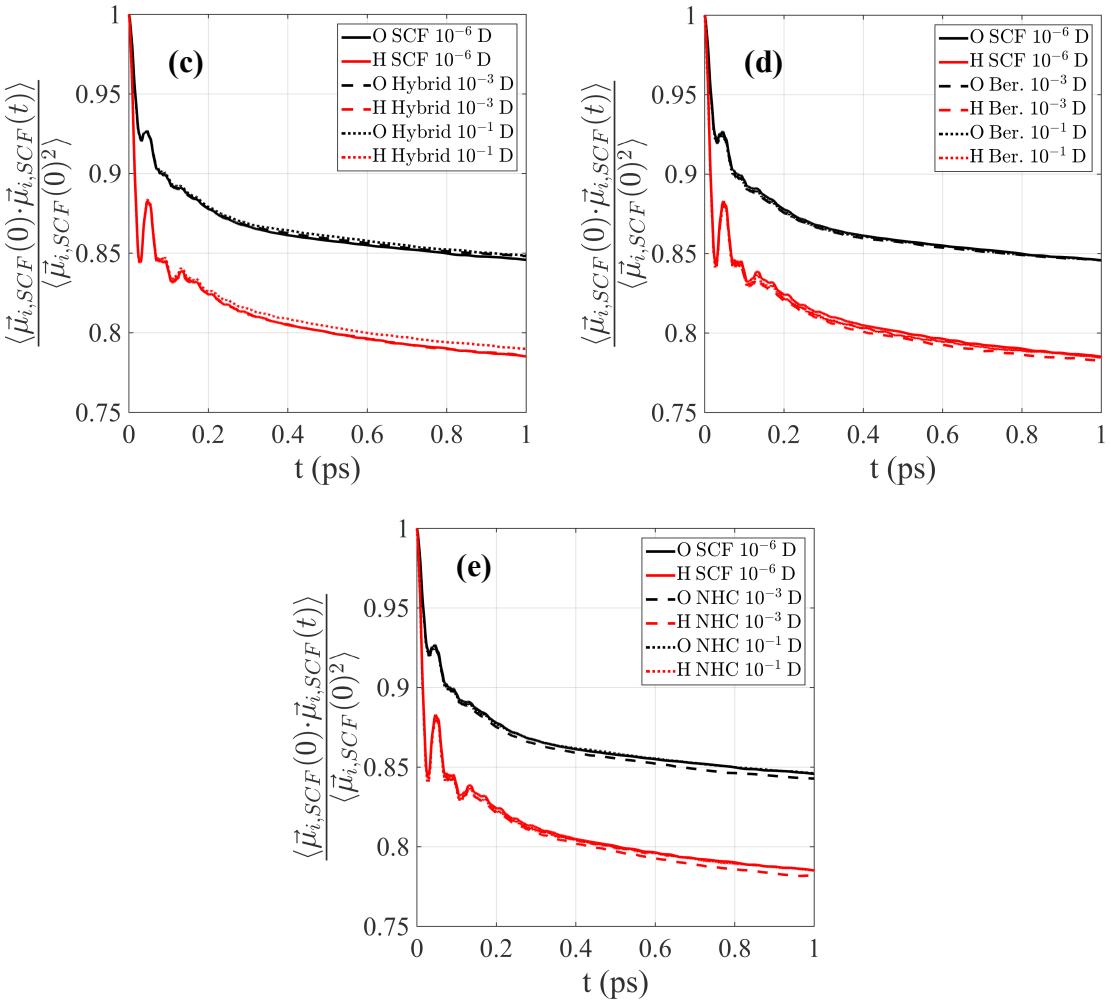


Figure 2.7: *Comparison of the normalized induced dipole time autocorrelation function. Using (a) standard PCG solver with predictor (SCF), (b) basic EL method, (c) the adapted EL/SCF method with no dissipation (hybrid), (d) the adapted EL/SCF method with pseudo temperature control of the auxiliary dipoles using Berendsen (Ber.) rescaling, and (e) the adapted EL/SCF method with pseudo temperature control of the auxiliary dipoles using a Nosé-Hoover thermostat (NHC). The continuous curves give the data obtained using base AMOEBA at a level of 10^{-6} D convergence.*

Table 2.2 reports the average potential energy and molecular dipole which were calculated from a single NVT simulation at 298.0 K, as well as the diffusion coefficient, calculated by taking independent snapshots from the 298.0 K NVT simulation and calculating the molecular mean squared displacement in the NVE ensemble. The SCF scheme gives correct average potential energies and molecular dipoles down to a relatively loose convergence of 10^{-2} D, a benefit of thermostatting in regards thermodynamic quantities, but gives incorrect diffusion coefficients starting at a much tighter tolerance of 10^{-4} D. For the EL method, even using a smaller time step than the other two approaches, we see more significant differences between the average potential energies and molecular dipoles compared to the best converged SCF result, and poor results for the diffusion coefficient. Across all convergence levels all EL/SCF schemes reproduce the thermodynamic and kinetic data quite well, even producing a diffusion coefficient within error bars of the gold standard at much looser convergence levels.

Although ultimately the computational cost of the increasing number of SCF cycles is early evidence for eventual numerical instability in the non-dissipative EL/SCF, the iEL/SCF method retains excellent property performance with a stable algorithm that conserves energy at loose convergence and minimizes the number of SCF cycles relative to the standard PCG with predictor method.

Table 2.2. *Average potential energy, average molecular dipole moment, and diffusion coefficients as a function of mutual induction convergence for PCG with predictor (standard SCF), standard EL, the hybrid EL/SCF scheme with no dissipation (standard EL/SCF), and a hybrid EL/SCF method that thermostats the auxiliary dipole velocities. Average potential energy and molecular dipole were calculated from NVT simulations at 298.0 K. Diffusion coefficients were averaged over multiple NVE simulation using independent snapshots from 298.0 K NVT simulations as the initial condition.*

Standard SCF			
Convergence (RMS Debye)	**Average Potential Energy (kcal/mol)**	**Average Molecular Dipole (Debye)**	**Diffusion Coefficient (10^5 cm^2/s)**
10^{-6}	-8.84±0.09	2.742±0.014	2.22±0.29
10^{-5}	-8.83±0.08	2.742±0.012	2.26±0.14
10^{-4}	-8.84±0.08	2.744±0.013	3.45±0.28
10^{-3}	-8.83±0.09	2.743±0.013	2.71±0.22
10^{-2}	-8.84±0.09	2.743±0.013	0.0019±0.00021
10^{-1}	-8.67±0.09	2.703±0.013	0.0020±0.00025
Standard EL			
Time step (fs)	**Average Potential Energy (kcal/mol)**	**Average Molecular Dipole (Debye)**	**Diffusion Coefficient (10^5 cm^2/s)**
0.25	-8.97±0.08	2.753±0.012	1.28±0.14
0.50	-8.92±0.08	2.746±0.012	1.22±0.17
Standard EL/SCF			
Convergence (RMS Debye)	**Average Potential Energy (kcal/mol)**	**Average Molecular Dipole (Debye)**	**Diffusion Coefficient (10^5 cm^2/s)**
10^{-6}	-8.83±0.08	2.742±0.013	2.39±0.21
10^{-5}	-8.84±0.09	2.743±0.013	2.25±0.17
10^{-4}	-8.83±0.09	2.742±0.014	2.16±0.18
10^{-3}	-8.83±0.09	2.742±0.013	2.23±0.16
10^{-2}	-8.83±0.09	2.743±0.013	2.25±0.13
10^{-1}	-8.84±0.09	2.743±0.013	2.09±0.12
Hybrid EL/SCF with Berendsen			
Convergence (RMS Debye)	**Average Potential Energy (kcal/mol)**	**Average Molecular Dipole (Debye)**	**Diffusion Coefficient (10^5 cm^2/s)**
10^{-6}	-8.84±0.09	2.744±0.013	2.17±0.15
10^{-5}	-8.84±0.09	2.743±0.013	2.25±0.15
10^{-4}	-8.83±0.09	2.742±0.013	2.21±0.16
10^{-3}	-8.83±0.09	2.743±0.013	2.28±0.13
10^{-2}	-8.84±0.08	2.743±0.013	2.17±0.13
10^{-1}	-8.83±0.08	2.742±0.013	2.28±0.13
Hybrid EL/SCF with Nosé-Hoover			
Convergence (RMS Debye)	**Average Potential Energy (kcal/mol)**	**Average Molecular Dipole (Debye)**	**Diffusion Coefficient (10^5 cm^2/s)**
10^{-6}	-8.83±0.09	2.742±0.013	2.36±0.14
10^{-5}	-8.83±0.09	2.743±0.013	2.30±0.14

10^{-4}	-8.84±0.08	2.743±0.013	2.27±0.15
10^{-3}	-8.84±0.09	2.743±0.013	2.43±0.18
10^{-2}	-8.83±0.08	2.742±0.013	2.37±0.20
10^{-1}	-8.84±0.08	2.744±0.013	2.17±0.11

2.5 Conclusions

I have presented a new adaptation of a hybrid EL/SCF scheme applied to induced dipole polarization in classical simulations that overcomes numerical instability problems that were also observed in the equations of motion for the auxiliary electronic degrees in BOMD simulations[73]. However, instead of using a dissipative force that can compromise time reversibility of the auxiliary induced dipole initial guess, I have introduced the use of thermostats applied to the auxiliary dipole velocities to control for the accumulation of a pseudo kinetic energy due to resonances. This approach simultaneously preserves energy conservation and numerical stability at loose real dipole convergence values of 10^{-1} D, thereby requiring fewer SCF cycles to describe polarization and system properties accurately when compared to a typical SCF convergence of 10^{-5} to 10^{-6} D. For the AMOEBA14 water model this also translates to an overall decrease in the required number of SCF iterations per time step from 5 to 3, resulting in some computational savings.

I have addressed a problem that arises from the fact that the dynamics of the auxiliary dipoles evolve on a faster timescale than the time evolution of the real induced dipoles, and that resonances arise due to their coupling through the potential energy term in Eq. (2.4). This leads to a build up of inertia in the auxiliary dipoles that can be solved by either reducing the time step for their equation of motion or by controlling it through pseudo temperature control as I have done here. My iEL/SCF approach is a simple velocity attenuation scheme that effectively removes the effect of these resonances while still maintaining time-reversibility (within that allowed by the numerical integration scheme) depending on the thermostat method chosen. I am thus able to have better energy conservation and can reproduce molecular properties, even at very loose levels of convergence with only a small number of SCF iterations required. The method is also simple to implement, require few modifications to existing code structures.

While the iEL/SCF method has been demonstrated on a classical inducible dipole model AMOEBA (and is available in the TINKER program), it has also been proven to be useful in the context of *ab initio* molecular dynamics. The inertial restraining approach of applying thermostats to auxiliary degrees of freedom has been applied to the density in a linear scaling density functional theory (DFT) implementation in the ONETEP software package[92]. In that work it was shown that using an iEL/SCF approach, as described in this chapter, and a Langevin-like friction dissipative approach for EL/SCF[73] provide enhancements over standard SCF solutions for *ab initio* molecular dynamics. One of the interesting outcomes in comparing the dissipative iEL/SCF approaches when applied to classical electric dipoles versus the quantum electron density is that while both schemes perform equally well in the latter, only the iEL/SCF method works in the case of classical polarization. More specifically, the small amount of broken time reversibility is negligible compared to the error in the DFT forces, whereas classical polarization forces are close to exact so that the undesirable dissipative forces that destroy time reversibility

are much more prominent. The primary point is that new approaches to solving for mutual polarization are held to a high level of stringency when implemented for classical polarization and suggest that what works classically should perform well in the context of the quantum electron density.

3.
iEL/0-SCF: Eliminating Iteration for Polarization

3.1 Introduction

In this chapter I present a solution for classical polarization that does not require any self-consistent field iterations[22], the aspect of classical polarization that makes it computationally expensive. The method builds upon a time-reversible extrapolation scheme for an auxiliary set of dipoles that drives the evolution of the set of real induced dipoles, iEL/SCF presented in Chapter 2, to achieve truly SCF-less dynamics. I therefore call the method iEL/0-SCF (inertial extended Lagrangian/no iteration self-consistent field). While iEL/SCF, presented in Chapter 2, is simple to implement it merely reduces iterations required for standard SCF procedures. The iEL/0-SCF method presented in this chapter is more sophisticated, as it requires more code modification, but eliminates the need for an SCF solver completely.

Both methods are fundamentally based on work by Niklasson and colleagues[70-79, 81], which starts from the broken time-reversal symmetry problem in Born-Oppenheimer molecular dynamics (BOMD) to derive a time-reversible extrapolation scheme for the electronic degrees of freedom. Equivalently, it can be formulated in the form of an extended Lagrangian in which an additional set of auxiliary electronic degrees of freedom are propagated alongside the nuclei with the purpose of generating either good quality time-reversal guesses for the SCF calculations, or as a stand-alone SCF-free extended Lagrangian formulation of BOMD. It is known that the numerical error in the SCF solutions (which are never exact due to incomplete convergence and/or an approximate numerical algebra) will leak back to the auxiliary degrees of freedom and give rise to instabilities in their equations of motion. To address this problem Niklasson and co-workers proposed a dissipative integration scheme for the equations of motion of the electronic degrees of freedom using a modified form of the Verlet algorithm that controls the numerical instability at the expense of a small amount of broken time-reversibility[66, 73]. In Chapter 2 I presented an alternative method that couples the auxiliary velocities to a thermostat to prevent buildup of inertia, thus preserving time-reversibility through the use of a time-reversible solution to Nosé-Hoover thermostat variables for the auxiliary degrees of freedom[23]. For condensed phase simulation this thermostatting approach proves to be superior[92]. This same technique is applied in this chapter for the iEL/0-SCF method.

I find that the iEL/0-SCF approach is as accurate as standard SCF approaches for pure water, dilute to concentrated salt solutions, and small peptides and large proteins in water. In the TINKER implementation of the iEL/0-SCF approach I show that it also scales better under OpenMP (shared memory) parallelization than the standard SCF solvers and offers significant computational savings over more expensive SCF calculations based on traditional single time scale 1.0 fs time step integration using

symplectic integrators, and is as fast as and more accurate than a reversible reference system propagator (RESPA) algorithm with an outer 2.0 fs time step.

3.2 Theory

To achieve a polarization scheme that does not require self-consistent field (SCF) iterations, one can start with a Lagrangian for a classical system of *N* atoms and examine how self-consistently iterated induced dipoles on each atom contribute to the potential, given in Eq. (3.1).

$$\mathcal{L}(\boldsymbol{r}, \dot{\boldsymbol{r}}) = \frac{1}{2}\sum_{i=1}^{N} m_i \dot{\boldsymbol{r}}_i^2 - U(\boldsymbol{r}) \tag{3.1}$$

In Eq. (1) $\boldsymbol{r}_i$ and $\dot{\boldsymbol{r}}_i$ are the position and velocity of the *i*-th atom, respectively, and $U(\boldsymbol{r})$ is the potential energy for the current atomic configuration. Here the potential energy can be broken into both a many-body polarization energy contribution, $U^{polar}(\boldsymbol{r})$, and all other contributions, $U^{other}(\boldsymbol{r})$, which can include non-bonded terms such as permanent electrostatic and van der Waals interactions and bonded valence terms like bond-stretching, angle-bending, torsionals, and others, as shown in Eq. (3.2).

$$U(\boldsymbol{r}) = U^{other}(\boldsymbol{r}) + U^{polar}(\boldsymbol{r}) \tag{3.2}$$

The component of the potential energy that depends on polarization is then given by Eq. (3.3).

$$U_{SCF}^{polar}(\boldsymbol{r}) = -\frac{1}{2}\boldsymbol{\mu}_{\boldsymbol{SCF}}^{\boldsymbol{T}}\boldsymbol{E} \tag{3.3}$$

Eq. (3.3) assumes complete SCF convergence, which meets the condition given by Eq. (3.4).

$$\boldsymbol{\mu}_{SCF,i} = \alpha_i \boldsymbol{E}_i + \alpha_i \sum_{j=1}^{N} \boldsymbol{T}'_{ij}\boldsymbol{\mu}_{SCF,j} \tag{3.4}$$

In Eq. (3.4) $\boldsymbol{E}_i$ is the electric field contribution at site *i* due to permanent electrostatics in the system (fixed charges, dipoles, quadrupoles, etc.), $\boldsymbol{T}'_{ij}$ is the dipole-dipole interaction matrix between sites *i* and *j*, and α_i, is the atomic polarizability of the *i*-th atom. Here Eq. (3.4) can be recast into Eq. (3.5), which uses a super matrix $\boldsymbol{C}$.

$$\boldsymbol{\mu}_{\boldsymbol{SCF}} = \boldsymbol{C}^{-\boldsymbol{1}}\boldsymbol{E} \tag{3.5}$$

In Eq. (3.5) the blocks of the super matrix $\boldsymbol{C}$ are given by $\boldsymbol{C}_{ij} = (\alpha_j^{-1}\delta_{ij} - \boldsymbol{T}'_{ij})$. While the induced dipoles can be solved exactly through matrix inversion, as is done for three-body interactions in the 3-AMOEBA model[84-85], in practice it is prohibitively

expensive to perform this calculation for all but the smallest systems as it involves inverting the 3*N* by 3*N* $\boldsymbol{C}$ super matrix.

Usually Eq. (3.5) is solved iteratively to obtain the set of self-consistent field induced dipoles, $\boldsymbol{\mu}_{SCF}$, through standard self-consistent methods to some tolerance, usually with a predictor to accelerate the convergence of the SCF problem. However, predictors that use information from previous steps break time reversibility in the context of molecular dynamics, leading to an inevitable degradation in energy conservation. An alternative to solving for the many-body electronic problem that avoids a self-consistent solution was originally addressed by Car and Parrinello[67]. In their seminal work they formulate an extended Lagrangian (EL) to evolve the electronic degrees of freedom dynamically by introducing fictitious masses, and the EL approach was extended to induced dipole polarization by Wodak and co-workers[68]. Overall, standard EL solutions to both quantum[67] and classical[42-43] electronic degrees of freedom are found to be stable and conserve energy, but only for reduced and sometimes very small time steps that precludes their practical use in molecular simulation[23].

Niklasson *et al.* tackled the problem of broken time-reversibility of the electronic dynamics through a distinct formalism[70] that can also be expressed in a Lagrangian formulation[73, 79], which was adapted for classical polarization in Chapter 2[23], as shown in Eq. (3.6).

$$\mathcal{L}'(\boldsymbol{r}, \dot{\boldsymbol{r}}, \boldsymbol{a}, \dot{\boldsymbol{a}}) = \frac{1}{2}\sum_{i=1}^{N} m_i \dot{\boldsymbol{r}}_i^2 + \frac{1}{2}\sum_{i=1}^{N} m_i^a \, \dot{\boldsymbol{a}}_i^2 - U(\boldsymbol{r}, \boldsymbol{\mu}_{SCF}) - \frac{1}{2}\omega^2 \sum_{i=1}^{N} m_i^a \left(\boldsymbol{\mu}_{SCF,i} - \boldsymbol{a}_i\right)^2 \tag{3.6}$$

In Eq. (3.6) $\boldsymbol{a}_i$ and $\dot{\boldsymbol{a}}_i$ are the auxiliary induced dipole and corresponding velocity for atom *i*, each of which has a fictitious mass associated with it, m_i^a. The additional terms in the Lagrangian are now a kinetic energy associated with the auxiliary dipoles and a harmonic potential that aims to keep the auxiliary dipoles close to the true SCF dipoles, which in turn is determined by the steepness of this harmonic well and characterized by a frequency ω. It has been shown that for a time-reversible Verlet integration algorithm with a finite integration time step Δt, the maximum stable value of ω is $\sqrt{2}/\Delta t$[79]. These auxiliary variables then serve as good initial guesses to the SCF solution for the real dipoles to reach the final solution, $\boldsymbol{\mu}_{SCF,i}$, and with notable improvements in energy conservation[23, 72]. However, in practice the iterative solution of the induced dipoles never reaches the exact solution, as the iteration is stopped at some convergence threshold. This error couples to the auxiliary degrees of freedom through the potential (last term in Eq. (3.6)), which corrupts the auxiliary dynamics such that they become more and more poor initial guesses for the real degrees of freedom, and the number of SCF cycles increases without bound, as demonstrated in Chapter 2[23].

Niklasson proposed the introduction of dissipation for this numerical error that achieves only a small amount of broken time-reversibility since it is introduced at an order commensurate with the integration error[73]. Unfortunately when the dissipative force is combined with the accurate classical polarization force, the dissipation scheme unambiguously exhibits energy drift. An alternative approach taken in Chapter 2 introduced a simple thermostatting scheme, illustrated using both Berendsen weak coupling[91] and Nosé-Hoover chain thermostats[65-66], applied to the auxiliary dipole velocities[23]. The latter case was shown to provide superior energy conservation with less

stringent convergence thresholds and a correspondingly small number of SCF cycles, to reproduce all properties of the classical polarization model in the NVT and NVE ensembles accurately.

Now turning to the problem of achieving iteration-free dynamics for the electronic degrees of freedom, given this background, I will introduce modifications to the Lagrangian in Eq. (3.6) in two important ways. Although for completely converged induced dipoles the requirement that $\boldsymbol{C\mu - E = 0}$ must be met, I introduce a general form for the polarization potential energy that does not assume convergence of the real induced dipoles, given in Eq. (3.7) [82-83].

$$U^{polar}(\boldsymbol{r}, \boldsymbol{a}) = \frac{1}{2}\boldsymbol{\mu^T C\mu - \mu^T E} \tag{3.7}$$

This means that if the real dipole differs from the exact self-consistent solution by a small amount $\boldsymbol{\delta}$, $\boldsymbol{\mu = \mu_{SCF} + \delta}$, the polarization potential differs from the exact self-consistent solution by an order of the square of the error, $|\boldsymbol{\delta}|^2$. To see this one can substitute $\boldsymbol{\mu = \mu_{SCF} + \delta}$ into Eq. (3.7), yielding Eq. (3.8).

$$U^{polar} = \tfrac{1}{2}(\boldsymbol{\mu_{SCF} + \delta})^T\boldsymbol{C}(\boldsymbol{\mu_{SCF} + \delta}) - (\boldsymbol{\mu_{SCF} + \delta})^T\boldsymbol{E} = -\tfrac{1}{2}\boldsymbol{\mu_{SCF}}^T\boldsymbol{E} + \tfrac{1}{2}\boldsymbol{\delta^T C\delta} \tag{3.8}$$

Eq. (3.8) makes use of the requirement that for completely converged induced dipoles $\boldsymbol{C\mu_{SCF} - E = 0}$. The difference between the general potential and the exact solution is given by Eq. (3.9).

$$U^{polar} - U^{polar}_{SCF} = \tfrac{1}{2}\boldsymbol{\delta^T C\delta} = O(|\boldsymbol{\delta}|^2) \tag{3.9}$$

Eq. (3.9) implies that for any small errors in the induced dipoles that result from no iteration, the error in the potential is pushed to second order. The same analysis applies to the gradient of this potential, which affects the forces for dynamics and is what is relevant in molecular dynamics. After differentiating Eq. (3.8) one obtains Eq. (3.10).

$$\frac{dU^{polar}}{\boldsymbol{dr}} = \tfrac{1}{2}\boldsymbol{\mu_{SCF}}^T\frac{\boldsymbol{\partial C}}{\boldsymbol{\partial r}}\boldsymbol{\mu_{SCF}} - \boldsymbol{\mu_{SCF}}^T\frac{\boldsymbol{\partial E}}{\boldsymbol{\partial r}} + \tfrac{1}{2}\boldsymbol{\delta^T}\frac{\boldsymbol{\partial C}}{\boldsymbol{\partial r}}\boldsymbol{\delta} \tag{3.10}$$

Taking the gradients of the fully converged potential and SCF condition yields Eqs. (3.11a) and (3.11b), respectively.

$$\frac{dU^{polar}_{SCF}}{\boldsymbol{dr}} = \tfrac{1}{2}\boldsymbol{\mu_{SCF}}^T\frac{\boldsymbol{\partial C}}{\boldsymbol{\partial r}}\boldsymbol{\mu_{SCF}} - \boldsymbol{\mu_{SCF}}^T\frac{\boldsymbol{\partial E}}{\boldsymbol{\partial r}} \tag{3.11a}$$

$$\left(\frac{\boldsymbol{\partial C}}{\boldsymbol{\partial r}}\boldsymbol{\mu_{SCF}} + \boldsymbol{C}\frac{\boldsymbol{\partial \mu_{SCF}}}{\boldsymbol{\partial r}} - \frac{\boldsymbol{\partial E}}{\boldsymbol{\partial r}}\right) = \boldsymbol{0} \tag{3.11b}$$

One can now confirm that errors in the forces are also second order by combining Eq. (3.10) and (3.11).

$$\frac{dU^{polar}}{d\boldsymbol{r}} - \frac{dU^{polar}_{SCF}}{d\boldsymbol{r}} = \frac{1}{2}\boldsymbol{\delta}^{\boldsymbol{T}}\frac{\partial \boldsymbol{C}}{\partial \boldsymbol{r}}\boldsymbol{\delta} = O(|\boldsymbol{\delta}|^2) \tag{3.12}$$

Like the iEL/SCF procedure, the real induced dipoles will depend explicitly on using the auxiliary dipoles as an initial guess, given by Eq. (3.13).

$$\boldsymbol{\mu_i} = \alpha_i \boldsymbol{E}_i + \alpha_i \sum_{j=1}^{N} \boldsymbol{T}'_{ij}\boldsymbol{a}_j \tag{3.13}$$

But in this case no SCF cycles are performed and instead $\boldsymbol{\mu}_{SCF,i}$ is estimated from simple linear mixing of the real and auxiliary induced dipoles via a local kernel approximation, given by Eq. (3.14) [79].

$$\boldsymbol{\mu}_{SCF,i} \approx \gamma\boldsymbol{\mu}_i + (1-\gamma)\boldsymbol{a}_i \tag{3.14}$$

In Eq. (3.14) γ is an adjustable mixing parameter that will need to be tuned. This is similar in spirit to a predictor-corrector scheme[63] in which Eq. (3.13) gives a prediction of the converged induced dipoles from the auxiliary dipoles, and then Eq. (3.14) serves as a correction by mixing the time reversible zeroth iteration solutions corresponding to the auxiliary variables with the first iteration solution for the real dipoles. While more complicated forms of approximating the SCF dipole are possible, such as a non-local kernel method[63], I find that this simple approximation works well for all cases shown in the Results (Section 3.4). To be clear Eq. (3.14) is not used for the calculation of the true polarization energy and forces, but only applies to the derivation of the auxiliary equation of motion as we show below.

By applying the Euler-Lagrange equation to the Lagrangian in Eq. (3.6) for both the real coordinates $\boldsymbol{r}$ and the auxiliary dipoles $\boldsymbol{a}$ one can obtain equations of motion for each, given by Eq. (3.15) and (3.16), respectively.

$$m_i\ddot{\boldsymbol{r}}_i = -\frac{dU(\boldsymbol{r},\boldsymbol{a})}{d\boldsymbol{r}_i}\Big|_{\boldsymbol{a}} \tag{3.15}$$

$$m_i^a\ddot{\boldsymbol{a}}_i = \omega^2 m_i^a\left(\boldsymbol{\mu}_{SCF,i} - \boldsymbol{a}_i\right) + \sum_{j=1}^{N}\left(\alpha_j^{-1}\delta_{ij} - \boldsymbol{T}'_{ij}\right)(\boldsymbol{\mu}_j - \boldsymbol{a}_j) \tag{3.16}$$

Eq. (3.15) is the familiar statement of Newton's law for the nuclei where the gradients of the various forms of the potential energy on the right hand side remain in their usual forms, except for the polarization potential energy, which now depends on the auxiliary dipoles directly as seen in form of Eq. (3.7). Eq. (3.17) gives the equation of motion for the auxiliary dipoles, which I note has an extra term proportional to $(\boldsymbol{\mu} - \boldsymbol{a})$. It can be formally shown that by assuming $m_i^a \sim 1/\omega^2$, and in the limit that $m_i^a \to 0$ and $\omega \to \infty$ such that ωm_i^a is a constant, one recovers an adiabatic Born-Oppenheimer-like approximation of the auxiliary dipole moment[79]. In practice I will show *a posteriori* that indeed this term is negligible and can (and will be) ignored. When Eq. (3.14) is

substituted into the auxiliary equation of motion in Eq. (3.16), one obtains the final form given by Eq. (3.17).

$$\ddot{\boldsymbol{\alpha}}_i = \omega^2 \gamma (\boldsymbol{\mu}_i - \boldsymbol{\alpha}_i) \tag{3.17}$$

Of course for dynamics one will also need the forms of the gradient to calculate forces. In this case the complete polarization gradient, Eq. (3.18), is used.

$$\frac{dU^{pol}(\boldsymbol{r^N})}{d\boldsymbol{r}} = \frac{\partial U^{pol}(\boldsymbol{r^N})}{\partial \boldsymbol{r}} + \frac{\partial U^{pol}(\boldsymbol{r^N})}{\partial \boldsymbol{\mu}} \frac{\partial \boldsymbol{\mu}}{\partial \boldsymbol{r}} \tag{3.18}$$

The first term on the right hand side of Eq. (3.18) is the nuclear term and the second term is the dipole response term, which goes to 0 in the limit of perfectly converged induced dipoles[82-83]. In practice, although perfect convergence can never be achieved, the dipole response term is typically not explicitly calculated with the understanding that this is a good approximation for a tight level of SCF convergence. With the iEL/0-SCF approach, however, the dipole response term is needed, since the analytical polarization potential requires it, unlike the iterative SCF solvers. Since there is have an explicit position dependence of our dipoles, defined by Eq. (3.13), it is straightforward to calculate these dipole response terms with no additional algorithmic expense. Substituting Eq. (3.7) and (3.13) into Eq. (3.18) the final form of the gradient used for dynamics is obtained, Eq. (3.19).

$$\frac{dU^{pol}(\boldsymbol{r^N})}{d\boldsymbol{r}} = -\frac{1}{2}\boldsymbol{\mu}^T \frac{\partial \boldsymbol{T'}}{\partial \boldsymbol{r}} \boldsymbol{\mu} - [(\alpha \boldsymbol{T}\boldsymbol{\mu})^T + (\alpha \boldsymbol{E})^T - \boldsymbol{\mu}^T] \frac{\partial \boldsymbol{T'}}{\partial \boldsymbol{r}} \boldsymbol{a} - [(\alpha \boldsymbol{T}\boldsymbol{\mu})^T + (\alpha \boldsymbol{E})^T] \frac{\partial \boldsymbol{TM}}{\partial \boldsymbol{r}} \tag{3.19}$$

Finally, the AMOEBA force field has additional features of electrostatic and polarization scaling that are further discussed in Appendix C, knowledge that would be required to reproduce the results shown subsequently.

3.3 Methods

The modifications to the polarization potential and gradient and the addition of auxiliary induced dipoles were introduced into the TINKER software package. For pure water systems I used the AMOEBA14 force field[34] which used Force Balance[97] to optimize the AMOEBA parameters for better description of water properties. For other systems, such as three concentrations of magnesium chloride salt solutions and the solvated dihydrofolate reductase (DHFR) protein, I used the AMOEBABIO09[32-33] force field, a general-purpose biomolecular parameterization of the AMOEBA functional form. For the solvated zwitterionic glycine peptide ($+NH_3$-C_αH-COO–), I used the AMOEBA parameters derived from recent work comparing its decomposed THz spectra with *ab initio* molecular dynamics[98].

Unless otherwise noted simulations were started from pre-equilibrated structures to achieve correct densities and temperatures at 298.0 K. The water system was comprised of a box of 512 water molecules with a 24.83 Å box length that conforms to 1.0 g/cc density. Magnesium chloride solutions were performed at concentrations of 0.30

M, 1.21 M, and 4.66 M, which had 2175, 2133, and 2013 atoms with box lengths of 28.18 Å, 28.00 Å, and 28.35 Å, respectively. For the zwitterionic glycine simulations I used 256 water molecules with an isotropic box with side length of 19.73 Å. A velocity Verlet integration scheme[88] was used to propagate the nuclear coordinates and velocities with a 1.0 fs time step. Production simulations were performed in the NVT ensemble with temperature control using a 4th-order Nosé-Hoover chain thermostat[66] for the nuclear degrees of freedom. Particle-mesh Ewald (PME) was used to evaluate polarization and electrostatic interactions with a real-space cutoff of 7.0 Å. Neighbor lists were employed for electrostatic, polarization, and van der Waals (vdW) interactions. Shared memory multi-core parallelism available in the TINKER software package was used.

The primary difference investigated in this work is how mutual polarization is treated. The iEL/0-SCF simulations were initiated with real and auxiliary dipoles that were solved to an excessively tightly converged solution of 10^{-9} RMS Debye. A velocity Verlet integration scheme was then used to propagate the auxiliary induced dipoles and dipole velocities with a 0.5 fs time step in the NVE ensemble and with a 1.0 fs time step in the NVT ensemble. To determine an optimal γ for a given system, short (~100 ps) NVE trajectories were run over the range of possible γ values (0 to 1.0) and the γ of the trajectory with the lowest energy drift was selected for production. For the NVT ensemble, auxiliary dipoles were thermostatted using a 4th-order Nosé-Hoover chain, like the nuclear degrees of freedom, as described in Chapter 2. The set point of the auxiliary dipole pseudo temperature was 5.3 $e^2Å^2/ps^2$, which I found to be the natural temperature of well-integrated, stable iEL/0-SCF simulations at small time steps. For comparison purposes, simulations performed with a standard preconditioned conjugate gradient SCF (PCG) [20] iterative procedure were calculated using an SCF convergence threshold of 10^{-6} RMS Debye, which is an order of magnitude tighter than the TINKER default in order to absolutely ensure accurate data for comparison purposes.

3.4 Results

To test whether the iEL/0-SCF approach reproduces all properties of any system when compared to the default PCG method used in TINKER, I have examined four different physical systems: (1) homogeneous bulk water, (2) heterogeneous $MgCl_2$ salt solutions, at three different concentrations of $MgCl_2$ in water: 0.30 M, 1.21 M, and 4.66 M, (3) a small biomolecule, a zwitterionic glycine peptide in water, and (4) a larger protein, DHFR, in water. As I show in what follows, the iEL/0-SCF method can reproduce all of aspects of polarization compared to the standard SCF approach, and thus all energetic, structural, and dynamical properties of these four very different cases. Additional validation data for iEL/0-SCF can be found in Appendix D.

3.4.1 *NVE and NVT Conservation Properties*

Figure 3.1a shows the results when optimizing for the one free parameter γ that is needed to estimate the $\boldsymbol{\mu}_{SCF,i}$ solution using Eq. (3.14), examined by running the iEL/0-SCF method in the NVE ensemble with a range of γ values from 0 to 1. From this analysis I determined that the maximum stable value of γ is 0.92, with values greater than that yielding unstable trajectories, and values of $\gamma < 0.5$ yielding poor energy conservation. The best value of $\gamma = 0.9$ gives energy drifts that are commensurate with

the standard PCG method. Figure 3.1b reports the 'pseudo-temperature' of the auxiliary dipoles in these simulations, defined as $T_{aux} = 1/3\langle \dot{\boldsymbol{a}}_i^2 \rangle$ since these dipoles do not have a mass, in which we see increases in inertia throughout the simulation, regardless of γ value. The rate of buildup of the auxiliary inertia degrades the real dipole dynamics through resonances as we have shown previously[23], therefore in the NVE ensemble we are restricted to time steps of 0.5 fs to realize good energy conservation for the real degrees of freedom.

For the NVT ensemble simulations a thermostat must be applied to both real and auxiliary degrees of freedom, and thus there is a new conservation law that depends directly on the extended system variables[66]. Although formally the fictitious auxiliary mass parameter, m^a, is 0 from our derivation, this quantity can be fit *a posteriori* to see what its value would be to minimize the drift of the extended conserved quantity. Figure 3.1c shows that this extended system quantity is well conserved for the pure water system when using a fitted fictitious mass of the auxiliary dipoles of 0.0103 g/mol/e^2 and which corresponds to a real temperature of 0.0656 K for the auxiliary dipoles, giving a natural separation of energy scales between the real and auxiliary degrees of freedom and allowing for an increased time step of 1.0 fs. The NVT energies and fluctuations for the real nuclear degrees of freedom of water for the iEL/0-SCF method is in perfect accord with the standard SCF benchmark over 3.0 ns (Table 3.1 and Figure D.1) and additional tests run out to tens of nanoseconds show that SCF and iEL/0-SCF solutions remained in quantitative agreement for all systems for long time scales.

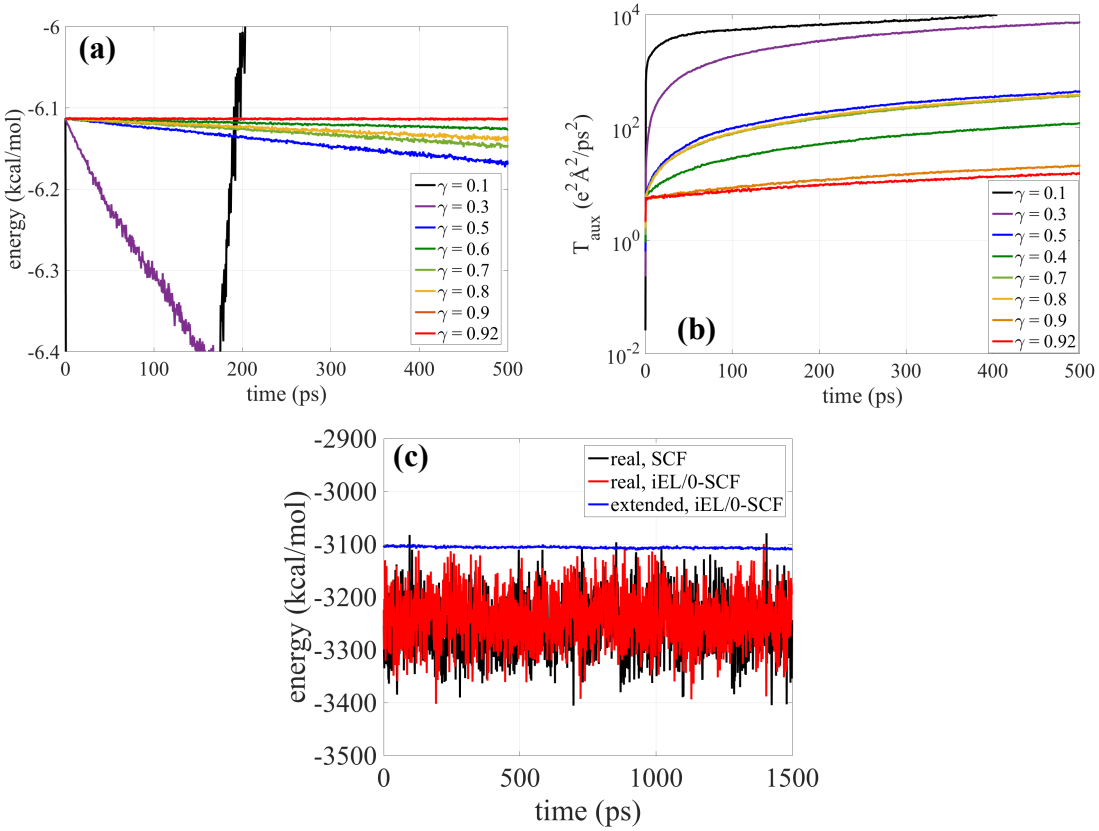


Figure 3.1: *Energy conservation in the NVE and NVT ensembles for water. (a) Energy along trajectories of the iEL/0-SCF method for a range of γ values. (b) The auxiliary pseudo-temperature,* $T_{aux} = 1/3\langle \dot{\boldsymbol{a}}_i^2 \rangle$,

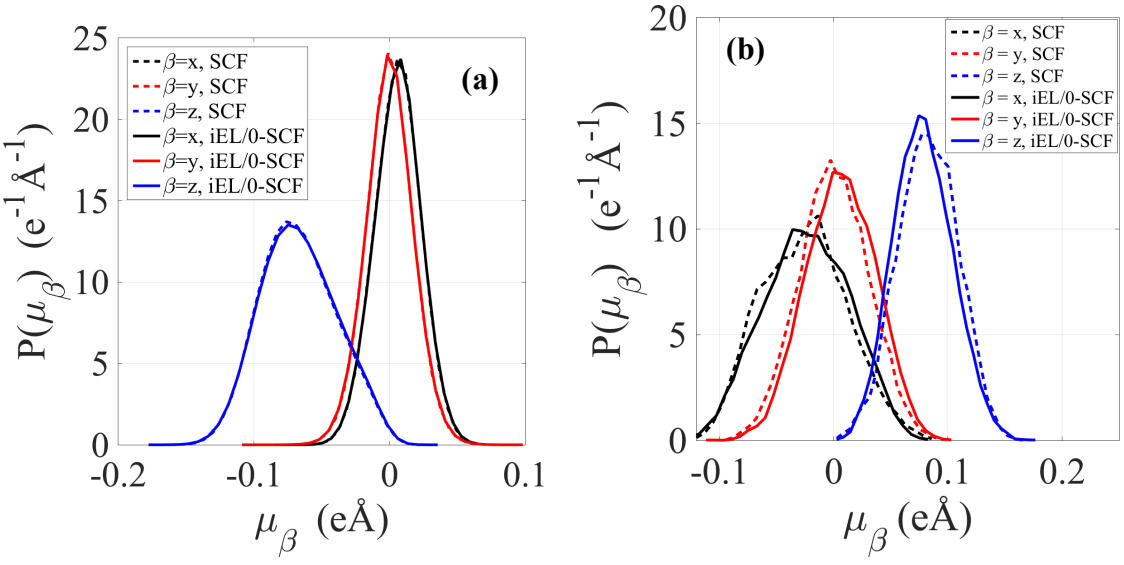
of the simulations from (a). In (a) and (b) all NVE simulations used a time step of 0.5 fs. (c) Real energy for iEL/0-SCF and standard PCG-SCF methods and extended energy for iEL/0-SCF, as well. We calculated the conserved energy of the extended system in the NVT ensemble (blue), along with the energy for the real degrees of freedom in the NVT ensemble of 512 water molecules using the iEL/0-SCF approach (red) and standard SCF approach (black). For the NVT simulations we used a time step of 1.0 fs, γ was set to 0.9, while the SCF solution used a PCG-SCF method with a convergence threshold of 10^{-6} RMS Debye.

Table 3.1: *Ensemble average total energy and polarization energies for bulk water, salt solutions, and glycine dipeptide in water. Average total potential energy, $\langle U \rangle$, and average polarization potential energy, $\langle U_{pol} \rangle$, for both a standard SCF method and the iEL/0-SCF approach. Data is generated from 3.0 ns trajectories in the NVT ensemble at 298.0 K.*

Bulk Water	Method	$\langle U \rangle$ (kcal/mol)	$\langle U_{pol} \rangle$ (kcal/mol)
	iEL/0-SCF	-4610+/-40	-2590+/-50
	SCF	-4620+/-40	-2570+/-50
Glycine in Water	iEL/0-SCF	-2420+/30	-1290+/-40
	SCF	-2430+/-30	-1300+/-30
0.3M $MgCl_2$	iEL/0-SCF	-8830+/-50	-4010+/-60
	SCF	-8850+/-50	-4040+/-70
1.21M $MgCl_2$	iEL/0-SCF	-15690+/-50	-5100+/-70
	SCF	-15710+/-50	-5070+/-70
4.66M $MgCl_2$	iEL/0-SCF	-43220+/-50	-9640+/-100
	SCF	-43220+/-50	-9490+/-90

3.4.2 *Polarization Properties*

For bulk water I found that the average molecular dipole moment for a condensed phase water molecule yields a value of 2.81 Debye with both the SCF and iEL/0-SCF methods. The corresponding probability distribution of real induced dipoles generated from the iEL/0-SCF solutions is in excellent agreement with the standard SCF solutions, regardless of system, as shown in Figure 3.2 and Figure D.2.



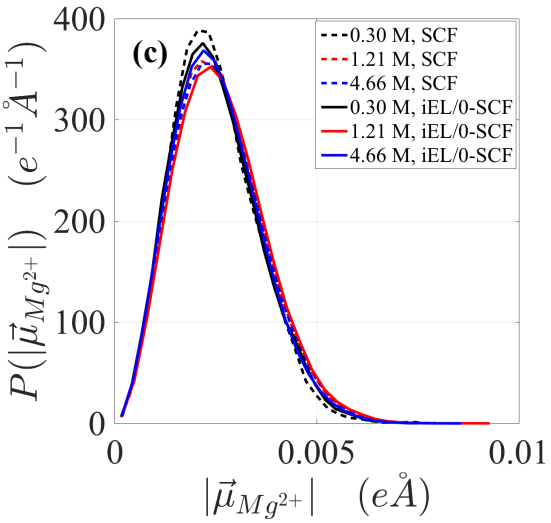


Figure 3.2: *Comparison of induced dipole probability distributions using the iEL/0-SCF and standard SCF methods. (a) Hydrogen in bulk water, (b) the glycine carbonyl oxygen, and (c) Mg^{2+} for all $MgCl_2$ salt concentration for a standard SCF method (dashed) and our iEL/0-SCF method (solid). For plots (a) and (b) we consider the x, y, and z dipole components. For plot (c) we consider three different $MgCl_2$ salt concentrations of 0.30 M (black), 1.21 M (red), and 4.66 M (blue). The induced dipole distributions for water and glycine use an internal coordinate frame; see[17] for details on the internal coordinates.*

3.4.3 *Dynamic Properties.*

Using the NVE ensemble simulation, I calculated the self-diffusion constant of water to be 1.90 +/- 0.13 using the iEL/0-SCF approach and 2.02 +/- 0.20 for the standard SCF procedure. I also reported the Mg^{2+} and Cl^- ion diffusion constants for the 4.66M concentration in Table D.1 that show the two methods are in quantitative agreement. A particularly stringent time-dependent property in the NVT ensemble, the autocorrelation function for induced dipoles, demonstrates that the iEL/0-SCF method does not suffer from any loss of accuracy, regardless of system (Figure 3.3 and Figure D.3). One can see that for both hydrogen and oxygen atoms for bulk water, Mg^{2+} and Cl^- for the salt solution (shown for 1.21M), and the backbone carbonyl carbon, carbonyl oxygen, nitrogen, and α-carbon of glycine dipeptide, that the iEL/0-SCF method reproduces the autocorrelations well, especially at the critical short time scale range less than about 0.2 ps over which there can be rapid changes in dipole direction (especially for bulk water).

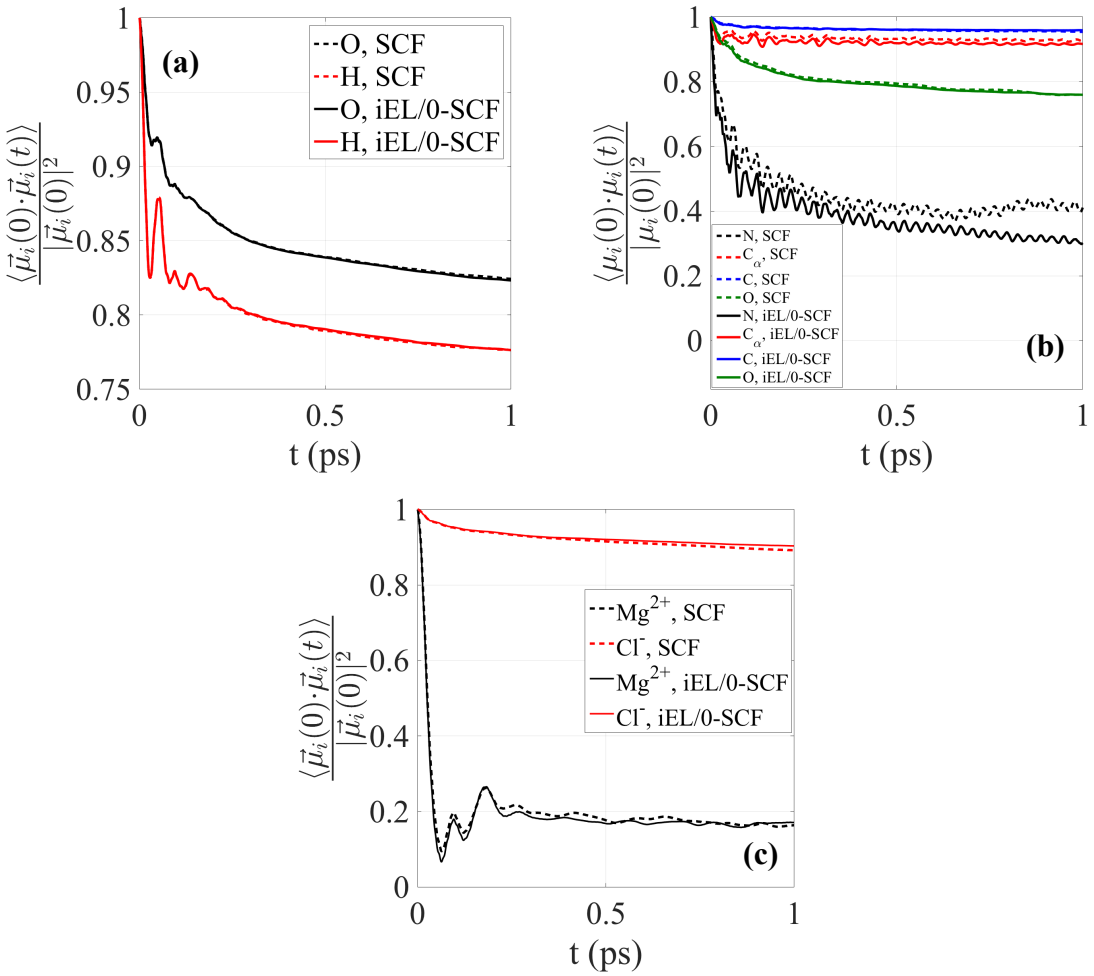


Figure 3.3: *Comparison of dynamic polarization properties using the iEL/0-SCF and standard SCF methods. Autocorrelation function of real dipole for (a) bulk water, (b) solvated glycine, and (c) Mg^{2+} and Cl^- for 1.21M salt concentrations with a standard SCF method (dashed) and our iEL/0-SCF method (solid). The induced dipole distributions for water and glycine used an internal coordinate frame; see[17] for details on the internal coordinates. All data is based on 30 ps trajectories in the NVT ensemble at 298.0 K.*

3.4.4 *Structural Properties*

The iEL/0-SCF method also reproduces structural properties for all systems. Figure 3.4 and Figure D.4 present the radial distribution functions for atom correlations of bulk water, the water oxygen and hydrogen correlations with the anion and cation of the salt solutions, and finally the water oxygen and hydrogen radial distribution functions with the glycine peptide backbone atoms. In summary, there is excellent agreement between the iEL/0-SCF method and the standard SCF solver.

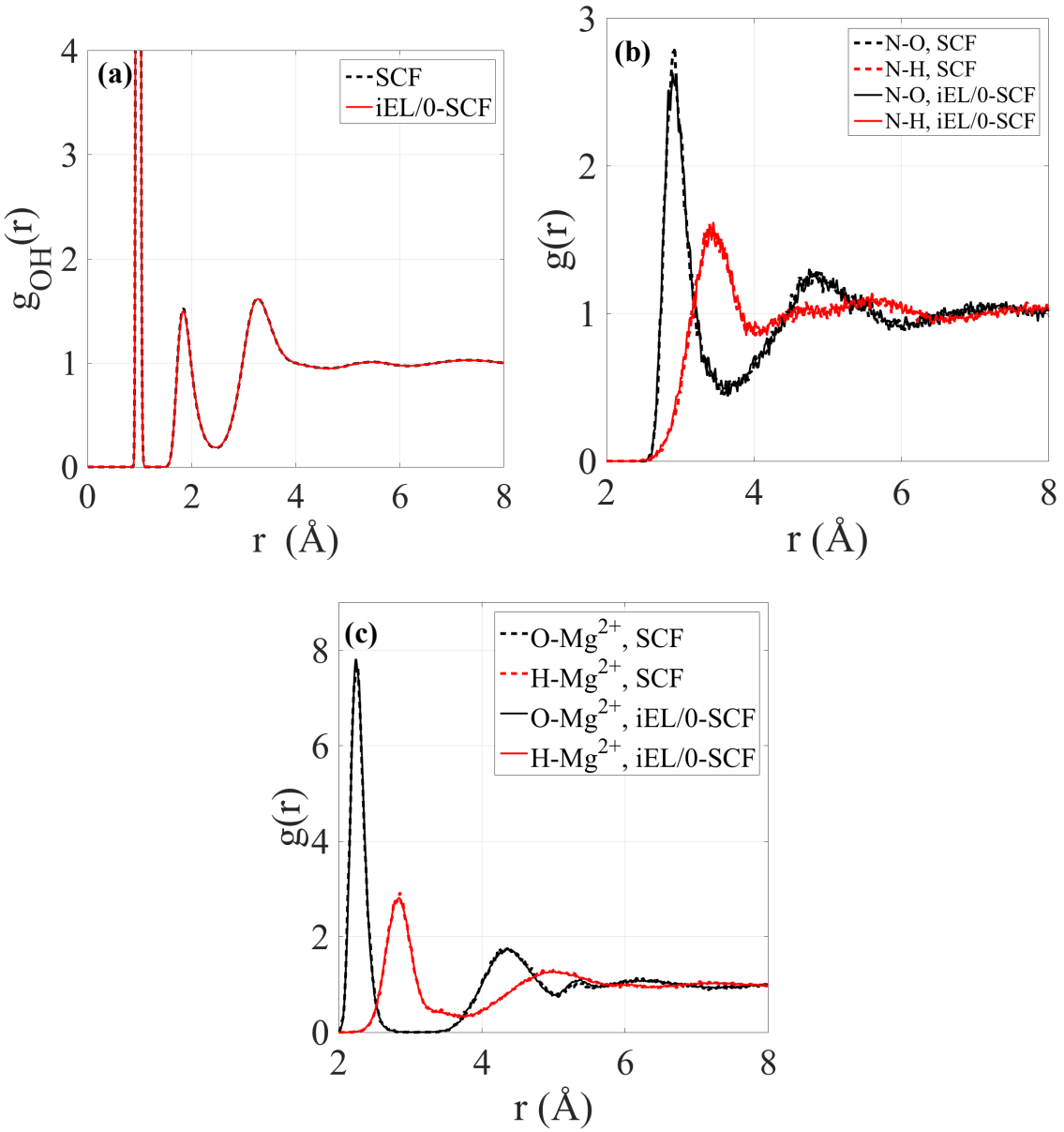


Figure 3.4: *Comparison of intermolecular structure using the iEL/0-SCF and standard SCF methods. Radial distrubtion functions for (a) oxygen-hydrogen for bulk water, (b) water oxygen-backbone nitrogen and water hydrogen-backbone nitrogen of glycine dipeptide, and (c) water oxygen-Mg^{2+} and water hydrogen- Mg^{2+} for the 0.30 M salt concentrations for a standard SCF method (dashed) and our iEL/0-SCF method (solid). The induced dipole distributions for water and glycine used an internal coordinate frame; see[17] for details on the internal coordinates. All data is based on 3.0 ns trajectories in the NVT ensemble at 298.0 K.*

3.4.5 *Timing Comparisons*

The iEL/0-SCF method was implemented in the TINKER software package and takes advantage of the shared-memory parallelism of that code. While I have shown that the iEL/0-SCF method can reproduce the property results for any of the above systems when compared to the standard PCG SCF solver in the TINKER code, I now examine the

computational efficiency of the iEL/0-SCF method in the same code and on the same hardware platform.

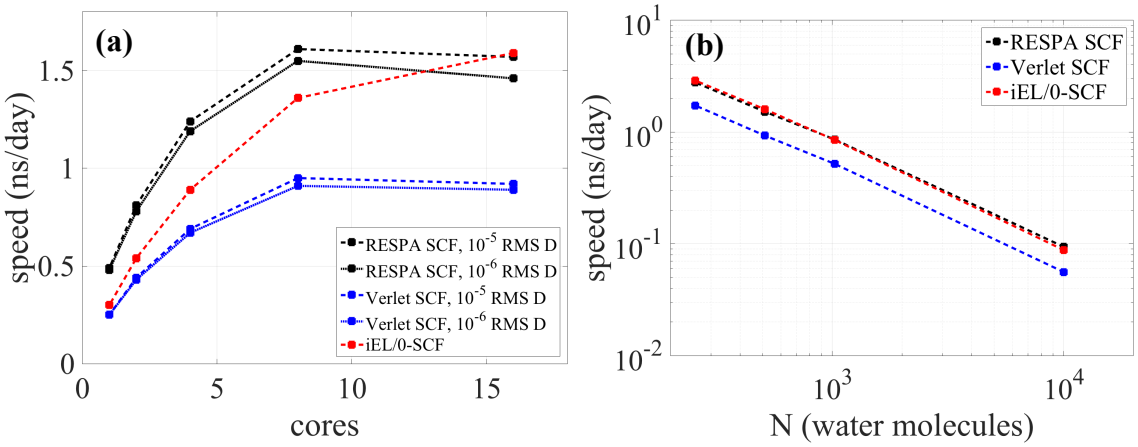


Figure 3.5: *Timing comparisons between iEL/0-SCF and standard iteration methods. Simulation speed-up in nanoseconds per day for (a) weak scaling as a function of the number of cores for a box of 512 water molecules in the NVT ensemble at 298.0 K and (b) strong scaling for increasing system size with the number of cores fixed at 16. The methods that are compared include the iEL/0-SCF method at a time step of 1.0 fs (red), a preconditioned conjugate gradient solver integrated with Verlet at a time step of 1.0 fs (blue), and with a RESPA method using an outer time step of 2.0 fs (black). The real dipoles of the standard SCF methods in (b) were converged to 10^{-5} RMS Debye (solid), the standard default in TINKER.*

Figure 3.5a shows the weak scaling results in units of ns/day of the SCF and iEL/0-SCF methods applied to bulk water in which both use the same Verlet integration scheme with a time step of 1.0 fs. For 8 cores one sees that the iEL/0-SCF method is 40% faster than the standard SCF solution, and since the iEL/0-SCF method also scales better to larger numbers of cores it rises to ~70% faster than the default polarization solution. In fact, when compared to a multi-time step reversible reference system propagator algorithms (RESPA) [26], which uses a 2.0 fs outer time step, the iEL/0-SCF method is slightly faster at 16 cores, although it uses Verlet integration with a 1.0 fs time step. Figure 3.5b shows the strong scaling results (as system size grows while holding the number of cores constant at 16 cores) for the standard SCF polarization result using either a RESPA integration scheme with a 2.0 fs outer time step, the Verlet integration with a 1.0 fs time step, and the iEL/0-SCF approach that uses the Verlet integration scheme with a time step of 1.0 fs. It is seen that the iEL/0-SCF approach is not only as fast as the RESPA scheme in timings, the iEL/0-SCF method is also more accurate as I show below for DHFR and for bulk water results in Figure D.1.

3.5 Discussion

In general, energies, dynamical properties, and polarizability and structural correlations, are accurately reproduced by my iEL/0-SCF method when compared to the standard SCF solver for a diverse set of simple systems ranging from bulk water, dilute to concentrated salt solutions, and a small peptide in water. I also providde an additional stress test for our new iEL/0-SCF solution to the polarization equations by simulating a

much larger biomolecular system, DHFR in water, the so-called Joint AMBER-CHARMM benchmark. Table 3.2 gives the average potential energy and average polarization energy for the aqueous DHFR system when I used a single γ value for both the protein and the water solvent when defining the local kernel approximation for $\boldsymbol{\mu}_{SCF,i}$ in Eq. (3.9). For γ values greater than 0.7 I found that the aqueous DHFR simulations became unstable (which also happened for the other water-based systems when $\gamma > 0.92$), and ultimately $\gamma = 0.7$ was the best value that we found from numerical testing. In this case the differences between the iEL/0-SCF and SCF energy values are on the order of 1.0%, although this small difference is nonetheless outside numerical error bars. However this error is much smaller than that found from simulations with the RESPA algorithm, as shown in Figure 3.6 which displays the time evolution of the DHFR-water system real total energy.

Table 3.2: *Average total potential energy, $\langle U \rangle$, and polarization energy, $\langle U_{pol} \rangle$, for the iEL/0-SCF method for a protein in water using different values of γ_S and γ_P for the local kernel definition for solvent and protein. Data is generated from 300 ps trajectories in the NVT ensemble at 298.0 K.*

Local Kernel	$\langle U \rangle$ **(kcal/mol)**	$\langle U_{pol} \rangle$ **(kcal/mol)**
$\gamma_S = \gamma_P = 0.7$	-66060+/-170	-31160+/-150
$\gamma_S = 0.9;\ \gamma_P = 0.1$	-66310+/-180	-31440+/-160
SCF	-66670+/-160	-31850+/-150

While this suggests that the iEL/0-SCF method is sensitive to the local kernel definition for γ, I found that I can use the same γ value for water of 0.9 (now referred to as γ_S with 's' for solvent) that I determined for all previous aqueous systems considered, and allow the protein to independently optimize to a new value of $\gamma_P = 0.1$ (Table 3.2). It is clear that basic properties of the protein system are reproduced; the protein remains stable as measured by its root mean square deviation (RMSD) over the course of the simulation. While the local kernel definition has proven reliable across a large diversity of systems, and the same γ values are resuseable across systems, it does motivate a generalization of Eq. (3.14) to a 'non-local' kernel[79] to improve the iEL/0-SCF method in future work, provided such methods do not add significantly to the computational cost.

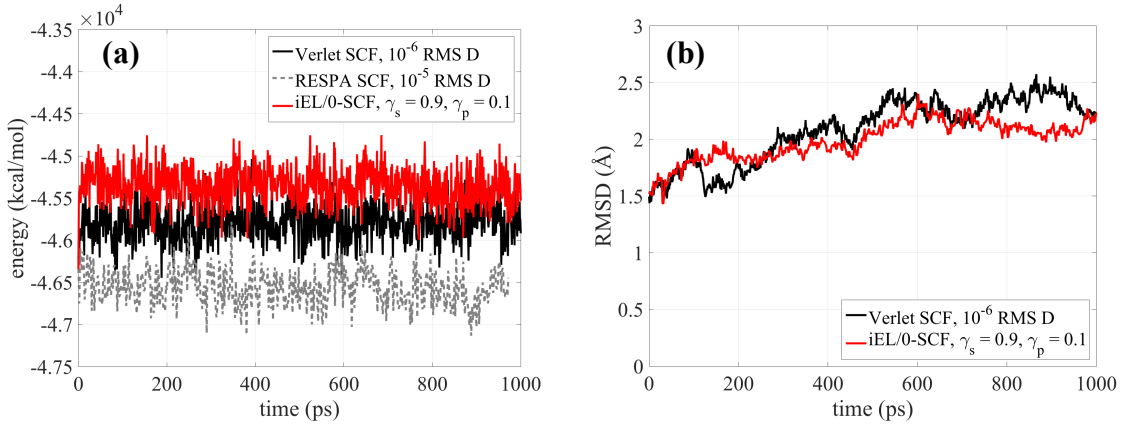


Figure 3.6: *Simulations of the DHFR protein in water using the iEL/0-SCF, standard SCF, and standard SCF with RESPA method. The real total energy (potential and kinetic) (a) and RMSD (b) as a function of simulation time. The values of $\gamma_S = 0.9$ and $\gamma_P = 0.1$ were used for the local*

kernel definition for solvent and protein. The SCF method (black) uses a Verlet with a 1.0 fs time step and PCG with 10^{-6} RMS Debye convergence, and the SCF method with RESPA (dashed gray) with a 2.0 fs outer time step and PCG with 10^{-5} RMS Debye convergence.

3.6 Conclusions

I have developed a new solution to classical polarization for molecular dynamics simulations that does not require any SCF iteration. The iEL/0-SCF method I have laid out here builds upon our iEL/SCF scheme of Chapter 2 in that it continues to use an auxiliary set of induced dipoles that are dynamically integrated in a time-reversible manner, and which again serve as an initial guess for the real induced dipoles, but for the iEL/0-SCF method no further SCF cycles are required. Across all relevant systems typically studied in molecular simulation: a homogeneous fluid, ionic solutions, and small and large solutes in aqueous media, all ensemble averages over polarization properties are in good agreement between the SCF and iEL/0-SCF methods, and therefore by extension all physical properties of any system examined here are reproduced accurately. The iEL/0-SCF method is numerically stable over long simulation times, and I offered a qualitative explanation as to why the iEL/0-SCF scheme works as well as it does. I speculate that, since the autocorrelation function for the dipoles decay on a 100 fs timescale (Figure D.3), then the repeated application of real dipole and auxiliary dipole updates every 1.0 fs, is effectively doing SCF iterations on the fly. The iEL/0-SCF approach also realizes significant computational speedups when compared to an SCF method that uses a single 1.0 fs time step Verlet integration, and is as fast and far more accurate than that obtained by using RESPA with a 2.0 fs outer time step. Since the method saves about 50% in computational effort by avoiding all SCF cycles, and the fact that standard SCF implementations do not scale well in TINKER the 100% factor improvement in timings make perfect sense.

4.
iEL/0-SCF for Drude Polarization

4.1 Introduction

The induced dipole model is arguably a more robust approach to polarization than the Drude model, since its foundation arises from an expansion of the electric potential about an applied electric field, and thus it can be systematically generalized to higher-order moments like induced quadrupoles, induced octupoles, etc. [54] By contrast, the Drude oscillator offers the advantage of simplicity, since it only requires additional point charge Coulombic interactions that are straightforwardly incorporated into existing molecular dynamics codes in conjunction with Ewald summation[99-101]. Due to its inherent simplicity the Drude model is still a popular choice as a polarization model[44, 52, 102], so in this chapter I expand the theoretical framework of the inertial extended Langrangian, self-consistent field iteration-free method (iEL/0-SCF), introduced for induced dipoles in Chapter 3, to the polarization model of a Drude oscillator.

Historically the polarization solution for the induced dipole model is solved through self-consistent field (SCF) iterative solvers, such as successive over-relaxation (SOR) [60], preconditioned conjugate gradient (PCG) [20], or direct inversion in the iterative subspace (DIIS) [21, 61] methods. By contrast, Drude models for polarization are typically solved through an extended Lagrangian (EL) formulation to treat polarization with negligible cost compared to the SCF approaches[42-43, 67]. In the case of Drude oscillators the EL equation of motion is based on a mass repartitioning between the parent atom and its Drude oscillator, with the goal of making the Drude mass small enough to obey some facsimile of the Born-Oppenheimer condition. Even so, a basic EL approach using thermalized "hot" Drude oscillators can be plagued with problems of accuracy since the effective polarization vector fluctuates around an average orientation that does not conform to the true electric field vector, and/or problems of stability in the context of a molecular dynamics trajectory that forces the reduction of the time step to be unacceptably short. To combat this problem, Sprik developed an EL approach whereby the polarization degrees of freedom are kept cold at a temperature T* relative to the temperature of the real degrees of freedom, T, such that $T^* << T$[69]. The dual temperature EL approach, EL(T,T*), has become a standard approach for treating Drude polarization due to its low cost, but can still exhibit unwanted instability.

In this chapter I present the theory that extends the iEL/0-SCF approach to Drude polarization, in which the need for any mass repartitioning is eliminated, greater stability is realized, and the correct thermodynamic ensembles are achieved. The advantages of iEL/0-SCF for Drude are illustrated with simulations for bulk water using the polarizable simple point charge (PSPC) Drude force field[42-43]. I show that the iEL/0-SCF method for Drude polarization is as stable as a tightly converged SCF calculation and more stable than the EL(T,T*) approach. This greater stability allows us to take molecular dynamics timesteps as large as 6.0 fs while still preserving the properties of the PSPC model in the NVT ensemble.

4.2 Theory

The PSPC model is a simple rigid water model[42-43], with partial charges on both hydrogens and oxygen, and a Lennard-Jones site and a Drude particle that is harmonically bound to each oxygen atom. The total energy of the PSPC model is then given by Eq. (4.1).

$$U(\boldsymbol{r}) = U_{LJ}(\boldsymbol{r}) + \sum_i \sum_{j>i} \frac{q_i q_j}{|\boldsymbol{r_i} - \boldsymbol{r_j}|} + \sum_i \sum_j \frac{q_i q_{D,j}}{|\boldsymbol{r_i} - \boldsymbol{r_{D,j}}|} + \sum_i \sum_{j>i} \frac{q_{D,i} q_{D,j}}{|\boldsymbol{r_{D,i}} - \boldsymbol{r_{D,j}}|} + \frac{1}{2} k_D \sum_i |\boldsymbol{r_i} - \boldsymbol{r_{D,i}}|^2 \tag{4.1}$$

In Eq. (4.1) $U_{LJ}(\boldsymbol{r})$ is the standard Lennard-Jones term, the second through fourth terms are the electrostatic energy broken down into charge-charge, charge-Drude, and Drude-Drude interactions, respectively. The final term is the harmonic spring between Drude particles and their parent oxygen atom, where a Drude spring force constant, k_D, of 1000 kcal/mol/Å is employed in the PSPC model.

The relationship between the Drude spring force constant, k_D, polarizability, α, and Drude charge, q_D, can be easily derived and is given by Eq. (4.2)[43].

$$\alpha = \frac{q_D^2}{k_D} \tag{4.2}$$

The energy minima with respect to Drude displacements from their parent atom, $\boldsymbol{d}_i = \boldsymbol{r}_{D,i} - \boldsymbol{r}_i$, is given by Eq. (4.3).

$$\frac{\partial U}{\partial \boldsymbol{r}_{D,i}} = k_D \boldsymbol{d}_i - q_{D,i} \boldsymbol{E}_{D,i}(\boldsymbol{d}^N) = \boldsymbol{0} \tag{4.3}$$

In Eq. (4.3) $\boldsymbol{E}_i(\boldsymbol{d}^N)$ is the electric field at site *i* due to the other atomic charges and Drude particles in the system. Eq. (4.3) states that the system's polarization energy is minimized when there is no net force on the Drude particles, and it defines the iterative equation that must be solved to determine the SCF solution for the Drude model.

Following generalized time reversible Born-Oppenheimer dynamics[79] and its application to classical dipolar polarization presented in previous chapter[22], I begin by defining a new extended Lagrangian for the Drude PSPC polarization model, given by Eq. (4.4).

$$\begin{aligned} \mathcal{L}(\boldsymbol{r}, \dot{\boldsymbol{r}}, \boldsymbol{a}_D, \dot{\boldsymbol{a}}_D) = {} & \frac{1}{2} \sum_i m_i |\dot{\boldsymbol{r}}_i|^2 - U_{LJ}(\boldsymbol{r}) - \sum_i \sum_{j>i} \frac{q_i q_j}{|\boldsymbol{r}_i - \boldsymbol{r}_j|} - \sum_i \sum_j \frac{q_i q_{D,j}}{|\boldsymbol{r}_i - \boldsymbol{r}_{D,j}(\boldsymbol{a}_D)|} - \sum_i \sum_{j>i} \frac{q_{D,i} q_{D,j}}{|\boldsymbol{r}_{D,i}(\boldsymbol{a}_D) - \boldsymbol{r}_{D,j}(\boldsymbol{a}_D)|} \\ & - \frac{1}{2} k_D \sum_i |\boldsymbol{r}_i - \boldsymbol{r}_{D,i}(\boldsymbol{a}_D)|^2 - \frac{1}{2} \sum_i m_i^a |\dot{\boldsymbol{a}}_{D,i}|^2 - \frac{1}{2} \omega^2 \sum_i m_i^a |\boldsymbol{r}_{D,i}^{SCF} - \boldsymbol{a}_{D,i}|^2 \end{aligned} \tag{4.4}$$

The first six terms (kinetic energy, Lennard-Jones potential, Coulombic terms, and harmonic force between the Drude and parent atom) largely recapitulate Eq. (4.1), but with several important exceptions when formulated within the iEL/SCF hybrid approach. First I introduce an extended system of auxiliary Drude positions, $\boldsymbol{a}_{D,i}$, as

dynamic degrees of freedom that have an associated fictitious mass, m_i^a. Second, these auxiliaries in turn give rise to a positional dependence of the real Drude particles on these auxiliary degrees of freedom, $\boldsymbol{r}_{D,i}(\boldsymbol{a}_D)$. Third, a new kinetic energy term for the auxiliary degrees of freedom is included in the Lagrangian. Finally, a harmonic coupling between the auxiliary positions and the true SCF converged position, $\boldsymbol{r}_{D,i}^{SCF}$ is introduced, as well. The strength of this coupling is controlled by the parameter ω, which should be as high as possible. It has been shown that the highest stable value of ω for a Verlet integration scheme is $\sqrt{2}/\Delta t$, where Δt is the time step[79].

One can now recognize that the most general form of the potential energy for the original Drude model from Eq. (4.4) is given as Eq. (4.5).

$$\begin{aligned} U(\boldsymbol{r},\boldsymbol{a}_D) = U_{LJ}(\boldsymbol{r}) &+ \sum_i \sum_{j>i} \frac{q_i q_j}{|\boldsymbol{r}_i - \boldsymbol{r}_j|} + \sum_i \sum_j \frac{q_i q_{D,j}}{|\boldsymbol{r}_i - \boldsymbol{r}_{D,j}(\boldsymbol{a}_D^N)|} + \sum_i \sum_{j>i} \frac{q_{D,i} q_{D,j}}{|\boldsymbol{r}_{D,i}(\boldsymbol{a}_D^N) - \boldsymbol{r}_{D,j}(\boldsymbol{a}_D^N)|} \\ &+ \frac{1}{2} k_D \sum_i |\boldsymbol{r}_i - \boldsymbol{r}_{D,i}(\boldsymbol{a}_D^N)|^2 \end{aligned} \tag{4.5}$$

Since the Drude positions depend explicitly on the auxiliary Drude positions, there is no assumption that the potential energy $U(\boldsymbol{r},\boldsymbol{a}_D)$ is at the ground state solution as per Eq. (4.3). The positional dependence of the real Drude particles on the auxiliary degrees of freedom, $\boldsymbol{r}_{D,i}(\boldsymbol{a}_D)$, is formulated by making one evaluation of the electric field using the auxiliary positions as the positions of the Drude particles, as shown in Eq. (4.6).

$$\boldsymbol{r}_{D,i}(\boldsymbol{a}_D) = \frac{q_{D,i}}{k_D} \boldsymbol{E}_{D,i}(\boldsymbol{a}_D) + \boldsymbol{r}_i \tag{4.6}$$

In terms of the PSPC model, specifically, Eq. (4.6) can be expanded into Eq. (4.7).

$$\boldsymbol{r}_{D,i}(\boldsymbol{a}_D) = -\frac{q_{D,i}}{k_D} \left[\sum_j \frac{q_j(\boldsymbol{a}_{D,i} - \boldsymbol{r}_j)}{|\boldsymbol{a}_{D,i} - \boldsymbol{r}_j|^3} + \sum_j \frac{q_j(\boldsymbol{a}_{D,i} - \boldsymbol{a}_{D,j})}{|\boldsymbol{a}_{D,i} - \boldsymbol{a}_{D,j}|^3} \right] + \boldsymbol{r}_i \tag{4.7}$$

It is important to emphasize why the Drude oscillator model is different than an induced dipole model in an SCF process. For the induced dipole model, the electric field due to other permanent charges or multipoles in the system is defined at an atomic center. This direct component of the electric field will not change throughout an iterative procedure because the atomic positions do not change. It is only the field due to other induced dipoles in the system that will change throughout an iterative procedure. This is not true of a Drude model. In a Drude model, the electric field at the Drude particles due to both the permanent multipoles and other Drude particles will change because the position where the electric field is evaluated (the Drude position) itself changes throughout the iterative process. This can be seen explicitly in Eq. (4.7) where the first term in the bracket can be thought of as a direct electric field and the second term as a mutual electric field, but for both terms they show that the optimized Drude positions depend specifically on the auxiliary Drude positions.

One can now use Eq. (4.5) to compress the description of the Lagrangian in Eq. (4.4), now given as Eq. (4.8).

$$\mathcal{L}(\boldsymbol{r}, \dot{\boldsymbol{r}}, \boldsymbol{a}_D, \dot{\boldsymbol{a}}_D) = \frac{1}{2}\sum_i m_i |\dot{\boldsymbol{r}}_i|^2 + \frac{1}{2}\sum_i m_i^a \left|\dot{\boldsymbol{a}}_{D,i}\right|^2 - \frac{1}{2}\omega^2 \sum_i m_i^a \left|\boldsymbol{r}_{D,i}^{SCF} - \boldsymbol{a}_{D,i}\right|^2 - U(\boldsymbol{r}, \boldsymbol{a}_D) \tag{4.8}$$

Eq. (4.8) can now be used to generate equations of motion by applying the Euler-Lagrange equation to both the real coordinates $\boldsymbol{r}$ and the auxiliary dipoles $\boldsymbol{a}_D$.

For the original iEL/SCF approach, the auxiliary positions would be used as an initial guess for a subsequent SCF calculation to the converged solution. This is undesirable especially for the Drude formulation since now the direct part of the electric field has to be evaluated at each iteration, as described above. In the iEL/0-SCF approach, SCF is avoided altogether by introducing a linear mixing of real and auxiliary Drude positions as an approximation to the ground state solution, as shown in Eq. (4.9).

$$\boldsymbol{r}_{D,i}^{SCF} \approx \gamma \boldsymbol{r}_{D,i} + (1-\gamma)\boldsymbol{a}_{D,i} \tag{4.9}$$

Eq. (4.9) is then used to derive the first equation of motion for the auxiliary degrees of freedom when taking the limit of $m_i^a \to 0$, given by Eq. (4.10).

$$\ddot{\boldsymbol{a}}_i = \gamma\omega^2(\boldsymbol{r}_{D,i} - \boldsymbol{a}_{D,i}) \tag{4.10}$$

Two points are worth mentioning at this juncture. First, the predictor-corrector form of the converged solution in Eq. (4.9) is not used for the calculation of the true polarization energy and forces, but only applies to the derivation of the auxiliary equation of motion in Eq. (4.10) [22]. Second, the Drude degrees of freedom now have no mass and thus there is no mass repartitioning in our approach.

The equation of motion for the auxiliaries is straightforward, simply taking the form of harmonic motion about the real Drude positions. This introduces one free parameter, γ, which controls the mixing of the real and auxiliary Drude positions, and can be easily determined as we show in the Methods (section 4.3) and Appendix E, which contains additional iEL/0-SCF Drude material. The equation of motion for the atomic centers looks like the usual real particle equation of motion and is given by Eq. (4.11).

$$m_i \ddot{\boldsymbol{r}}_i = -\left.\frac{dU(\boldsymbol{r}, \boldsymbol{a}_D)}{d\boldsymbol{r}_i}\right|_{\boldsymbol{a}_D} \tag{4.11}$$

In examining Eq. (4.11) note again that unlike the usual EL schemes the atomic mass m_i has not been reduced through repartitioning. The total potential derivative with respect to position is shown in Eq. (4.12).

$$\frac{dU(\boldsymbol{r}, \boldsymbol{a}_D)}{d\boldsymbol{r}_i} = \frac{\partial U(\boldsymbol{r}, \boldsymbol{a}_D)}{\partial \boldsymbol{r}_i} + \sum_j \frac{\partial U(\boldsymbol{r}, \boldsymbol{a}_D)}{\partial \boldsymbol{r}_{D,j}} \frac{\partial \boldsymbol{r}_{D,j}}{\partial \boldsymbol{r}_i} \tag{4.12}$$

In Eq. (4.12) the first term corresponds to the usual gradient terms for a standard EL scheme and the second term is a Drude 'response' term, which is normally not evaluated under an SCF approach since it goes to 0 in the limit of complete convergence. However, since the form of the potential, $U(\boldsymbol{r}, \boldsymbol{a}_D)$, is now general, in the sense that it does not assume complete convergence, this response term must be explicitly evaluated to ensure that the potential and potential gradient are commensurate. Since there are no iterations this is just a small cost overhead.

The force vectors $\frac{\partial U(\boldsymbol{r},\boldsymbol{a}_D)}{\partial \boldsymbol{r}_i}$ and $\frac{\partial U(\boldsymbol{r},\boldsymbol{a}_D)}{\partial \boldsymbol{r}_{D,j}}$ are already evaluated in a standard Drude oscillator EL simulation, so it is easy to calculate these force vectors. The new additional term, $\frac{\partial \boldsymbol{r}_{D,j}}{\partial \boldsymbol{r}_i}$, when expanded, is shown in Eq. (4.13).

$$\begin{aligned}\frac{\partial \boldsymbol{r}_{D,i}}{\partial \boldsymbol{r}_k} &= -\frac{q_{D,i}}{k_D}\left[\sum_j \frac{\partial}{\partial \boldsymbol{r}_k}\left(\frac{q_j(\boldsymbol{a}_{D,i}-\boldsymbol{r}_j)}{|\boldsymbol{a}_{D,i}-\boldsymbol{r}_j|^3}\right)+\sum_j \frac{\partial}{\partial \boldsymbol{r}_k}\left(\frac{q_j(\boldsymbol{a}_{D,i}-\boldsymbol{a}_{D,j})}{|\boldsymbol{a}_{D,i}-\boldsymbol{a}_{D,j}|^3}\right)\right]+\frac{\partial \boldsymbol{r}_i}{\partial \boldsymbol{r}_k} \\ &= \frac{q_{D,i}}{k_D}\sum_j q_j \frac{\partial}{\partial \boldsymbol{r}_k}\left(-\frac{(\boldsymbol{a}_{D,i}-\boldsymbol{r}_j)}{|\boldsymbol{a}_{D,i}-\boldsymbol{r}_j|^3}\right)+\delta_{ik}\boldsymbol{I}\end{aligned} \tag{4.13}$$

One can recognize that the derivative term within the sum of Eq. (4.13) is actually the definition of a dipole-dipole interaction tensor, $\boldsymbol{T}'$, between the k atomic center, $\boldsymbol{r}_k$, and the i-th auxiliary Drude position, $\boldsymbol{a}_{D,i}$. Using this fact Eq. (4.13) can be simplified to yield Eq. (4.14a), where Eq. (4.14b) gives the definition of the dipole-dipole interaction tensor.

$$\frac{\partial \boldsymbol{r}_{D,i}}{\partial \boldsymbol{r}_k} = \frac{q_{D,i}}{k_D}\boldsymbol{T}'_{a_D i,k} q_k + \delta_{ik}\boldsymbol{I} \tag{4.14a}$$

$$\boldsymbol{T}'_{a_D i,k} = \begin{bmatrix} \frac{3(a_{D,i}^x - r_k^x)^2}{|\boldsymbol{a}_{D,i}-\boldsymbol{r}_k|^5} - \frac{1}{|\boldsymbol{a}_{D,i}-\boldsymbol{r}_k|^3} & \frac{3(a_{D,i}^x - r_k^x)(a_{D,i}^y - r_k^y)}{|\boldsymbol{a}_{D,i}-\boldsymbol{r}_k|^5} & \frac{3(a_{D,i}^x - r_k^x)(a_{D,i}^z - r_k^z)}{|\boldsymbol{a}_{D,i}-\boldsymbol{r}_k|^5} \\ \frac{3(a_{D,i}^x - r_k^x)(a_{D,i}^y - r_k^y)}{|\boldsymbol{a}_{D,i}-\boldsymbol{r}_k|^5} & \frac{3(a_{D,i}^y - r_k^y)^2}{|\boldsymbol{a}_{D,i}-\boldsymbol{r}_k|^5} - \frac{1}{|\boldsymbol{a}_{D,i}-\boldsymbol{r}_k|^3} & \frac{3(a_{D,i}^y - r_k^y)(a_{D,i}^z - r_k^z)}{|\boldsymbol{a}_{D,i}-\boldsymbol{r}_k|^5} \\ \frac{3(a_{D,i}^x - r_k^x)(a_{D,i}^z - r_k^z)}{|\boldsymbol{a}_{D,i}-\boldsymbol{r}_k|^5} & \frac{3(a_{D,i}^y - r_k^y)(a_{D,i}^z - r_k^z)}{|\boldsymbol{a}_{D,i}-\boldsymbol{r}_k|^5} & \frac{3(a_{D,i}^z - r_k^z)^2}{|\boldsymbol{a}_{D,i}-\boldsymbol{r}_k|^5} - \frac{1}{|\boldsymbol{a}_{D,i}-\boldsymbol{r}_k|^3} \end{bmatrix} \tag{4.14b}$$

Substituting Eq. (4.14) into Eq. (4.12) one can obtain Eq. (4.15).

$$\frac{dU(\boldsymbol{r},\boldsymbol{a}_D)}{d\boldsymbol{r}_i} = \frac{\partial U(\boldsymbol{r},\boldsymbol{a}_D)}{\partial \boldsymbol{r}_i} + \sum_j \frac{\partial U(\boldsymbol{r},\boldsymbol{a}_D)}{\partial \boldsymbol{r}_{D,j}}\left(\frac{q_{D,j}q_i}{k_D}\boldsymbol{T}'_{a_D j,i} + \delta_{ji}\boldsymbol{I}\right) \tag{4.15}$$

Eq. (4.15) is the working definition of the atomic center gradients that are used to drive the molecular dynamics updates of positions and velocities for the atomic and Drude particles. It should be noted that the introduction of the dipole interaction tensor, $\boldsymbol{T}'$, that follows from our theory is not always available from community codes that are currently restricted to point charges and point charge-only Ewald calculations. For this work we modified the TINKER software package to use Drude polarization instead of its

usual induced point dipole formalism. TINKER already has code to handle the dipole interaction tensor and higher-order multipolar Ewald sums[101, 103], meaning that much of the software infrastructure to properly evaluate Eq. (4.15) is already in place.

4.3 Methods

Three different methods for solving the Drude polarization condition in Eq. (4.3) were then implemented for comparison purposes- an SCF solver, the iEL/0-SCF method described above, and the EL(T,T*) method based on the dual temperature NVT,T* ensemble described by Lamoureux and Roux for the PSPC model[43]. The SCF solver is a naïve successive over-relaxation (SOR) method[60] that iterates Eq. (4.3) to within some defined tolerance in terms of the root mean square change in force on the Drude particles between successive iterations. A range of tolerances were tested with the tightest being 10^{-6} RMS kcal/mol/Å, which I define as the 'gold standard' for accuracy in this study.

The EL(T,T*) method for polarization was originated by Sprik[69], and was used by Wodak and colleagues for point induced dipoles[68] and by Lamoureux and Roux for Drude polarization[43]. Under the EL(T,T*) scheme the Drude particles themselves are given a portion of the mass of their parent atom, $m_{D,i}$, and to conserve mass the parent atom now has a mass of $m_i - m_{D,i}$. The Lagrangian of the system then becomes Eq. (4.16).

$$\begin{aligned}\mathcal{L}(\boldsymbol{r},\dot{\boldsymbol{r}},\boldsymbol{r}_D,\dot{\boldsymbol{r}}_D) = \frac{1}{2}\sum_i (m_i - m_{D,i})|\dot{\boldsymbol{r}}_i|^2 + \frac{1}{2}\sum_i m_{D,i}|\dot{\boldsymbol{r}}_{D,i}|^2 - U_{LJ}(\boldsymbol{r}) - \sum_i \sum_{j>i} \frac{q_i q_j}{|\boldsymbol{r}_i - \boldsymbol{r}_j|} \\ - \sum_i \sum_j \frac{q_i q_{D,j}}{|\boldsymbol{r}_i - \boldsymbol{r}_{D,j}|} - \sum_i \sum_{j>i} \frac{q_{D,i} q_{D,j}}{|\boldsymbol{r}_{D,i} - \boldsymbol{r}_{D,j}|} - -\frac{1}{2} k_D \sum_i |\boldsymbol{r}_i - \boldsymbol{r}_{D,i}|^2\end{aligned} \tag{4.16}$$

In Eq. (4.16) the last five terms give the potential energy of the system and the Lagrangian can be simplified as shown in Eq. (4.17).

$$\mathcal{L}(\boldsymbol{r},\dot{\boldsymbol{r}},\boldsymbol{r}_D,\dot{\boldsymbol{r}}_D) = \frac{1}{2}\sum_i (m_i - m_{D,i})|\dot{\boldsymbol{r}}_i|^2 + \frac{1}{2}\sum_i m_{D,i}|\dot{\boldsymbol{r}}_{D,i}|^2 - U(\boldsymbol{r},\boldsymbol{r}_D) \tag{4.17}$$

Applying the Euler-Lagrange equation of motion one can now obtain equations of motion for atomic centers and Drude particles in Eqs. (18a) and (18b), respectively.

$$(m_i - m_{D,i})\ddot{\boldsymbol{r}}_i = -\frac{dU}{d\boldsymbol{r}_i} \tag{4.18a}$$

$$m_D \ddot{\boldsymbol{r}}_{D,i} = -\frac{dU}{d\boldsymbol{r}_{D,i}} \tag{4.18b}$$

In short, the dynamics of the atomic centers are driven by their usual forces, but with a smaller mass, and a net force now drives the dynamics of the Drude particles given by the harmonic spring attaching it to its parent atom and an electrostatic force through interactions with its charge and the charges of the rest of the system.

In the NVT,T* or NPT,T* ensemble of the EL(T,T*) method, the atomic temperature T is defined by the total atomic mass m_i and the center of mass velocity of the Drude-parent atom pair *i*, $\dot{\boldsymbol{R}}_i$.

$$T = \frac{1}{3k_B N_{atom}} \sum_{i \in atom} \frac{\left| m_i \dot{\boldsymbol{r}}_i + m_{D,i} \dot{\boldsymbol{r}}_{D,i} \right|^2}{m_i} = \frac{1}{3k_B N_{atom}} \sum_{i \in atom} m_i \dot{\boldsymbol{R}}_i^2 \tag{4.19}$$

In Eq. (4.19) N_{atom} is the number of atomic centers. The Drude particles are kept at a lower temperature T*, defined by the reduced mass and the relative velocity of the Drude-parent atom pair, m_i' and $\dot{\boldsymbol{d}}_i$, respectively.

$$T^* = \frac{1}{3k_B N_{Drude}} \sum_{i \in Drude} m_D \left(1 - \frac{m_D}{m_i}\right) \left| \dot{\boldsymbol{r}}_{D,i} - \dot{\boldsymbol{r}}_i \right|^2 = \frac{1}{3k_B N_{Drude}} \sum_{i \in Drude} m_i' \dot{\boldsymbol{d}}_i^2 \tag{4.20}$$

In Eq. (4.20) N_{Drude} is the number of Drude centers. By keeping the Drude temperature low, kinetic energy is bled out to better satisfy the Born-Oppenheimer condition.

For the iEL/0-SCF and EL(T,T*) methods the Drude positions and auxiliaries were initialized to those of a tight SCF solution, and all production simulations were started from pre-equilibrated restart files. Unless noted otherwise all simulations (NVT or NVE) are initialized at 298.0 K. For all methods all equations of motion were integrated with a velocity Verlet scheme[88] and I used particle-mesh Ewald summation to treat all electrostatic interactions[101, 103] with a real-space cutoff of 9.0 Å. Non-bonded interactions were evaluated using neighbor lists and intra-molecular geometry of the PSPC model was constrained using the RATTLE algorithm[59]. When thermostats are used, either to perform NVT simulations or to couple to auxiliaries (for iEL/0-SCF) or relative Drude-parent atom harmonic motion (for EL(T,T*)), a time-reversible 4th-order Nosé-Hoover chain[65-66] was used with a thermostat time scale parameter of 0.005 ps (expect for the EL method at a time step of 4.0 fs, which uses a time scale parameter of 0.01 ps, as 0.005 ps is unstable at that condition) . All three methods were tested over a range of time steps, ranging from 0.5 fs to 6.0 fs depending on method.

The iEL/0-SCF method requires the determination of an optimal value of γ, the mixing parameter that allows us to approximate a ground state polarization solution. In Chapter 3 I showed the iEL/0-SCF is numerically stable over long simulation times because the general form of the polarization energy and forces (Eqs. (4.5) and (4.12) for the Drude model) ensures that small deviations from the converged solution only give rise to errors on the order of $O(\delta^2)$ in energy and forces, as discussed in Chapter 3[22]. Thus the linear mixing parameter γ will control the level of accuracy of the polarization solution and corresponding stability of the equations of motion.

In Chapter 3 I determined γ by looking for a value that best conserved the real system energy while minimizing the auxiliary pseudo temperature drift rate that arises from resonances. For AMOEBA, I found an optimal $\gamma = 0.9$ value for water using a 0.5 fs time step in the NVE ensemble, with values of the time step greater than 0.5 fs and $\gamma > 0.9$ displaying unstable trajectories, and values of $\gamma < 0.5$ yielding very poor real system energy conservation regardless of time step. As shown in Figures E.1-E.3, the

Drude model allows for a far greater range of γ values and a time step of up to 2.0 fs for which the NVE trajectory is both stable and exhibits good energy conservation comparable to or better than the SCF solution. I adopted a value of $\gamma = 1.0$ for all simulations reported in the Results (section 4.4).

4.4 Results

For the iEL/SCF methods, the equations of motion of the Drude particles given in Eq. (4.10) are defined in the limit $m_i^a \to 0$, so there are no contributions to the real system energy due to the kinetic energy of the Drude particles. This is equivalent to an SCF scheme under tight convergence (which I take to be 10^{-6} RMS kcal/mol/Å), but is unlike an EL method where there is mass repartitioning between the Drude particle and its parent atom. For EL, the assignment of small Drude masses are preferred for energy conservation by minimizing the Drude kinetic energy (the Born Oppenheimer condition) but the simulation is restricted to short time steps due to their higher frequency oscillations (since $\omega = \sqrt{k_D/m_D}$). By contrast larger masses in the EL scheme that allow for longer time steps tend to conserve energy more poorly due to the increasing Drude kinetic energy contribution. In what follows I report results in the EL method using the lowest stable Drude mass, m_D, for a given time step (see Fig. 4.3a).

In the NVE ensemble, since I am not enforcing temperature control through thermostats, the two extended Lagrangian approaches simplify to just the equivalent of EL/0-SCF (no inertial restraints) and EL methods, as opposed to iEL/0-SCF and EL(T,T*), respectively. Figure E.4 shows that the system energy, which is the sum of all kinetic and potential energy of atomic and Drude particle contributions, for both EL approaches is as well conserved as the SCF gold standard at a time step of 1.0 fs over a 2.0 ns timescale. Figure 4.1a shows that while the EL/0-SCF method also conserves energy using a 2.0 fs time step (in fact better than the SCF method over the 2.0 ns trajectory) the EL method conserves energy poorly after ~250 ps at the lowest Drude mass where stable trajectories are possible. By restricting to the first several hundred picoseconds for the EL method, I can collect values of the diffusion constant at 298.0 K that are in good agreement between the methods, 4.00 x10^{-5} cm^2/s (± 0.08, SCF), 4.06 x10^{-5} cm^2/s (± 0.08, EL/0-SCF), and 3.99 x10^{-5} cm^2/s (± 0.23, EL), and all of which are in agreement with previous reported results for the PSPC model[42-43].

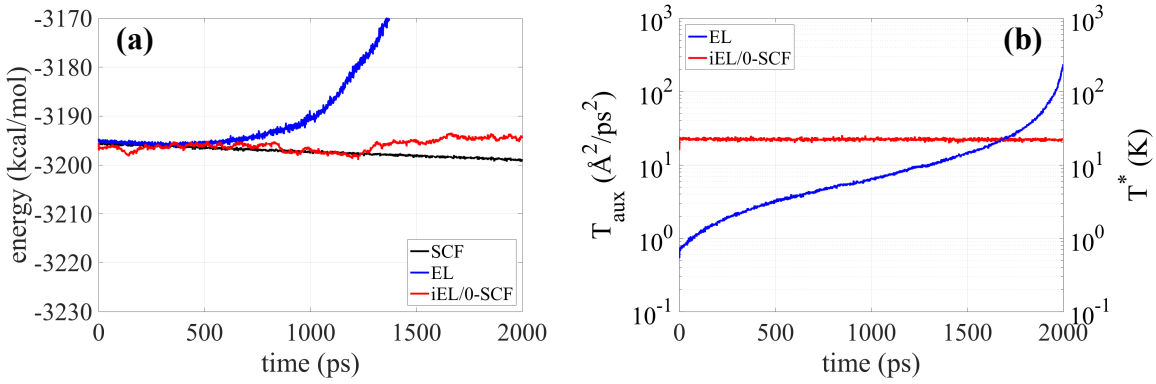


Figure 4.1: *NVE energy properties of the iEL/0SCF and EL(T,T*=0) method for the Drude PSPC model compared to the SCF reference. All simulations were performed with a time step of 2.0 fs on a test system of 512 water molecules with no thermostats. (a) Total real system energy for SCF (black), iEL/0-SCF (red), and EL (blue) (b) auxiliary variables pseudo temperature for iEL/0-SCF (red, left axis) and temperature of the Drude-parent atom motion of the EL method (blue, right axis). Temperature is placed on a log scale.*

The energy instability of EL after about 250 ps arises from the fact that even for small values of $m_{D,i}$ the Drude particles will eventually reach thermal equilibrium with the rest of the system, and the kinetic energy of the Drude particles (which is formally zero for the SCF and EL/0-SCF methods) will increase over time for the EL method. To show this, Figure 4.1b reports the Drude temperature for the relative Drude-parent atom motion for the SCF and EL methods and the 'pseudo temperature' of the auxiliary dipoles, defined as $T_{aux} = 1/3\langle \dot{\boldsymbol{a}}_i^2 \rangle$, for the EL/0-SCF method at the same 2.0 fs time step and simulated over 2.0 ns. The EL method shows exponential degradation, rising from near 0.0 K at the beginning of a simulation to almost the atomic temperature after 2.0 ns. This will be the case for the EL method even for smaller time steps where the same kinetic energy change will be evident but on a longer simulation time scale. By contrast the EL/0-SCF and SCF methods show linear increases in inertia throughout the simulation due to accumulation of integration error. Of course the point of the EL(T,T*) approach is to arrest this energy redistribution by keeping the Drude thermostat cold to stay near the Born-Oppenheimer surface. But the NVE results are instructive as they are a harbinger for the fact the EL/0-SCF is intrinsically more stable than the EL method at a given time step and length of simulation, and comparable in quality to a well converged SCF solution.

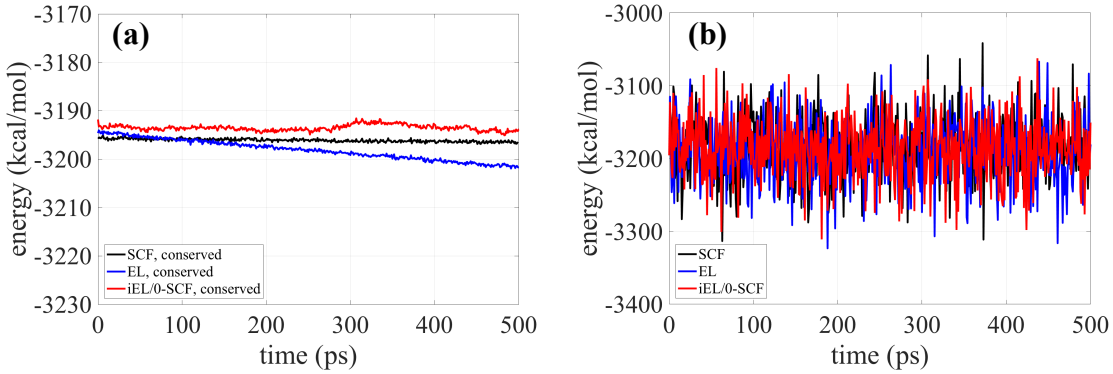


Figure 4.2: *NVT energy properties of the EL, SCF and iEL/0-SCF method for the Drude PSPC water model. All simulations were performed with a time step of 2.0 fs on a test system of 512 water molecules. (a) The conserved quantity for the NVT extended system for the three methods, (b) the real system energy in the NVT ensemble at 298.0 K for SCF (black), EL(T*,T) (blue), and iEL/0-SCF (red). The system energy is the sum of the atomic kinetic and potential energies.*

I show that this continues to hold in the NVT ensemble by next considering the polarization properties of the PSPC model, where the atomic degrees of freedom are coupled to a Nosé-Hoover thermostat to maintain a temperature of 298.0 K. Figure 4.2a shows the extended system conserved energy quantity, which differs for the three methods, using a 2.0 fs timestep. The conserved quantity for the SCF gold standard corresponds to all potential and kinetic energy terms of the atomic degrees of freedom

and the atomic thermostats[66]. For the EL(T,T*) method the ensemble is NV(T,T*), since a cold T* thermostat of 1.0K is coupled to the Drude-parent atom relative motion, and thus its conserved energy quantity now includes both sets of thermostats. The iEL/0-SCF method also couples the pseudo temperature of the auxiliaries to a thermostat. So there is an extended system energy that includes the potential energy and kinetic energies of all the particles in the system, all of the auxiliary variable contributions, and the potential and kinetic energy of all the thermostats (atomic and auxiliary). To make the contributions from auxiliary terms commensurate with the real terms we can fit an effective fictitious auxiliary mass, $m_{a,eff}$, although it is formally 0 from the derivation of the method. In essence the extended system energy conservation is enforced by construction with errors only associated with uncertainty in the linear fit. It should be noted that these fitted auxiliary masses are purely for analysis purposes and do not contribute to the actual simulations, which formally partition no mass to the Drude or auxiliary Drude particles.

Table 4.1 shows that the EL(T,T*) method has more drift in the conserved energy than does the SCF solution, although the real system energy and molecular dipole properties among the three methods are well reproduced at a 2.0 fs time step. However, the greater drift in the conserved energy quantity for EL(T,T*) indicates that it is becoming more sensitive to integration errors and thus limitations on allowed time steps compared to the SCF and iEL/0-SCF approaches.

Table 4.1. *Conserved extended energy drift, average system potential energy, and average molecular dipole of the PSPC model in the NVT ensemble for the SCF, EL and iEL/0-SCF methods at different integration time steps (fs).* $\langle U \rangle$ *is the average system potential energy in kcal/mol,* $\langle \mu_{mol} \rangle$ *is the average molecular dipole in units of Debye, and* $|\Delta E|/\Delta t$ *is the extended system conserved energy quantity in kcal/mol/ps. For iEL/0-SCF the conserved quantity is enforced by construction. Results are from 1.0 ns simulations of 512 water molecules at 298.0 K.*

	SCF			**EL(T,T*)**			**iEL/0-SCF**	
Δt	$\langle \mu_{mol} \rangle$	$\langle U \rangle$	$\lvert\Delta E\rvert/\Delta t$	$\langle \mu_{mol} \rangle$	$\langle U \rangle$	$\lvert\Delta E\rvert/\Delta t$	$\langle \mu_{mol} \rangle$	$\langle U \rangle$
1.0	2.71 ± 0.01	-4110 ± 35	-0.0011	2.71 ± 0.01	-4110 ± 35	-0.0023	2.71 ± 0.01	-4110 ± 35
2.0	2.71 ± 0.01	-4100 ± 35	-0.0017	2.71 ± 0.01	-4100 ± 35	-0.014	2.71 ± 0.01	-4100 ± 35
3.0	2.71 ± 0.01	-4090 ± 35	-0.0037	2.71 ± 0.01	-4080 ± 35	-1.8	2.71 ± 0.01	-4080 ± 35
4.0	2.71 ± 0.01	-4070 ± 35	-0.0014	2.70 ± 0.01	-4070 ± 35	-81	2.70 ± 0.01	-4050 ± 35
5.0	2.70 ± 0.01	-4040 ± 35	-0.0027	---	---	---	2.70 ± 0.01	-4020 ± 35
6.0	2.69 ± 0.01	-4010 ± 35	+0.057	---	---	---	2.69 ± 0.01	-3990 ± 30
7.0	2.69 ± 0.01	-3970 ± 35	+1.0	---	---	---	2.69 ± 0.01	-3950 ± 35

This is confirmed in Figure 4.3a which plots the minimum Drude mass required to maintain a stable simulation as a function of increasing time step from 1.0 fs to 4.0 fs for the EL(T,T*) method where T = 298.0 K and using a Drude-parent atom relative motion temperature T*= 1.0 K; beyond 4.0 fs the EL(T,T*) approach becomes unstable under any amount of mass repartitioning. As the integration time step increases, the repartitioned Drude mass must also increase, thereby becoming more and more untenable for accurately preserving the Born-Oppenheimer condition, which would manifest as evident degradation in polarization properties. This is seen in Figure 4.3b and 4.3c (as well as Figures E.11 and E.12), which report the induced dipole distributions and induced dipole autocorrelation function at the 4.0 fs time step (in which about 25% of the mass of

the parent oxygen atom is partitioned to the Drude oscillator) are in significant disagreement with the SCF result.

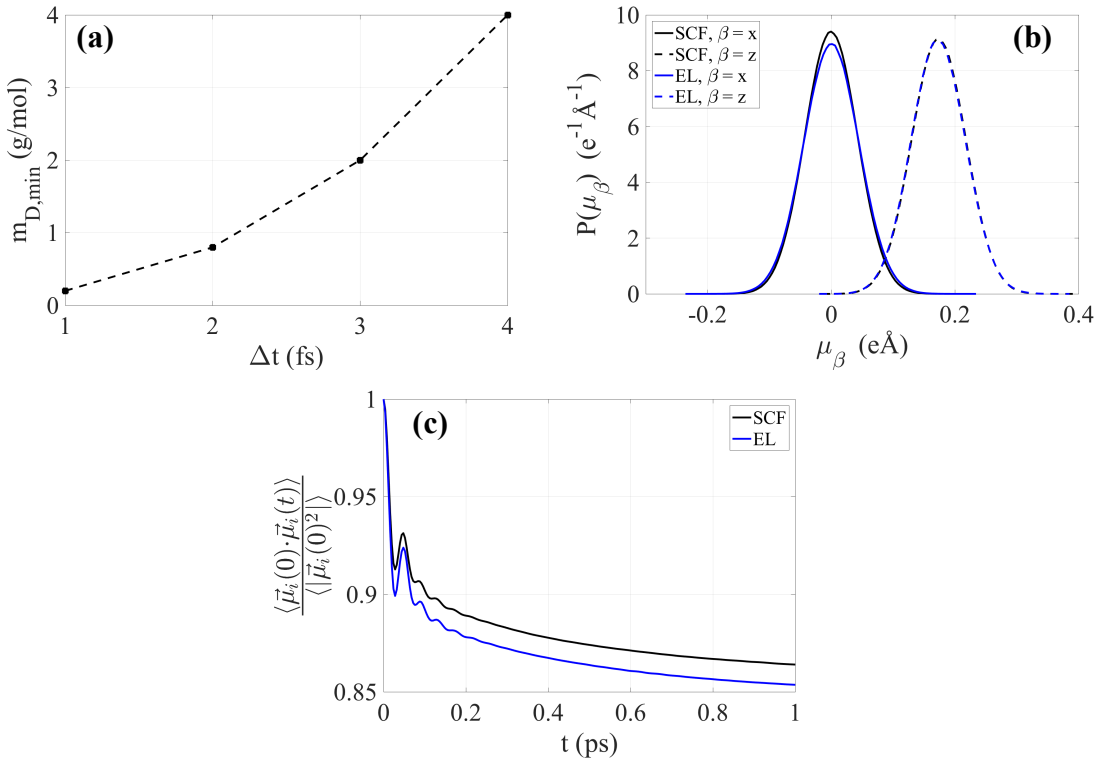


Figure 4.3: *Molecular dynamics trajectory stability and polarization properties of the PSPC model using the EL(T,T*) method. (a) the minimum Drude mass required to maintain a stable MD simulation as a function of time step using a Drude-parent atom relative motion temperature set point of T* = 1.0 K; the EL(T,T*) method is unstable above 4.0 fs. (b) probability density distributions and (c) autocorrelation function for the oxygen induced dipole,* $\boldsymbol{\mu}_i = q_{D,i}(\boldsymbol{r}_{D,i} - \boldsymbol{r}_i)$ *from the EL(T,T*) method compared to SCF at a 4.0 fs time step. All simulations were performed in the NVT ensemble at 298.0 K. All calculations presented in this figure use an internal coordinate system where the z-direction is given by the H-O-H bisector, the y-direction is out of the H-O-H plane, and the x-direction is orthogonal to z and y (see*[31] *for details).*

By contrast, Figure 4.4a shows that the effective, fitted mass of the auxiliary Drude centers for the iEL/0-SCF method is about an order of magnitude smaller than the repartitioned masses of the EL(T,T*) method, such that the effective temperature of the auxiliary degrees of freedom remain cold (< 10.0 K) for up to a 6.0 fs time step. Figures 4.4b and 4.4c show that the iEL/0-SCF method tracks the SCF result accurately for the induced dipole distributions and induced dipole autocorrelation function at the 4.0 fs time step, unlike the EL(T,T*) method. In fact, polarization properties are as well reproduced as the SCF solution all the way up to the 6.0 fs time step (Figure E.10 and E.11) and other properties generated by iEL/0-SCF, such as the radial distribution functions of the liquid, match the SCF result at this largest time step (Figure 4.5).

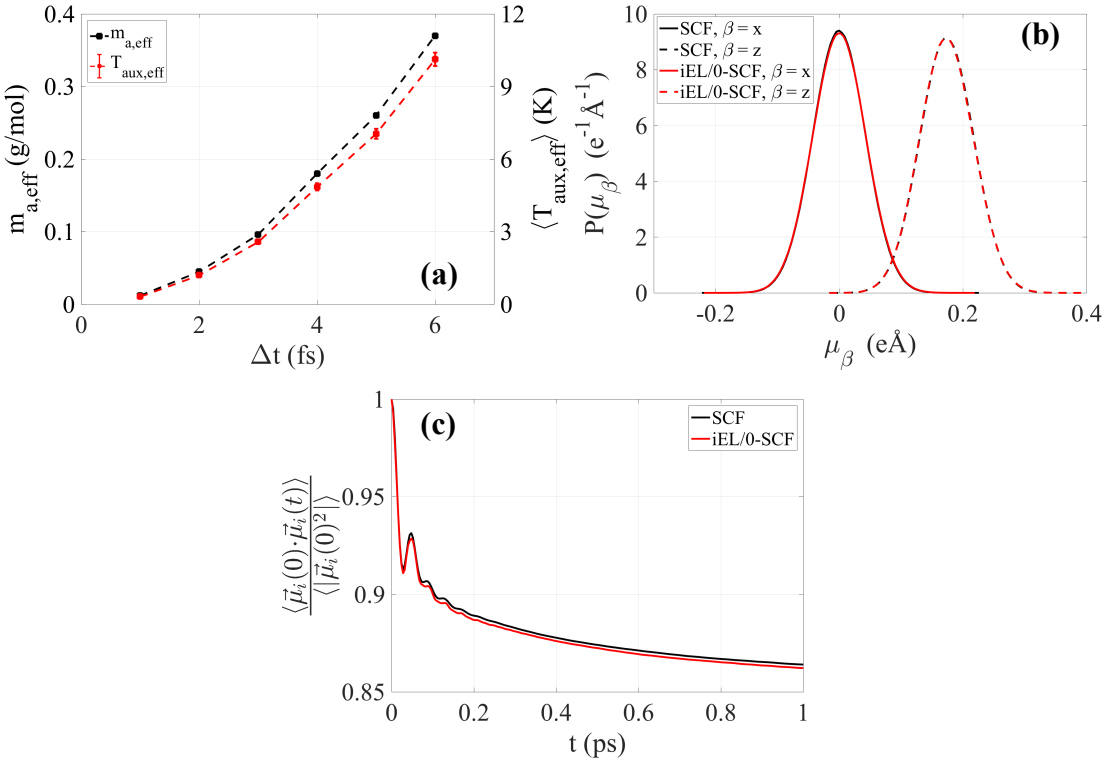
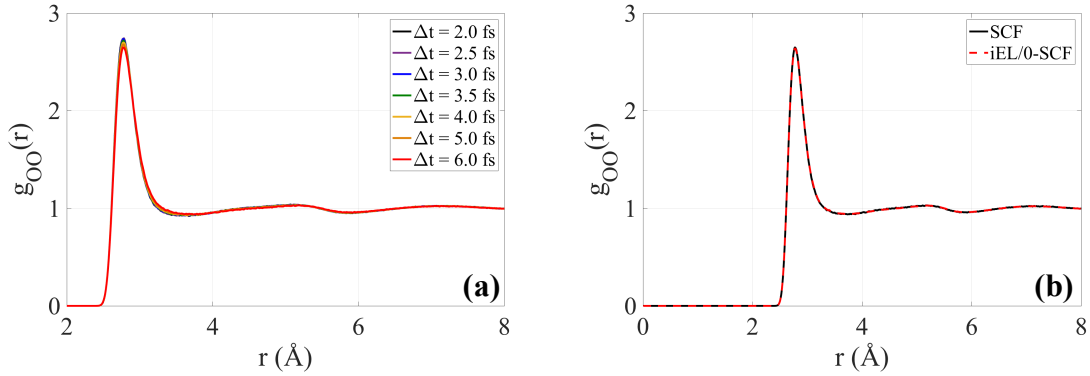


Figure 4.4: *Molecular dynamics trajectory stability and polarization properties of the PSPC model using the iEL/0-SCF method. (a) The effective auxiliary mass, $m_{a,eff}$ (left axis) and the resultant auxiliary mean temperatures (right axis) using the effective masses to convert from pseudo temperature to real temperature for the iEL/0-SCF method. (b) probability density distributions and (c) autocorrelation function for the oxygen induced dipole, $\boldsymbol{\mu}_i = q_{D,i}(\boldsymbol{r}_{D,i} - \boldsymbol{r}_i)$ from the iEL/0-SCF method compared to SCF at a 4.0 fs time step. See Figure 4.3 for remaining details.*



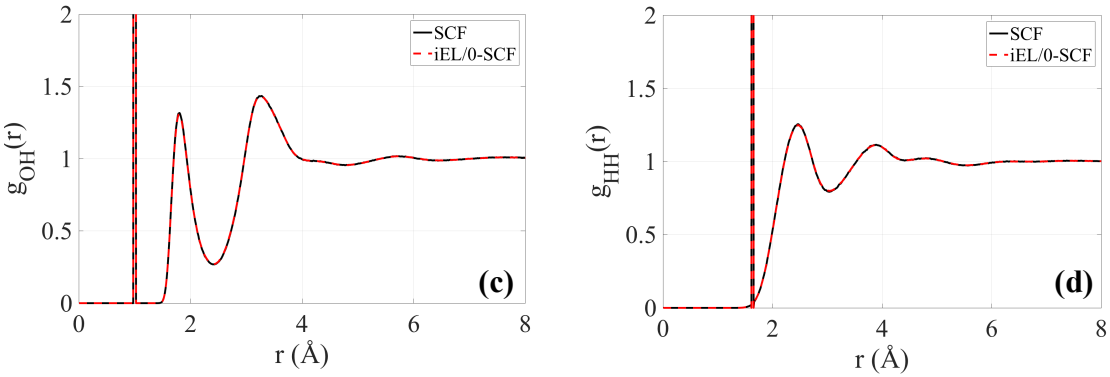


Figure 4.5: *Radial distribution functions (RDF) of the PSPC Drude water model using the SCF and iEL/0-SCF methods. (a) oxygen-oxygen RDF as a function of time step for the SCF method showing small changes in the first peak as time step increases. Oxygen-oxygen RDF (b), oxygen-hydrogen RDF (c), and hydrogen-hydrogen RDF for the SCF method and iEL/0-SCF method. Simulations were run in the NVT ensemble at 298.0 K and were at least 0.5 ns in length. Panels (b), (c), and (d) used a time step of 6.0 fs.*

4.5 Discussion

I would now like to consider why the iEL/0-SCF method is as good as an SCF solution although it requires no explicit SCF calculations at each time step. The hypothesis is that the induced dipole updates, or equivalently the position updates of the real Drude particles, by making one evaluation of the electric field using the current auxiliary positions is performing a tightly converged SCF calculation on the fly. To bolster this argument, Figure 4.6a shows the autocorrelation function for the real dipoles and the single exponential fit to the initial decay, while Figure 4.6b reports the fitted time constant, as a function of the integration time step. Since the decay constant is on the order of a couple of hundred femtoseconds, the numerical integration of the auxiliary dipoles can "keep up" for adjusting the real polarization degrees of freedom for time steps up to 6.0 fs to follow that longer time scale decay. I found that a 7.0 fs time step remained stable but properties of the PSPC model were beginning to degrade, and at 8.0 fs the simulation ultimately developed numerical instabilities after about 1.0 ns.

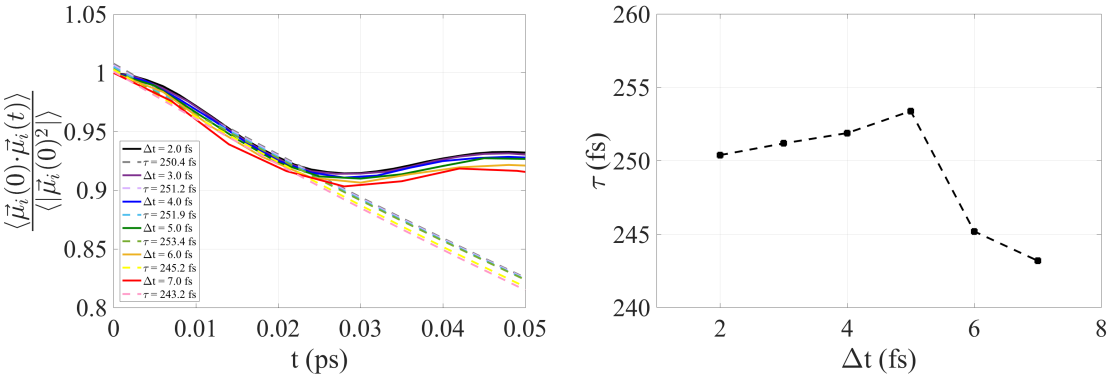


Figure 4.6: *The initial time scales for decay of the oxygen time autocorrelation for the real induced dipoles for the iEL/0-SCF method. (a) single exponential fits (dotted lines) to the initial decay of the dipole created between the Drude and its parent atom, $\boldsymbol{\mu}_{D,i} = q_{D,i}(\boldsymbol{r}_{D,i} - \boldsymbol{r}_i)$ (solid lines), for a range of time steps. (b) the time constant of the exponential fit to the initial decay from (a) as a function of time step.*

It is also useful to analyze why the AMOEBA[31, 34] polarizable model for water is restricted to a time step that is a factor of 6 times smaller than the PSPC model using the iEL/0-SCF approach, which I attribute to the following. First, the PSPC has a rigid and not a flexible geometry like AMOEBA. Second, PSPC does not include a Drude particle on the light hydrogen centers, unlike the AMOEBA model, which includes induced dipoles on hydrogens. Third, PSPC has no intramolecular polarization and AMOEBA does. Finally, PSPC does not have short range, rapidly varying forces arising from dipole and quadrupole permanent electrostatics, as AMOEBA does. Thus the simplicity of the PSPC model is better able to conform its motion on the Born-Oppenheimer surface with much longer time steps based on a simpler design choice of the polarizable and electrostatic model.

It would also be expected that for a polarizable Drude model for a protein with flexible bonds or a 4-site Drude polarizable water model that places Drude particles on light hydrogens that the time step of 6.0 fs possible with the PSPC model presented here would need to decrease. However, there is a possibility that a time reversible integration of an auxiliary system could nonetheless help boot strap the real degrees of freedom toward larger time steps even for these cases.

4.6. Conclusions

In this chapter I have presented a Drude implementation of the iEL/0-SCF method, which proves to be as accurate and stable for both dynamic and thermodynamic properties for the PSPC water model when compared to a tightly converged SCF solution, and a significant improvement over the two temperature EL(T,T*) approaches by increasing the allowable integration time step by a factor of 2-3 with recommended time steps as large as 6.0 fs. For iEL/0-SCF the better stability likely comes from the exact analytical agreement between the energy and gradients compared to SCF, and the avoidance of any mass repartitioning when compared to EL so that the Born-Oppenheimer condition is easily satisfied. I therefore conclude that using iEL/0-SCF for Drude gains the benefits of the minimal computational expense of an EL approach but with greatly increased time steps and the accuracy and robustness of an SCF method. As I noted in the Theory section (section 4.2), the addition of the iEL/0-SCF method to community codes would need to handle the dipole-dipole interaction tensor (Eq. (4.15)) and higher-order multipolar Ewald sums. Given that Amber and CHARMM have implemented the AMOEBA model, the software to implement the iEL/0-SCF approach is largely in place in these community codes.

5.
General-Order Potentials, Integration, and Dissipation for iEL/0-SCF Methods

5.1 Introduction

Under the Born-Oppenheimer approximation[64] 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. Over the last decade Niklasson and colleagues[70-81] 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 is termed a "shadow potential" (which conceptually is closely related to a shadow Hamiltonian[104]). 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[105]. Fundamentally these XLBOMD approaches form the basis for the iEL/SCF and iEL/0-SCF methods presented in Chapters 2, 3, and 4. While all such 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) [73]. Alternatively, and more effectively for large classical systems[92], one can use the auxiliary thermostatting scheme presented in Chapter 2 to tackle this issue, as well.

In this chapter I show that the iEL/0-SCF approach introduced in Chapter 3 is part of a general class of approaches whose success rests on the proper formulation of a shadow potential functional, that is 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[79, 81] previous approaches to building such potentials tended to be *ad hoc*[22, 79, 81] 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, I combine the shadow potential formalism with Langevin-like friction[73-75] 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[75], in this chapter I develop a general structure for systematically increasing the order of integration and dissipation to better satisfy time reversibility and energy conservation. I will 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 iEL/0-SCF formalism, as illustrated on a simple fluctuating charge model, but trivially extensible to induced formalism for polarization. Finally, I will show how the iEL/0-SCF method fits within this general framework, whose success 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.

5.2 Theory

Here I 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 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 one can choose or construct shadow potentials, dissipation, and integration, each to any order, allowing one to control accuracy and energy drift to an arbitrary degree. I present the theory in terms of a fluctuating charge model, for which I have formulated numerical experiments in the Results (section 5.3), as well as the small variations needed to formulate the same theory for an inducible dipole model described in Appendix F.

5.2.1 *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. (5.1).

$$U^{el}(\boldsymbol{r},\boldsymbol{q}) = \boldsymbol{\chi}^{\boldsymbol{T}}\boldsymbol{q} + \frac{1}{2}\boldsymbol{q}^{\boldsymbol{T}}\boldsymbol{C}\boldsymbol{q} \tag{5.1}$$

In Eq. (5.1) $\boldsymbol{\chi}$ 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 ($\boldsymbol{q} = [q_1 \dots q_N]^T$), $\boldsymbol{r}$ denotes the positions of the N atoms in the system, and $\boldsymbol{C} = \boldsymbol{C}(\boldsymbol{r})$ 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 ($C_{ii} = 2H_i$)[47]. Ignoring enforcement of charge neutrality for simplicity, the set of charges that minimizes the potential surface in Eq. (5.1) is thus the ground state Born-Oppenheimer solution for the system as given by Eq. (5.2).

$$\frac{\partial U^{el}}{\partial \boldsymbol{q}} = 0 = \boldsymbol{\chi}^{\boldsymbol{T}} + \boldsymbol{C}\boldsymbol{q} \quad \Longrightarrow \quad \boldsymbol{q}_{min} = -\boldsymbol{C}^{-1}\boldsymbol{\chi} \tag{5.2}$$

The potential energy surface is then given as the variationally optimized equilibrated charge minimum, that is $U^{el}(\boldsymbol{r}^N) = U^{el}(\boldsymbol{r}^N, \boldsymbol{q}_{min})$ from Eq. (5.2). Note that solving for $\boldsymbol{C}^{-1}$ through matrix inversion or via an SCF procedure is the costly step that needs to be avoided.

I now introduce a dynamical auxiliary variable $\boldsymbol{X}$, which is designed to be a good approximation to $\boldsymbol{C}^{-1}$. Next a 'shadow' potential functional is built around $\boldsymbol{X}$ using some function $f(\boldsymbol{X})$ that is required to give a better approximation of $\boldsymbol{C}^{-1}$ than $\boldsymbol{X}$ itself, and for which one can control the degree to which this shadow potential matches the reference potential of Eq. (5.1). This is presented in Eq. (5.3).

$$U^{el}_{shadow}(\boldsymbol{r}, \boldsymbol{X}, \boldsymbol{q}) = \boldsymbol{\chi}^{\boldsymbol{T}}\boldsymbol{q} + \frac{1}{2}\boldsymbol{q}^{\boldsymbol{T}}[f(\boldsymbol{X})]^{-1}\boldsymbol{q} \tag{5.3}$$

Eq. (5.4) 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 $m \rightarrow \infty$ whose general form was introduced by Niklasson[106] as a generalization of the Schulz method originally derived for the case of $m = 2$[107].

$$f(\boldsymbol{X}) = \boldsymbol{C}^{-1}[\boldsymbol{I} - (\boldsymbol{I} - \boldsymbol{C}\boldsymbol{X})^m] \tag{5.4}$$

The set of charges $\boldsymbol{q}$ that minimizes this new form of the shadow potential functional can be determined by substituting Eq. (5.4) into Eq. (5.3) and differentiating, as shown in Eq. (5.5).

$$\frac{\partial U^{el}_{shadow}(\boldsymbol{r}^N, \boldsymbol{X}, \boldsymbol{q})}{\partial \boldsymbol{q}} = 0 = \boldsymbol{\chi}^{\boldsymbol{T}} + [f(\boldsymbol{X})]^{-1}\boldsymbol{q} \tag{5.5a}$$

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

$$U^{el}_{shadow}(\boldsymbol{r}^N, \boldsymbol{X}) = \frac{1}{2}\boldsymbol{\chi}^{\boldsymbol{T}}\boldsymbol{q}_{min} = -\frac{1}{2}\boldsymbol{\chi}^{\boldsymbol{T}}\boldsymbol{C}^{-1}[\boldsymbol{I} - (\boldsymbol{I} - \boldsymbol{C}\boldsymbol{X})^m]\boldsymbol{\chi} = -\frac{1}{2}\boldsymbol{\chi}^{\boldsymbol{T}}f(\boldsymbol{X})\boldsymbol{\chi} \tag{5.5b}$$

From Eq. (5.5a) one 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. (5.5b), then, gives various equivalent forms of this shadow potential energy surface at the variationally optimized minimum described by Eq. (5.5a) for different values of m. Note that while $\boldsymbol{C}^{-1}$ does appear explicitly in Eq. (5.5) it is annihilated when multiplied into the expansion and does not appear when writing $f(\boldsymbol{X})$ for some finite m (Table 5.1).

Following the previous work of Niklasson and colleagues[79, 81] one can now define an extended Lagrangian for a system with the shadow potential of Eq. (5.5b). Instead of introducing auxiliary degrees of freedom that represent the electronic degrees of freedom themselves (auxiliaries representing $\boldsymbol{q}$) we now 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 dynamic degree of freedom with a fictitious mass, m_X. This extended Lagrangian is given by Eq. (5.6).

$$\mathcal{L}\left(\boldsymbol{r}^N, \dot{\boldsymbol{r}}^N, \boldsymbol{X}, \dot{\boldsymbol{X}}\right) = \frac{1}{2}\sum_{i=1}^{N} m_i \dot{\boldsymbol{r}}_i^2 + \frac{1}{2} m_X \mathrm{tr}\left[\dot{\boldsymbol{X}}^2\right] - U(\boldsymbol{r}^N) + \frac{1}{2}\boldsymbol{\chi}^{\boldsymbol{T}} f(\boldsymbol{X})\boldsymbol{\chi} - \frac{1}{2} m_X \omega^2 \mathrm{tr}\left[\left(\widetilde{\boldsymbol{C}}^{-1} - \boldsymbol{X}\right)^2\right] \tag{5.6}$$

In Eq. (5.6) $U(\boldsymbol{r})$ are other non-electrostatic potential energy functions that may be in the system such as bonds, angles, van der Waals interactions, etc. that are independent of the fluctuating charges. The final term in Eq. (5.6) 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 is to be avoided, 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 5.1: *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_{shadow}^{el}, 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_{shadow}^{el}	$-\frac{1}{2}\boldsymbol{\chi}^T(2\boldsymbol{X} - \boldsymbol{XCX})\boldsymbol{\chi}$	$-\frac{1}{2}\boldsymbol{\chi}^T(3\boldsymbol{X} - 3\boldsymbol{XCX} + \boldsymbol{XCXCX})\boldsymbol{\chi}$	$-\frac{1}{2}\boldsymbol{\chi}^T(4\boldsymbol{X} - 6\boldsymbol{XCX} + 4\boldsymbol{XCXCX} - \boldsymbol{XCXCXCX})\boldsymbol{\chi}$
$m_i\ddot{\boldsymbol{r}}_i = -\frac{\partial U(\boldsymbol{r}^N)}{\partial \boldsymbol{r}_i} + \cdots$	$-\frac{1}{2}\boldsymbol{\chi}^T\boldsymbol{X}^T\frac{\partial \boldsymbol{C}}{\partial \boldsymbol{r}_i}\boldsymbol{X}\boldsymbol{\chi}$	$-\frac{1}{2}\boldsymbol{\chi}^T\boldsymbol{X}^T\left(3\frac{\partial \boldsymbol{C}}{\partial \boldsymbol{r}_i} - \frac{\partial \boldsymbol{C}}{\partial \boldsymbol{r}_i}\boldsymbol{XC} - \boldsymbol{CX}\frac{\partial \boldsymbol{C}}{\partial \boldsymbol{r}_i}\right)\boldsymbol{X}\boldsymbol{\chi}$	$-\frac{1}{2}\boldsymbol{\chi}^T\boldsymbol{X}^T\left(6\frac{\partial \boldsymbol{C}}{\partial \boldsymbol{r}_i} - 4\frac{\partial \boldsymbol{C}}{\partial \boldsymbol{r}_i}\boldsymbol{XC} + \frac{\partial \boldsymbol{C}}{\partial \boldsymbol{r}_i}\boldsymbol{XCXC} - 4\boldsymbol{CX}\frac{\partial \boldsymbol{C}}{\partial \boldsymbol{r}_i} + \boldsymbol{CXCX}\frac{\partial \boldsymbol{C}}{\partial \boldsymbol{r}_i} + \boldsymbol{CX}\frac{\partial \boldsymbol{C}}{\partial \boldsymbol{r}_i}\boldsymbol{XC}\right)\boldsymbol{X}\boldsymbol{\chi}$
$\ddot{\boldsymbol{X}} =$	$\omega^2(\boldsymbol{X} - \boldsymbol{XCX})$	$\omega^2(2\boldsymbol{X} - 3\boldsymbol{XCX} + \boldsymbol{XCXCX})$	$\omega^2(3\boldsymbol{X} - 6\boldsymbol{XCX} + 4\boldsymbol{XCXCX} - \boldsymbol{XCXCXCX})$

Applying the Euler-Lagrange equation to the Lagrangian of Eq. (5.6) one can derive the equations of motion for the system where Eq. (5.7a) is the equation of motion for the *i*-th real particle and Eq. (5.7b) for the auxiliary matrix, $\boldsymbol{X}$.

$$m_i\ddot{\boldsymbol{r}}_i = -\frac{\partial U(\boldsymbol{r}^N)}{\partial \boldsymbol{r}_i} + \frac{1}{2}\boldsymbol{\chi}^T\frac{\partial f(\boldsymbol{X})}{\partial \boldsymbol{r}_i}\boldsymbol{\chi} - m_X\omega^2(\widetilde{\boldsymbol{C}}^{-1} - \boldsymbol{X})\frac{\partial \widetilde{\boldsymbol{C}}^{-1}}{\partial \boldsymbol{r}_i} \tag{5.7a}$$

$$m_X\ddot{\boldsymbol{X}} = \frac{1}{2}\frac{\partial}{\partial \boldsymbol{X}}[\boldsymbol{\chi}^T f(\boldsymbol{X})\boldsymbol{\chi}] + m_X\omega^2\left(\widetilde{\boldsymbol{C}}^{-1} - \boldsymbol{X}\right) - m_X\omega^2(\widetilde{\boldsymbol{C}}^{-1} - \boldsymbol{X})\frac{\partial \widetilde{\boldsymbol{C}}^{-1}}{\partial \boldsymbol{X}} \tag{5.7b}$$

From the equations of motion in Eq. (5.7) one can 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. (5.8), where $\frac{\partial f(\boldsymbol{X})}{\partial \boldsymbol{r}_i}$ has been explicitly evaluated and Eq. (5.4) is used for $f(\boldsymbol{X})$.

$$m_i \ddot{\boldsymbol{r}}_i = -\frac{\partial U(\boldsymbol{r}^N)}{\partial \boldsymbol{r}_i} - \frac{1}{2}\boldsymbol{\chi}^T \boldsymbol{C}^{-1}\left[\frac{\partial \boldsymbol{C}}{\partial \boldsymbol{r}_i}\boldsymbol{C}^{-1}(\boldsymbol{I} - (\boldsymbol{I} - \boldsymbol{C}\boldsymbol{X})^m) - \sum_{j=0}^{m-1}(\boldsymbol{I} - \boldsymbol{C}\boldsymbol{X})^j \frac{\partial \boldsymbol{C}}{\partial \boldsymbol{r}_i}\boldsymbol{X}(\boldsymbol{I} - \boldsymbol{C}\boldsymbol{X})^{m-j-1}\right]\boldsymbol{\chi} \quad (5.8a)$$

$$\ddot{\boldsymbol{X}} = \omega^2\left(\widetilde{\boldsymbol{C}}^{-1} - \boldsymbol{X}\right) = \omega^2(\boldsymbol{C}^{-1}[\boldsymbol{I} - (\boldsymbol{I} - \boldsymbol{C}\boldsymbol{X})^m] - \boldsymbol{X}) \quad (5.8b)$$

Eq. (5.8a) 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. (5.8b) defines an equation of motion that will propagate the auxiliary $\boldsymbol{X}$, choosing $\widetilde{\boldsymbol{C}}^{-1} = f(\boldsymbol{X})$, and guarantees that the energy is variationally minimized by construction with a shadow potential that matches the exact reference potential from Eq. (5.1) to an increasing degree of accuracy as m increases. Notice that for both Eq. (5.8a) and (5.8b), 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. (5.8), but this inverse is annihilated in the expansions of these equations for finite m, as seen in Table 5.1. Table 5.1 also gives explicit expressions for the equations of motion (Eq. (5.8)), $f(\boldsymbol{X})$, 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. (5.8) one can now build iEL/0-SCF based schemes that are time reversible, have exact analytical agreement between forces and energies, and yield ground state fluctuating 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[82-83] 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, here I obtain an exact minimization of an approximate potential energy, which matches the true potential energy to an arbitrary degree.

5.2.2 *Higher Order Dissipation*.

In practice the equation of motion of the auxiliary degree of freedom (Eq. (5.8b)) 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[23, 73], discussed extensively in Chapter 2. 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[73] 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 Chapter 2 I introduced a method whereby the auxiliaries are coupled to thermostat control to alleviate this issue. Here I will take a different tack and introduce some amount of dissipation into the Verlet integration of the auxiliary equation of motion, Eq. (5.8b), this is introduced

with a friction-like term that is dependent on the history of the auxiliary matrix $\boldsymbol{X}$, as given in Eq. (5.9).

$$\boldsymbol{X}(t+\Delta t) = 2\boldsymbol{X}(t) - \boldsymbol{X}(t-\Delta t) + \kappa\left[\boldsymbol{C^{-1}}\left[\boldsymbol{I} - [\boldsymbol{I} - \boldsymbol{C}\boldsymbol{X}(t)]^m\right] - \boldsymbol{X}(t)\right] + \alpha\sum_{k=0}^{K} c_k \boldsymbol{X}(t-k\Delta t) \qquad (5.9)$$

In Eq. (5.9) Δt is the time step of the simulation, α controls the strength of the dissipation, and K is the order of the dissipation. I have also introduced $\kappa = \omega^2\Delta t^2$, a dimensionless parameter that now controls the frequency of the auxiliary harmonic oscillation. While Eq. (5.9) 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[73]. These coefficients are reproduced for several orders governed by K in Table 5.2.

Table 5.2: *Coefficients for friction-like dissipation*[73].

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

5.2.3 *Higher Order Integrators*

The integration of the equation of motion given in Eq. (5.8), and combined with dissipation, is typically done by Verlet integration, shown in Eq. (5.9). 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. (5.10).

$$\dot{\boldsymbol{X}}\left(t + \sum_{j=1}^{l} b_j\Delta t\right) = \dot{\boldsymbol{X}}\left(t + \sum_{j=1}^{l-1} b_j\Delta t\right) + b_l\Delta t\ddot{\boldsymbol{X}}\left(t + \sum_{j=1}^{l-1} a_j\Delta t\right), \quad l = 1,2,\dots,L \qquad (5.10a)$$

$$\boldsymbol{X}\left(t + \sum_{j=1}^{l} a_j\Delta t\right) = \boldsymbol{X}\left(t + \sum_{j=1}^{l-1} a_j\Delta t\right) + a_l\Delta t\dot{\boldsymbol{X}}\left(t + \sum_{j=1}^{l} b_j\Delta t\right), \quad l = 1,2,\dots,L \qquad (5.10b)$$

For an integrator with L intermediate integration steps, Eq. (5.10) 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. (5.8b), was not well-defined in the

context of higher-order integrators of the form of Eq. (5.10) due to the use of velocities as intermediates for higher order integrators[75].

Here I define a generalization of the friction-like dissipation terms of the modified Verlet integration scheme, Eq. (5.9), to the higher order integrator schemes of Eq. (5.10). As a general integration scheme this combination of friction-like dissipation and general order integration takes the form of Eq. (5.11). Eq. (5.11) introduces dissipation as a friction-like force term as the last term in Eq. (5.11a), 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.

$$\dot{\boldsymbol{X}}\left(t+\sum_{j=1}^{l} b_j \Delta t\right) = \dot{\boldsymbol{X}}\left(t+\sum_{j=1}^{l-1} b_j \Delta t\right) + \ddot{\boldsymbol{X}}\left(t+\sum_{j=1}^{l-1} a_j \Delta t\right) + \frac{b_l \alpha}{\Delta t}\sum_{k=0}^{K} c_k \boldsymbol{X}\left(t-k\Delta t+\sum_{j=1}^{l-1} a_j \Delta t\right) \tag{5.11a}$$

$$\boldsymbol{X}\left(t+\sum_{j=1}^{l} a_j \Delta t\right) = \boldsymbol{X}\left(t+\sum_{j=1}^{l-1} a_j \Delta t\right) + a_l \Delta t \dot{\boldsymbol{X}}\left(t+\sum_{j=1}^{l} b_j \Delta t\right) \tag{5.11b}$$

Applying the general integration of Eq. (5.11) to the auxiliary equation of motion, Eq. (5.8b), then takes the following form of Eq. (5.12).

$$\begin{aligned}\dot{\boldsymbol{X}}\left(t+\sum_{j=1}^{l} b_j \Delta t\right) = \; & \dot{\boldsymbol{X}}\left(t+\sum_{j=1}^{l-1} b_j \Delta t\right) + \frac{b_l \kappa}{\Delta t}\left[\boldsymbol{C}^{-1}\left[\boldsymbol{I}-\left[\boldsymbol{I}-\boldsymbol{C}\boldsymbol{X}\left(t+\sum_{j=1}^{l-1} a_j \Delta t\right)\right]^{m}\right]-\boldsymbol{X}\left(t+\sum_{j=1}^{l-1} a_j \Delta t\right)\right] \\ & + \frac{b_l \alpha}{\Delta t}\sum_{k=0}^{K} c_k \boldsymbol{X}\left(t-k\Delta t+\sum_{j=1}^{l-1} a_j \Delta t\right)\end{aligned} \tag{5.12a}$$

$$\boldsymbol{X}\left(t+\sum_{j=1}^{l} a_j \Delta t\right) = \boldsymbol{X}\left(t+\sum_{j=1}^{l-1} a_j \Delta t\right) + a_l \Delta t \dot{\boldsymbol{X}}\left(t+\sum_{j=1}^{l} b_j \Delta t\right) \tag{5.12b}$$

Table 5.3 gives the optimal integration parameters a_l and b_l to minimize the error of the integration for a velocity Verlet method[88] and also several optimal higher-order methods as described by McLachlan and Atela[96].

Table 5.3: *Integration schemes of various order and their parameters[96]. 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.*

Integrator Name	L	a_l	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	$a_1 = 0.919661523017399857$	$b_1 = a_3$ $b_2 = a_2$

		$a_2 = \frac{1}{4a_1} - \frac{a_1}{2}$ $a_3 = 1 - a_1 - a_2$	$b_3 = a_1$
4th–order optimal	4	$a_1 = 0.5153528374311229364$ $a_2 = -0.085782019412973646$ $a_3 = 0.4415830236164665242$ $a_4 = 0.1288461583653841854$	$b_1 = 0.1344961992774310892$ $b_2 = -0.224819030794208058$ $b_3 = 0.7563200005156682911$ $b_4 = 0.3340036032863214255$
5th–order optimal	6	$a_1 = 0.339839625839110000$ $a_2 = -0.088601336903027329$ $a_3 = 0.5858564768259621188$ $a_4 = -0.603039356536491888$ $a_5 = 0.3235807965546976394$ $a_6 = 0.4423637942197494587$	$b_1 = 0.1193900292875672758$ $b_2 = 0.6989273703824752308$ $b_3 = -0.1713123582716007754$ $b_4 = 0.4012695022513534480$ $b_5 = 0.0107050818482359840$ $b_6 = -0.0589796254980311632$

5.2.4 *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 5.1 and driving the equations of motion in Eq. (5.8). One can also select a dissipative scheme that will only break time reversibility up to some chosen order by selecting from Table 5.2. Finally one can choose an integration scheme that is correct to some order in Δt by integrating with Eq. (5.12) and choosing a method from Table 5.3. This sets forth a general framework for extended Lagrangian Born-Oppenheimer dynamics for classical polarization. While here I have illustrated this formalism with a classical fluctuating charge model I also present the formalism adapted for an induced dipole model in Appendix F.

Table 5.4: *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,* $[\rho(\boldsymbol{T})]_{min}$*, is also given. Details on the derivation of these values are given in Appendix G.*

Integrator Name	L	K	κ	α	$[\rho(\boldsymbol{T})]_{min}$
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
4^{th}-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
5^{th}-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

5.2.5 *Optimal Parameters*

The integration described by Eq. (5.11) is dependent on two key parameters not yet discussed, κ and α, which are given in Table 5.4. 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 G examines what the optimal sets of κ and α parameters are (summarized in Table 5.4), which rounds out the necessary information to build a general iEL/0-SCF method.

5.3 Results

To give a clear illustration of the general approach allow me to choose the lowest order of shadow potential, dissipation, and integration by selecting $m = 2$ from Table 5.1, $K = 3$ from Table 5.2, and $L = 2$ (second-order optimal) from Table 5.3. The equations of motion become Eq. (5.13a) for the particles and Eq. (5.13b) for the auxiliary, $\boldsymbol{X}$.

$$m_i \ddot{\boldsymbol{r}}_i = -\frac{\partial U(\boldsymbol{r}^N)}{\partial \boldsymbol{r}_i} - \frac{1}{2}\boldsymbol{\chi}^T \boldsymbol{X}^T \frac{\partial \boldsymbol{C}}{\partial \boldsymbol{r}_i} \boldsymbol{X}\boldsymbol{\chi} \tag{5.13a}$$

$$\ddot{\boldsymbol{X}} = \omega^2(\boldsymbol{X} - \boldsymbol{X}\boldsymbol{C}\boldsymbol{X}) \tag{5.13b}$$

If Eq. (5.13) is now integrated with 2^{nd}-order optimal integration and 3^{rd}–order dissipation, a single full integration time step would look like Eq. (5.14).

$$\dot{\boldsymbol{X}}(t+b_1\Delta t) = \dot{\boldsymbol{X}}(t) + \frac{b_1\kappa}{\Delta t}[\boldsymbol{X}(t) - \boldsymbol{X}(t)\boldsymbol{C}\boldsymbol{X}(t)] + \frac{b_1\alpha}{\Delta t}[c_0\boldsymbol{X}(t) + c_1\boldsymbol{X}(t-\Delta t) + c_2\boldsymbol{X}(t-2\Delta t) + c_3\boldsymbol{X}(t-3\Delta t)] \tag{5.14a}$$

$$\boldsymbol{X}(t+a_1\Delta t) = \boldsymbol{X}(t) + a_1\Delta t\dot{\boldsymbol{X}}(t+b_1\Delta t) \tag{5.14b}$$

$$\dot{\boldsymbol{X}}(t+\Delta t) = \dot{\boldsymbol{X}}(t+b_1\Delta t) + \frac{b_2\kappa}{\Delta t}[\boldsymbol{X}(t+a_1\Delta t) - \boldsymbol{X}(t+a_1\Delta t)\boldsymbol{C}\boldsymbol{X}(t+a_1\Delta t)] + \frac{b_2\alpha}{\Delta t}[c_0\boldsymbol{X}(t+a_1\Delta t) + c_1\boldsymbol{X}(t+a_1\Delta t-\Delta t) + c_2\boldsymbol{X}(t+a_1\Delta t-2\Delta t) + c_3\boldsymbol{X}(t+a_1\Delta t-3\Delta t)] \tag{5.14c}$$

$$\boldsymbol{X}(t+\Delta t) = \boldsymbol{X}(t+a_1\Delta t) + a_2\Delta t\dot{\boldsymbol{X}}(t+\Delta t) \tag{5.14d}$$

For brevity the time dependence of the $\boldsymbol{C}$ matrices has been dropped, but these need to be updated with the positions, $\boldsymbol{r}$. 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. (5.15), which would be interleaved with Eq. (5.14) at equal time intervals.

$$\dot{\boldsymbol{r}}_i(t+b_1\Delta t) = \dot{\boldsymbol{r}}_i(t) + \frac{b_1}{m_i\Delta t}\left[-\frac{\partial U(\boldsymbol{r}^N(t))}{\partial \boldsymbol{r}_i(t)} - \frac{1}{2}\boldsymbol{\chi}^T\boldsymbol{X}^T(t)\frac{\partial \boldsymbol{C}}{\partial \boldsymbol{r}_i}\boldsymbol{X}(t)\boldsymbol{\chi}\right] \tag{5.15a}$$

$$\boldsymbol{r}_i(t+a_1\Delta t) = \boldsymbol{r}_i(t) + a_1\Delta t\dot{\boldsymbol{r}}_i(t+b_1\Delta t) \tag{5.15b}$$

$$\dot{\boldsymbol{r}}_i(t+\Delta t) = \dot{\boldsymbol{r}}_i(t+a_1\Delta t) + \frac{b_2}{m_i\Delta t}\left[-\frac{\partial U(\boldsymbol{r}^N(t+a_1\Delta t))}{\partial \boldsymbol{r}_i(t+a_1\Delta t)} - \frac{1}{2}\boldsymbol{\chi}^T\boldsymbol{X}^T(t+a_1\Delta t)\frac{\partial \boldsymbol{C}}{\partial \boldsymbol{r}_i}\boldsymbol{X}(t+a_1\Delta t)\boldsymbol{\chi}\right] \tag{5.15c}$$

$$\boldsymbol{r}_i(t+\Delta t) = \boldsymbol{r}_i(t+a_1\Delta t) + a_2\Delta t\dot{\boldsymbol{r}}_i(t+\Delta t) \tag{5.15d}$$

I next test the theory presented above using a dimensionless charge equilibration model, involving three particles with electrostatic interactions described by Eq. (5.1) 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[92].

5.3.1 *Dissipation*

To illustrate why the iEL/0-SCF schemes require dissipation, I introduce a perturbation after 1000 time steps into the simulation, by swapping the auxiliary matrix, $\boldsymbol{X}$, with its value from three time steps previous. This modeled perturbation simulates a spike in numerical noise or a resonance instability one may encounter in a more complex system. Figure 5.1 shows the scaled energy deviation from the initial system energy over the course of a simulation. For good energy conservation one would expect this quantity to stay close to 0 with little drift. In Figure 5.1 it can be seen that this perturbation is quickly corrected when a dissipative scheme is used (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.

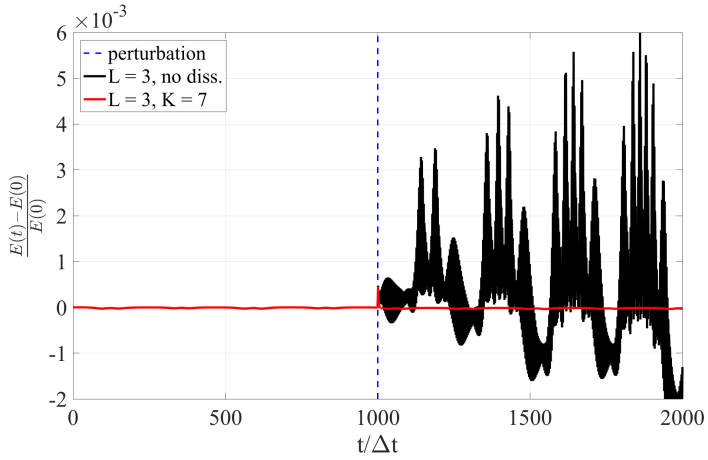


Figure 5.1*: The response of the scaled energy deviation to a perturbation for a dissipative (red) and non-dissipative (black) integration scheme. Both schemes used a 3rd-order optimal integrator ($L = 3$) and a 7th-order dissipative scheme ($K = 7$). The introduced perturbation (dotted blue line) is 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 a dissipative scheme is used as described in section 5.2.2 one can achieve low energy drift rates by using higher order dissipation, where the dissipative c_k $(k = 0, 1, \dots, K)$ coefficients are designed to only break time reversibility up to some order K in the time step[73, 75]. In Figure 5.2a I give an example of the energy drift over the course of a trajectory for a specific integration order and shadow potential order where it can be seen 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 5.2b, where I report the fitted energy drift rates as a function of dissipative order.

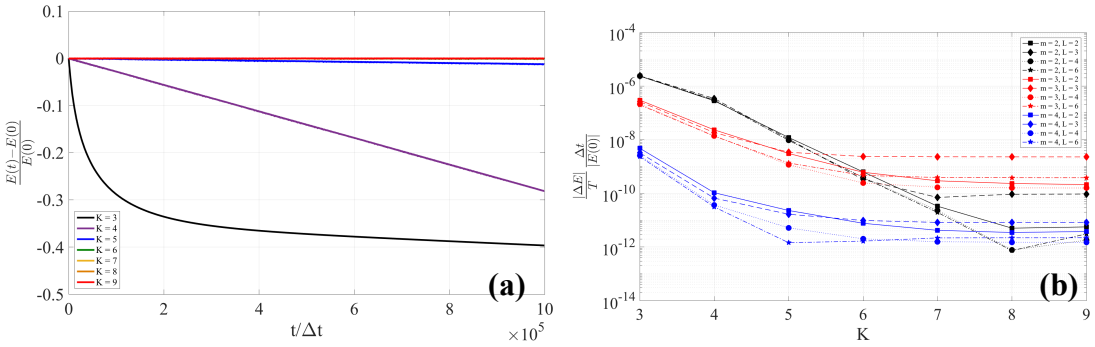


Figure 5.2: *Energy conservation properties as a function of dissipation. (a) The scaled energy deviation from the total initial energy over trajectories using the second order optimal integration method ($L = 2$) and various dissipative orders, K. (b) The fitted energy drifts as a function of dissipative 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, $\Delta E/T$, and non-dimensionalized by the time step and initial energy.

I note that the dissipation schemes presented here in section 5.2.2 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, as introduced an proven in Chapter 2. While this method works well for large systems where the temperature better manifests as a macroscopic quantity, it has been shown that it does not perform well if system sizes are too small[92], where the temperature becomes quite variable and hard to control, such as in this 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. Auxiliary temperature control also performs better for classical condensed phase simulation than dissipation discussed in this Chapter.

5.3.2 *Integration*

Using the formulation I developed in Eq. (5.11) one 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 5.3 for a second order ($m = 2$) shadow potential. Figures 5.3a and 5.3b show 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 a fixed cost is maintained (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 5.3c 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 I 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 5.3d and 5.3e 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.

5.3.3 *Shadow Potentials*

Since the potential is variationally optimized at each time step by construction with Eq. (5.5b), with no explicit matrix inversion, this system is at the ground state of the shadow potential at every time step. Similar to integrator order, in the limit of high shadow potential order the true self-consistent solution to the reference potential is obtained, which is the highest possible level of accuracy one could obtain in terms of potential energy convergence.

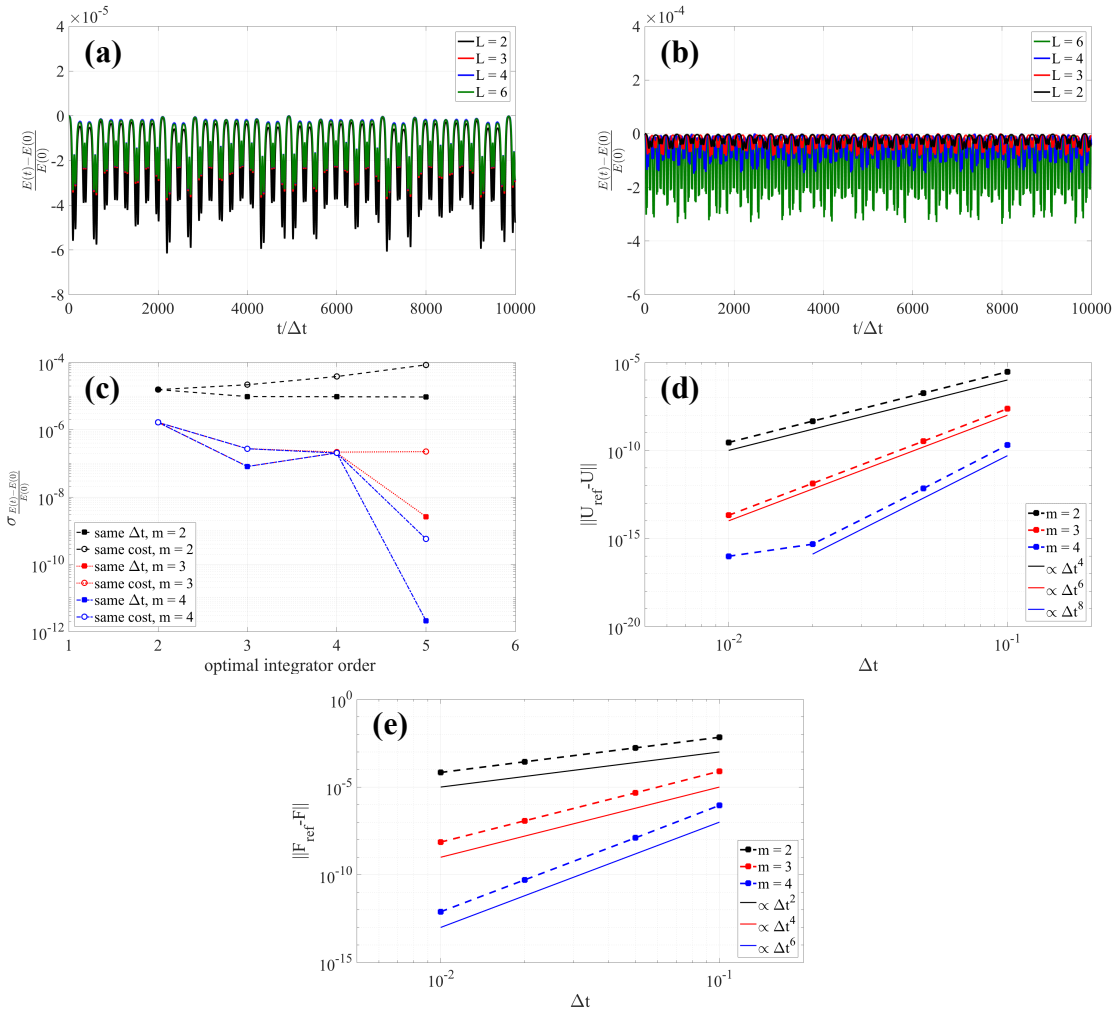


Figure 5.3: *Scaled energy deviation and standard deviation over the course of a trajectory using fifth order dissipation ($K = 5$) with a second order shadow potential ($m = 2$) for various integration orders (L) and scaling of errors in potential energy and forces with time step. (a) For each order 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 5.4a gives an example of this, where I present 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 5.4b shows that one 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 ($L = 6$) a fourth order shadow potential ($m = 4$) is needed to achieve a level of convergence akin to the fully converged reference potential. Therefore, one 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[24], as discussed in Chapter 4, then the shadow potential may not need to be of a high order to accurately follow the dynamics.

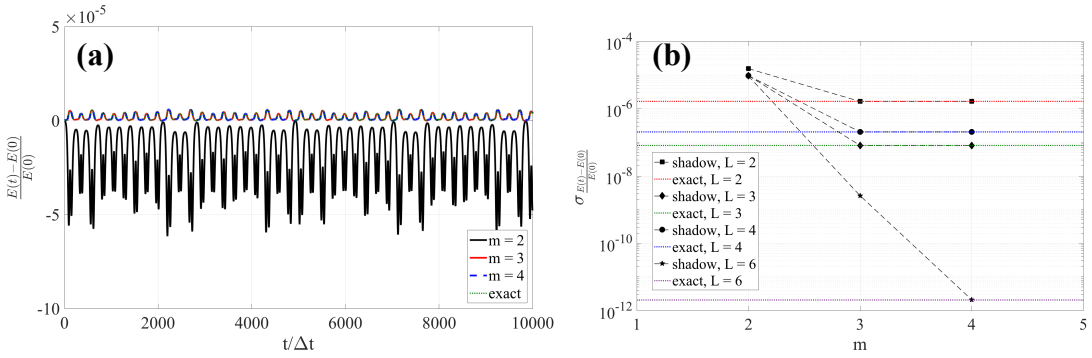


Figure 5.4*: (a) Scaled energy deviation and (b) standard deviation over the course of a trajectory using fifth order dissipation* ($K = 5$) *with a second order optimal integration* ($L = 2$) *for various shadow potential orders* (m). *In both plots the 'exact' dynamics are given as a point of comparison where the* $\boldsymbol{C}$ *matrix is inverted directly and used with the reference potential, Eq. (1) to drive the dynamics for a given integrator order.*

Figure 5.5 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.

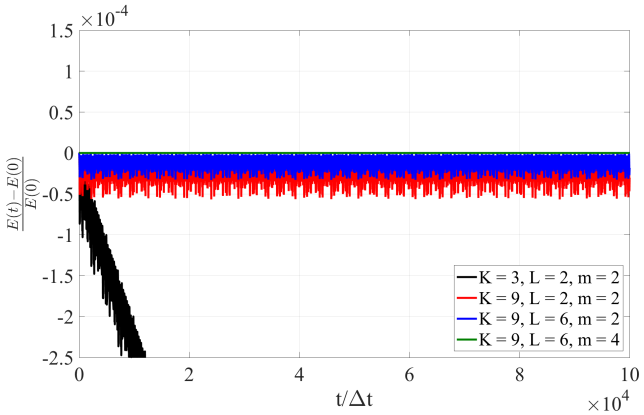


***Figure 5.5**: Scaled energy deviations for selected combinations of dissipation order, K; integration order, L; and shadow potential order, m. The combinations are low dissipation, low integration, low shadow potential (black); high dissipation, low integration, low shadow potential (red); high dissipation, high integration, low potential (blue); and high dissipation, high integration, and high shadow potential (green).*

5.4 Discussion

The methods discussed in this chapter present a generalized starting point from which to work with for iEL/0-SCF methods. As discussed, using $f(\boldsymbol{X})$ in Eq. (5.4) as the approximate inverse matrix, $\widetilde{\boldsymbol{C}}^{-1}$, in Eq. (5.7) introduces expensive matrix-matrix multiplication into the auxiliary equation of motion. 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. Here I show that the iEL/0-SCF method introduced in Chapter 3 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. (5.16)

$$m_i \ddot{\boldsymbol{r}}_i = -\frac{\partial U(\boldsymbol{r}^N)}{\partial \boldsymbol{r}_i} - \frac{1}{2}\boldsymbol{\chi}^T \boldsymbol{X}^T \frac{\partial \boldsymbol{C}}{\partial \boldsymbol{r}_i} \boldsymbol{X}\boldsymbol{\chi} \tag{5.16a}$$

$$\ddot{\boldsymbol{X}} = \omega^2(2\boldsymbol{X} - \boldsymbol{X}\boldsymbol{C}\boldsymbol{X} - \boldsymbol{X}) \tag{5.16b}$$

From a computational efficiency perspective there is no great expense in the particle equation of motion, Eq. (5.16a), as one only ever has to perform matrix-vector multiplications which are $O(N^2)$. 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. (5.16b), 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. One practical remedy that can be applied for the $m = 2$ case is to perform a change of variables and actually 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 Eq. (5.17).

$$m_i \ddot{\boldsymbol{r}}_i = -\frac{\partial U(\boldsymbol{r}^N)}{\partial \boldsymbol{r}_i} - \frac{1}{2}\boldsymbol{a}^T \frac{\partial \boldsymbol{C}}{\partial \boldsymbol{r}_i}\boldsymbol{a} \tag{5.17a}$$

$$\ddot{\boldsymbol{a}} = \omega^2(\boldsymbol{q}[\boldsymbol{a}] - \boldsymbol{a}) \tag{5.17b}$$

In Eq. (17b) I have made use of the identity in Eq. (5.5b), $\boldsymbol{q} = -f(\boldsymbol{X})\boldsymbol{\chi}$. In this specific case, $\boldsymbol{a}$ is now a vector and costly matrix-matrix multiplication is avoided. 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. For $m > 2$, $\boldsymbol{X}$ is not always multiplied into $\boldsymbol{\chi}$ like it is in particle equation of motion, Eq. (5.16a), so this substitution trick is specific to $m = 2$, but does represent a special case were integration of the auxiliary equation of motion could be made more efficient.

This result is now very similar to previously described methods presented in Chapters 2, 3, and 4, which can now be seen 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. (5.18a), which is a single iteration using $\boldsymbol{a}$ as an initial guess. A further refinement for the auxiliary equation of motion, Eq. (5.18b) is introduced as a corrector-like step with a tunable correction parameter, γ, which can range from 0 to 1.

$$\boldsymbol{q}[\boldsymbol{a}] = -\boldsymbol{D}^{-1}\boldsymbol{\chi} - \boldsymbol{D}^{-1}(\boldsymbol{L} + \boldsymbol{U})\boldsymbol{a} \tag{5.18a}$$

$$\boldsymbol{q}_{SCF}[\boldsymbol{a}] \approx \gamma \boldsymbol{q}[\boldsymbol{a}] - (1-\gamma)\boldsymbol{a} \tag{5.18b}$$

In Eq. (5.18) $\boldsymbol{D}$, $\boldsymbol{L}$, 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, one 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. (5.19).

$$m_i \ddot{\boldsymbol{r}}_i = -\frac{\partial U(\boldsymbol{r}^N)}{\partial \boldsymbol{r}_i} - \frac{1}{2}\boldsymbol{q}^T[\boldsymbol{a}]\frac{\partial \boldsymbol{C}}{\partial \boldsymbol{r}_i}\boldsymbol{q}[\boldsymbol{a}] + (\boldsymbol{\chi}^T + \boldsymbol{q}^T[\boldsymbol{a}]\boldsymbol{C})\boldsymbol{D}^{-1}\frac{\partial \boldsymbol{C}}{\partial \boldsymbol{r}_i}\boldsymbol{a} \tag{5.19a}$$

$$\ddot{\boldsymbol{a}} = \gamma\omega^2(\boldsymbol{q}[\boldsymbol{a}] - \boldsymbol{a}) \tag{5.19b}$$

This, then, is a specific example of how one might use this general formalism, in which the formal methods can be computationally expensive, as a starting point to introduce approximations to reduce cost. While the general method of Eq. (5.16) involves expensive, poorly scaling matrix-matrix multiplication, subsequent approximations that led to Eq. (5.19) gives the iEL/0-SCF method, which was proven effective for large, condensed phase systems in Chapters 3 and 4. 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, as introduced in Chapter 2. While the auxiliary thermostatting 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 presented in this chapter may be suitable.

5.5 Conclusions

I have presented a general and flexible framework on which to build iEL/0-SCF 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 I 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 general, greater dissipation order will lead to less energy drift and greater integration and shadow potential order will lead to better accuracy. I illustrated these results with a small, simple fluctuating charge system.

From this chapter there is now a better understanding as to why the previously introduced iEL/0-SCF method from Chapters 3 and 4 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 ($m = 2$), one 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[23] is largely analogous to implementation of a very high order dissipation scheme.

6.
Combining iEL/0-SCF with Stochastic Isokinetic Integration

6.1 Introduction

Recently Tuckerman and colleagues have developed a numerical integration scheme for molecular dynamics that allows for longer time steps that would be unstable using standard approaches[27, 108-109]. The method is stochastic-isokinetic (SI) integration and works by constraining the amount of kinetic energy in each degree of freedom in a simulation so that unphysical flows of energy between fast and slow modes, known as resonance, is prevented. By combining this type of integration with a reversible replica system propagator algorithm (RESPA) method[26], which integrates fast and slow modes on different time scales, long time steps are possible. SI combined with RESPA increases the efficiency of simulations by increasing the time step and iEL/0-SCF, introduced in Chapter 3, increases the efficiency of polarizable simulations by eliminating the need for self-consistent field (SCF) iterations[22]. Since these methods are complementary in their approaches to accelerating molecular dynamics simulations in this chapter I will examine how they can be combined to achieve cutting edge speeds for classical polarizable simulations.

In Section 6.2 I present the theoretical foundation for combining iEL/0-SCF with SI and RESPA. In Section 6.3 I show that isokinetic integration of both the atomic and iEL/0-SCF auxiliary degrees of freedom is possible with SI. I further show that by combining these methods further with RESPA one can achieve time steps of 9.0 fs, compared to a standard time step of 1.0 fs, without sacrificing stability or much accuracy. Overall this leads to speed up of 2.6 times relative to standard SCF polarization methods and integration.

6.2 Methods and Theory

This chapter is focused on combining several advanced molecular dynamics methods- iEL/0-SCF[22], SI[108-109], and RESPA[26]. This section will briefly review these approaches and present the modifications necessary for all of these methods to be combined.

6.2.1 iEL/0-SCF

The iEL/0-SCF method for an induced dipole polarization model, presented in Chapter 3, introduces a set of auxiliary induced dipoles, $\boldsymbol{a}$, which are then driven dynamically by an equation of motion along with the atomic degrees of freedom. The equations of motion for the atoms and auxiliary dipoles are reproduced in Eqs. (6.1a) and (6.1b), respectively.

$$m_i\ddot{\boldsymbol{r}}_i = -\frac{dU(\boldsymbol{r},\boldsymbol{a})}{d\boldsymbol{r}_i}\Big|_{\boldsymbol{a}} \tag{6.1a}$$

$$\ddot{\boldsymbol{a}}_i = \omega^2\gamma(\boldsymbol{\mu}_i - \boldsymbol{a}_i) \tag{6.1b}$$

In Eq. (6.1) $\boldsymbol{r}_i$ is the position of the *i*-th atom, $\boldsymbol{r}$ is the set of all these positions, $\boldsymbol{\mu}_i$ and $\boldsymbol{a}_i$ are the *i*-th induced dipole and associated auxiliary induced dipole, respectively, $\boldsymbol{a}$ is the set of all the auxiliary dipoles, ω is the frequency of the auxiliary harmonic potential and is set to its maximum stable value for Verlet integration $\sqrt{2}/\Delta t$ where Δt is the simulation time step Niklasson, 2014 #145}, and γ is a tunable parameter that comes from estimating a ground state SCF solution from a simple linear mixing of real and auxiliary dipoles. Eq. (6.1b) drives the auxiliary dipoles according to a harmonic potential that seeks to keep the auxiliaries close to a ground state solution. Eq. (6.1a) is the familiar Newtonian equation of motion for the atoms where a general polarization potential is used, as is given in Eq. (6.2).

$$U(\boldsymbol{r},\boldsymbol{a}) = U^{other}(\boldsymbol{r}) + U^{polar}(\boldsymbol{r},\boldsymbol{a}) = U^{other}(\boldsymbol{r}) + \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\boldsymbol{\mu}_i^T\boldsymbol{C}_{ij}\boldsymbol{\mu}_j - \sum_{i=1}^{N}\boldsymbol{\mu}_i^T\boldsymbol{E}_i \tag{6.2a}$$

$$\boldsymbol{\mu}_i = \alpha_i\boldsymbol{E}_i + \alpha_i\sum_{j=1}^{N}\boldsymbol{T}'_{ij}\boldsymbol{a}_j \tag{6.2b}$$

Eq. (6.2a) gives the general potential where a distinction is made between the polarization potential, which is dependent on the auxiliaries and uses a general form, and the other potential terms, $U^{other}(\boldsymbol{r})$, which can include bonds, angles, van der Waals, and permanent electrostatic terms. Eq. (6.2b) then gives the relationship between the real induced dipoles, $\boldsymbol{\mu}_i$, and the auxiliary dipoles, $\boldsymbol{a}_i$, where the real dipoles are essentially a first pass iteration using the auxiliaries as an initial guess. In Eq. (6.2) $\boldsymbol{E}_i$ is the permanent electrostatic field (the field due to non-inducible moments in the system), α_i is the polarizability of the *i*-th atom, $\boldsymbol{T}'_{ij}$ is the dipole-dipole interaction matrix between sites i and j, and $\boldsymbol{C}_{ij} = \alpha_j^{-1}\delta_{ij} - \boldsymbol{T}'_{ij}$. Coupling the integration of the auxiliary dipoles, Eq. (6.1a), to an auxiliary thermostat, as described in Chapter 2, such that the auxiliary pseudo temperature, $T_{aux} = 1/3\langle\dot{\boldsymbol{a}}_i^2\rangle$, is controlled to a set point then rounds out the iEL/0-SCF method.

6.2.2 Stochastic-Isokinetic Integration

Stochastic-isokinetic integration works by constraining the total kinetic energy in each degree of freedom to be a fixed amount that can then transfer between the real degree of freedom and thermostat variables associated with each real degree of freedom. The equations of motion are given in Eq. (6.3)[109].

$$dr_{i,\alpha} = v_{i,\alpha}dt \tag{6.3a}$$

$$dv_{i,\alpha} = \left[\frac{F_{i,\alpha}(\boldsymbol{r})}{m_i} - \lambda_{i,\alpha}v_{i,\alpha}\right]dt \tag{6.3b}$$

$$dv_{1,i,\alpha}^{(k)} = -\lambda_{i,\alpha}v_{1,i,\alpha}^{(k)}dt - v_{2,i,\alpha}^{(k)}v_{1,i,\alpha}^{(k)}dt \tag{6.3c}$$

$$dv_{2,i,\alpha}^{(k)} = \frac{Q_1(v_{1,i,\alpha}^{(k)})^2 - k_B T}{Q_2} dt - \gamma_{SI} v_{2,i,\alpha}^{(k)} dt + \sigma dw_{i,\alpha}^{(k)} \tag{6.3d}$$

In Eq. (6.3) i denotes the particular atom from 1 to N and $\alpha \in x, y, z$ gives the particular component, as the stochastic-isokinetic method works on individual degrees of freedom. In this notation then $r_{i,\alpha}$ is the α component of the i-th atom's position and $v_{i,\alpha}$ and $F_{i,\alpha}(\boldsymbol{r})$ are the associated velocity and force, respectively, and m_i is the mass of the i-th atom. In this way positions are updated directly from velocities, as one may expect, via Eq. (6.3a). Velocities are then updated with forces and with respect to an isokinetic constraint, as well, as given by Eq. (6.3b). This isokinetic constraint is enforced using a Lagrange multiplier for each degree of freedom, $\lambda_{i,\alpha}$. This Lagrange multiplier is given in Eq. (6.4). The integration also introduces two sets of thermostat variables, one that couples to directly to the real velocities, $v_{1,i,\alpha}^{(k)}$, and another set, $v_{2,i,\alpha}^{(k)}$, that couples to the first set of thermostat variables and a stochastic process. These are Nosé-Hoover (NH) thermostat variables and a chain of L variables is used[65-66] where k goes from 1 to L. As a result there are $6NL$ Nosé-Hoover variables, which can represent a significant algorithmic overhead. However, as each degree of freedom is treated independently the problem is trivially parallelizable. In examining the updates for the first set of Nosé-Hoover variables, Eq. (6.3c), one can see that this set of variables couples to the isokinetic constraint and also to the second set of Nosé-Hoover variables. Eq. (6.3d), the equation of motion for the second set of NH variables couples to a driving force created by the difference between the actual and set point temperature, T. They also couple to a stochastic Ornstein-Uhlenbeck (OU) process $dw_{i,\alpha}^{(k)}$ to ensure ergodicity, and a friction term with friction constant γ_{SI}. In these equations the NH masses are defined in the usual way, $Q_1 = k_B T \tau_1^2$ and $Q_2 = k_B T \tau_2^2$ with time scale parameters τ_1 and τ_2 and $\sigma = \sqrt{k_B T \gamma_{SI} / Q_2}$.

$$\lambda_{i,\alpha} = \frac{v_{i,\alpha} F_{i,\alpha}(\boldsymbol{r}) - \frac{L}{L+1} \sum_{k=1}^{L} Q_1 (v_{1,i,\alpha}^{(k)})^2 v_{2,i,\alpha}^{(k)}}{m_i v_{i,\alpha}^2 + \frac{L}{L+1} \sum_{k=1}^{L} Q_1 (v_{1,i,\alpha}^{(k)})^2} \tag{6.4}$$

The isokinetic constraint given in Eq. (6.5) is then what is enforced by the Lagrange multiplier of Eq. (6.4).

$$m_i v_{i,\alpha}^2 + \frac{L}{L+1} \sum_{k=1}^{L} Q_1 (v_{1,i,\alpha}^{(k)})^2 = L k_B T \tag{6.5}$$

From Eq. (6.5) one can see that a total amount of kinetic energy, Lk_BT, is given to each degree of freedom i, α, but this kinetic energy can fluctuate between the real velocity $v_{i,\alpha}$ and the NH velocities $v_{1,i,\alpha}^{(k)}$.

When combining iEL/0-SCF with this stochastic-isokinetic integration the atomic degrees of freedom are integrated with the equations of motion of Eq. (6.3) and the

auxiliary induced dipoles also need to be integrated in a similar manner. Adapting Eq. (6.3) to the case of auxiliary dipoles yield Eq. (6.6).

$$da_{i,\alpha} = v^a_{i,\alpha} dt \tag{6.6a}$$

$$dv^a_{i,\alpha} = \left[\gamma\omega^2(\mu_{i,\alpha} - a_{i,\alpha}) - \lambda^a_{i,\alpha} v_{i,\alpha}\right] dt \tag{6.6b}$$

$$dv^{a(k)}_{1,i,\alpha} = -\lambda^a_i v^{a(k)}_{1,i,\alpha} dt - v^{a(k)}_{2,i,\alpha} v^{a(k)}_{1,i,\alpha} dt \tag{6.6c}$$

$$dv^{a(k)}_{2,i,\alpha} = \frac{Q^a_1 v^{(k)\ 2}_{1,i,\alpha} - T_{aux}}{Q^a_2} dt - \gamma^a_{SI} v^{a(k)}_{2,i,\alpha}\, dt + \sigma dw^{a(k)}_{i,\alpha} \tag{6.6d}$$

Eq. (6.3a) gives the updates of the α component of the i-th auxiliary dipole $a_{i,\alpha}$ according to its respective velocity, $v^a_{i,\alpha}$. Eq. (6.4a) then gives the update of the auxiliary velocity according to its 'force' given by the general iEL/0-SCF auxiliary equation of motion, Eq. (6.1b), and its isokinetic constraint, enforced with the i,α degree of freedom's Lagrange multiplier, $\lambda^a_{i,\alpha}$. The analytic expression for this Lagrange multiplier is given by Eq. (6.7). Two sets of auxiliary NH thermostat variables are introduced for each i,α auxiliary degree of freedom, $v^{a(k)}_{1,i,\alpha}$ and $v^{a(k)}_{2,i,\alpha}$, each in a chain of order L with k running from 1 to L. The first set, $v^{a(k)}_{1,i,\alpha}$, is driven by Eq. (6.6c) where they couple to the isokinetic constraint and the second set. The second set, $v^{a(k)}_{2,i,\alpha}$, is driven by Eq. (6.6d) where they couple to a driving force between the actual and set point auxiliary temperature T_{aux}, a friction term, and an OU stochastic process. Here $Q^a_1 = T_{aux}(\tau^a_1)^2$ and $Q^a_2 = T_{aux}(\tau^a_2)^2$, γ^a_{SI} controls the friction, and $\sigma = \sqrt{T_{aux}\gamma^a_{SI}/Q_2}$. The Lagrange multiplier used in Eq. (6.6) is given in Eq. (6.7).

$$\lambda^a_{i,\alpha} = \frac{v^a_{i,\alpha}\gamma\omega^2(\mu_{i,\alpha} - a_{i,\alpha}) - \frac{L}{L+1}\sum_{k=1}^{L} Q^a_1 (v^{a(k)}_{1,i,\alpha})^2 v^{a(k)}_{2,i,\alpha}}{\left(v^a_{i,\alpha}\right)^2 + \frac{L}{L+1}\sum_{k=1}^{L} Q_1 (v^{a(k)}_{1,i,\alpha})^2} \tag{6.7}$$

With the Lagrange multiplier the pseudo isokinetic constraint for each auxiliary degree of freedom is given by Eq. (6.8), analogous to the case of the atomic constraint, Eq. (6.5).

$$\left(v^a_{i,\alpha}\right)^2 + \frac{L}{L+1}\sum_{k=1}^{L} Q^a_1 (v^{a(k)}_{1,i,\alpha})^2 = LT_{aux} \tag{6.5}$$

6.2.3 RESPA Multiple Time Stepping

Stochastic isokinetic integration eliminates unphysical flow of energy between fast and slow modes. To take full advantage of this development one can use a multiple time-scale integration like RESPA[26]. In RESPA different components of the force field potential energy are integrated with different time steps. In molecular dynamics the highest frequency (fastest) components of a force field are the bonded terms, which represent fast bond, angle, and other geometric fluctuations. The non-bonded through-space components like van der Waals, electrostatic, and polarization components vary

more slowly. The bonded interactions are computationally cheap, whereas the bonded interactions are computationally expensive, but RESPA takes advantage of this time scale separation to integrate the bonded components of the force field with short time steps and the non-bonded components with long time steps, leading to less frequent calculations of non-bonded forces and an overall improvement in efficiency.

In this work I use a three timescale RESPA scheme where the non-bonded forces are evaluated at a short time step of Δt_s with corresponding forces $\boldsymbol{F}_s$. The non-bonded interactions are then partitioned between a medium timescale integrated at a time step of Δt_m, and a longer (outer) time scale with a time step of Δt, each with corresponding forces $\boldsymbol{F}_m$ and $\boldsymbol{F}_l$, respectively. The long time scale time step must be an integer multiple of both the intermediate and short time steps and the intermediate time step needs to be an integer multiple of the short time step, that is $\Delta t = m\Delta t_m = ms\Delta t_s$ for integers m and s. The three non-bonded interactions (polarization, electrostatics, and van der Waals interactions) are placed in the intermediate and long time scales in various combinations to determine optimum efficiency and accuracy, as discussed in the results in Section 6.3.

The details of the RESPA algorithms are discussed in Appendix I, but the important aspect to note is that the short time scale forces are evaluated ms times per long time step Δt, the intermediate forces are updated m times per Δt, and the long timescale forces are evaluated once per Δt. When the polarization is included in the intermediate timescale then the auxiliary equations of motion are updated with an outer time step of Δt_m (so $\omega = \sqrt{2}/\Delta t_m$) and when polarization is included with the long timescale forces the auxiliary equations of motion have an outer time step of Δt (so $\omega = \sqrt{2}/\Delta t$).

6.2.4 Simulation Details

The combined iEL/0-SCF, SI, and RESPA methodology was applied to test systems of 512 water molecules and a zwitterionic glycine molecule solvated with 256 water molecules. The water test case used the latest AMOEBA force field parameterization for water[34]. The glycine system used a reparameterization of the AMOEBA force field[98]. The atomic system set point temperature T was 298.0 K and the auxiliary set point pseudo temperature was 5.3 $e^2Å^2/ps^2$. For the auxiliary stochastic isokinetic integration I found that the best auxiliary thermostat time scale parameters were 0.1 ps and 0.001 ps for τ_1^a and τ_2^a, respectively. A mixing parameter, γ of 0.9 was used for both test systems, as discussed in Chapter 3. The auxiliary integration also used a friction parameter, γ_{SI}^a, of 100.0 ps^{-1}. For the atomic integration a value of 0.1 ps was used for τ_1 and τ_2 and a value of 0.01 ps^{-1} for γ_{SI}. For both the atomic and auxiliary integrations a 4th-order Nosé-Hoover chain was used, that is $L = 4$, and was integrated with a 3rd-order Suzuki-Yoshida decomposition[110-111]. While the values of the intermediate and outer time steps is varied for analysis purposes the short timescale time step was fixed to 0.25 fs. All electrostatics were treated with particle-mesh Ewald summation[101] with a real space cutoff of 7.0 Å. Simulations using self-consistent field iteration (SCF) for comparison purposes used a preconditioned conjugate gradient (PCG) solver[20] with a convergence threshold of 10^{-6} D. Simulations that do not use SI integration for comparison purposes instead use velocity Verlet integration[88] and Nosé-Hoover thermostats for temperature control[66].

6.3 Results

I performed simulations with a pure water system and a system of a solvated glycine molecule. In the first set of simulations no RESPA multiple time step integration was used so as to validate that integrating the iEL/0-SCF auxiliary equations of motion with stochastic isokinetic integration is possible. Thereafter I introduced RESPA multiple time stepping into the scheme to ascertain which combinations of intermediate and long time steps and forces optimize the accuracy-computational cost tradeoff. These results are presented in the subsequent sections, respectively.

6.3.1 Validation

Figure 6.1 gives the polarization properties collected from bulk water simulations using a standard method (SCF and velocity Verlet integration), a stochastic isokinetic method with SCF, and a stochastic isokinetic method with iEL/0-SCF. From Figures 6.1a and 6.1b it is clear that the static induced dipole distributions match well across all methods. Figure 6.1c shows the induced dipole time autocorrelations. As the time autocorrelation is a dynamic property the use of stochastic isokinetic integration deviates slightly from the base correlation since stochastic isokinetic integration is only canonical in positions, but not velocities[108] so one can expect some deviation in dynamic properties. In Figure 6.1c, however, it is clear that the use of iEL/0-SCF with SI does not alter the results and that iEL/0-SCF can reproduce SI results using a standard SCF procedure.

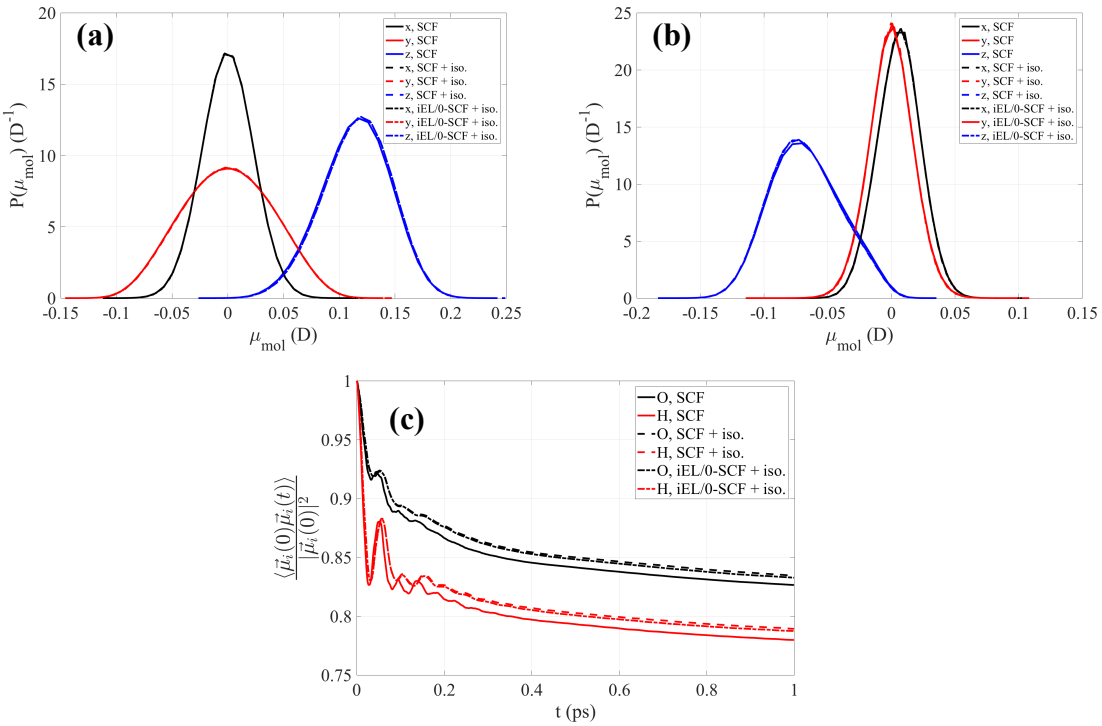


Figure 6.1: *Induced dipole probability distributions for water oxygen (a) and hydrogen (b) and their corresponding induced dipole autocorrelations (c). Three methods are presented- SCF with standard velocity Verlet integration, SCF with SI integration, and iEL/0-SCF with SI integration. All simulations were performed with a single 1.0 fs time step and at a temperature of 298.0 K.*

I also examined a solvated glycine system to determine that SI integration was compatible with iEL/0-SCF. The polarization properties of this system are presented in Figure 6.2. In Figure 6.2a – 6.2c the distributions from the components of the glycine carbonyl oxygen show excellent agreement between a standard approach, SI integration with SCF, and, most importantly, SI with iEL/0-SCF. This proves, again, that one can safely use iEL/0-SCF with SI. Again as before, Figure 6.2d shows that the autocorrelations of SI integration are slightly offset from those of standard Verlet integration, but both SI methods using SCF and iEL/0-SCF are in good agreement. This suggests that iEL/0-SCF performs as well as SCF and is not influenced by the underlying integration method. Further validation data for other atomic species in the glycine system are given in Figures J.1 – J.3 in Appendix J.

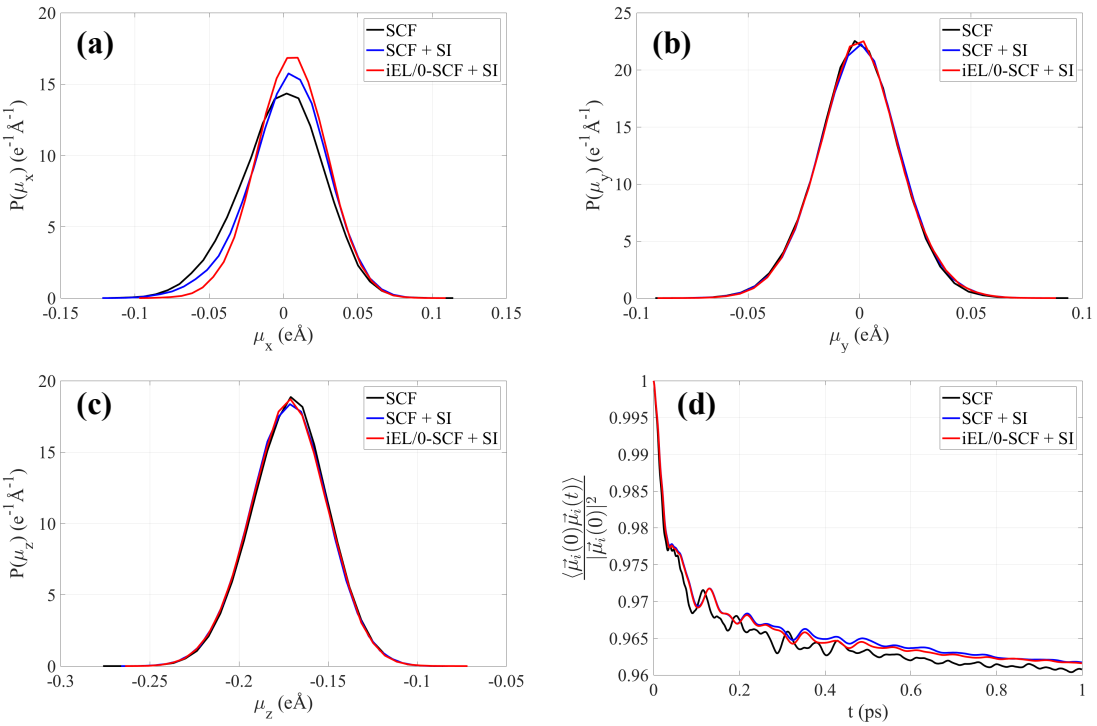


Figure 6.2: *Induced dipole probability distributions for glycine carbonyl carbon x-component (a), y-component (b), z-component (c), and the corresponding induced dipole autocorrelation (d). Three methods are presented- SCF with standard velocity Verlet integration, SCF with SI integration, and iEL/0-SCF with SI integration. All simulations were performed with a single 1.0 fs time step and a temperature of 298.0 K.*

From the data in Figures 6.1, 6.2, and Appendix J it is clear that SI integration and iEL/0-SCF are compatible and yield consistent results when used together. In the subsequent section I will examine using this combination along with RESPA multiple time stepping to increase the time step of these simulations.

6.3.2 iEL/0-SCF with SI and RESPA

The purpose of using SI is to eliminate resonances between fast and slow modes. Using RESPA to then separate the integration of such modes takes full advantage of the benefits of SI in terms of computational efficiency. As a first pass of using all three methods (iEL/0-SCF, SI, and RESPA) together the bonded forces are placed in a short

time scale integrated at a time step of $\Delta t_s = 1.0$ fs. All of the non-bonded forces (electrostatic, polarization, van der Waals) are then treated at a variable long time scale time step. The data for both iEL/0-SCF and SCF using SI and this two time scale RESPA implementation are given in Figure 6.3.

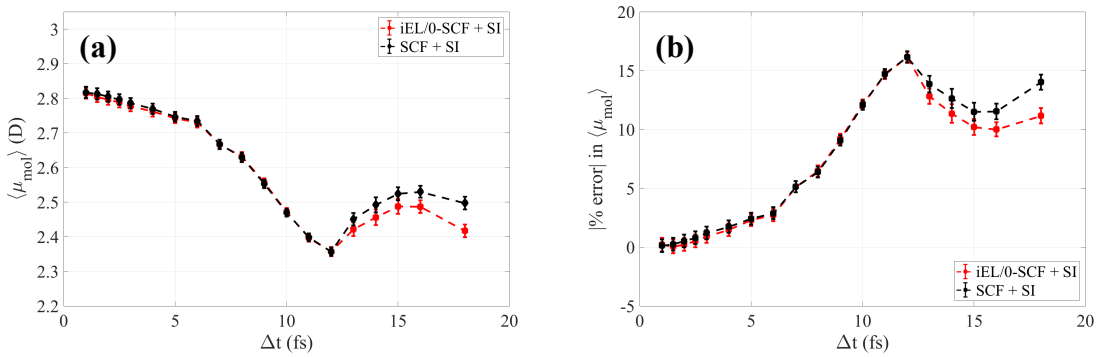


Figure 6.3: *Two time scale RESPA used in conjunction with SI and iEL/0-SCF (red) and SCF (black) on a system of pure water. The average molecular dipole of water $\langle \mu_{mol} \rangle$ as a function of the outer RESPA time step, Δt, is reported (a) along with the error of the average molecular dipole (b) with respect to the vale of an SCF simulation with standard Verlet integration at a 1.0 fs time step (2.81 D).*

Now introducing a third time scale to the RESPA method, an intermediate time scale with a time step of Δt_m, components of the non-bonded forces can then be partitioned between the long and intermediate time scales. This then leads to six possible algorithmic combinations, given in Table 6.1

Table 6.1: Possible combinations of non-bonded forces between the intermediate and long time scales using RESPA.

Name	Intermediate (Δt_m)	Intermediate (Δt)
pol. mid.	polarization	electrostatics, van der Waals
pol. long	electrostatics, van der Waals	polarization
el. mid.	electrostatics	polarization, van der Waals
el. long	polarization, van der Waals	electrostatics
vdW mid.	van der Waals	polarization, electrostatics
vdW long	polarization, electrostatics	van der Waals

Results, including computational timings, for the el. mid. and vdW long methods are presented in Figures 6.4 and 6.5, respectively. These methods seem to give the best returns in terms of accuracy versus cost. The results for the remainder of the methods can be found in Figures J.4 – J.7 in Appendix J.

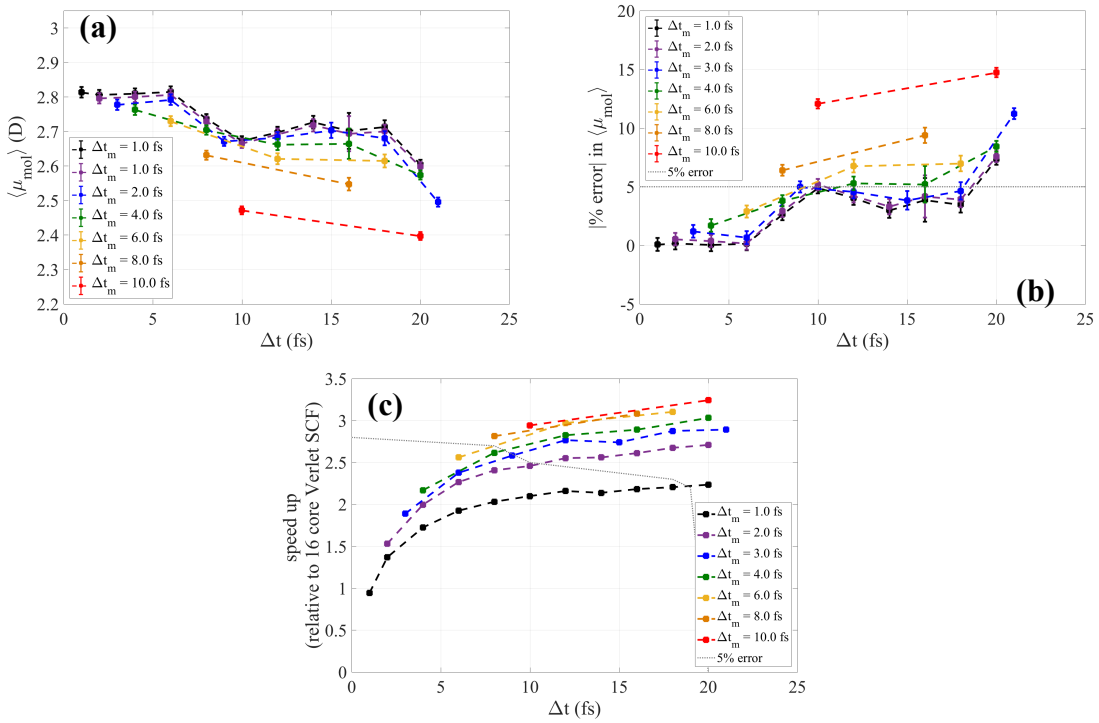


Figure 6.4: *Three time scale el. mid. RESPA used in conjunction with SI and iEL/0-SCF on a system of pure water. The average molecular dipole of water $\langle \mu_{mol} \rangle$ as a function of the outer RESPA time step, Δt, is reported (a) along with the error of the average molecular dipole (b) with respect to the vale of an SCF simulation with standard Verlet integration at a 1.0 fs time step (2.81 D). Different colored curves represent different intermediate time steps, Δt_m. Also reported are the computational speed ups with respect to standard 1.0 fs Verlet integration (c). All calculations were performed with 16 CPUs and shared-memory parallelization. The dotted gray line in (b) represents 5% error and all simulations below that error in (b) are within the 5% error envelope in (c).*

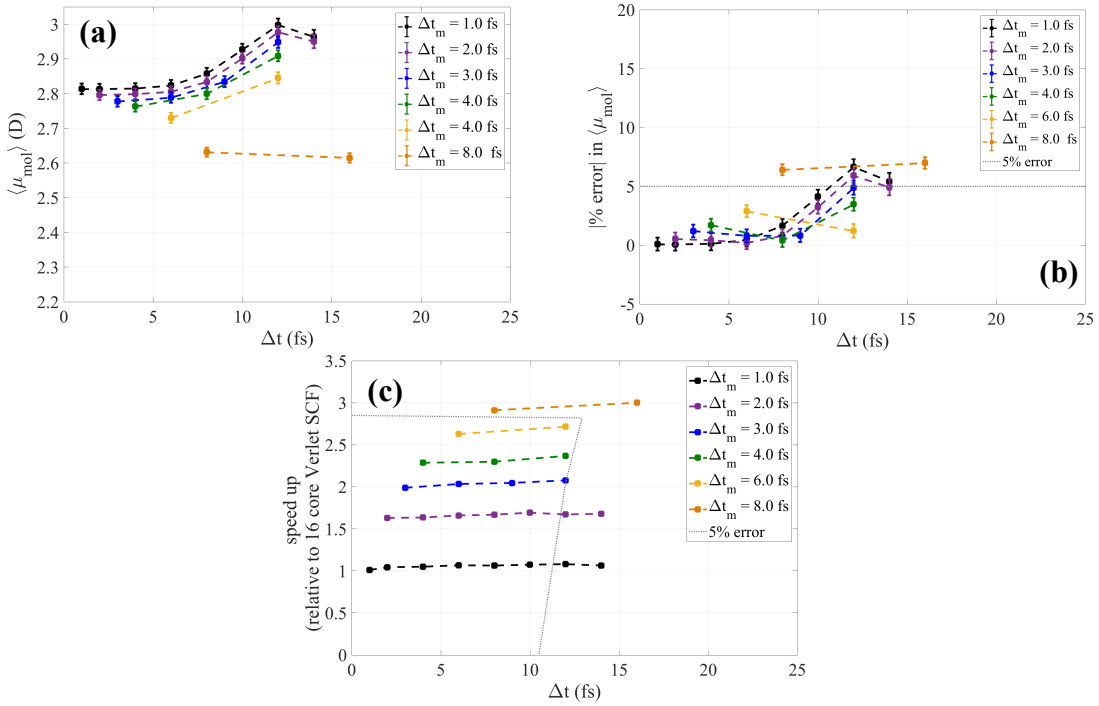


Figure 6.5: *Three time scale vdW long RESPA used in conjunction with SI and iEL/0-SCF on a system of pure water. The average molecular dipole of water* $\langle \mu_{mol} \rangle$ *as a function of the outer RESPA time step,* Δt*, is reported (a) along with the error of the average molecular dipole (b) with respect to the vale of an SCF simulation with standard Verlet integration at a 1.0 fs time step (2.81 D). Different colored curves represent different intermediate time steps,* Δt_m*. Also reported are the computational speed ups with respect to standard 1.0 fs Verlet integration (c). All calculations were performed with 16 CPUs and shared-memory parallelization. The dotted gray line in (b) represents 5% error and all simulations below that error in (b) are within the 5% error envelope in (c).*

From Figure 6.4 one can see that using an intermediate time step of 1.0 fs outer time steps as large as 18.0 fs can be used while the simulations remain stable and accurate ($< 5\%$ error). For this el. mid. method, though, the fastest simulations that are also stable and accurate are three combinations $\Delta t_m = 6.0$ fs and $\Delta t = 6.0$ fs, $\Delta t_m = 4.0$ fs and $\Delta t = 8.0$ fs, and $\Delta t_m = 3.0$ fs and $\Delta t = 9.0$ fs all of which yield speedups of about 2.6, which is superior to the reported speed up of only iEL/0-SCF in Chapter 3 of about 1.6. Interestingly, for the case of $\Delta t_m = 6.0$ fs and $\Delta t = 6$ fs the partitioning of non-bonded interactions between the intermediate and long time scales in immaterial since the intermediate and long time scale time steps are the same and the intermediate and long time scale forces would therefore be evaluated simultaneously and their partitioning is the same across all methods in Table 6.1. This particular time step setting is the optimal in terms of accuracy and speed up for the pol. mid., pol. long, el. long, and vdW mid. methods, as shown in Figures J.4 – J.7. From Figure 6.5 one can see that for vdW long the optimal setting is $\Delta t_m = 6.0$ fs and $\Delta t = 12.0$ fs, which has a speed up of 2.7. To further demonstrate the accuracy of this significantly more efficient methods, their radial distribution functions are given in Figure 6.6.

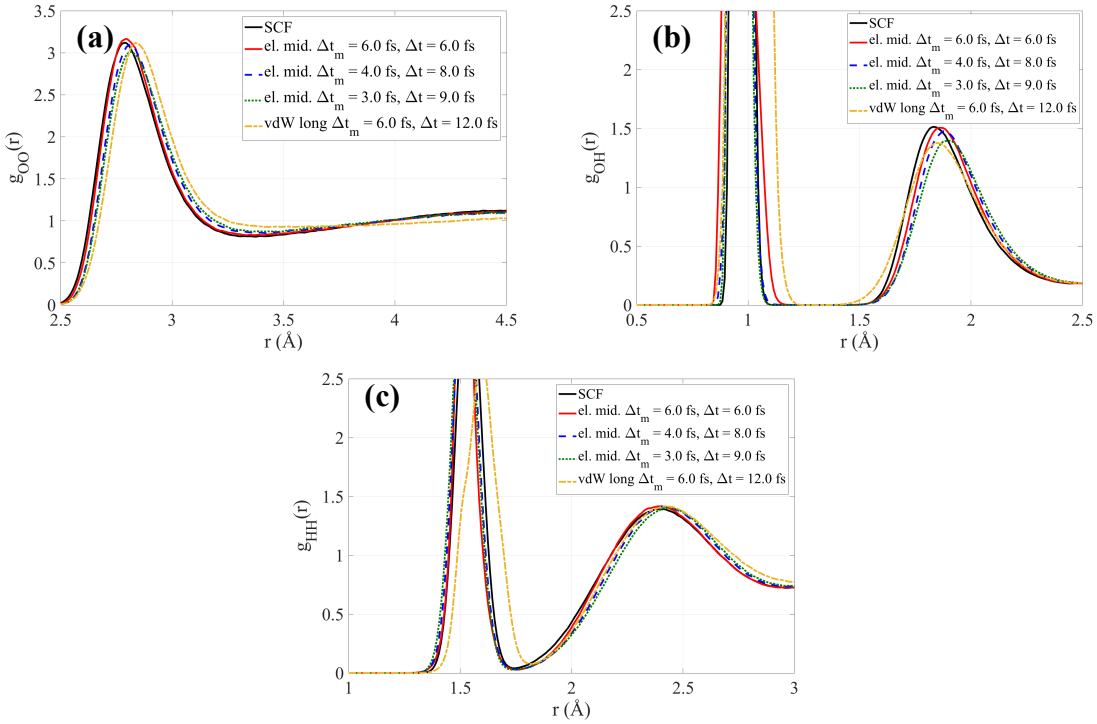


Figure 6.6: *Radial distribution functions for iEL/0-SCF with SI and several flavors of RESPA- el. mid. with* $\Delta t_m = 6.0$ *fs and* $\Delta t = 6.0$ *fs (red, solid), el. mid. with* $\Delta t_m = 4.0$ *fs and* $\Delta t = 8.0$ *fs (blue, dashed),*

el. mid. with $\Delta t_m = 3.0$ fs and $\Delta t = 9.0$ fs (green, dotted), and vdW long with $\Delta t_m = 6.0$ fs and $\Delta t = 1.02$ fs (yellow, dashed-dotted). These are compared with a standard SCF procedure using Verlet integration with $\Delta t = 1.0$ fs (black, solid).

From Figure 6.6 it is clear that, in general, as the outer time step increases the radial distribution functions become less accurate. For the el. mid. methods the radial distribution functions are quite accurate across the board and the graphs are framed to show slight differences between the methods around the first solvation shells. One can conclude that using the $\Delta t_m = 6.0$ fs and $\Delta t = 6.0$ fs method (which is equivalent across all RESPA flavors of Table 6.1) gives us excellent accuracy in the radial distribution function with a computational speed that is 2.6 times greater than a standard velocity Verlet SCF method.

6.4 Discussion

As the data that gives timing results show (Figures. 6.4c, 6.5c, J.4c – J.7c), when polarization is included in the long time scale the speed ups improve as the long time scale time step increases, but when polarization is in the intermediate time scale the speed ups increase much less with increasing long time step. Polarization is therefore rate limiting step for the simulations. As iEL/0-SCF eliminates the need for SCF iteration the cause is likely the $6NL$ Nosé-Hoover auxiliary thermostat variables that must be updated at the same time scale as the polarization. Even after parallelization this update clearly serves as a significant cost. This revelation argues for the development of global isokinetic control, where the overall pseudo kinetic energy of all auxiliary degrees of freedom is controlled, instead of massive (local) isokinetic control of the auxiliary degrees of freedom. In the case of global isokinetic control only 6 auxiliary thermostat variables would be needed.

Additionally, to take full advantage of RESPA and SI one can divide non-bonded forces into short- and long-range components where the overall potential (be it polarization, electrostatics, or van der Waals) is split by a switching function at some distance[109]. In this way all non-bonded interactions have an intermediate and long range time scale component whereas in this work entire non-bonded potentials were placed in either the intermediate or long time scales. By employing such switching functions long time scale time steps of 100.0 fs are possible[109] and speed ups even greater than those presented here would be possible. This would be a logical next step for the combined iEL/0-SCF, SI, and RESPA method.

6.5 Conclusions

In this chapter I have given the theory and methods that would allow one to combine iEL/0-SCF, SI, and RESPA. iEL/0-SCF is a method for treating polarization without iteration. SI is an integration scheme resolves resonances between fast and slow modes and RESPA takes advantage of SI to make large time steps possible. These methods are complementary in terms of computational efficiency and I have shown that these methods can be combined to give speedups of about 2.6 times relative to a standard single time scale Verlet integration with a 1.0 fs time step and SCF iteration to treat

polarization. Such combined methods can realize long scale time steps of up to 9.0 fs, while maintaining simulation stability and accuracy.

As with most simulation protocols the combined method has a cost-accuracy tradeoff. In these analyses I have assumed a 5% error to be acceptable, but using the data presented here one could choose simulation parameters (intermediate and long time steps and non-bonded force partitioning between them) that can yield lower error with correspondingly lower speed ups or greater error with a corresponding greater efficiency. Using the data presented here a user could make such a choice based on the needs of a given simulation.

Overall, the combined methodology presented in the chapter brings several advanced simulation techniques into compatibility and represents a thrust toward remarkable computational efficiency for polarizable classical molecular dynamics. By taking advantage of possible future improvements such as global isokinetic control and non-bonded potential splitting with switching functions, one can imagine simulations with truly astounding efficiency.

References

1. Shi, Y.; Ren, P.; Schnieders, M.; Piquemal, J. P., Polarizable Force Fields for Biomolecular Modeling. *Rev. Comp. Chem.* **2015,** *28*, 51.
2. Albaugh, A.; Boateng, H. A.; Bradshaw, R. T.; Demerdash, O.; Dziedzic, J.; Mao, Y.; Margul, D. T.; Swails, J.; Zeng, Q.; Case, D. A.; Eastman, P.; Essex, J. W.; Head-Gordon, M.; Pande, V. S.; Ponder, J. W.; Shao, Y.; Skylaris, C. K.; Todorov, I. T.; Tuckerman, M. E.; Head-Gordon, T., Advanced Potential Energy Surfaces for Molecular Simulation. *J. Phys. Chem. B* **2016,** *120*, 9811.
3. Stone, A. J., *The Theory of Intermolecular Forces*. 1996.
4. Demerdash, O.; Yap, E. H.; Head-Gordon, T., Advanced Potential Energy Surfaces for Condensed Phase Simulation. *Annu. Rev. Phys. Chem.* **2014,** *65*, 149.
5. Gresh, N.; Cisneros, G. A.; Darden, T. A.; Piquemal, J. P., Anisotropic, Polarizable Molecular Mechanics Studies of Inter- and Intramoecular Interactions and Ligand-Macromolecule Complexes. A Bottom-up Strategy. *J. Chem. Theory Comput.* **2007,** *3*, 1960.
6. Piquemal, J. P.; Williams-Hubbard, B.; Fey, N.; Deeth, R. J.; Gresh, N.; Giessner-Prettre, C., Inclusion of the Ligand Field Contribution in a Polarizable Molecular Mechanics: SIBFA-LF. *J. Comput. Chem.* **2003,** *24*, 1963.
7. Jungwirth, P.; Winter, B., Ions at Aqueous Interfaces: From Water Surface to Hydrated Proteins. *Annu. Rev. Phys. Chem.* **2008,** *59*, 343.
8. Vazdar, M.; Pluharova, E.; Mason, P. E.; Vacha, R.; Jungwirth, P., Ions at Hydrophobic Aqueous Interfaces: Molecular Dynamics with Effective Polarization. *J. Phys. Chem. Lett.* **2012,** *3*, 2087.
9. Bhowmick, A.; Sharma, S. C.; Head-Gordon, T., The Importance of the Scaffold for de Novo Enzymes: a case study with Kemp Eliminase. *Journal of the American Chemical Society* **2017,** *139* (16), 5793-5800.
10. Fried, S. D.; Wang, L. P.; Boxer, S. G.; Ren, P.; Pande, V. S., Calculations of the Electric Fields in Liquid Solutions. *J. Phys. Chem. B* **2013,** *117*, 16236.
11. Harder, E.; Kim, B. C.; Friesner, R. A.; Berne, B. J., Efficient Simulation Method for Polarizable Protein Force Fields: Application to the Simulation of BPTI in Liquid. *J. Chem. Theory Comput.* **2005,** *1*, 169.
12. Thompson, M. A.; Schenter, G. K., Excited-States of the Bacteriochlorophyll-B Dimer of Rhodopseudomonas-Viridis - a QM/MM Study of the Photosynthetic Reaction-Center That Includes MM Polarization. *J. Phys. Chem.* **1995,** *99*, 6374.
13. Lindert, S.; Bucher, D.; Eastman, P.; Pande, V.; McCammon, J. A., Accelerated molecular dynamics simulations with the AMOEBA polarizable force field on graphics processing units. *Journal of chemical theory and computation* **2013,** *9* (11), 4684-4691.
14. Johnson, M.; Malardier-Jugroot, C.; Murarka, R.; Head-Gordon, T., Hydration Water Dynamics near Biological Interfaces. *J. Phys. Chem. B* **2009,** *113*, 4082.
15. Johnson, M. E.; Malardier-Jugroot, C.; Head-Gordon, T., Effects of Co-Solvents on Peptide Hydration Water Structure and Dynamics. *Phys. Chem. Chem. Phys.* **2010,** *12*, 393.
16. Lopes, P. E.; Roux, B.; Mackerell, A. D., Molecular Modeling and Dynamics Studies with Explicit Inclusion of Electronic Polarizability. Theory and Applications. *Theor. Chem. Acc.* **2009,** *124*, 11.
17. Ren, P. Y.; Ponder, J. W., Consistent Treatment of Inter- and Intramolecular Polarization in Molecular Mechanics Calculations. *J. Comput. Chem.* **2002,** *23*, 1497.
18. Albaugh, A.; Boateng, H. A.; Bradshaw, R. T.; Demerdash, O. N.; Dziedzic, J.; Mao, Y.; Margul, D. T.; Swails, J.; Zeng, Q.; Case, D. A.; Eastman, P.; Wang, L. P.; Essex, J. W.; Head-Gordon, M.; Pande, V. S.; Ponder, J. W.; Shao, Y.; Skylaris, C. K.; Todorov, I. T.; Tuckerman, M. E.; Head-Gordon, T., Advanced Potential Energy Surfaces for Molecular Simulation. *J. Phys. Chem. B* **2016,** *120* (37), 9811.
19. Wang, L. P.; Head-Gordon, T.; Ponder, J. W.; Ren, P.; Chodera, J. D.; Eastman, P. K.; Martinez, T. J.; Pande, V. S., Systematic Improvement of a Classical Molecular Model of Water. *J. Phys. Chem. B* **2013,** *117*, 9956.
20. Wang, W.; Skeel, R. D., Fast Evaluation of Polarizable Forces. *J. Chem. Phys.* **2005,** *123*, 164107.
21. Lipparini, F.; Lagardère, L.; Stamm, B.; Cancès, É.; Schnieders, M.; Ren, P.; Maday, Y.; Piquemal, J. P., Scalable Evaluation of Polarization Energy and Associated Forces in Polarizable Molecular

Dynamics: I. Toward Massively Parallel Direct Space Computations. *J. Chem. Theory Comput.* **2014,** *10*, 1638.
22. Albaugh, A.; Niklasson, A. M. N.; Head-Gordon, T., Accurate Classical Polarization Solution with No Self-Consistent Field Iterations. *The Journal of Physical Chemistry Letters* **2017,** *8* (8), 1714-1723.
23. 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.
24. Albaugh, A.; Head-Gordon, T., A New Method for Treating Drude Polarization in Classical Molecular Simulation. *Journal of Chemical Theory and Computation* **2017,** *13* (11), 5207-5216.
25. Albaugh, A.; Head-Gordon, T.; Niklasson, A. M. N., Higher-Order Extended Lagrangian Born–Oppenheimer Molecular Dynamics for Classical Polarizable Models. *Journal of Chemical Theory and Computation* **2018,** *14* (2), 499-511.
26. Tuckerman, M.; Berne, B. J.; Martyna, G. J., Reversible Multiple Time Scale Molecular Dynamics. *J. Chem. Phys.* **1992,** *97*, 1990.
27. Minary, P.; Tuckerman, M. E.; Martyna, G. J., Long Time Molecular Dynamics for Enhanced Conformational Sampling in Biomolecular Systems. *Phys. Rev. Lett.* **2004,** *93*, 150201.
28. Warshel, A.; Levitt, M., Theoretical studies of enzymatic reaction: Dielectric, electrostatic, and steric stabilization of carbonium-ion in reaction of lysozyme. *J. Mol. Biol.* **1976,** *103*, 227.
29. Applequist, J.; Carl, J. R.; Fung, K. K., Atom Dipole Interaction Model for Molecular Polarizability - Application to Polyatomic-Molecules and Determination of Atom Polarizabilities. *J. Am. Chem. Soc.* **1972,** *94* (9), 2952.
30. Thole, B. T., Molecular Polarizabilities Calculated with a Modified Dipole Interaction. *Chem. Phys.* **1981,** *59* (3), 341.
31. Ren, P. Y.; Ponder, J. W., Polarizable Atomic Multipole Water Model for Molecular Mechanics Simulation. *J. Phys. Chem. B* **2003,** *107*, 5933.
32. Ponder, J. W.; Wu, C.; Ren, P.; Pande, V. S.; Chodera, J. D.; Schnieders, M. J.; Haque, I.; Mobley, D. L.; Lambrecht, D. S.; DiStasio, R. A., Current Status of the AMOEBA Polarizable Force Field. *J. Phys. Chem. B* **2010,** *114*, 2549.
33. Ren, P. Y.; Wu, C. J.; Ponder, J. W., Polarizable Atomic Multipole-Based Molecular Mechanics for Organic Molecules. *J. Chem. Theory Comput.* **2011,** *7*, 3143.
34. Laury, M. L.; Wang, L. P.; Pande, V. S.; Head-Gordon, T.; Ponder, J. W., Revised Parameters for the AMOEBA Polarizable Atomic Multipole Water Model. *J. Phys. Chem. B* **2015,** *119*, 9423.
35. Piquemal, J. P.; Gresh, N.; Giessner-Prettre, C., Improved Formulas for the Calculation of the Electrostatic Contribution to the Intermolecular Interaction Energy from Multipolar Expansion of the Electronic Distribution. *J. Phys. Chem. A* **2003,** *107*, 10353.
36. Engkvist, O.; Astrand, P. O.; Karlstrom, G., Accurate intermolecular potentials obtained from molecular wave functions: Bridging the gap between quantum chemistry and molecular simulations. *Chem. Rev.* **2000,** *100* (11), 4087.
37. Palmo, K.; Mannfors, B.; Mirkin, N. G.; Krimm, S., Inclusion of charge and polarizability fluxes provides needed physical accuracy in molecular mechanics force fields. *Chem. Phys. Lett.* **2006,** *429* (4–6), 628.
38. Xie, W. S.; Pu, J. Z.; MacKerell, A. D.; Gao, J. L., Development of a polarizable intermolecular potential function (PIPF) for liquid amides and alkanes. *J. Chem. Theory Comput.* **2007,** *3* (6), 1878.
39. Jorgensen, W. L.; Jensen, K. P.; Alexandrova, A. N., Polarization Effects for Hydrogen-Bonded Complexes of Substituted Phenols with Water and Chloride Ion. *J. Chem. Theory Comput.* **2007,** *3*, 1987.
40. Jacucci, G.; McDonald, I. R.; Singer, K., Introduction of Shell-Model of Ionic Polarizability into Molecular-Dynamics Calculations. *Phys. Lett. A* **1974,** *50* (2), 141.
41. Mitchell, P. J.; Fincham, D., Shell-Model Simulations by Adiabatic Dynamics. *J. Phys.: Condens. Matter* **1993,** *5* (8), 1031.
42. Lamoureux, G.; MacKerell, A. D.; Roux, B., A Simple Polarizable Model of Water Based on Classical Drude Oscillators. *J. Chem. Phys.* **2003,** *119*, 5185.
43. Lamoureux, G.; Roux, B., Modeling Induced Polarization with Classical Drude Oscillators: Theory and Molecular Dynamics Simulation Algorithm. *J. Chem. Phys.* **2003,** *119*, 3025.

44. Lopes, P. E. M.; Huang, J.; Shim, J.; Luo, Y.; Li, H.; Roux, B.; MacKerell, A. D., Polarizable Force Field for Peptides and Proteins Based on the Classical Drude Oscillator. *J. Chem. Theory Comput.* **2013,** *9*, 5430.
45. Yu, H. B.; Hansson, T.; van Gunsteren, W. F., Development of a simple, self-consistent polarizable model for liquid water. *J. Chem. Phys.* **2003,** *118* (1), 221.
46. Geerke, D. P.; van Gunsteren, W. F., Calculation of the Free Energy of Polarization: Quantifying the Effect of Explicitly Treating Electronic Polarization on the Transferability of Force-Field Parameters. *J. Phys. Chem. B* **2007,** *111*, 6425.
47. Rick, S.; Stuart, S.; Berne, B., Dynamical fluctuating charge force-fields: Application to liquid water. *J. Chem. Phys.* **1994,** *101* (7), 6141.
48. Rick, S. W.; Stuart, S. J.; Bader, J. S.; Berne, B. J., Fluctuating Charge Force-Fields for Aqueous-Solutions. *J. Mol. Liq.* **1995,** *65–66*, 31.
49. Stern, H. A.; Kaminski, G. A.; Banks, J. L.; Zhou, R. H.; Berne, B. J.; Friesner, R. A., Fluctuating Charge, Polarizable Dipole, and Combined Models: Parameterization from Ab Initio Quantum Chemistry. *J. Phys. Chem. B* **1999,** *103*, 4730.
50. 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 Chemistry. *J. Phys. Chem. A* **2004,** *108*, 621.
51. Patel, S.; Brooks, C. L., CHARMM Fluctuating Charge Force Field for Proteins: I Parameterization and Application to Bulk Organic Liquid Simulations. *J. Comput. Chem.* **2004,** *25*, 1.
52. Patel, S.; Mackerell, A. D.; Brooks, C. L., CHARMM fluctuating charge force field for proteins: II - Protein/solvent properties from molecular dynamics simulations using a nonadditive electrostatic model. *J. Comput. Chem.* **2004,** *25* (12), 1504.
53. Sprik, M.; Klein, M. L., A Polarizable Model for Water Using Distributed Charge Sites. *J. Chem. Phys.* **1988,** *89*, 7556.
54. Dykstra, C. E., Electrostatic interaction potentials in molecular force fields. *Chemical Reviews* **1993,** *93* (7), 2339-2353.
55. Drude, P., Zur elektronentheorie der metalle; II. Teil. galvanomagnetische und thermomagnetische effecte. *Annalen der Physik* **1900,** *308* (11), 369-402.
56. Drude, P., Zur elektronentheorie der metalle. *Annalen der Physik* **1900,** *306* (3), 566-613.
57. Halgren, T. A., The representation of van der Waals (vdW) interactions in molecular mechanics force fields: potential form, combination rules, and vdW parameters. *Journal of the American Chemical Society* **1992,** *114* (20), 7827-7843.
58. Stone, A., Distributed multipole analysis, or how to describe a molecular charge distribution. *Chemical Physics Letters* **1981,** *83* (2), 233-239.
59. Andersen, H. C., Rattle: A "velocity" version of the shake algorithm for molecular dynamics calculations. *J. Comput. Phys.* **1983,** *52* (1), 24.
60. Young, D., *Iterative Solutions of Large Linear Systems*. 1971.
61. Lagardère, L.; Lipparini, F.; Polack, É.; Stamm, B.; Cancès, É.; Schnieders, M.; Ren, P.; Maday, Y.; Piquemal, J. P., Scalable Evaluation of Polarization Energy and Associated Forces in Polarizable Molecular Dynamics: II. Toward Massively Parallel Computations Using Smooth Particle Mesh Ewald. *J. Chem. Theory Comput.* **2015,** *11*, 2589.
62. Aviat, F.; Levitt, A.; Stamm, B.; Maday, Y.; Ren, P.; Ponder, J. W.; Lagardère, L.; Piquemal, J. P., Truncated Conjugate Gradient: An Optimal Strategy for the Analytical Evaluation of the Many-Body Polarization Energy and Forces in Molecular Simulations. *J. Chem. Theory Comput.* **2017,** *13* (1), 180.
63. Kolafa, J., Time-reversible always stable predictor–corrector method for molecular dynamics of polarizable molecules. *J. Comput. Chem.* **2004,** *25* (3), 335.
64. Marx, D.; Hutter, J., *Ab Initio Molecular Dynamics*. 2009.
65. Martyna, G. J.; Klein, M. L.; Tuckerman, M., Nosé–Hoover Chains: The Canonical Ensemble Via Continuous Dynamics. *J. Chem. Phys.* **1992,** *97*, 2635.
66. Martyna, G. J.; Tuckerman, M. E.; Tobias, D. J.; Klein, M. L., Explicit Reversible Integrators for Extended Systems Dynamics. *Mol. Phys.* **1996,** *87*, 1117.
67. Car, R.; Parrinello, M., Unified Approach for Molecular-Dynamics and Density-Functional Theory. *Phys. Rev. Lett.* **1985,** *55* (22), 2471.
68. 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.

69. Sprik, M., Computer simulation of the dynamics of induced polarization fluctuations in water. *J. Phys. Chem.* **1991,** *95* (6), 2283.
70. Niklasson, A. M.; Tymczak, C.; Challacombe, M., Time-Reversible Born-Oppenheimer Molecular Dynamics. *Phys. Rev. Lett.* **2006,** *97*, 123001.
71. Niklasson, A. M.; Tymczak, C.; Challacombe, M., Time-Reversible Ab Initio Molecular Dynamics. *J. Chem. Phys.* **2007,** *126*, 144103.
72. Niklasson, A. M., Extended Born-Oppenheimer Molecular Dynamics. *Phys. Rev. Lett.* **2008,** *100*, 123004.
73. 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.
74. Steneteg, P.; Abrikosov, I. A.; Weber, V.; Niklasson, A. M. N., Wave function extended lagrangian born-oppenheimer molecular dynamics. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2010,** *82*, 075110.
75. 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.
76. 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.
77. Niklasson, A. M. N.; Cawkwell, M. J., Fast method for quantum mechanical molecular dynamics. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2012,** *86*, 174308.
78. 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.
79. Niklasson, A. M.; Cawkwell, M. J., Generalized Extended Lagrangian Born-Oppenheimer Molecular Dynamics. *J. Chem. Phys.* **2014,** *141*, 164123.
80. Souvatzis, P.; Niklasson, A. M. N., First principles molecular dynamics without self-consistent field optimization. *J. Chem. Phys.* **2014,** *140*, 044117.
81. Niklasson, A. M. N., Next generation extended lagrangian first principles molecular dynamics. *J. Chem. Phys.* **2017,** *147*, 054103.
82. 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.
83. 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* (16), 164101.
84. Demerdash, O.; Head-Gordon, T., Parallel implementation of approximate atomistic models of the AMOEBA polarizable model. *Chemical Physics Letters* **2016,** *664*, 191-198.
85. Demerdash, O. N.; Head-Gordon, T., Convergence of the Many-Body Expansion for Energy and Forces for Classical Polarizable Models in the Condensed Phase. *J. Chem. Theory Comput.* **2016,** *12*, 3884.
86. Xie, W.; Pu, J.; Gao, J., A coupled polarization-matrix inversion and iteration approach for accelerating the dipole convergence in a polarizable potential function. *The Journal of Physical Chemistry A* **2009,** *113* (10), 2109-2116.
87. Pulay, P., Convergence Acceleration of Iterative Sequences. The Case of SCF Iteration. *Chem. Phys. Lett.* **1980,** *73*, 393.
88. 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 Physical Clusters of Molecules: Application to Small Water Clusters. *J. Chem. Phys.* **1982,** *76*, 637.
89. T., S.; M., M.; D., S. R.; K., S., *J. Comput. Phys.* **1998,** *140*, 1.
90. A., M. J.; E., M. T.; M., C.; J., B. B., *J. Chem. Phys.* **2011,** *134*, 014103.
91. Berendsen, H. J.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J., Molecular Dynamics with Coupling to an External Bath. *J. Chem. Phys.* **1984,** *81*, 3684.
92. 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. *The Journal of Chemical Physics* **2017,** *146* (12), 124115.
93. van Duijnen, P. T.; Swart, M., *J. Phys. Chem. A* **1998,** *102*, 2399.

94. Császár, P.; Pulay, P., Geometry optimization by direct inversion in the iterative subspace. *J. Mol. Struct.* **1984,** *114*, 31.
95. Sagui, C.; Pedersen, L. G.; Darden, T. A., Towards an Accurate Representation of Electrostatics in Classical Force Fields: Efficient Implementation of Multipolar Interactions in Biomolecular Simulations. *J. Chem. Phys.* **2004,** *120*, 73.
96. McLachlan, R. I.; Atela, P., The accuracy of symplectic integrators. *Nonlin.* **1992,** *5*, 541.
97. Wang, L. P.; Chen, J.; van Voorhis, T., Systematic Parametrization of Polarizable Force Fields from Quantum Chemistry Data. *J. Chem. Theory Comput.* **2013,** *9*, 452.
98. Esser, A.; Belsare, S.; Marx, D.; Head-Gordon, T., Mode specific THz spectra of solvated amino acids using the AMOEBA polarizable force field. *Phys. Chem. Chem. Phys.* **2017,** *19* (7), 5579.
99. Darden, T.; York, D.; Pedersen, L., Particle Mesh Ewald: An Nlog(N) Method for Ewald Sums in Large Systems. *J. Chem. Phys.* **1993,** *98*, 10089.
100. Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G., A Smooth Particle Mesh Ewald Method. *J. Chem. Phys.* **1995,** *103*, 8577.
101. Toukmaji, A.; Sagui, C.; Board, J.; Darden, T., Efficient particle-mesh Ewald based approach to fixed and induced dipolar interactions. *J. Chem. Phys.* **2000,** *113* (24), 10913.
102. Brooks, B. R.; Brooks, C. L.; Mackerell, A. D.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S., CHARMM: The Biomolecular Simulation Program. *J. Comput. Chem.* **2009,** *30*, 1545.
103. Nymand, T. M.; Linse, P., Ewald Summation and Reaction Field Methods for Potentials with Atomic Charges, Dipoles, and Polarizabilities. *J. Chem. Phys.* **2000,** *112*, 6152.
104. Gans, J.; Shalloway, D., Shadow mass and the relationship between velocity and momentum in symplectic numerical integration. *Phys. Rev. E: Stat. Phys., Plasmas, Fluids, Relat. Interdiscip. Top.* **2000,** *61*, 4587.
105. Hairer, E.; Lubich, C.; Wanner, G., Geometric numerical integration illustrated by the störmer–verlet method. *Acta Numer.* **2003,** *12*, 399.
106. Niklasson, A. M. N., Iterative refinement method for the approximate factorization of a matrix inverse. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2004,** *70*, 193102.
107. Schulz, G., Iterative berechung der reziproken matrix. *Z. Angew. Math. Mech.* **1933,** *13*, 57.
108. Leimkuhler, B.; Margul, D. T.; Tuckerman, M. E., Stochastic, Resonance-Free Multiple Time-Step Algorithm for Molecular Dynamics with Very Large Time Steps. *Mol. Phys.* **2013,** *111*, 3579.
109. Margul, D. T.; Tuckerman, M. E., A Stochastic, Resonance-Free Multiple Time-Step Algorithm for Polarizable Models That Permits Very Large Time Steps. *J. Chem. Theory Comput.* **2016,** *12*, 2170.
110. Yoshida, H., Construction of higher order symplectic integrators. *Physics letters A* **1990,** *150* (5-7), 262-268.
111. Suzuki, M., General theory of fractal path integrals with applications to many‐body theories and statistical physics. *Journal of Mathematical Physics* **1991,** *32* (2), 400-407.

Appendix

Appendix A: iEL/SCF Numerical Analysis

Following the numerical stability analysis of Niklasson and colleagues[73] I have developed a Verlet-like recursion of the Berendsen rescaling applied to auxiliary induced dipoles, $\boldsymbol{a}$, for analysis purposes. First two Taylor series expansions of the auxiliary dipole "positions", given by Eq. (A.1), are added.

$$\boldsymbol{a}_{n+2} = \boldsymbol{a}_{n+1} + \dot{\boldsymbol{a}}_{n+1}\,\Delta t + \frac{1}{2}\ddot{\boldsymbol{a}}_{n+1}\,\Delta t^2 \tag{A.1a}$$

$$\boldsymbol{a}_{n} = \boldsymbol{a}_{n+1} - \dot{\boldsymbol{a}}_{n}\Delta t - \frac{1}{2}\ddot{\boldsymbol{a}}_{n}\Delta t^2 \tag{A.1b}$$

Combining Eq. (A.1a) and (A.1b), Eq. (A.2) is obtained.

$$\boldsymbol{a}_{n+2} = 2\boldsymbol{a}_{n+1} - \boldsymbol{a}_{n} + \Delta t(\dot{\boldsymbol{a}}_{n+1} - \dot{\boldsymbol{a}}_{n}) + + \frac{1}{2}\Delta t^2(\ddot{\boldsymbol{a}}_{n+1} - \ddot{\boldsymbol{a}}_{n}) \tag{A.2}$$

$$\dot{\boldsymbol{a}}_{n+1} = \alpha_{n+1}[\dot{\boldsymbol{a}}_{n} + \frac{1}{2}\Delta t(\ddot{\boldsymbol{a}}_{n+1} + \ddot{\boldsymbol{a}}_{n})] \tag{A.3}$$

Using a scaled velocity from the velocity Verlet recursion, given by Eq. (A.3) where α_{n+1} is a velocity rescaling factor given by Eq. (2.6), one can substitute for $\dot{\boldsymbol{a}}_n$ in Eq. (A.2) to obtain Eq. (A.4).

$$\boldsymbol{a}_{n+2} = 2\boldsymbol{a}_{n+1} - \boldsymbol{a}_{n} + \Delta t\dot{\boldsymbol{a}}_{n+1}\left(1 - \frac{1}{\alpha_{n+1}}\right) + \Delta t^2\ddot{\boldsymbol{a}}_{n+1} \tag{A.4}$$

$$\dot{\boldsymbol{a}}_{n+1} = \frac{1}{2\,\Delta t}(\boldsymbol{a}_{n+2} - \boldsymbol{a}_{n}) \tag{A.5}$$

Finally an approximation for $\dot{\boldsymbol{a}}_{n+1}$ is given by a finite difference, shown in Eq. (A.5) and Eq. (2.5b) is substituted for $\ddot{\boldsymbol{a}}_{n+1}$ to obtain the recursion of Eq. (A.6).

$$\frac{1}{2}\left(1 + \frac{1}{\alpha_n}\right)\boldsymbol{a}_{n+1} = 2\boldsymbol{a}_{n} - \frac{1}{2}\left(3 - \frac{1}{\alpha_n}\right)\boldsymbol{a}_{n+1} + \Delta t^2\omega^2(\boldsymbol{\mu}_n^{SCF} - \boldsymbol{a}_{n}) \tag{A.6}$$

In Eq. (A.6) $\boldsymbol{a}_n = \boldsymbol{a}_i(t_0 + n\Delta t)$ and $\alpha_n = \alpha(t_0 + n\Delta t)$ is the instantaneous velocity rescaling factor. Note that when there is no pseudo temperature control ($\alpha_n = 1$) we recover the Verlet recursion, as expected. To analyze the numerical stability of the velocity rescaling scheme Eq. (A.6) has a characteristic equation given by Eq. (A.7).

$$\frac{1}{2}\left(1+\frac{1}{\alpha_n}\right)\lambda^{n+1}+(\Delta t^2\omega^2-\gamma\Delta t^2\omega^2-2)\lambda^n+\frac{1}{2}\left(3-\frac{1}{\alpha_n}\right)\lambda^{n-1}=0 \tag{A.7}$$

In Eq. (A.7) λ corresponds to the roots of the characteristic equation and γ is the largest eigenvalue of the iterative response matrix acting on the difference between $\boldsymbol{a}_n$ and the true solution of the induced dipoles. Thus Eq. (A.7) allows us to explore various levels of convergence of the solution of the auxiliary dipoles over the range $-1<\gamma<1$, where in the limit $\gamma\to 0$ corresponds to the exact solution, while the magnitude of the largest root, $|\lambda|_{max}$, will determine the stability of the recursion, with $|\lambda|_{max}>1$ giving exponentially growing solutions, $|\lambda|_{max}<1$ giving exponentially decaying solutions, with an exactly stable solution occurring at $|\lambda|_{max}=1$. I note that $\alpha\neq 1$ corresponds to increasing time irreversibility in the equations of motion, which should be avoided when possible. Time-reversibility is preserved in the iEL/SCF method when $\alpha=1$, whereas a related coupling parameter that multiplies the dissipation force, α_{diss}, discussed in[73] remains time reversible when $\alpha_{diss}=0$.

Similarly, I can show that for a Nosé-Hoover chain (NHC) thermostat (taking into account rescaling that happens twice per time step now) the Verlet-like recursion is given by Eq. (A.8).

$$\begin{aligned}\frac{1}{2}\left(1+\frac{1}{\alpha_n\alpha_{n-1}}\right)\boldsymbol{a}_{n+1}=\;&2\boldsymbol{a}_n-\frac{1}{2}\left(3-\frac{1}{\alpha_n\alpha_{n-1}}\right)\boldsymbol{a}_{n-1}+\frac{1}{2}\left(\frac{1}{\alpha_{n-1}}+1\right)\Delta t^2\omega^2(\boldsymbol{\mu}_n^{SCF}-\boldsymbol{a}_n)\\&+\frac{1}{2}\left(\frac{1}{\alpha_{n-1}}-1\right)\Delta t^2\omega^2(\boldsymbol{\mu}_{n-1}^{SCF}-\boldsymbol{a}_{n-1})\end{aligned} \tag{A.8}$$

In Eq. (A.8) α_n and α_{n-1} now represent the velocity scaling factors that come from a time-reversible Nosé-Hoover method at the end and beginning of the *n*-th recursion. These two α-values mask the complexity of the Nosé-Hoover chains which are also a function of additional extended system variables; see Martyna *et al.* for details[66]. Again, one can construct the characteristic equation of Eq. (A.8) for analysis, given in Eq. (A.9).

$$\begin{aligned}\frac{1}{2}\left(1+\frac{1}{\alpha_n\alpha_{n-1}}\right)\lambda^{n+1}+\frac{1}{2}&\left(\frac{\Delta t^2\omega^2}{\alpha_{n-1}}+\Delta t^2\omega^2-\frac{\gamma\Delta t^2\omega^2}{\alpha_{n-1}}-\gamma\Delta t^2\omega^2-2\right)\lambda^n\\&+\frac{1}{2}\left(3-\frac{1}{\alpha_n\alpha_{n-1}}+\frac{\Delta t^2\omega^2}{\alpha_{n-1}}-\Delta t^2\omega^2-\frac{\gamma\Delta t^2\omega^2}{\alpha_{n-1}}+\gamma\Delta t^2\omega^2\right)\lambda^{n-1}=0\end{aligned} \tag{A.9}$$

Figure A.1a gives $|\lambda|_{max}$ as a function of the SCF convergence, γ, for various values of α (assuming $\alpha_n=\alpha_{n-1}$ for the NHC case). Intuitively for $\alpha>1$ then $|\lambda|_{max}>1$ and the equations of motion are unstable which would correspond to an accumulation in the pseudo kinetic energy as observed in Figure A.1b. For $\alpha<1$ the increasing dissipation will realize stable solutions ($|\lambda|_{max}<1$) but at the expense of time-reversibility as α decreases. Thus both the Berendsen weak coupling and NHC iEL/SCF schemes have the desirable property that the equations of motion can be made stable under incomplete SCF convergence in the full γ interval for α values that are close to that needed for time reversibility. We would like to note that Eqs. (A.6) and (A.8) are presented for analysis purposes only and the method is truly implemented using a velocity Verlet scheme with thermostat action being applied at the appropriate points within such a scheme.

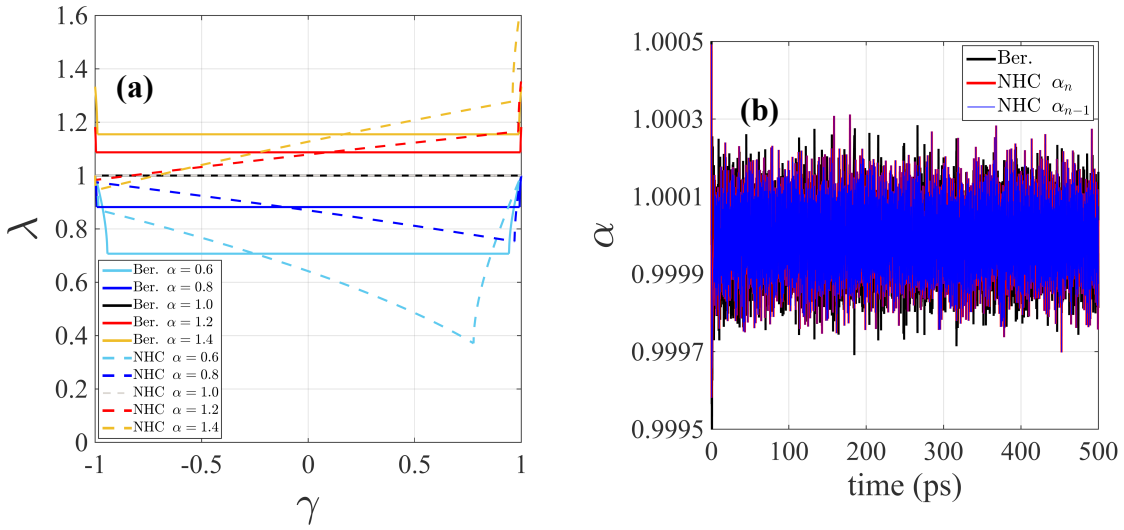


Figure A.1: *The use of thermostats for the hybrid EL/SCF scheme using Berendsen weak coupling velocity scaling. (a) Roots of the characteristic equation for Berendsen rescaling (Eq. A.7) and Nosé-Hoover thermostating (Eq. A.9) as a function of the degree of SCF convergence, γ, for various velocity rescaling factors, α. Note that for this figure the range of α is exaggerated with observed α values ranging from 0.9997 and 1.0003 in the course of a typical simulation. (b) The time trajectory for the velocity rescaling factors, α, for the Berendsen scheme and the Nosé-Hoover scheme. The Nosé-Hoover scheme scales the velocity at the beginning and end of a single step, hence the 'n-1' and 'n', respectively.*

Figure A.1b shows the simulated trajectory of the rescaling parameter α during the course of our weak coupling Berendsen velocity rescaling as well as Nosé-Hoover scheme which is shown to range from 0.9997 to 1.0003 with an average of ~0.99999 such that the simulations are essentially close to the exact time–reversible solution. While velocity rescaling using weak coupling Berendsen formally breaks the time-reversibility of the integration scheme, an α-value so close to 1.0 corresponds to only a slight disturbance of this reversibility while dissipating the integration error that causes divergence in $\langle \dot{\boldsymbol{a}}_i{}^2 \rangle/3$ due to resonances. In any event, errors in time reversibility are formally circumvented through use of NHC thermostats, also shown in Figure A.1b, although, at least for our test system of bulk water, the practical differences are largely unimportant. Thus my diagnosis of the problem in the original hybrid EL/SCF scheme arises from resonances in the auxiliary equations of motion that can be controlled by a simple velocity rescaling scheme that prevents the accumulation of a pseudo kinetic energy for these degrees of freedom.

Appendix B: Additional iEL/SCF Validation Data

Table B.1. *Average potential energy and average molecular dipole moment, and diffusion coefficients as a function of mutual induction convergence for PCG with predictor (standard SCF), standard EL, the hybrid EL/SCF scheme with no dissipation (hybrid EL/SCF), and a hybrid EL/SCF method which thermostats the auxiliary dipole velocities using either Berendsen rescaling or a Nosé-Hoover thermostat. Average potential energy and molecular dipole were calculated from NVE simulations.*

Standard SCF		
Convergence (RMS Debye)	**Average Potential Energy (kcal/mol)**	**Average Molecular Dipole (Debye)**
10^{-6}	-8.78±0.05	2.738±0.011
10^{-5}	-8.79±0.05	2.739±0.011
10^{-4}	-7.79±0.25	2.678±0.036
10^{-3}	-8.27±0.02	2.700±0.012
10^{-2}	-12.73±0.01	3.019±0.001
10^{-1}	-12.70±0.01	3.037±0.001
Standard EL		
Time step (fs)	**Average Potential Energy (kcal/mol)**	**Average Molecular Dipole (Debye)**
0.25	-9.09±0.16	2.761±0.017
0.50	-9.41±0.22	2.777±0.024
Hybrid EL/SCF		
Convergence (RMS Debye)	**Average Potential Energy (kcal/mol)**	**Average Molecular Dipole (Debye)**
10^{-6}	-8.74±0.05	2.735±0.011
10^{-5}	-8.74±0.05	2.734±0.011
10^{-4}	-8.74±0.05	2.735±0.011
10^{-3}	-8.74±0.05	2.734±0.011
10^{-2}	-8.76±0.05	2.736±0.011
10^{-1}	-8.89±0.05	2.748±0.012
Hybrid EL/SCF with Berendsen		
Convergence (RMS Debye)	**Average Potential Energy (kcal/mol)**	**Average Molecular Dipole (Debye)**
10^{-6}	-8.75±0.05	2.735±0.011
10^{-5}	-8.75±0.05	2.735±0.011
10^{-4}	-8.75±0.05	2.735±0.011
10^{-3}	-8.75±0.05	2.735±0.011
10^{-2}	-8.75±0.05	2.735±0.011
10^{-1}	-8.74±0.05	2.734±0.011
Hybrid EL/SCF wth Nosé-Hoover		
Convergence (RMS Debye)	**Average Potential Energy (kcal/mol)**	**Average Molecular Dipole (Debye)**
10^{-6}	-8.75±0.05	2.735±0.011
10^{-5}	-8.75±0.05	2.735±0.011
10^{-4}	-8.75±0.05	2.735±0.011
10^{-3}	-8.75±0.05	2.736±0.011
10^{-2}	-8.75±0.05	2.735±0.011
10^{-1}	-8.74±0.05	2.734±0.011

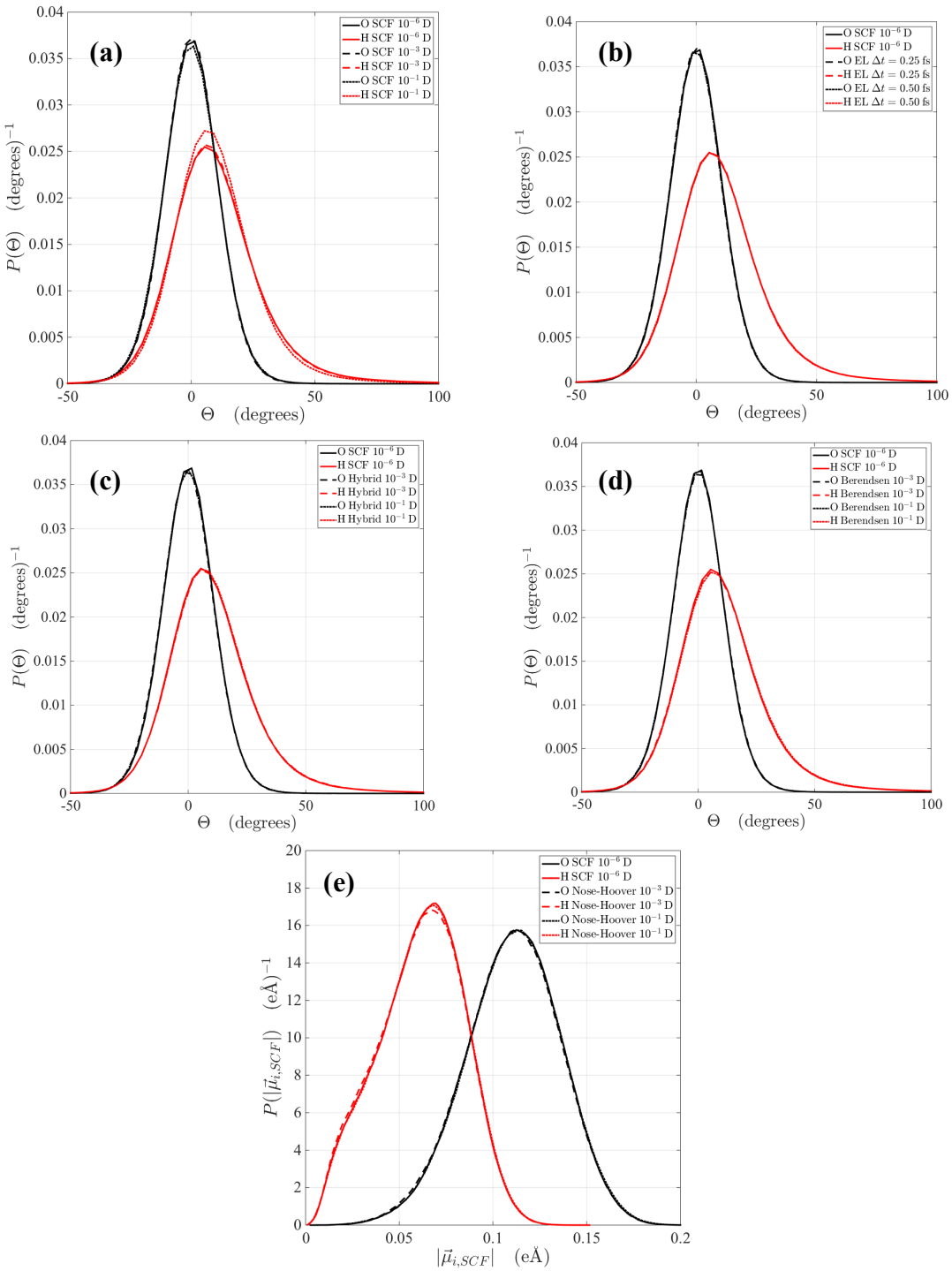


Figure B.1: *Comparison of the NVT ensemble averaged probability distributions for in-plane induced dipole angle, θ, using (a) standard SCF solver (PCG with predictor), (b) basic EL method, (c) the adapted EL/SCF method with no dissipation (hybrid), (d) the adapted EL/SCF method with pseudo temperature control of the auxiliary dipoles using Berendsen rescaling, and (e) the adapted EL/SCF method with pseudo temperature control of the auxiliary dipoles using a Nosé-Hoover thermostat. The continuous curves give the data obtained using base AMOEB14A at a level of 10^{-6} RMS Debye convergence. The in-plane angle, θ, is defined as 0 at the H-O-H bisector for oxygen and 0 along the O-H axis for hydrogen, with $+\theta$ being in the direction of the H-O-H bisector. In-plane angles range from -180 to 180 degrees.*

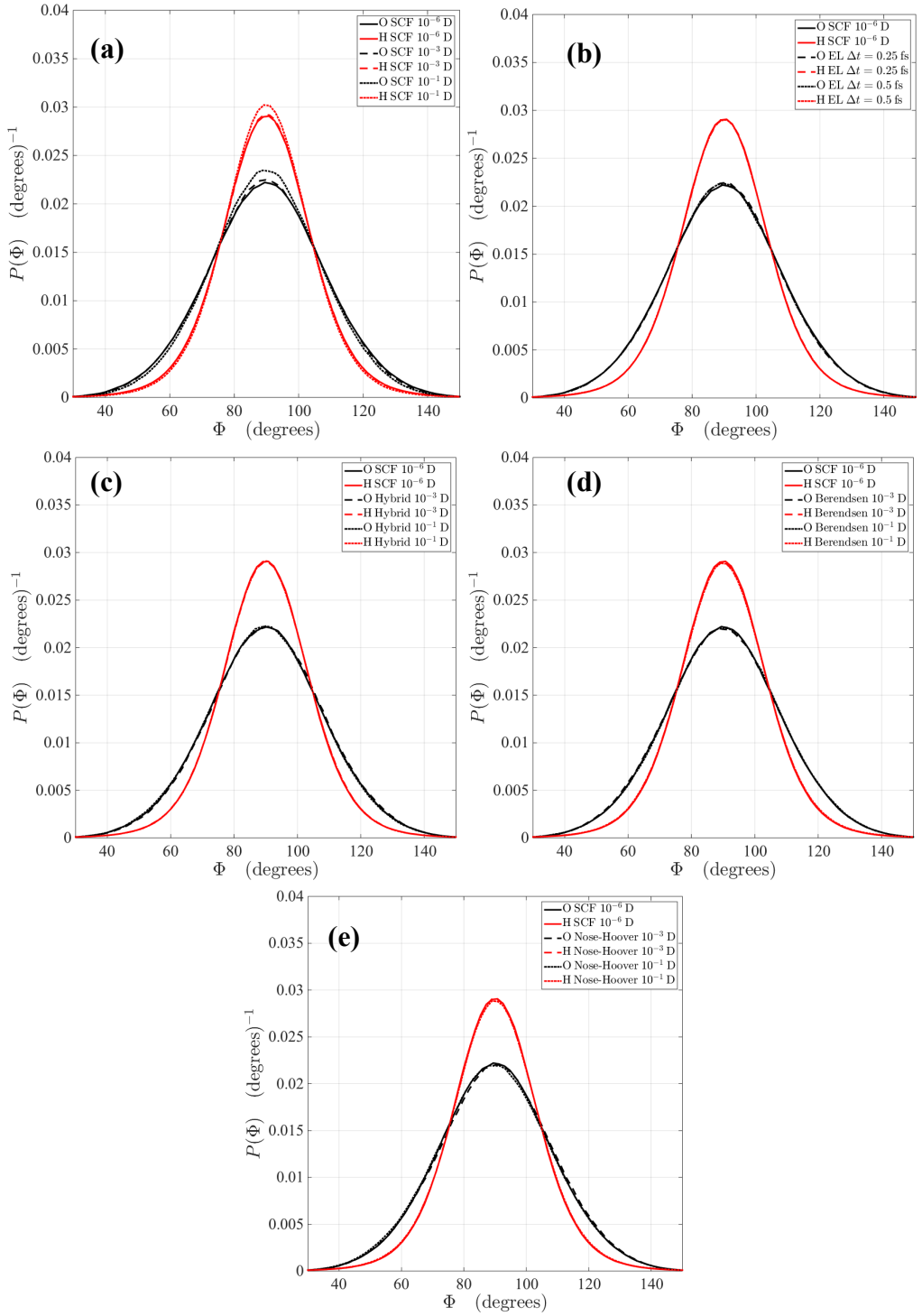


Figure B.2: *Comparison of the NVT ensemble averaged probability distributions for out-of-plane induced dipole angle, φ, using (a) standard SCF solver (PCG with predictor), (b) basic EL method, (c) the adapted EL/SCF method with no dissipation (hybrid), (d) the adapted EL/SCF method with pseudo temperature control of the auxiliary dipoles using Berendsen rescaling, and (e) the adapted EL/SCF method with pseudo temperature control of the auxiliary dipoles using a Nosé-Hoover thermostat. The continuous curves give the data obtained using base AMOEBA14 at a level of 10^{-6} RMS Debye convergence. The out-of-plane angle, φ, is defined as 0 perpendicular to the H-O-H plane and ranges from 0 to 180 degrees.*

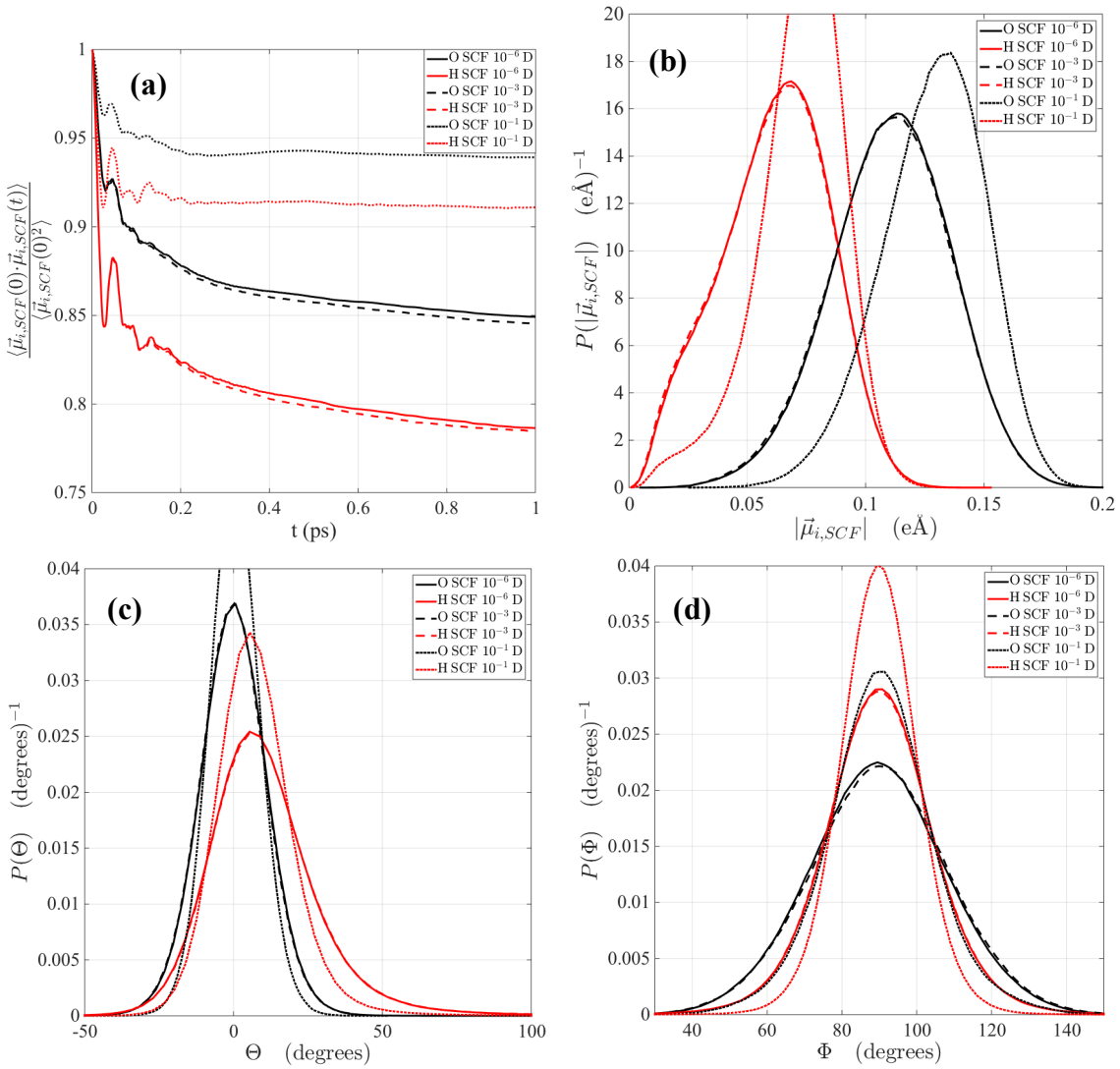


Figure B.3: *Comparison of the NVE ensemble averaged probability distributions for induced dipole properties using standard SCF (PCG with predictor). Normalized induced dipole autocorrelations (a) and probability distributions for induced dipole magnitude (b), in-plane angle (c), and out-of-plane angle (d) with various levels of SCF convergence for AMOEBA14 water.*

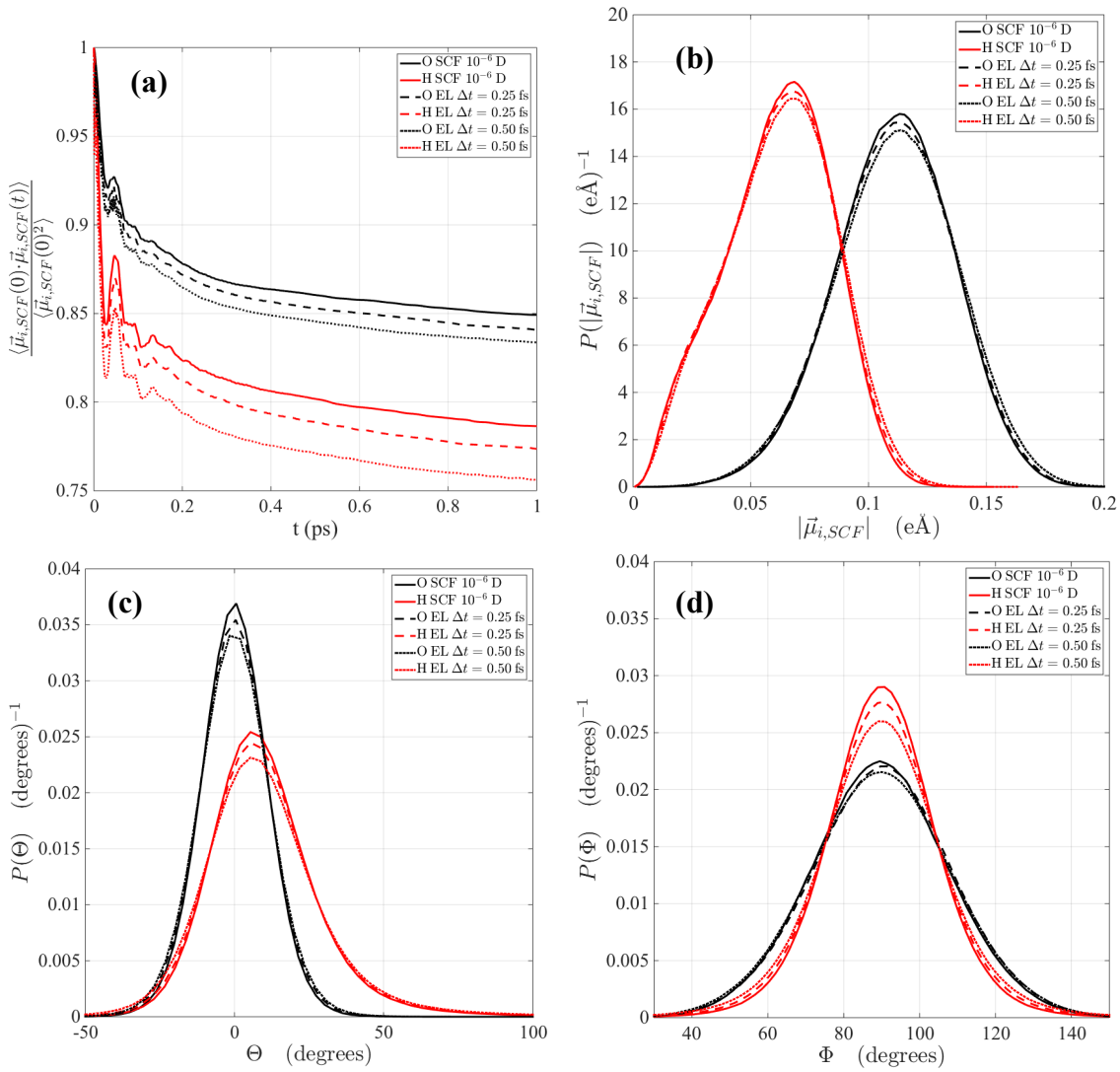


Figure B.4: *Comparison of the NVE ensemble averaged probability distributions for induced dipole properties using standard EL. Normalized induced dipole autocorrelations (a) and probability distributions for induced dipole magnitude (b), in-plane angle (c), and out-of-plane angle (d) with various levels of SCF convergence for AMOEBA14 water. A reference SCF result (10^{-6} RMS Debye) is shown as the solid curves.*

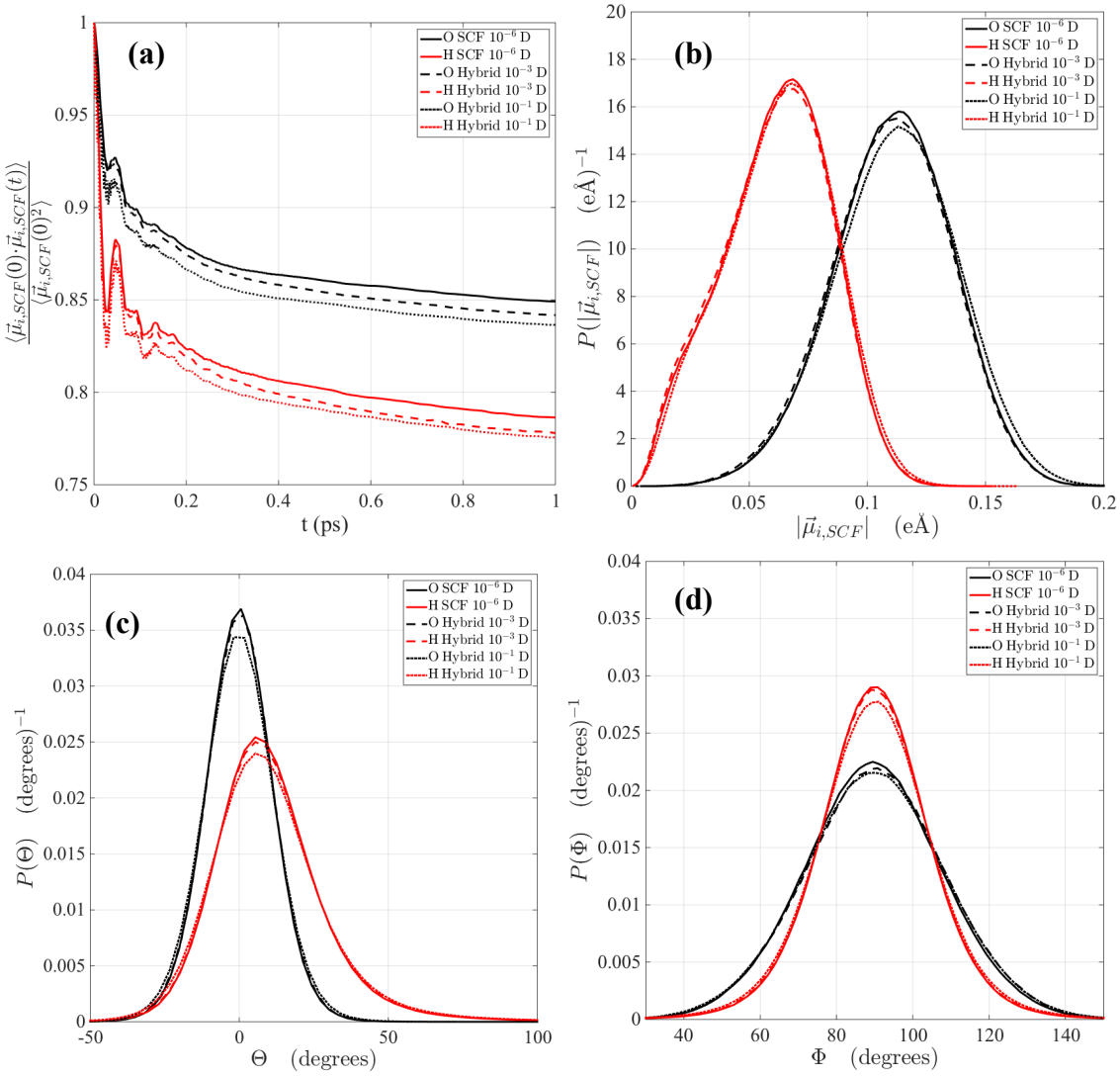


Figure B.5: *Comparison of the NVE ensemble averaged probability distributions for induced dipole properties using the hybrid EL/SCF scheme. Induced dipole autocorrelations (a) and probability distributions for induced dipole magnitude (b), in-plane angle (c), and out-of-plane angle (d) with various levels of SCF convergence for AMOEBA14 water. A reference SCF result (10^{-6} RMS Debye) is shown as the solid curves.*

.

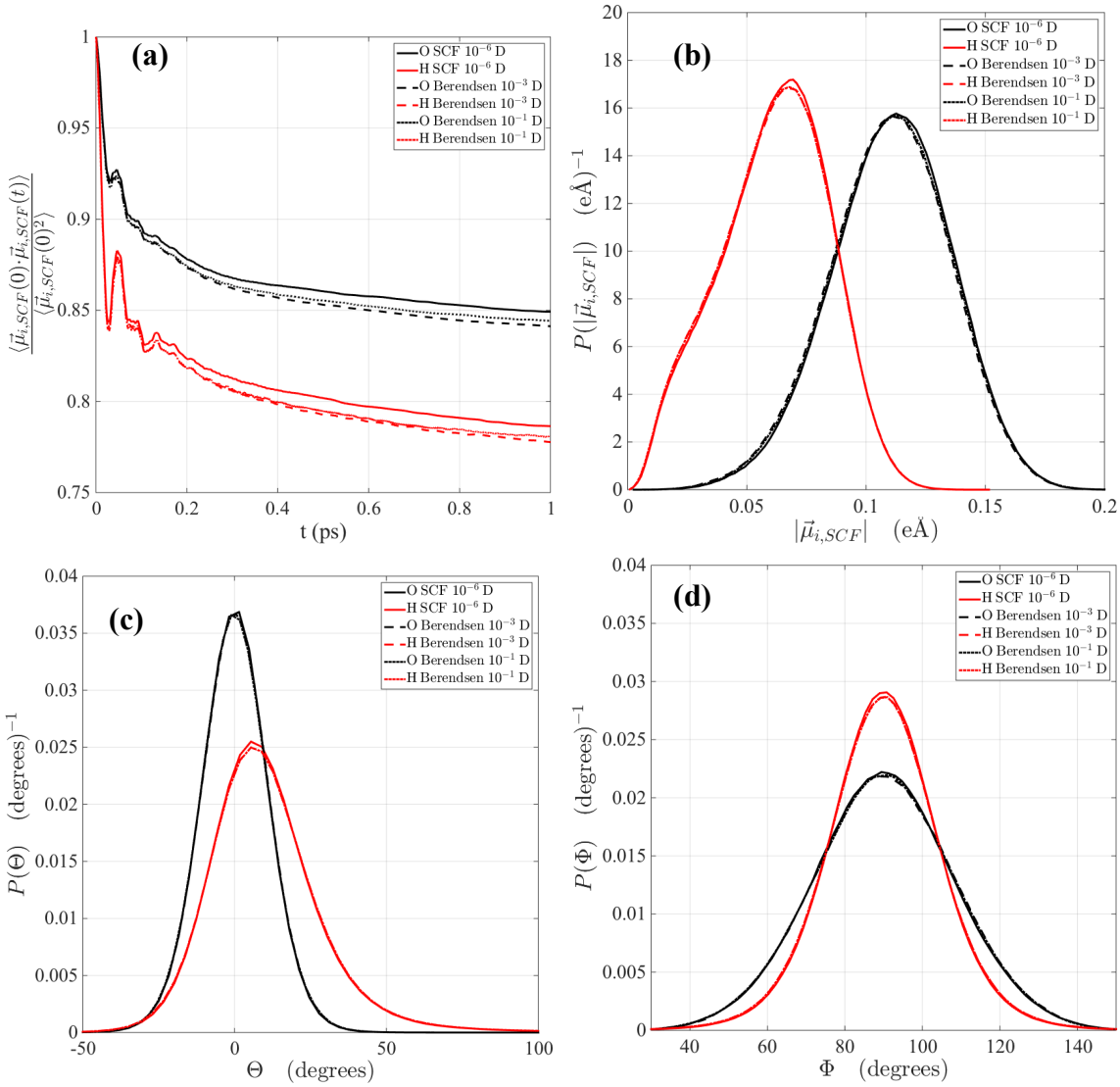


Figure S6: *Comparison of the NVE ensemble averaged probability distributions for induced dipole properties using the hybrid EL/SCF scheme with pseudo temperature control of the auxiliary dipoles using Berendsen rescaling. Induced dipole autocorrelations (a) and probability distributions for induced dipole magnitude (b), in-plane angle (c), and out-of-plane angle (d) with various levels of SCF convergence for AMOEBA14 water. A reference SCF result (10^{-6} RMS Debye) is shown as the solid curves.*

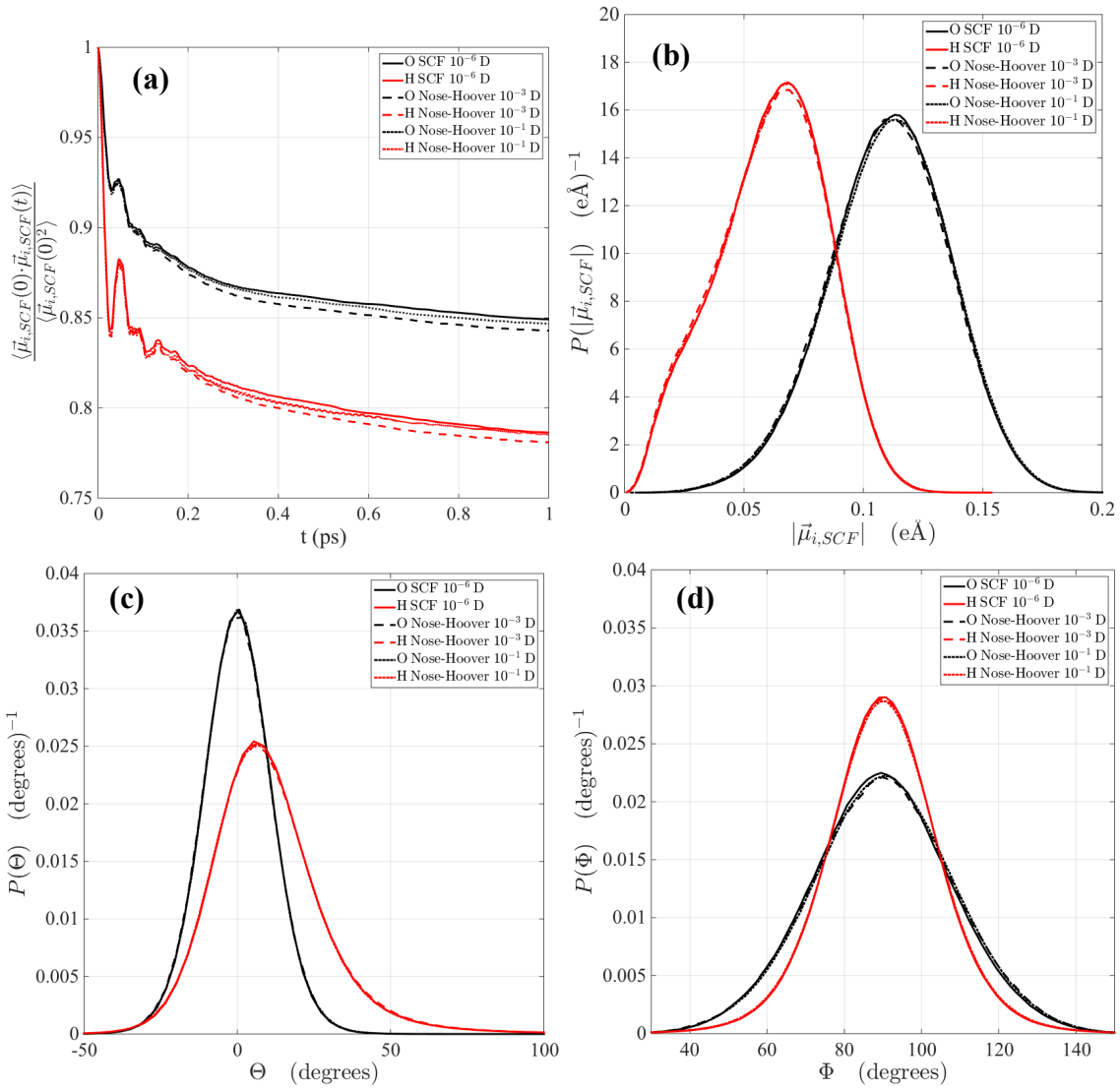


Figure B.7: *Comparison of the NVE ensemble averaged probability distributions for induced dipole properties using the hybrid EL/SCF scheme with pseudo temperature control of the auxiliary dipoles using a Nosé-Hoover thermostat. Induced dipole autocorrelations (a) and probability distributions for induced dipole magnitude (b), in-plane angle (c), and out-of-plane angle (d) with various levels of SCF convergence for AMOEBA14 water. A reference SCF result (10^{-6} RMS Debye) is shown as the solid curves.*

Appendix C: AMOEBA Force Field p- and d-Scaling

The AMOEBA force field uses scaling to exclude certain electrostatics based on bond separation, as these terms are modeled implicitly in bonded valence terms like bonds, angles, bond-angle, torsion, etc[33]. The forms of the electrostatic interactions are scaled differently depending on their form and type. Permanent-permanent interactions are scaled by a factor known as 'm-scaling' and induced-induced interactions are scaled by 'u-scaling', both of which are determined by the number of bonds separating the species to be scaled. Permanent-induced interactions, however, are scaled differently depending on whether the calculation is determining the induced dipoles or calculating the polarization energy. The former calculation relies on the polarization group the species are in and is referred to as 'd-scaling' while the latter relies on bond separation and is referred to as 'p-scaling'. For small species like water and ions there is no difference between p-scaling and d-scaling because the entire molecule is a single polarization group. When this is the case the equations presented in the body of the text are sufficient, but when larger molecules with multiple polarization groups are considered we need to adapt our equations accordingly.

For a general case with distinct p- and d-scaling one can start with a general definition of the polarization energy in Eq. (C.1).

$$U^{pol}(\boldsymbol{r^N}) = \frac{1}{2}\boldsymbol{\mu}^d \boldsymbol{C} \boldsymbol{\mu}^p - \frac{1}{2}\boldsymbol{\mu}^d \boldsymbol{E}^p - \frac{1}{2}\boldsymbol{\mu}^p \boldsymbol{E}^d \tag{C.1}$$

Here, again, $\boldsymbol{\mu}$ is vector of all of the individual induced dipoles, $\boldsymbol{\mu}_i$, and $\boldsymbol{C}$ is a super-matrix made up of individual matrices $\boldsymbol{C}_{ij} = (\alpha_j^{-1}\delta_{ij} - \boldsymbol{T}'_{ij})$ where $\boldsymbol{T}'_{ij}$ is the dipole-dipole interaction tensor between sites i and j and α_i is the polarizability of site i. $\boldsymbol{E}$ represents a vector of all of electric fields at each site, $\boldsymbol{E}_i$, and the superscripts p and d refer to whether this electric field was calculated with the p-scaled or d-scaled permanent-induced interaction tensor $\boldsymbol{T}$, $\boldsymbol{E}^p = \boldsymbol{T}^p\boldsymbol{M}$ or $\boldsymbol{E}^d = \boldsymbol{T}^d\boldsymbol{M}$, respectively. The two sets of induced dipoles, then, correspond to dipoles calculated with the p- or d-scaled permanent fields, $\boldsymbol{\mu}^p = \alpha(\boldsymbol{E}^p + \boldsymbol{T}'\boldsymbol{a}^p)$ or $\boldsymbol{\mu}^d = \alpha(\boldsymbol{E}^d + \boldsymbol{T}'\boldsymbol{a^d})$, respectively. It should be noted that in the completely converged SCF limit $\boldsymbol{\mu}^d = \boldsymbol{C}^{-1}\boldsymbol{E}^d$ and $\boldsymbol{\mu}^p = \boldsymbol{C}^{-1}\boldsymbol{E}^p$ and Eq. (C.1) reduces to the familiar $U^{pol} = -\frac{1}{2}\boldsymbol{\mu}^d\boldsymbol{E}^p$. In practice, then, Eq. (C.1) is used in place of Eq. (3.7).

One also need to evaluate the potential gradient so that we can determine forces. Including both the geometric term and dipole response terms yields Eq. (C.2).

$$\left.\frac{dU^{pol}(\boldsymbol{r^N}, \boldsymbol{a^N})}{d\boldsymbol{r}}\right|_{\boldsymbol{a}} = \frac{\partial U^{pol}(\boldsymbol{r^N}, \boldsymbol{a^N})}{\partial \boldsymbol{r}} + \frac{\partial U^{pol}(\boldsymbol{r^N}, \boldsymbol{a^N})}{\partial \boldsymbol{\mu}^p}\frac{\partial \boldsymbol{\mu}^p}{\partial \boldsymbol{r}} + \frac{\partial U^{pol}(\boldsymbol{r^N}, \boldsymbol{a^N})}{\partial \boldsymbol{\mu}^d}\frac{\partial \boldsymbol{\mu}^d}{\partial \boldsymbol{r}} \tag{C.2}$$

The full gradient can be evaluated since the real dipoles are evaluated using only a single electrostatic field calculation with the auxiliary dipoles as the initial guess whose position dependence is known, that is $\boldsymbol{\mu}^p = \alpha(\boldsymbol{E}^p + \boldsymbol{T}'\boldsymbol{a}^p)$ and $\boldsymbol{\mu}^d = \alpha(\boldsymbol{E}^d + \boldsymbol{T}'\boldsymbol{a^d})$. Substituting this into Eq. (C.2) gives the polarization gradient for the real degrees of freedom, Eq. (C.3).

$$\left.\frac{dU^{pol}(\boldsymbol{r}^N,\boldsymbol{a}^N)}{d\boldsymbol{r}}\right|_{\boldsymbol{a}} = -\frac{1}{2}[\,\boldsymbol{\mu}^d]^T\frac{\partial \boldsymbol{T}'}{\partial \boldsymbol{r}}\boldsymbol{\mu}^p + \frac{1}{2}[\,\boldsymbol{\mu}^d - \alpha\boldsymbol{T}\boldsymbol{\mu}^d - \alpha\boldsymbol{E}^d]^T\frac{\partial \boldsymbol{T}'}{\partial \boldsymbol{r}}\boldsymbol{a}^p$$
$$-\frac{1}{2}[\alpha\boldsymbol{T}\boldsymbol{\mu}^d + \alpha\boldsymbol{E}^d]^T\frac{\partial \boldsymbol{T}^p\boldsymbol{M}}{\partial \boldsymbol{r}} + \frac{1}{2}[\,\boldsymbol{\mu}^p - \alpha\boldsymbol{T}\boldsymbol{\mu}^p - \alpha\boldsymbol{E}^p]^T\frac{\partial \boldsymbol{T}'}{\partial \boldsymbol{r}}\boldsymbol{a}^d - \frac{1}{2}[\alpha\boldsymbol{T}\boldsymbol{\mu}^p + \alpha\boldsymbol{E}^p]^T\frac{\partial \boldsymbol{T}^d\boldsymbol{M}}{\partial \boldsymbol{r}} \tag{C.3}$$

Again, in the SCF limit, ($\boldsymbol{a} \rightarrow \boldsymbol{\mu} \rightarrow \boldsymbol{\mu}_{SCF}$), Eq. (C.3) reduces to the familiar form from the AMOEBA potential and in the limit that p-scaling and d-scaling are identical and Eq. (3.19) is recovered.

Appendix D: Additional iEL/0-SCF Validation Data

Table D.1: *Comparison of the diffusion constants using the iEL/0-SCF method and a standard SCF approach. Shown for pure water and a 4.66 M $MgCl_2$ solution. Starting points for NVE simulations were taken from NVT simulations at 298.0 K and then simulated with no thermostatting or auxiliary thermostatting and with time steps of 0.5 fs for both methods. Mean squared displacement of atoms over 100 ps were fit to determine the diffusion coefficients and then averaged over several independent NVE simulations. The SCF method used a convergence of 10^{-6} RMS Debye.*

System	iEL/0-SCF	SCF
Pure Water		
D_{H2O} (cm^2/s)	1.90 +/- 0.13	2.02 +/- 0.20
4.66 M $MgCl_2$ solution		
D_{H2O} (cm^2/s)	0.54 +/- 0.08	0.56 +/- 0.08
D_{Mg2+} (cm^2/s)	0.060 +/- 0.035	0.043 +/- 0.028
D_{Cl-} (cm^2/s)	0.090 +/- 0.034	0.086 +/- 0.032

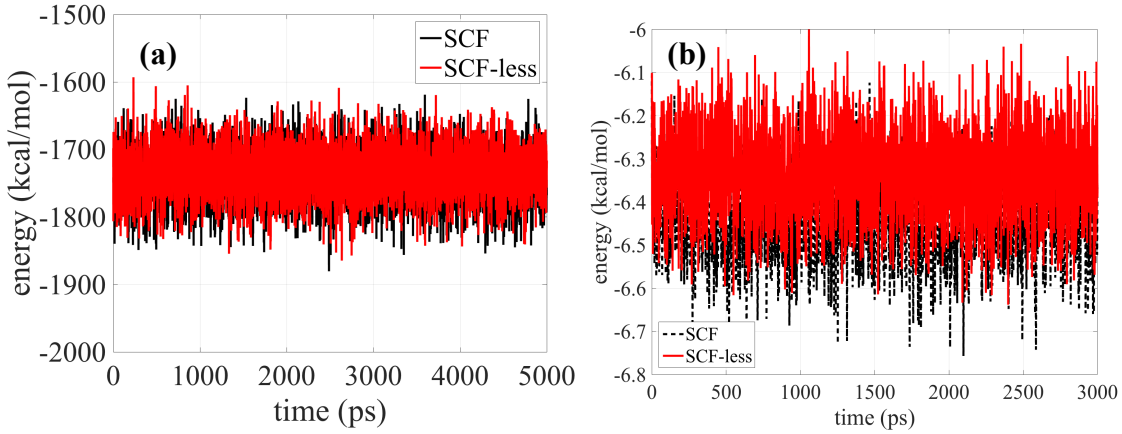


Figure D.1: *Comparison of ensemble average energy properties using the iEL/0-SCF and standard SCF methods. The average energy and fluctuations for the real degrees of freedom in the NVT ensemble of (a) glycine dipeptide using the iEL/0-SCF approach (red) and standard SCF approach (black) and (b) water using the iEL/0-SCF method (red) and standard SCF with RESPA (black); note the systematic offset for the RESPA approach). For iEL/0-SCF the γ value was set to 0.9, while the SCF solution used a PCG method with a convergence threshold of 10^{-5} RMS Debye. All simulations use a time step of 1.0 fs.*

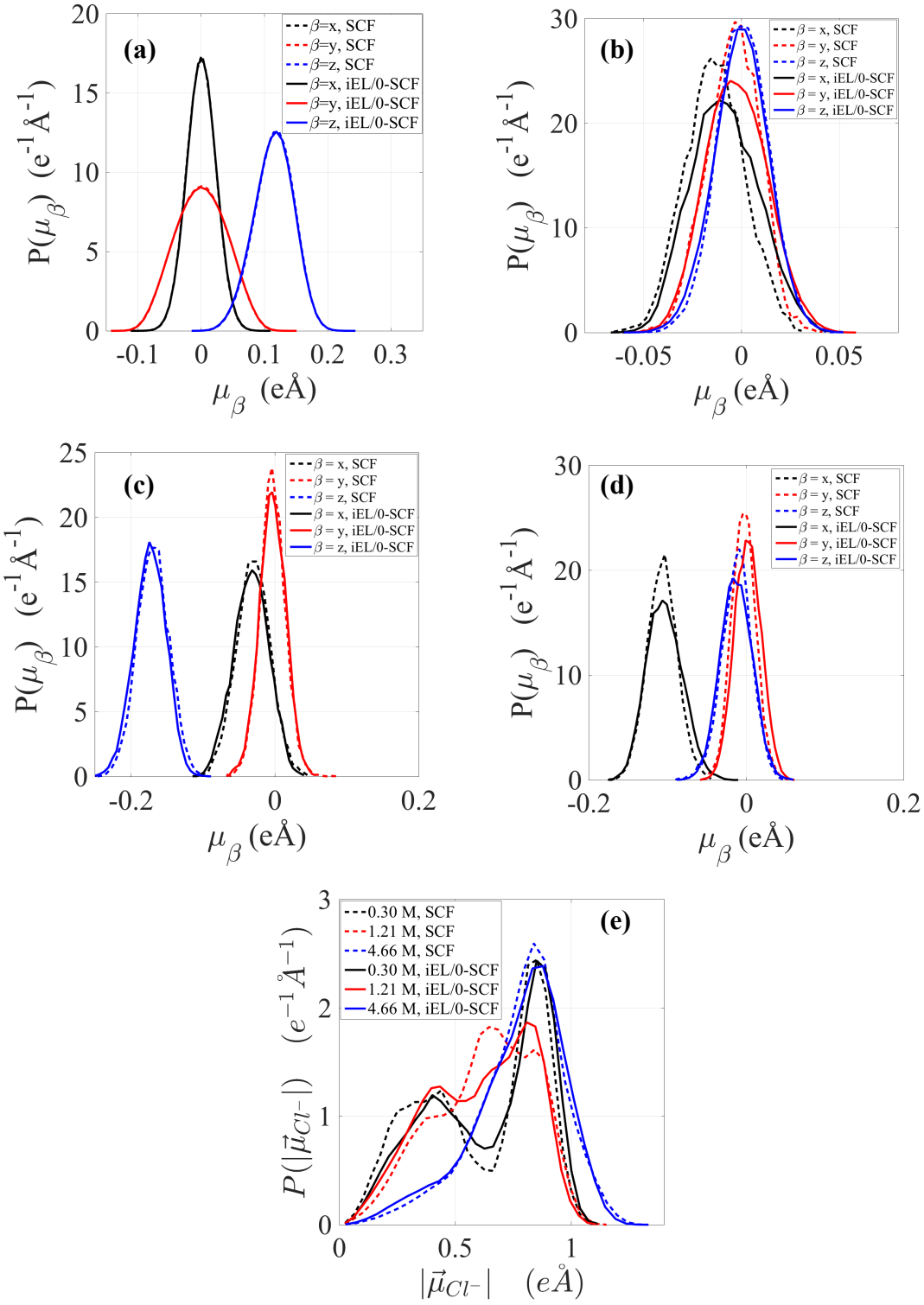


Figure D.2: *Comparison of ensemble averages of the induced dipole probability distributions using the iEL/0-SCF and standard SCF methods. Induced dipole probability distributions for (a) oxygen in bulk water; the glycine (b) nitrogen, (c) α-carbon, and (d) carbonyl carbon; and (e) Cl^- for $MgCl_2$ salt concentrations for a standard SCF method (dashed) and my iEL/0-SCF method (solid).*

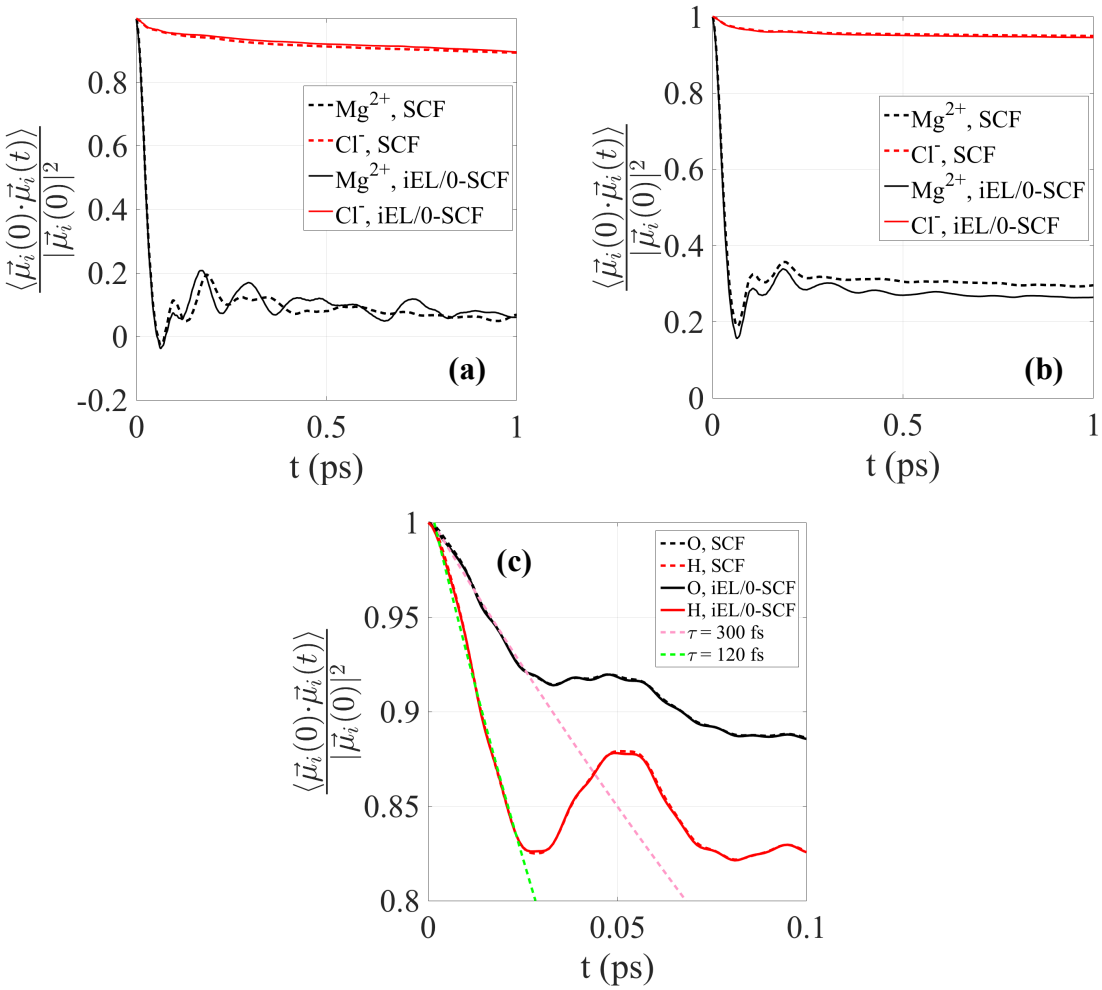
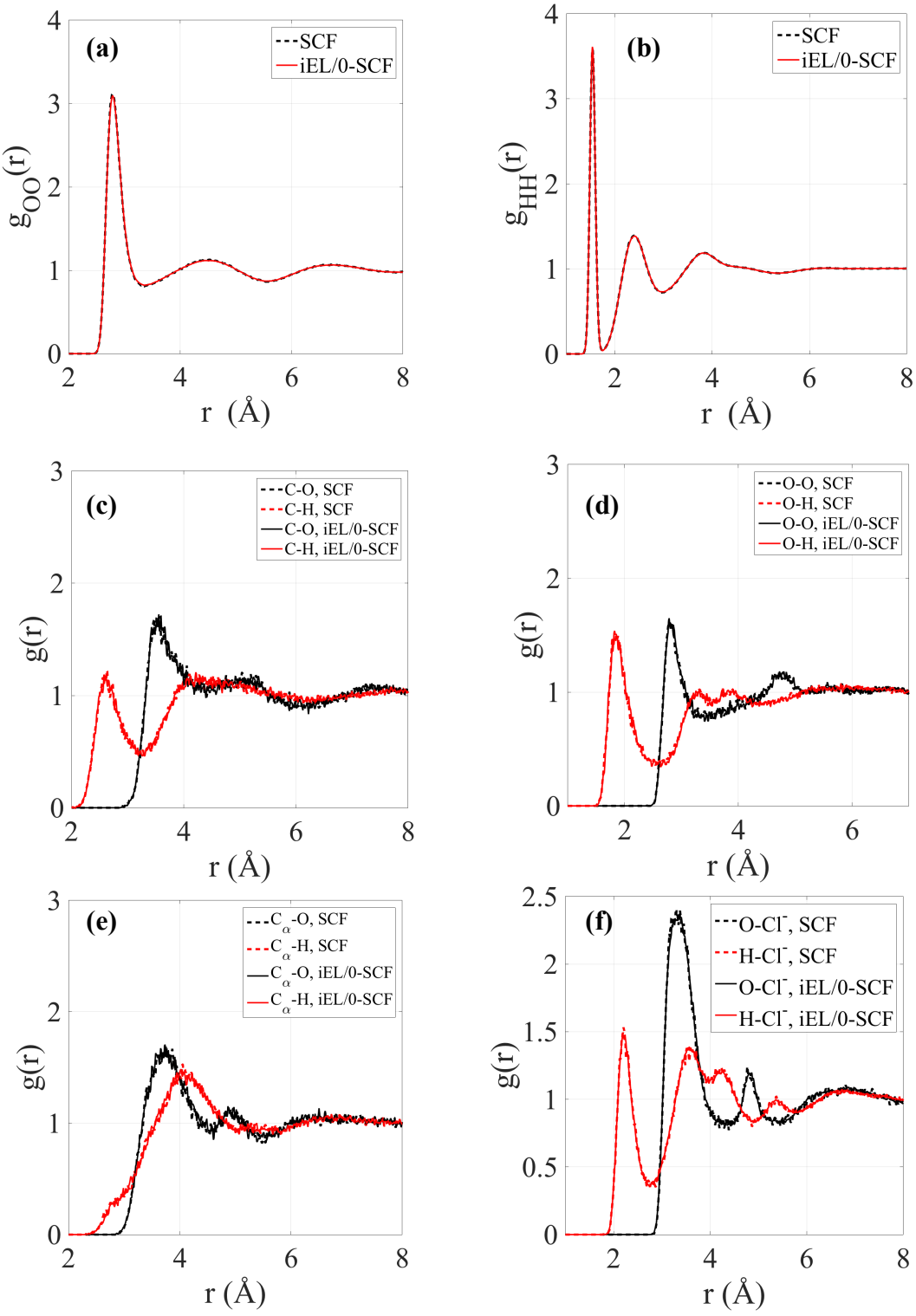


Figure D.3: *Comparison of ensemble average dynamical properties using the iEL/0-SCF and standard SCF methods. Autocorrelation function of real dipole for Mg^{2+} and Cl^- for (a) 0.30 M and (b) 4.66 M salt concentrations for a standard SCF method (dashed) and our IEL/0-SCF method (solid). Time scale of induced dipole moment decay for oxygen and hydrogen in water (c) when fit with a single exponential, $Aexp\left(-\frac{t}{\tau}\right)$, when fit to the short time scale (< 0.02 ps). For water the induced dipole distributions used an internal coordinate frame; see[17] for details on the internal coordinates. All data is taken from 30 ps trajectories in the NVT ensemble at 298.0 K.*


(a)
SCF
iEL/0-SCF
$g_{OO}(r)$
r (Å)
(b)
SCF
iEL/0-SCF
$g_{HH}(r)$
r (Å)
(c)
C-O, SCF
C-H, SCF
C-O, iEL/0-SCF
C-H, iEL/0-SCF
g(r)
r (Å)
(d)
O-O, SCF
O-H, SCF
O-O, iEL/0-SCF
O-H, iEL/0-SCF
g(r)
r (Å)
(e)
C_α-O, SCF
C_α-H, SCF
C_α-O, iEL/0-SCF
C_α-H, iEL/0-SCF
g(r)
r (Å)
(f)
O-Cl$^-$, SCF
H-Cl$^-$, SCF
O-Cl$^-$, iEL/0-SCF
H-Cl$^-$, iEL/0-SCF
g(r)
r (Å)

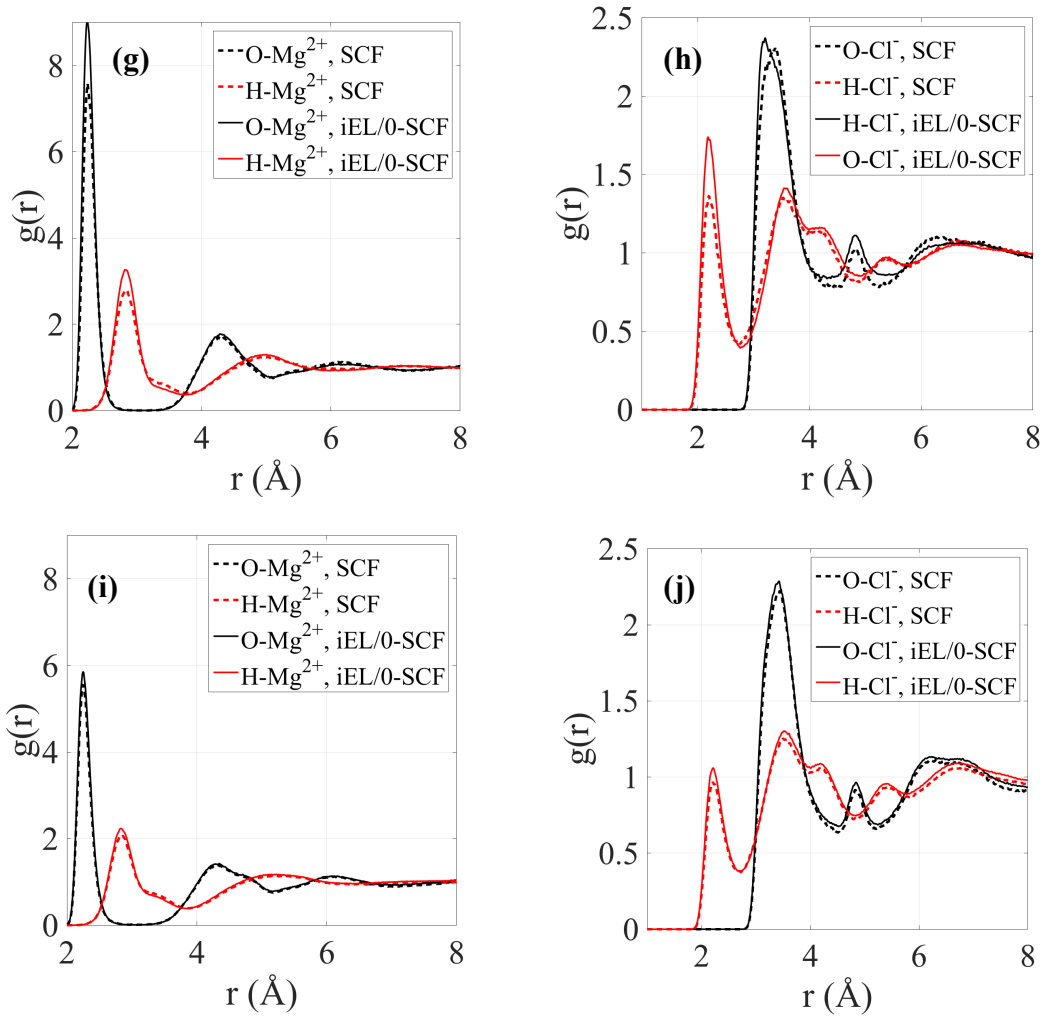


Figure D.4: *Comparison of ensemble average structural properties using the IEL/0-SCF and standard SCF methods. (a) g_{OO} and (b) g_{HH} for bulk water; water oxygen and hydrogen correlations with the backbone (c) carbonyl carbon, (d) carbonyl oxygen, and (e) α-carbon of glycine dipeptide; water oxygen and hydrogen correlations with (f) Cl^- for 0.30 M salt concentration, and (g) Mg^{2+} and (h) Cl^- for 1.21 M salt concentration, and (i) Mg^{2+} and (j) Cl^- for 4.66 M salt concentration, for a standard SCF method (black dashed) and our IEL/0-SCF method (red solid). All data is based on 3.0 ns trajectories in the NVT ensemble at 298.0 K.*

Appendix E: Additional iEL/0-SCF Drude Validation Data

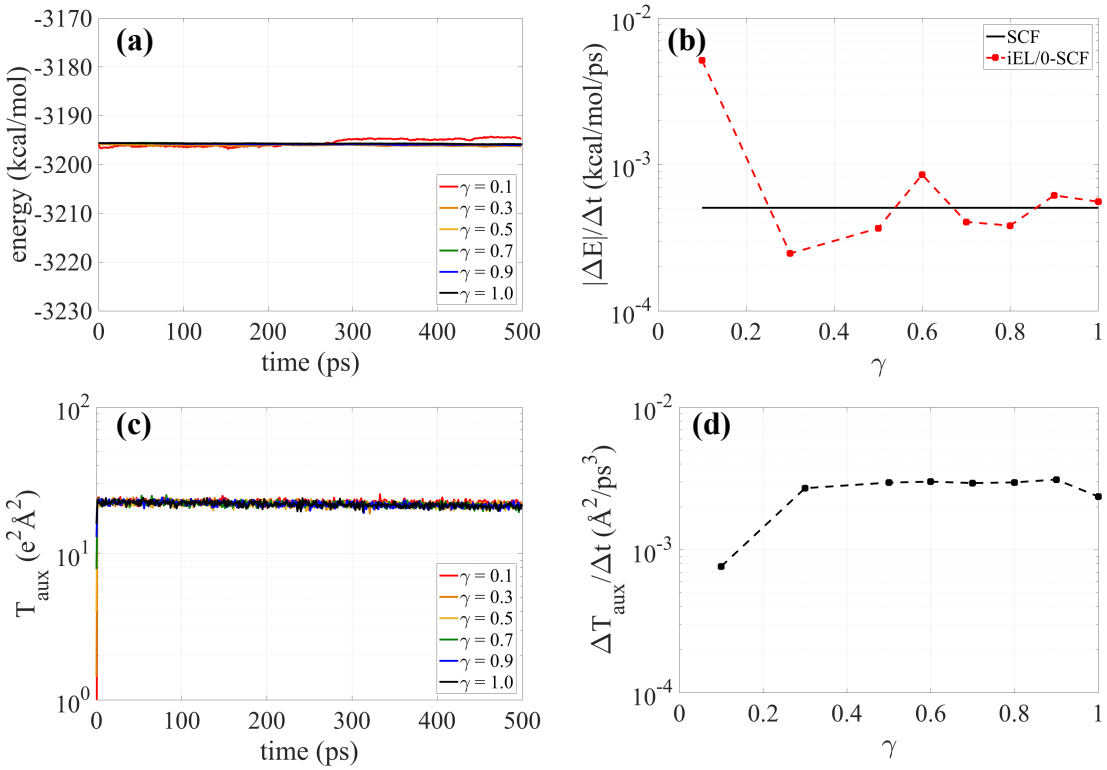


Figure E.1: *NVE energy properties of the iEL/0SCF method for the Drude PSPC model as a function of mixing parameter γ. All simulations were performed with a time step of 0.5 fs on a test system of 512 water molecules with no thermostatting. (a) Total real system energy (sum of the atomic kinetic and potentials energies) for iEL/0-SCF; (b) absolute real energy drift; (c) auxiliary degrees of freedom pseudo temperature; (d) drift of the auxiliary variables pseudo temperature.*

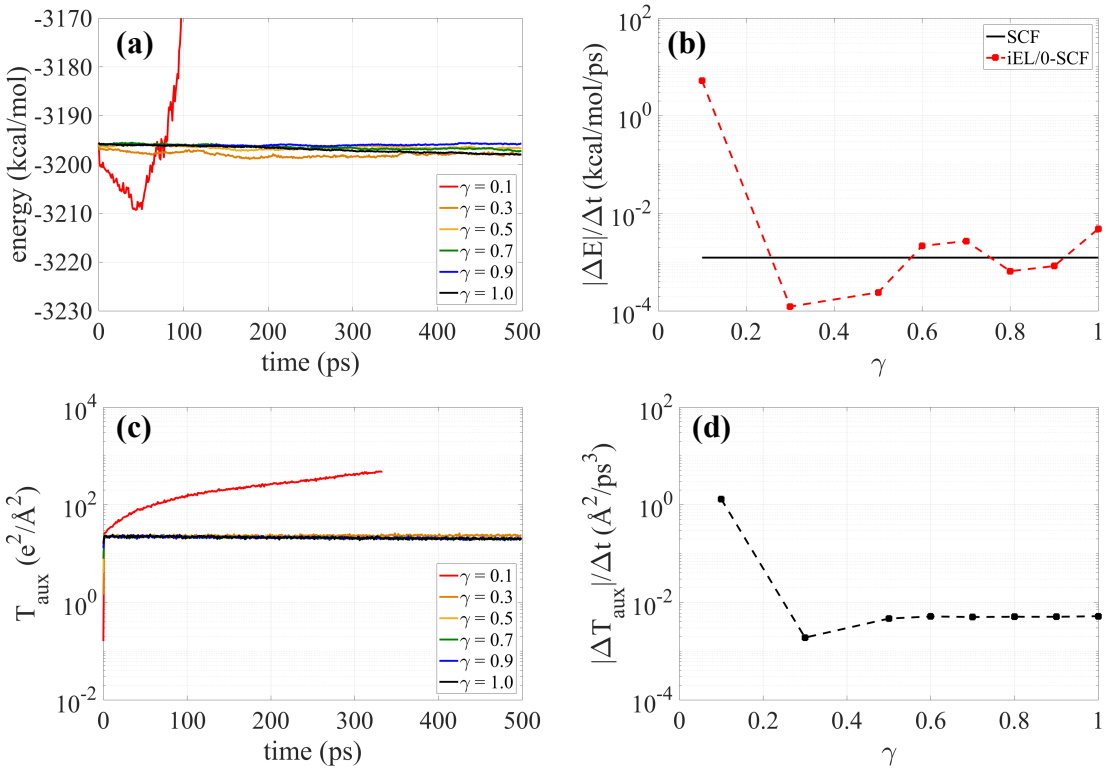


Figure E.2: *NVE energy properties of the iEL/0SCF method for the Drude PSPC model as a function of mixing parameter γ. All simulations were performed with a time step of 1.0 fs on a test system of 512 water molecules with no thermostatting. (a) Total real system energy (sum of the atomic kinetic and potentials energies) for iEL/0-SCF; (b) absolute real energy drift; (c) auxiliary degrees of freedom pseudo temperature; (d) drift of the auxiliary variables pseudo temperature.*

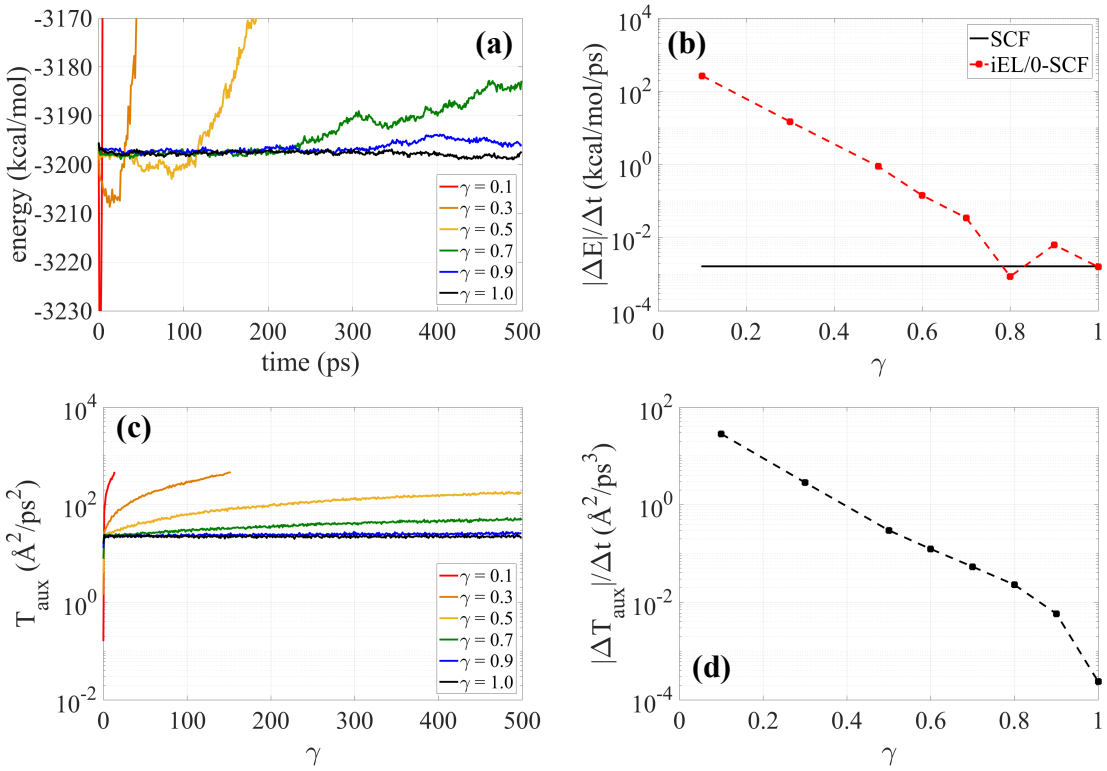


Figure E.3: *NVE energy properties of the iEL/0SCF method for the Drude PSPC model as a function of mixing parameter γ. All simulations were performed with a time step of 2.0 fs on a test system of 512 water molecules with no thermostatting for 0.5 ns. (a) Total real system energy (sum of the atomic kinetic and potentials energies) for iEL/0-SCF; (b) absolute real energy drift; (c) auxiliary variables pseudo temperature; (d) drift of the auxiliary degrees of freedom pseudo temperature.*

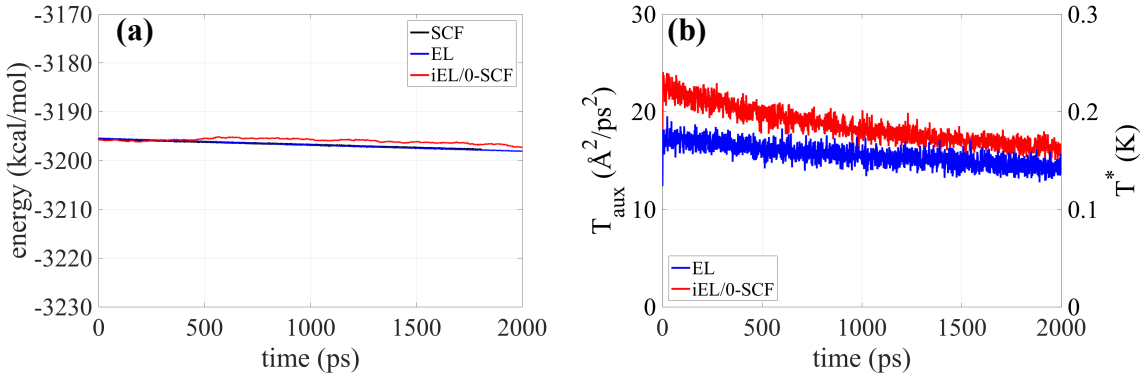


Figure E.4: *NVE energy properties of the iEL/0SCF and EL(T,T*=0) method for the Drude PSPC model compared to the SCF reference. All simulations were performed with a time step of 1.0 fs on a test system of 512 water molecules with no thermostats. (a) Total real system energy for SCF (black), iEL/0-SCF (red), and EL (blue) (b) auxiliary variables pseudo temperature for iEL/0-SCF (red, left axis) and temperature of the Drude-parent atom motion of the EL method (blue, right axis).*

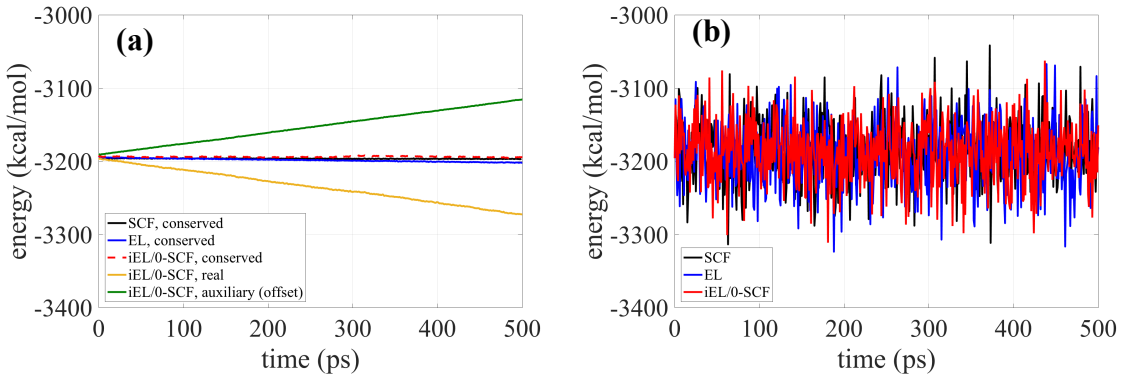


Figure E.5: *NVT energy properties of the SCF, EL, and iEL/0SCF method for the Drude PSPC model using a 2.0 fs time step. All simulations were all performed at 298.0 K. For iEL/0-SCF* $\gamma = 1.0$. *(a) For iEL/0-SCF: the real system energy (yellow, sum of the atomic kinetic and potentials energies plus the kinetic and potential energies associated to the thermostats controlling the atomic temperature), the auxiliary energy (green, sum of the auxiliary kinetic and potential energies plus the auxiliary thermostat and potential energies and artificially shifted downward by the time average of the total energy to be on the same scale) and their sum which is the conserved quantity (red). The corresponding conserved energy for SCF at a convergence of* 10^{-6} *RMS kcal/mol/Å (black) and that of the EL method with the minimum stable Drude mass, 0.8 g/mol, (blue). (b) The system energy is the sum of the kinetic and potentials energies of the atomic and Drude particles. The auxiliary mass used for the conserved iEL/0-SCF quantity is the slope value of the auxiliary energy given in (a), which for* $\gamma = 1.0$ *is 0.0017 g/mol at this time step.*

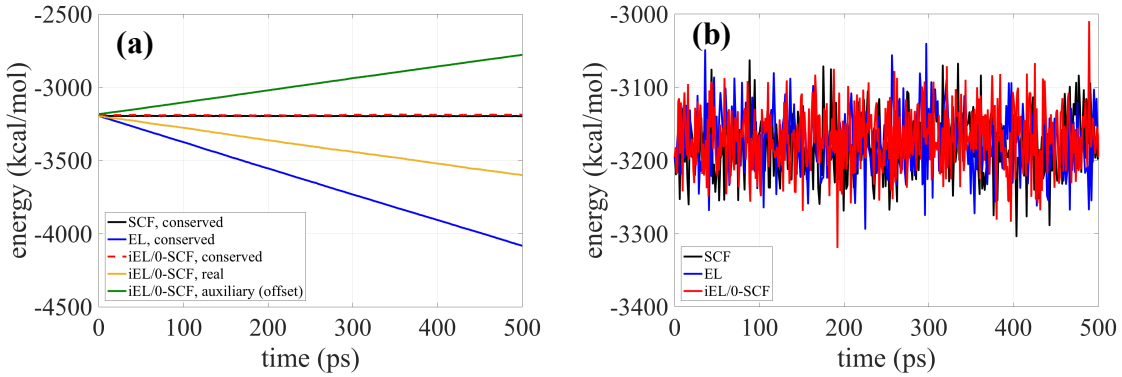


Figure E.6: *NVT energy properties of the SCF, EL, and iEL/0SCF method for the Drude PSPC model using a 3.0 fs time step. (a) For iEL/0-SCF: the real system energy (yellow), the auxiliary energy (green), and their sum which is the conserved quantity (red). The corresponding conserved energy for SCF at a convergence of* 10^{-6} *RMS kcal/mol/Å (black) and that of the EL method with the minimum stable Drude mass, 2.0 g/mol, (blue). (b) The system energy is the sum of the atomic kinetic and potentials energies. The auxiliary mass used for the conserved iEL/0-SCF quantity is the slope value of the auxiliary energy given in (a), which for* $\gamma = 1.0$ *is 0.0022 g/mol at this time step. Refer to Figure E.5 for further details.*

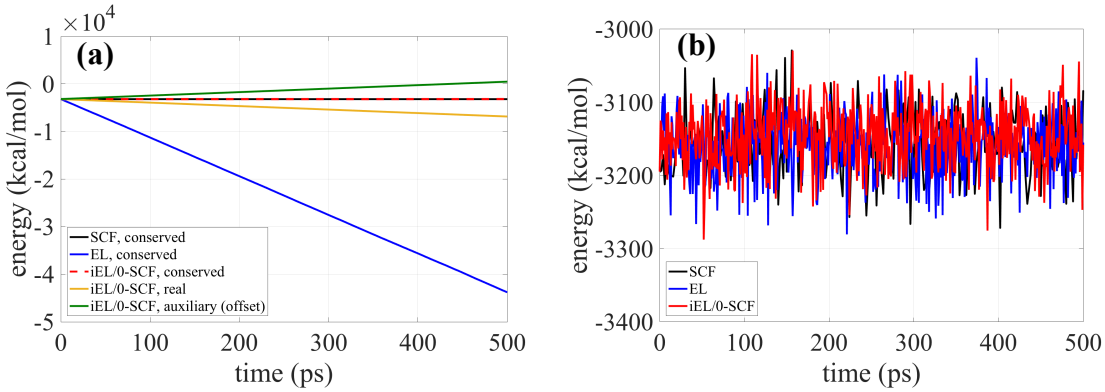


Figure E.7: *NVT energy properties of the SCF, EL, and iEL/0SCF method for the Drude PSPC model using a 4.0 fs time step. (a) For iEL/0-SCF: the real system energy (yellow), the auxiliary energy (green), and their sum which is the conserved quantity (red). The corresponding conserved energy for SCF at a convergence of 10^{-6} RMS kcal/mol/Å (black) and that of the EL method with the minimum stable Drude mass, 4.0 g/mol, (blue). (b) The system energy is the sum of the atomic kinetic and potentials energies. The auxiliary mass used for the conserved iEL/0-SCF quantity is the slope value of the auxiliary energy given in (a), which for $\gamma = 1.0$ is 0.0072 g/mol at this time step. Refer to Figure E.5 for further details.*

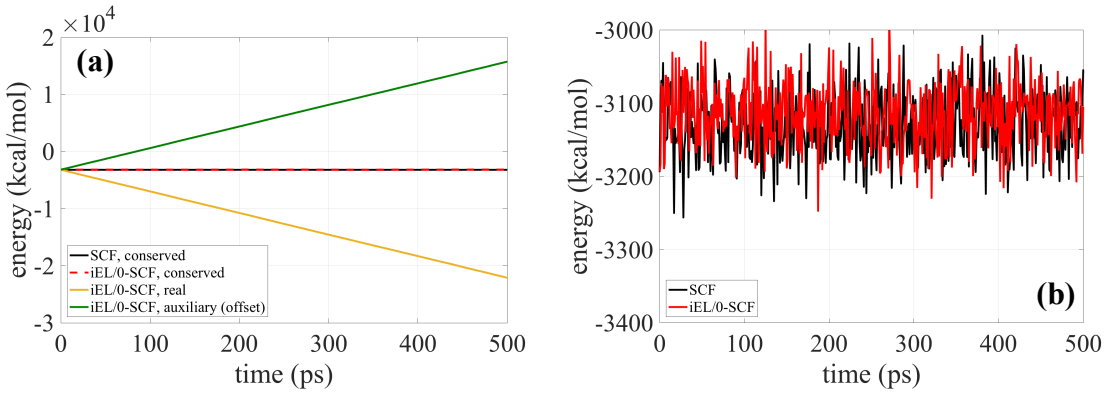


Figure E.8: *NVT energy properties of the SCF and iEL/0SCF method for the Drude PSPC model using a 5.0 fs time step. (a) For iEL/0-SCF: the real system energy (yellow), the auxiliary energy (green), and their sum which is the conserved quantity (red). The corresponding conserved energy for SCF at a convergence of 10^{-6} RMS kcal/mol/Å (black). (b) The system energy is the sum of the atomic kinetic and potentials energies. The auxiliary mass used for the conserved iEL/0-SCF quantity is the slope value of the auxiliary energy given in (a), which for $\gamma = 1.0$ is 0.0097 g/mol at this time step. EL is unstable at this time step for all values of Drude mass. Refer to Figure E.5 for further details.*

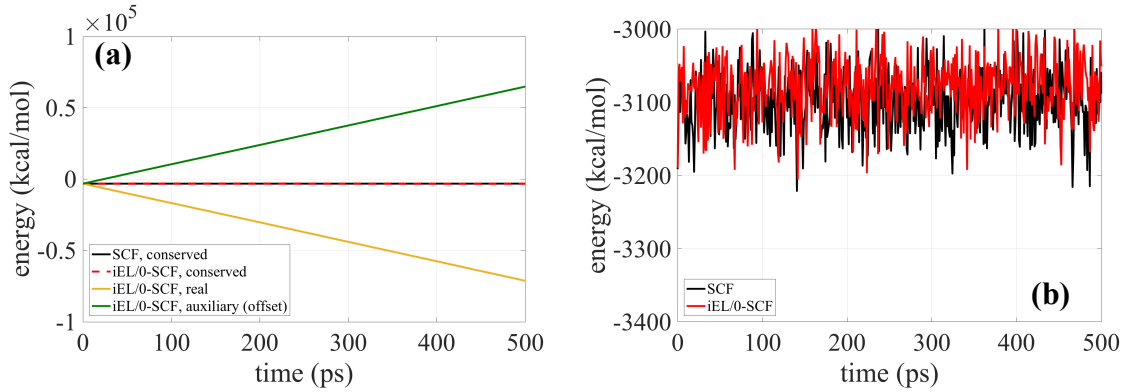
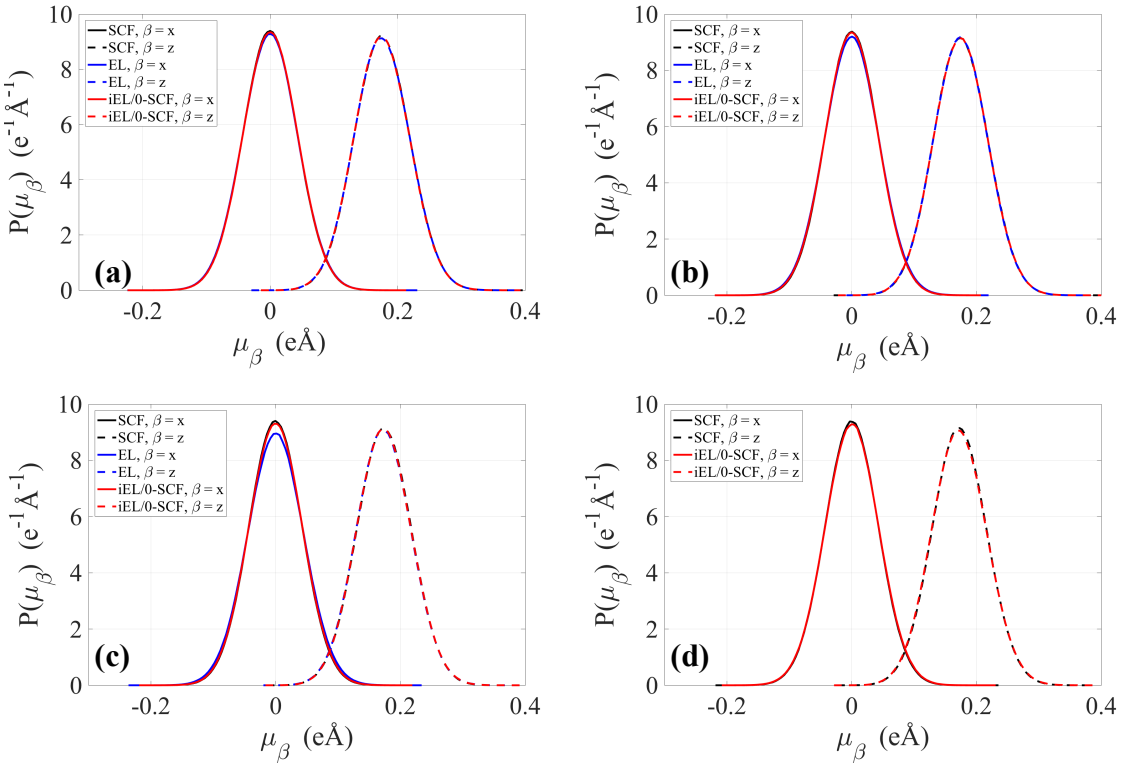


Figure E.9: *NVT energy properties of the SCF and iEL/0SCF method for the Drude PSPC model using a 6.0 fs time step. (a) For iEL/0-SCF: the real system energy (yellow), the auxiliary energy (green), and their sum which is the conserved quantity (red). The corresponding conserved energy for SCF at a convergence of 10^{-6} RMS kcal/mol/Å (black). (b) The system energy is the sum of the atomic kinetic and potentials energies. The auxiliary mass used for the conserved iEL/0-SCF quantity is the slope value of the auxiliary energy given in (a), which for $\gamma = 1.0$ is 0.0128 g/mol at this time step. EL is unstable at this time step for all values of Drude mass. Refer to Figure E.5 for further details.*



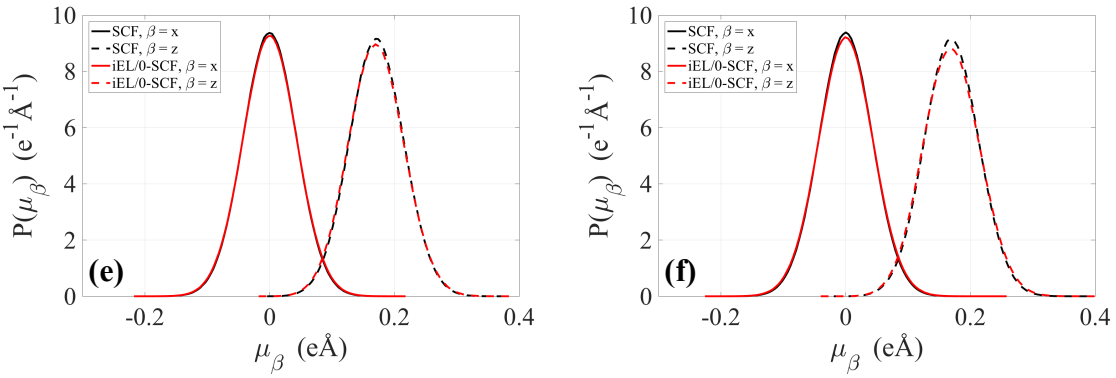
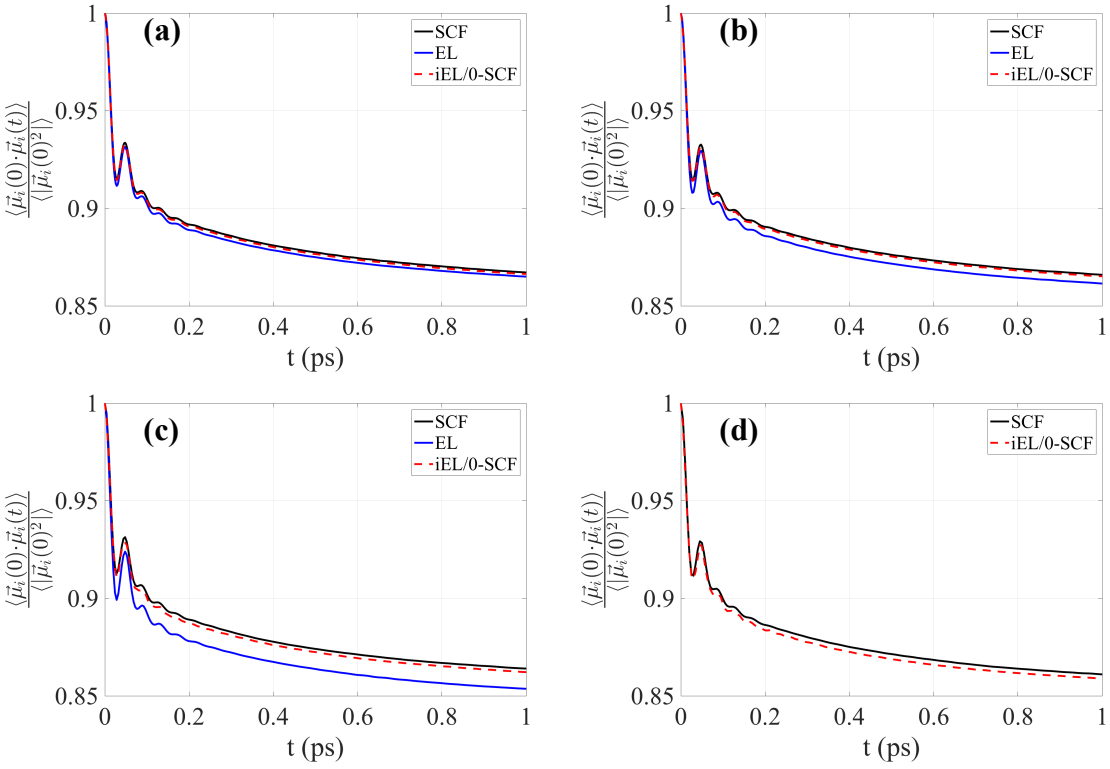


Figure E.10: *The oxygen induced dipole, $\boldsymbol{\mu}_i = q_{D,i}(\mathbf{r}_{D,i} - \mathbf{r}_i)$, probability density distributions for the PSPC model using the SCF, EL(T,T*) and iEL/0-SCF methods at different time steps. (a) 2.0 fs, (b) 3.0 fs, (c) 4.0 fs, (d) 5.0 fs, (e) 6.0 fs, and (f) 7.0 fs. All simulations were performed in the NVT ensemble at 298.0 K. All calculations presented in this figure use an internal coordinate system where the z-direction is given by the H-O-H bisector, the y-direction is out of the H-O-H plane, and the x-direction is orthogonal to z and y (see[31] for details).*



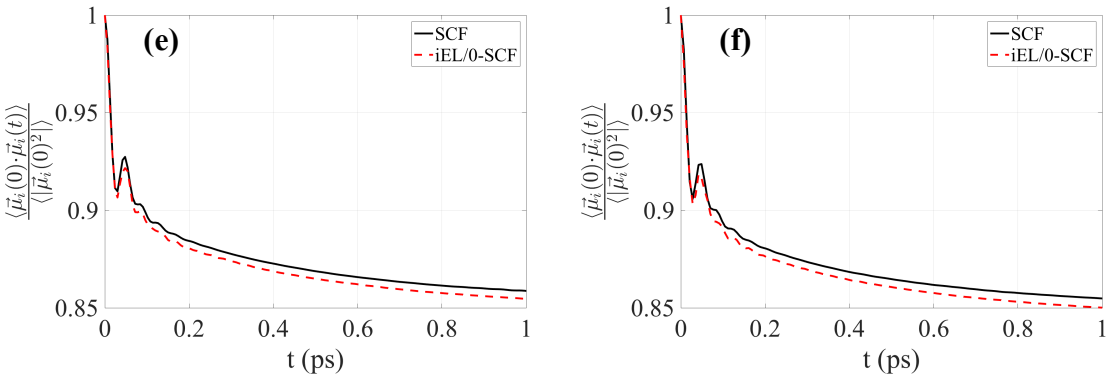


Figure E.11: *The oxygen induced dipole, $\boldsymbol{\mu}_i = q_{D,i}(\boldsymbol{r}_{D,i} - \boldsymbol{r}_i)$, time autocorrelation for the PSPC model using the SCF, EL(T,T*) and iEL/0-SCF methods at different time steps. (a) 2.0 fs, (b) 3.0 fs, (c) 4.0 fs, (d) 5.0 fs, (e) 6.0 fs, and (f) 7.0 fs. All simulations were performed in the NVT ensemble at 298.0 K. All calculations presented in this figure use an internal coordinate system where the z-direction is given by the H-O-H bisector, the y-direction is out of the H-O-H plane, and the x-direction is orthogonal to z and y (see[31] for details).*

Appendix F: General-Order Shadow Potentials for Induced Dipole Polarization

While the dissipation and integrations schemes laid out in the Chapter 5 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 I 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 Chapter 5 (section 5.2) so I will focus on brevity here.

The potential energy surface for an inducible dipole model is given by the constrained minimization of Eq. (F.1).

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

In Eq. (F.1), $\boldsymbol{\mu}$ 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 $\boldsymbol{C} = \boldsymbol{\alpha}^{-1} - \boldsymbol{T}'$ where $\boldsymbol{\alpha}$ is a diagonal matrix with the values of atomic polarizability for each atom, α_i, on the diagonal and $\boldsymbol{T}'$ is the dipole-dipole interaction matrix. The ground state (self-consistent field) solution of Eq. (F.1) is the set of induced dipoles that minimizes the polarization energy, which is given in Eq. (F.2).

$$\frac{\partial U^{pol}}{\partial \boldsymbol{\mu}} = 0 = -\boldsymbol{E} + \boldsymbol{C}\boldsymbol{\mu} \quad \Longrightarrow \quad \boldsymbol{\mu}_{min} = -\boldsymbol{C}^{-1}\boldsymbol{E} \tag{F.2}$$

As before we can now introduce a dynamically driven auxiliary matrix, $\boldsymbol{X}$, which should follow the behavior of $\boldsymbol{C}^{-1}$ and build a shadow potential functional around this auxiliary matrix as shown in Eq. (F.3).

$$U^{pol}_{shadow}(\boldsymbol{r}, \boldsymbol{X}, \boldsymbol{\mu}) = \frac{1}{2}\boldsymbol{\mu}^{\boldsymbol{T}}[f(\boldsymbol{X})]^{-1}\boldsymbol{\mu} - \boldsymbol{E}^{T}\boldsymbol{\mu} \tag{F.3}$$

Here $f(\boldsymbol{X})$ is the same as given in section 5.2 and is reproduced in Eq. (F.4). This form, again, will produce a better estimate of $\boldsymbol{C}^{-1}$ from $\boldsymbol{X}$.

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

One can also show that the shadow potential functional that is constructed in Eq. (F.3) is necessarily minimized for the set of induced dipoles given in Eq. (F.5), which is simply the function of Eq. (F.4) dotted into the permanent electrostatic field.

$$\begin{aligned} \frac{\partial U^{pol}_{shadow}(\boldsymbol{r}^{N}, \boldsymbol{X}, \boldsymbol{\mu})}{\partial \boldsymbol{\mu}} = 0 = \left[\boldsymbol{C}^{-1}[\boldsymbol{I} - (\boldsymbol{I} - \boldsymbol{C}\boldsymbol{X})^{m}]\right]^{-1}\boldsymbol{\mu} - \boldsymbol{E} \quad \Longrightarrow \\ \boldsymbol{\mu}_{min} = \boldsymbol{C}^{-1}[\boldsymbol{I} - (\boldsymbol{I} - \boldsymbol{C}\boldsymbol{X})^{m}]\boldsymbol{E} = f(\boldsymbol{X})\boldsymbol{E} \end{aligned} \tag{F.5}$$

Making the substitution of Eq. (F.5) into Eq. (F.3) we can now build a Lagrangian with this shadow potential and $\boldsymbol{X}$ as an additional dynamical degree of freedom, shown in Eq. (F.6).

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

In Eq. (F.6), $U(\boldsymbol{r})$ 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. (F.6) we obtain Eq. (F.7).

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

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

From Eq. (F.7) a classical adiabatic separation between particle and auxiliary motion is assumed 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. (F.8), where the derivate $\frac{\partial f(\boldsymbol{X})}{\partial \boldsymbol{r}_i}$ has been evaluated explicitly and Eq. (F.4) is used for $f(\boldsymbol{X})$.

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

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

I now choose $\widetilde{\boldsymbol{C}}^{-1} = f(\boldsymbol{X})$, as shown in the second right hand side of Eq. (F.8b), 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. (F.8) can be integrated and dissipation can be introduced as laid out in section 5.2, as before. One may be concerned that $\boldsymbol{C}^{-1}$ appears explicitly in Eq. (F.8), however, for some finite m the expansion in m and subsequent algebraic simplification removes any $\boldsymbol{C}^{-1}$ terms. Here now is a shadow potential for dipole polarization that can match the base potential to an arbitrary degree, yet is guaranteed to be minimized via Eq. (F.5) at any point with no iteration.

Appendix G: Optimization of Higher-Order Integration 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 $\boldsymbol{C}^{-1}$ one can describe $\widetilde{\boldsymbol{C}}^{-1}$ to be given through an approximate linearized optimization process $\boldsymbol{\Gamma}$ acting on $\boldsymbol{X}$, $\widetilde{\boldsymbol{C}}^{-1} = \boldsymbol{\Gamma X}$ [73-74]. For a convergent optimization one expects the eigenvalues of $\boldsymbol{\Gamma}$ to be $|\gamma| < 1$. By replacing $\boldsymbol{\Gamma}$ by its maximum eigenvalue, γ, one can examine the stability of the integration under incomplete convergence to the ground state reference potential. The equation of motion in Eq. (5.8b) of the text then becomes Eq. (G.1)

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

We can now examine the behavior of the integration 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. (G.1), using our combined higher order dissipation and integration schemes, given by Eq. (5.11) in the Chapter 5. This integration can be described as a mapping of $\boldsymbol{X}$ and its velocity $\dot{\boldsymbol{X}}$ at one time step to the next. Such a mapping is given by Eq. (G.2)

$$\begin{bmatrix} \dot{\boldsymbol{X}}(t+\Delta t) \\ \boldsymbol{X}(t+\Delta t) \\ \boldsymbol{X}(t) \\ \vdots \\ \boldsymbol{X}(t-(K+1)\Delta t) \end{bmatrix} = \boldsymbol{T} \begin{bmatrix} \dot{\boldsymbol{X}}(t) \\ \boldsymbol{X}(t) \\ \boldsymbol{X}(t-\Delta t) \\ \vdots \\ \boldsymbol{X}(t-K\Delta t) \end{bmatrix} \tag{G.2}$$

In Eq. (G.2) $\boldsymbol{T}$ is a matrix that describes the mapping (integration) process and is a function of a_l, b_l, κ, α, and γ (see Appendix H for more specifics). In Eq. (G.2) an implicit variable substitution $\dot{\boldsymbol{X}} \to \Delta t\dot{\boldsymbol{X}}$ is used 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 one 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 G.1 shows the analysis of the $\boldsymbol{T}$ matrix for the specific case of $K = 3$ and $L = 2$ (2nd-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 -1.0 to 1.0 is tested and the maximum spectral radius of $\boldsymbol{T}$ from that range of γ is recorded, which is shown in Figure G.1a. 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 G.1c. Using this set of κ and corresponding α_{max} values one can also look at their specific spectral radius behavior as a function of γ, shown in Figure G.1b. The lower the spectral radius is the more

dissipation that particular combination of α_{max} and κ achieves. The maximum dissipation is defined as the spectral radius at $\gamma = 0$, which is called $[\rho(\boldsymbol{T})]_{min}$ and is plotted as a function of κ in Figure G.1c, as well.

While κ needs to be maximized, the dissipation, $[\rho(\boldsymbol{T})]_{min}$, needs to maximized, as well (lower values of $[\rho(\boldsymbol{T})]_{min}$ correspond to more dissipation). Figure G.1c 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. I choose this point to be the optimal combination of large κ and maximum dissipation. In this specific case I 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 G.1 is repeated for a wide combination of K and L values and the optimal κ and α for each is given in Table 5.4.

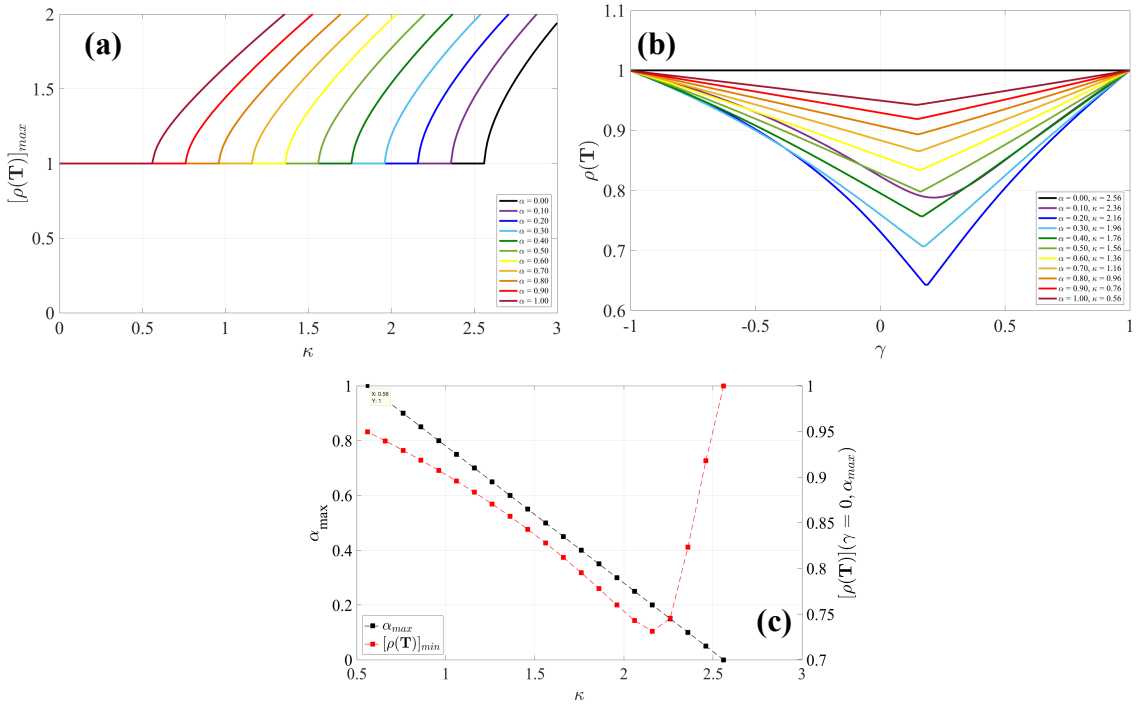


***Figure G.1**: Stability analysis of the integration schemes. Analysis of the **T** matrix (Eq. G.2) for the specific case of $K = 3$ and $L = 2$. (a) The maximum spectral radius of **T**, $[\rho(\boldsymbol{T})]_{max}$, as a function of κ for a range of α values. (b) The point at which $[\rho(\boldsymbol{T})]_{max}$ becomes greater than 1.0 defines a pair of κ and α_{max} and for those sets of points we plot the spectral radius, $\rho(\boldsymbol{T})$, as a function of γ.(c) We define the maximum dissipation as the spectral radius at $\gamma = 0$ and call this $[\rho(\boldsymbol{T})]_{min}$, which is a function of κ and α_{max}. Both α_{max} (black) and $[\rho(\boldsymbol{T})]_{min}$ (red) are given as function of κ.*

Appendix H: Integration Tensors for Spectral Radius Analysis

As discussed in Appendix G, the integration of Eq. (G.1) can be written in the form of Eq. (G.2), where the details of the $\boldsymbol{T}$ matrix depend on the specific integrator being used. For the general class of integrators described by Eq. (5.11) one can build $\boldsymbol{T}$, the integration that moves from t to $t + \Delta t$, as a successive application of velocity and position updates for each l^{th} intermediate update of an L order integrator, given in Eq. (H.1).

$$\boldsymbol{T} = \prod_{l=1}^{L} \boldsymbol{T}_l^X \boldsymbol{T}_l^V \tag{H.1}$$

In Eq. (H.1) $\boldsymbol{T}_l^X$ is the matrix giving the l^{th} position update and $\boldsymbol{T}_l^V$ is the matrix giving the l^{th} velocity update. For a 2nd-order integrator ($L = 2$) and a general dissipative order, K, the $\boldsymbol{T}_l^X$ and $\boldsymbol{T}_l^V$ matrices are given in Eq. (H.2).

$$\boldsymbol{T}_1^V = \begin{bmatrix} 1 & b_1\kappa(\gamma-1)+b_1c_0\alpha & b_1c_1\alpha & \dots & b_1c_{K-1}\alpha & b_1c_K\alpha & 0 & 0 & \dots & 0 & 0 \\ 0 & 1 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 & 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 1 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 1 \end{bmatrix} \tag{H.2a}$$

$$\boldsymbol{T}_1^X = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 1 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & 1 & 0 & 0 & \dots & 0 & 0 \\ a_1 & 1 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 & 1 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \ddots & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \ddots & 1 & 0 \end{bmatrix} \tag{H.2b}$$

$$\boldsymbol{T}_2^V = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 & 0 & b_2\kappa(\gamma-1)+b_2c_0\alpha & b_2c_1\alpha & \dots & b_2c_{K-1}\alpha & b_2c_K\alpha \\ 0 & 1 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 & 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 1 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 1 \end{bmatrix} \tag{H.2c}$$

$$\boldsymbol{T}_2^X = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 \\ a_2 & 0 & 0 & \dots & 0 & 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \ddots & 0 & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & \ddots & 1 & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 & 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 1 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 1 \end{bmatrix} \tag{H.2d}$$

For a 3rd-order integrator ($L = 3$) and a general dissipative order, K, the $\boldsymbol{T}_l^X$ and $\boldsymbol{T}_l^V$ matrices are given in Eq. (H.3).

$$\boldsymbol{T}_1^V = \begin{bmatrix} 1 & b_1\kappa(\gamma-1)+b_1c_0\alpha & b_1c_1\alpha & \dots & b_1c_{K-1}\alpha & b_1c_K\alpha & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 1 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & 1 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 & 1 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 & 0 & 1 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 1 & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 1 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 1 \end{bmatrix} \tag{H.3a}$$

$$\boldsymbol{T}_1^X = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 1 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & 1 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 \\ a_1 & 1 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 & 1 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \ddots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \ddots & 0 & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \ddots & 1 & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 1 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 1 \end{bmatrix} \tag{H.3b}$$

$$\boldsymbol{T}_2^V = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 & 0 & b_2\kappa(\gamma-1)+b_2c_0\alpha & b_2c_1\alpha & \dots & b_2c_{K-1}\alpha & b_2c_K\alpha & 0 & 0 & \dots & 0 & 0 \\ 0 & 1 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & 1 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 & 1 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 & 0 & 1 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 1 & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 1 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 1 \end{bmatrix} \tag{H.3c}$$

$$\boldsymbol{T}_2^X = \begin{bmatrix}
1 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 \\
0 & 1 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 \\
0 & 0 & 1 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & 0 & \dots & 1 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 \\
0 & 0 & 0 & \dots & 0 & 1 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 \\
0 & 0 & 0 & \dots & 0 & 0 & 1 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 \\
0 & 0 & 0 & \dots & 0 & 0 & 0 & 1 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 1 & 0 & 0 & 0 & \dots & 0 & 0 \\
0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 1 & 0 & 0 & \dots & 0 & 0 \\
a_2 & 0 & 0 & \dots & 0 & 0 & 1 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 \\
0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 1 & 0 & \dots & 0 & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\
0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \ddots & 0 & 0 \\
0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \ddots & 1 & 0
\end{bmatrix} \tag{H.3d}$$

$$\boldsymbol{T}_3^V = \begin{bmatrix}
1 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 & b_3\kappa(\gamma-1)+b_3c_0\alpha & b_3c_1\alpha & \dots & b_3c_{K-1}\alpha & b_3c_K\alpha \\
0 & 1 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 \\
0 & 0 & 1 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & 0 & \dots & 1 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 \\
0 & 0 & 0 & \dots & 0 & 1 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 \\
0 & 0 & 0 & \dots & 0 & 0 & 1 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 \\
0 & 0 & 0 & \dots & 0 & 0 & 0 & 1 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 1 & 0 & 0 & 0 & \dots & 0 & 0 \\
0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 1 & 0 & 0 & \dots & 0 & 0 \\
0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 1 & 0 & \dots & 0 & 0 \\
0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 1 & \dots & 0 & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 1 & 0 \\
0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 1
\end{bmatrix} \tag{H.3e}$$

$$\boldsymbol{T}_3^X = \begin{bmatrix}
1 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 \\
a_3 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 1 & 0 & \dots & 0 & 0 \\
0 & 1 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 \\
\vdots & \ddots & \ddots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & 0 & \ddots & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 \\
0 & 0 & 0 & \ddots & 1 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 \\
0 & 0 & 0 & \dots & 0 & 0 & 1 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 \\
0 & 0 & 0 & \dots & 0 & 0 & 0 & 1 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 1 & 0 & 0 & 0 & \dots & 0 & 0 \\
0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 1 & 0 & 0 & \dots & 0 & 0 \\
0 & 0 & 0 & \dots & 0 & 0 & 1 & 0 & \dots & 0 & 0 & 1 & 0 & \dots & 0 & 0 \\
0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 1 & \dots & 0 & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 1 & 0 \\
0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 1
\end{bmatrix} \tag{H.3f}$$

These $\boldsymbol{T}$ matrices are then analyzed as described in the section 5.2 and Appendix G. Higher order matrices can be constructed in a similar manner keeping in mind that L sets of K previous positions need to be stored for dissipation and every time the position is updated its corresponding history at integer multiples of full previous time steps must also be incremented in history. This incrementing process can be seen in the $\boldsymbol{T}_l^X$ as diagonal blocks of 1 that do not lie on the main diagonal.

Appendix I: Three Timescale RESPA with SI Integration

Translating the equations of motion, Eq. (6.3) and Eq. (6.6), into a multiple time step integration scheme requires some care and is discussed extensively by Tuckerman and colleagues[108-109]. A brief overview is presented here. The constructing of an integration algorithm begins by first examining the Liouville operator Eq. (6.3), which can be written as Eq. (I.1).

$$iL = iL_r + iL_v^s + iL_v^m + iL_v^l + iL_N + iL_{OU} \tag{I.1}$$

In Eq. (I.1) the velocity component, iL_v, is broken into short, medium, and long timescales. Specifically the components of Eq. (I.1) are given by Eq. (I.2), using the definitions presented in Chapter 6. The component iL_{OU} applies a Ornstein-Uhlenbeck stochastic process to every $v_{2,i,\alpha}^{(k)}$.

$$iL_r = \sum_{i,\alpha} v_{i,\alpha} \frac{\partial}{\partial r_{i,\alpha}} \tag{I.2a}$$

$$iL_v^s = \sum_{i,\alpha} \left[\left(\frac{F_{i,\alpha}^s(\boldsymbol{r})}{m_i} - \lambda_{i,\alpha}^s v_{i,\alpha} \right) \frac{\partial}{\partial v_{i,\alpha}} - \lambda_{i,\alpha}^s \sum_{k=1}^{L} v_{1,i,\alpha}^{(k)} \frac{\partial}{\partial v_{1,i,\alpha}^{(k)}} \right] \tag{I.2b}$$

$$iL_v^m = \sum_{i,\alpha} \left[\left(\frac{F_{i,\alpha}^m(\boldsymbol{r})}{m_i} - \lambda_{i,\alpha}^m v_{i,\alpha} \right) \frac{\partial}{\partial v_{i,\alpha}} - \lambda_{i,\alpha}^m \sum_{k=1}^{L} v_{1,i,\alpha}^{(k)} \frac{\partial}{\partial v_{1,i,\alpha}^{(k)}} \right] \tag{I.2c}$$

$$iL_v^l = \sum_{i,\alpha} \left[\left(\frac{F_{i,\alpha}^l(\boldsymbol{r})}{m_i} - \lambda_{i,\alpha}^l v_{i,\alpha} \right) \frac{\partial}{\partial v_{i,\alpha}} - \lambda_{i,\alpha}^l \sum_{k=1}^{L} v_{1,i,\alpha}^{(k)} \frac{\partial}{\partial v_{1,i,\alpha}^{(k)}} \right] \tag{I.2d}$$

$$\begin{aligned} iL_N = -\sum_{i,\alpha} [& \lambda_{i,\alpha}^N v_{i,\alpha} \frac{\partial}{\partial v_{i,\alpha}} + \lambda_{i,\alpha}^N \sum_{k=1}^{L} v_{1,i,\alpha}^{(k)} \frac{\partial}{\partial v_{1,i,\alpha}^{(k)}} \\ & + \sum_{k=1}^{L} v_{2,i,\alpha}^{(k)} v_{1,i,\alpha}^{(k)} \frac{\partial}{\partial v_{1,i,\alpha}^{(k)}} - \sum_{k=1}^{L} \frac{Q_1 (v_{1,i,\alpha}^{(k)})^2 - k_B T}{Q_2} \frac{\partial}{\partial v_{2,i,\alpha}^{(k)}}] \end{aligned} \tag{I.2e}$$

The various decompositions of the Lagrange multiplier used in Eq. (I.2) are given in Eq. (I.3).

$$\lambda_{i,\alpha}^N = \frac{\frac{L}{L+1} \sum_{k=1}^{L} Q_1 (v_{1,i,\alpha}^{(k)})^2 v_{2,i,\alpha}^{(k)}}{m_i v_{i,\alpha}^2 + \frac{L}{L+1} \sum_{k=1}^{L} Q_1 (v_{1,i,\alpha}^{(k)})^2} \tag{I.3a}$$

$$\lambda_{i,\alpha}^s = \frac{v_{i,\alpha} F_{i,\alpha}^s(\boldsymbol{r})}{m_i v_{i,\alpha}^2 + \frac{L}{L+1} \sum_{k=1}^{L} Q_1 (v_{1,i,\alpha}^{(k)})^2} \tag{I.3b}$$

$$\lambda_{i,\alpha}^{m} = \frac{v_{i,\alpha} F_{i,\alpha}^{m}(\boldsymbol{r})}{m_i v_{i,\alpha}^2 + \frac{L}{L+1}\sum_{k=1}^{L} Q_1 (v_{1,i,\alpha}^{(k)})^2} \tag{I.3c}$$

$$\lambda_{i,\alpha}^{l} = \frac{v_{i,\alpha} F_{i,\alpha}^{l}(\boldsymbol{r})}{m_i v_{i,\alpha}^2 + \frac{L}{L+1}\sum_{k=1}^{L} Q_1 (v_{1,i,\alpha}^{(k)})^2} \tag{I.3d}$$

Many possible Trotter factorizations of the Liouville operator $\exp(iL\,\Delta t)$ are possible, but for computational efficiency I chose the 'XO' factorization[109], where the deterministic (non-stochastic) Nosé-Hoover variable updates occur at the outer time step because this less frequent updating of the NH variables increases the computational efficiency. The factorization for this algorithm is given by Eq. (I.4).

$$\begin{aligned}\exp(iL\,\Delta t) = \exp\left(iL_N \frac{\Delta t}{2}\right) \exp\left(iL_v^l \frac{\Delta t}{2}\right) \{\exp\left(iL_v^m \frac{\Delta t_m}{2}\right) \\ \times \left[\exp\left(iL_v^s \frac{\Delta t_s}{2}\right) \exp\left(iL_r \frac{\Delta t_s}{2}\right) \exp(iL_{OU}\Delta t) \exp\left(iL_r \frac{\Delta t_s}{2}\right) \exp\left(iL_v^s \frac{\Delta t_s}{2}\right)\right]^s \\ \times \exp\left(iL_v^m \frac{\Delta t_m}{2}\right)\}^m \exp\left(iL_v^l \frac{\Delta t}{2}\right) \exp\left(iL_N \frac{\Delta t}{2}\right)\end{aligned} \tag{I.4}$$

The process is now repeated for the auxiliary equation of motion, Eq. (6.6). Recognizing that there is only one component to the auxiliary acceleration, given by Eq. (6.1b), one can see that there is no need to break the auxiliary factorization into separate timescales as there is only one possible time scale, yielding Eq. (I.5). Again iL_{OU}^{a} applies a Ornstein-Uhlenbeck stochastic process to every $v^{a}{}_{2,i,\alpha}^{(k)}$.

$$iL^a = iL_r^a + iL_v^a + iL_N^a + iL_{OU}^a \tag{I.5}$$

The components of Eq. (I.5) are given in Eq. (I.6).

$$iL_r^a = \sum_{i,\alpha} v_{i,\alpha}^a \frac{\partial}{\partial a_{i,\alpha}} \tag{I.6a}$$

$$iL_v^a = \sum_{i,\alpha}\left[\left(\gamma\omega^2(\mu_{i,\alpha} - a_{i,\alpha}) - \lambda_{i,\alpha}^a v_{i,\alpha}^a\right)\frac{\partial}{\partial v_{i,\alpha}^a} - \lambda_{i,\alpha}^a \sum_{k=1}^{L} v^{a}{}_{1,i,\alpha}^{(k)} \frac{\partial}{\partial v^{a}{}_{1,i,\alpha}^{(k)}}\right] \tag{I.6b}$$

$$\begin{aligned} iL_N^a = -\sum_{i,\alpha}[\lambda_{i,\alpha}^{a,N} v_{i,\alpha}^a \frac{\partial}{\partial v_{i,\alpha}^a} + \lambda_{i,\alpha}^{a,N} \sum_{k=1}^{L} v^{a}{}_{1,i,\alpha}^{(k)} \frac{\partial}{\partial v^{a}{}_{1,i,\alpha}^{(k)}} \\ + \sum_{k=1}^{L} v^{a}{}_{2,i,\alpha}^{(k)} v^{a}{}_{1,i,\alpha}^{(k)} \frac{\partial}{\partial v^{a}{}_{1,i,\alpha}^{(k)}} - \sum_{k=1}^{L} \frac{Q_1^a (v^{a}{}_{1,i,\alpha}^{(k)})^2 - T_{aux}}{Q_2^a} \frac{\partial}{\partial v^{a}{}_{2,i,\alpha}^{(k)}}] \end{aligned} \tag{I.6e}$$

The various Lagrange multipliers of Eq. (1.6) are then given by Eq. (1.7).

$$\lambda_{i,\alpha}^{a,N} = \frac{\frac{L}{L+1}\sum_{k=1}^{L} Q_1^a (v^{a}{}_{1,i,\alpha}^{(k)})^2 v^{a}{}_{2,i,\alpha}^{(k)}}{\left(v_{i,\alpha}^a\right)^2 + \frac{L}{L+1}\sum_{k=1}^{L} Q_1^a (v^{a}{}_{1,i,\alpha}^{(k)})^2} \tag{I.7a}$$

$$\lambda_{i,\alpha}^{a} = \frac{v_{i,\alpha}^a \gamma \omega^2 (\mu_{i,\alpha} - a_{i,\alpha})}{\left(v_{i,\alpha}^a\right)^2 + \frac{L}{L+1}\sum_{k=1}^{L} Q_1^a (v^{a}{}_{1,i,\alpha}^{(k)})^2} \tag{I.7b}$$

Now, if the polarization is evaluated in the outer time step then $\omega = \sqrt{2}/\Delta t$ and the factorization of the Liouville operator becomes Eq. (1.8).

$$\begin{aligned} \exp(iL^a\,\Delta t) &= \exp\left(iL_N^a \frac{\Delta t}{2}\right) \exp\left(iL_v^a \frac{\Delta t}{2}\right) \\ &\times \left[\exp\left(iL_r^a \frac{\Delta t_s}{2}\right) \exp(iL_{OU}^a \Delta t) \exp\left(iL_r^a \frac{\Delta t_s}{2}\right)\right]^{ms} \exp\left(iL_v^a \frac{\Delta t}{2}\right) \exp\left(iL_N^a \frac{\Delta t}{2}\right) \end{aligned} \tag{I.8}$$

If the polarization is evaluated in the intermediate time step then $\omega = \sqrt{2}/\Delta t_m$ and the factorization of the Liouville operator becomes Eq. (1.9).

$$\begin{aligned} \exp(iL^a\,\Delta t) &= \exp\left(iL_N^a \frac{\Delta t}{2}\right) \{\exp\left(iL_v^a \frac{\Delta t_m}{2}\right) \\ &\times \left[\exp\left(iL_r^a \frac{\Delta t_s}{2}\right) \exp(iL_{OU}^a \Delta t) \exp\left(iL_r^a \frac{\Delta t_s}{2}\right)\right]^{s} \exp\left(iL_v^a \frac{\Delta t_m}{2}\right)\}^m \exp\left(iL_N^a \frac{\Delta t}{2}\right) \end{aligned} \tag{I.9}$$

Thus when polarization is treated at the intermediate time step there are more frequent updates of the auxiliary NH variables, $\exp\left(iL_v^a \frac{\Delta t_m}{2}\right)$, leading to a decrease in efficiency. Combining the integration of the atomic and auxiliary degrees of freedom yields the full factorization from which a numerical integration is easily built. For the case of polarization in the outer time step the factorization becomes Eq. (I.10).

$$\begin{aligned} \exp(iL\,\Delta t) &= \exp\left(iL_N \frac{\Delta t}{2}\right) \exp\left(iL_N^a \frac{\Delta t}{2}\right) \exp\left(iL_v^l \frac{\Delta t}{2}\right) \exp\left(iL_v^a \frac{\Delta t}{2}\right) \{\exp\left(iL_v^m \frac{\Delta t_m}{2}\right) \\ &\times \left[\exp\left(iL_v^s \frac{\Delta t_s}{2}\right) \exp\left(iL_r \frac{\Delta t_s}{2}\right) \exp\left(iL_r^a \frac{\Delta t_s}{2}\right) \exp(iL_{OU} \Delta t) \exp(iL_{OU}^a \Delta t) \exp\left(iL_r \frac{\Delta t_s}{2}\right) \exp\left(iL_r^a \frac{\Delta t_s}{2}\right) \exp\left(iL_v^s \frac{\Delta t_s}{2}\right)\right]^{s} \\ &\times \exp\left(iL_v^m \frac{\Delta t_m}{2}\right)\}^m \exp\left(iL_v^l \frac{\Delta t}{2}\right) \exp\left(iL_v^a \frac{\Delta t}{2}\right) \exp\left(iL_N \frac{\Delta t}{2}\right) \exp\left(iL_N^a \frac{\Delta t}{2}\right) \end{aligned} \tag{I.10}$$

When polarization is treated in the intermediate time step then the algorithm factorization is given by Eq. (I.11).

$$\begin{aligned} \exp(iL\,\Delta t) &= \exp\left(iL_N \frac{\Delta t}{2}\right) \exp\left(iL_N^a \frac{\Delta t}{2}\right) \exp\left(iL_v^l \frac{\Delta t}{2}\right) \{\exp\left(iL_v^m \frac{\Delta t_m}{2}\right) \exp\left(iL_v^a \frac{\Delta t_m}{2}\right) \\ &\times \left[\exp\left(iL_v^s \frac{\Delta t_s}{2}\right) \exp\left(iL_r \frac{\Delta t_s}{2}\right) \exp\left(iL_r^a \frac{\Delta t_s}{2}\right) \exp(iL_{OU} \Delta t) \exp(iL_{OU}^a \Delta t) \exp\left(iL_r \frac{\Delta t_s}{2}\right) \exp\left(iL_r^a \frac{\Delta t_s}{2}\right) \exp\left(iL_v^s \frac{\Delta t_s}{2}\right)\right]^{s} \\ &\times \exp\left(iL_v^m \frac{\Delta t_m}{2}\right) \exp\left(iL_v^a \frac{\Delta t_m}{2}\right)\}^m \exp\left(iL_v^l \frac{\Delta t}{2}\right) \exp\left(iL_N \frac{\Delta t}{2}\right) \exp\left(iL_N^a \frac{\Delta t}{2}\right) \end{aligned} \tag{I.11}$$

Appendix J: Additional iEL/0-SCF with SI Data

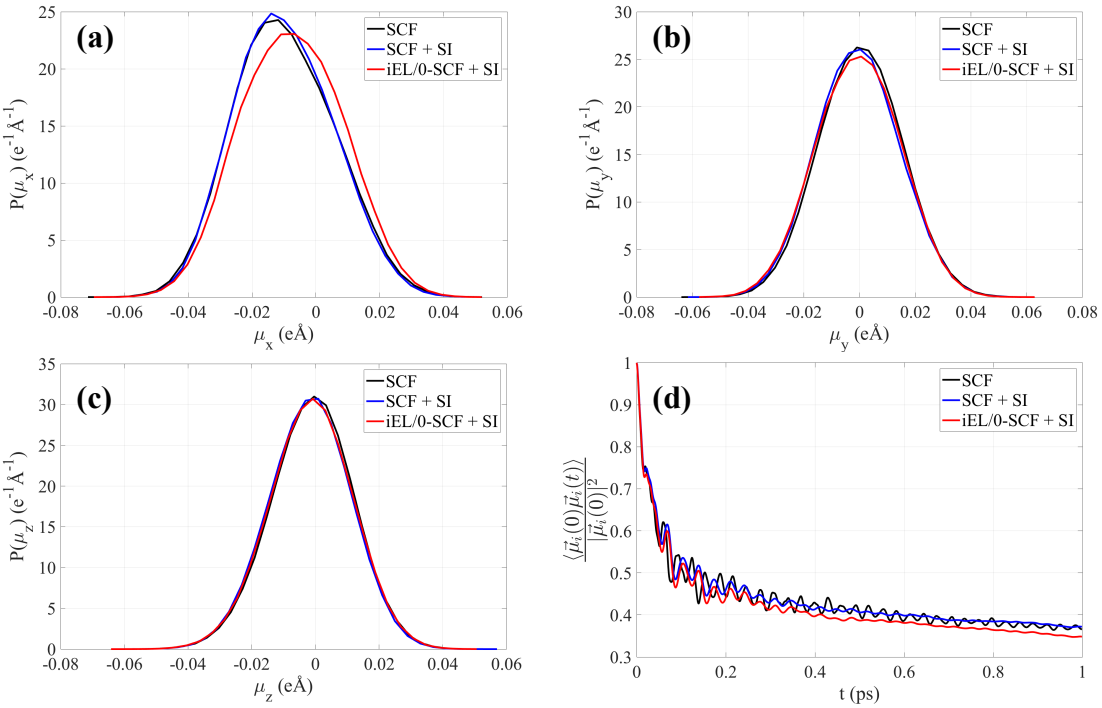


Figure J.1: *Induced dipole probability distributions for glycine N-terminus nitrogen x-component (a), y-component (b), z-component (c), and the corresponding induced dipole autocorrelation (d). Three methods are presented- SCF with standard velocity Verlet integration, SCF with SI integration, and iEL/0-SCF with SI integration. All simulations were performed with a single 1.0 fs time step and a temperature of 298.0 K.*

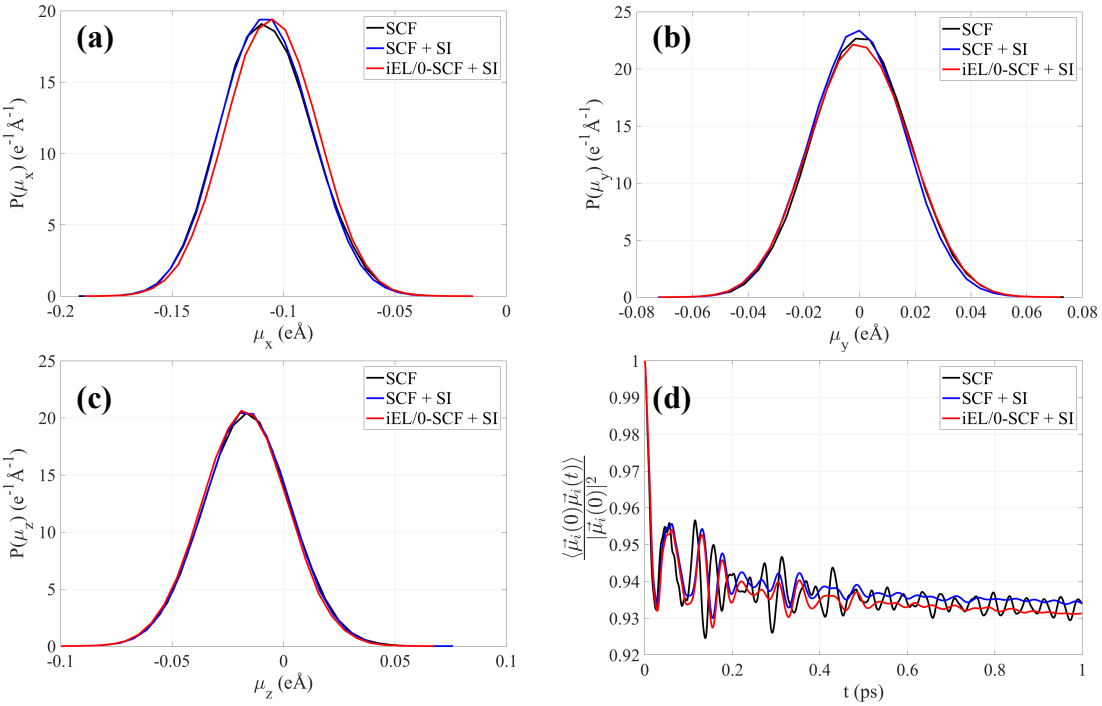


Figure J.2: *Induced dipole probability distributions for glycine α-carbon x-component (a), y-component (b), z-component (c), and the corresponding induced dipole autocorrelation (d). Three methods are*

presented- SCF with standard velocity Verlet integration, SCF with SI integration, and iEL/0-SCF with SI integration. All simulations were performed with a single 1.0 fs time step and a temperature of 298.0 K.

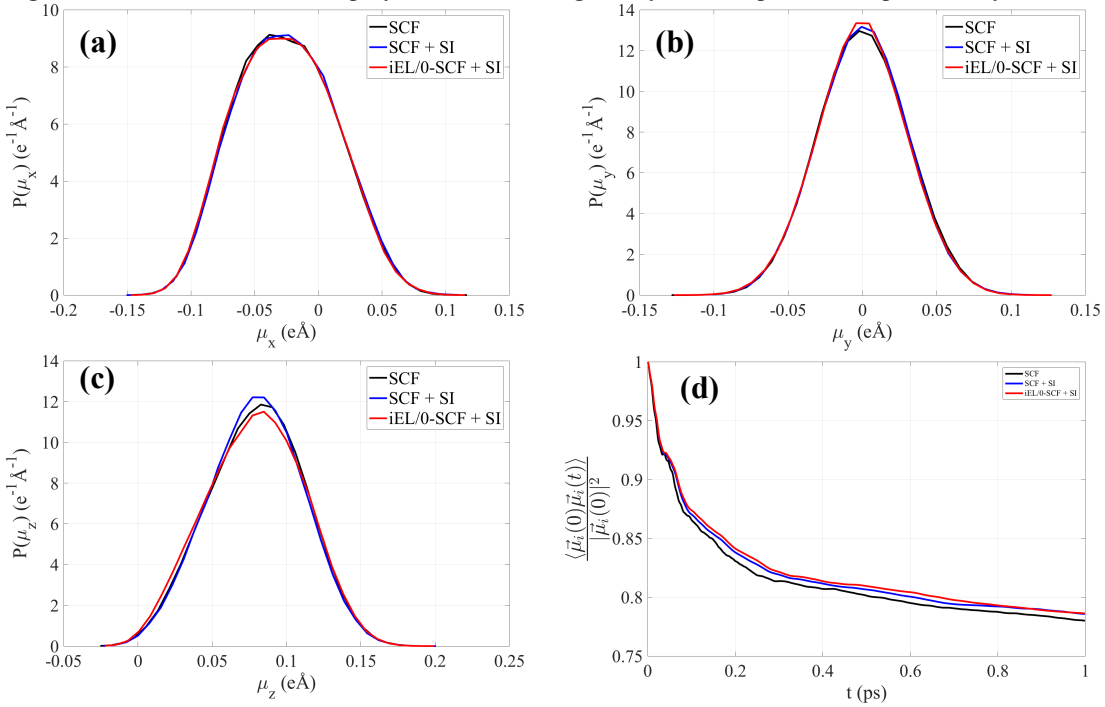


Figure J.3: *Induced dipole probability distributions for glycine carbonyl oxygen x-component (a), y-component (b), z-component (c), and the corresponding induced dipole autocorrelation (d). Three methods are presented- SCF with standard velocity Verlet integration, SCF with SI integration, and iEL/0-SCF with SI integration. All simulations were performed with a single 1.0 fs time step and a temperature of 298.0 K.*

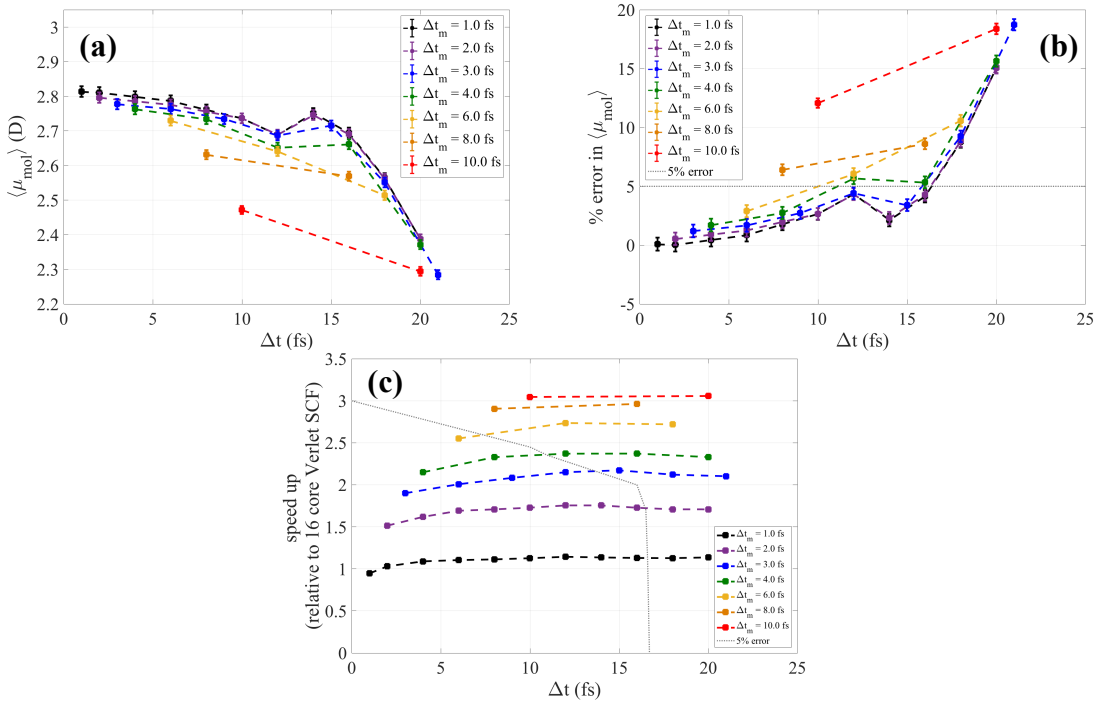


Figure J.4: *Three time scale pol. mid. RESPA used in conjunction with SI and iEL/0-SCF on a system of pure water. The average molecular dipole of water $\langle \mu_{mol} \rangle$ as a function of the outer RESPA time step, Δt, is reported (a) along with the error of the average molecular dipole (b) with respect to the vale of an SCF*

simulation with standard Verlet integration at a 1.0 fs time step (2.81 D). Different colored curves represent different intermediate time steps, Δt_m. Also reported are the computational speed ups with respect to standard 1.0 fs Verlet integration (c). All calculations were performed with 16 CPUs and shared-memory parallelization. The dotted gray line in (b) represents 5% error and all simulations below that error in (b) are within the 5% error envelope in (c).

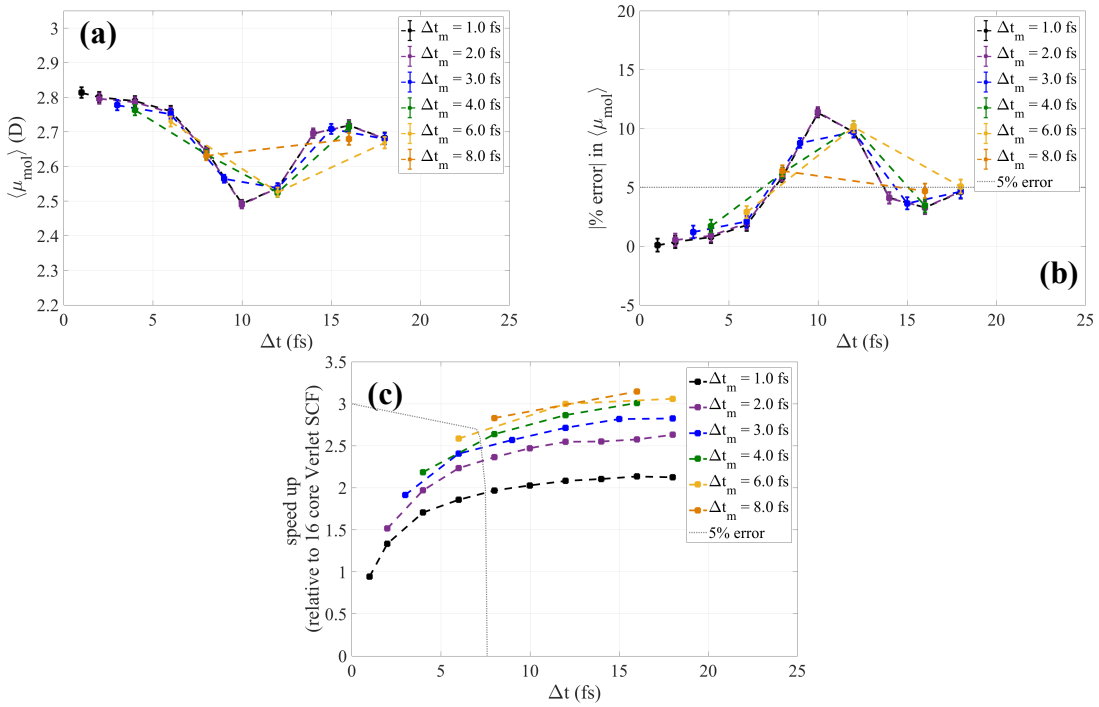
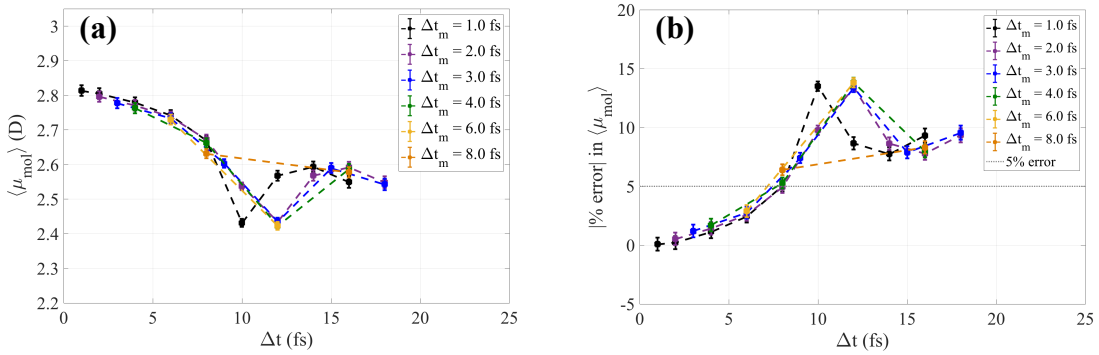


Figure J.5: *Three time scale pol. long RESPA used in conjunction with SI and iEL/0-SCF on a system of pure water. The average molecular dipole of water $\langle \mu_{mol} \rangle$ as a function of the outer RESPA time step, Δt, is reported (a) along with the error of the average molecular dipole (b) with respect to the vale of an SCF simulation with standard Verlet integration at a 1.0 fs time step (2.81 D). Different colored curves represent different intermediate time steps, Δt_m. Also reported are the computational speed ups with respect to standard 1.0 fs Verlet integration (c). All calculations were performed with 16 CPUs and shared-memory parallelization. The dotted gray line in (b) represents 5% error and all simulations below that error in (b) are within the 5% error envelope in (c).*



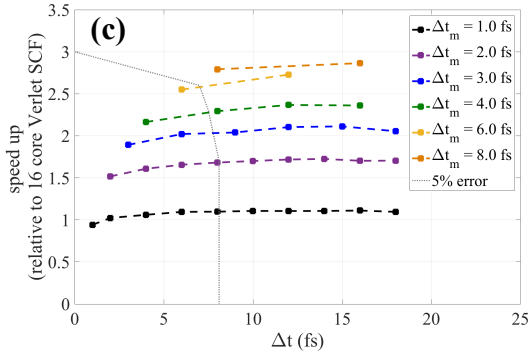


Figure J.6: *Three time scale el. long RESPA used in conjunction with SI and iEL/0-SCF on a system of pure water. The average molecular dipole of water $\langle \mu_{mol} \rangle$ as a function of the outer RESPA time step, Δt, is reported (a) along with the error of the average molecular dipole (b) with respect to the vale of an SCF simulation with standard Verlet integration at a 1.0 fs time step (2.81 D). Different colored curves represent different intermediate time steps, Δt_m. Also reported are the computational speed ups with respect to standard 1.0 fs Verlet integration (c). All calculations were performed with 16 CPUs and shared-memory parallelization. The dotted gray line in (b) represents 5% error and all simulations below that error in (b) are within the 5% error envelope in (c).*

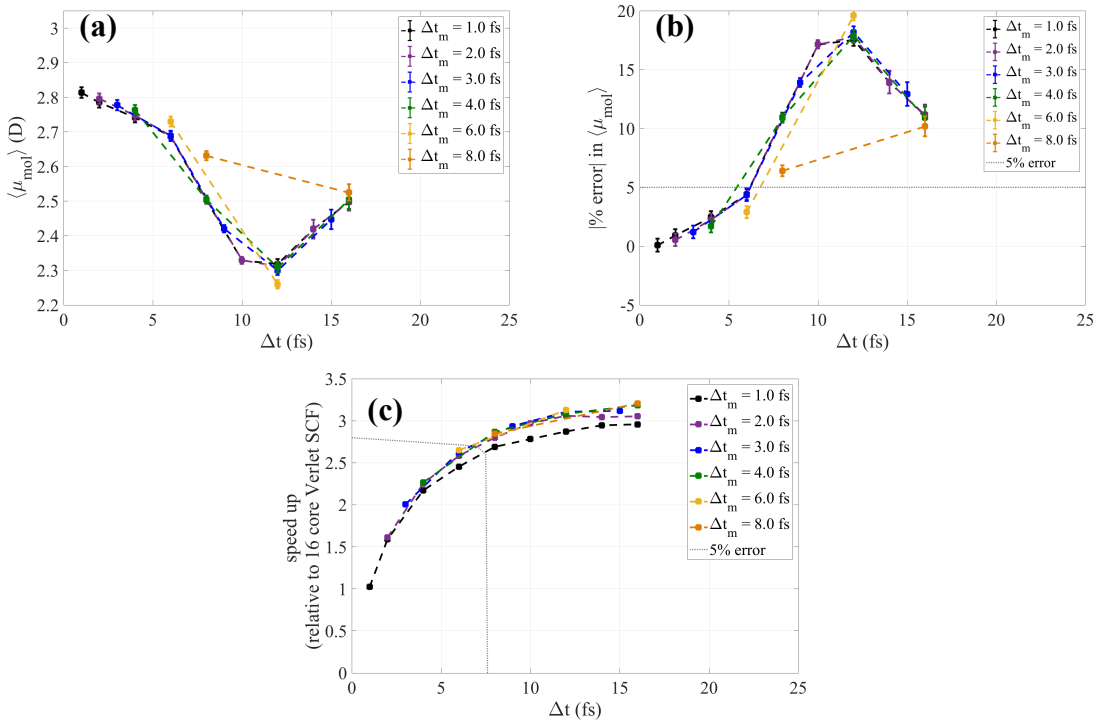


Figure J.7: *Three time vdW mid. RESPA used in conjunction with SI and iEL/0-SCF on a system of pure water. The average molecular dipole of water $\langle \mu_{mol} \rangle$ as a function of the outer RESPA time step, Δt, is reported (a) along with the error of the average molecular dipole (b) with respect to the vale of an SCF simulation with standard Verlet integration at a 1.0 fs time step (2.81 D). Different colored curves represent different intermediate time steps, Δt_m. Also reported are the computational speed ups with respect to standard 1.0 fs Verlet integration (c). All calculations were performed with 16 CPUs and shared-memory parallelization. The dotted gray line in (b) represents 5% error and all simulations below that error in (b) are within the 5% error envelope in (c).*