

Solvent Effects for Vertical Ionization Processes in Liquid  
Water and at the Liquid-Vapor Interface

Dissertation

Presented in Partial Fulfillment of the Requirements for the Degree  
Doctor of Philosophy in the Graduate School of The Ohio State  
University

By

Marc P. L. Coons, M.S.

Graduate Program in Chemistry

The Ohio State University

2017

Dissertation Committee:

Prof. John M. Herbert, Advisor

Prof. James V. Coe

Prof. Sherwin J. Singer

Prof. Laura Podalsky

© Copyright by

Marc P. L. Coons

2017

## Abstract

Liquid microjet photoelectron spectroscopy is a state-of-the-art experimental technique that provides avenues for investigating ultrafast charge and energy transfer processes in liquid water and aqueous solutions. The emphasis of this work is placed on making contact with recent liquid microjet experiments of hydrated electrons,  $e^-(aq)$ , in liquid water and at the liquid-vapor interface from a computational perspective. Since its discovery nearly 55 years ago,  $e^-(aq)$  has attracted significant attention from both experimental and theoretical communities due to its crucial role in radiation chemistry and its relatively elusive structure in liquid water at ambient conditions. Historically, the primary observable of hydrated electrons has been a well-characterized optical absorption spectrum, but with the advent of liquid microjet spectroscopy, this has shifted to measurements of its relative binding energy below vacuum level. Some experiments utilizing this methodology have been interpreted to suggest that  $e^-(aq)$  at the liquid-vapor interface is energetically dissimilar than in liquid water, and that, unlike the more strongly bound species in liquid, it can potentially undergo destructive reactions with solvated DNA molecules.

A variety of computational strategies are employed to demonstrate that the spectroscopic properties of hydrated electrons in liquid water are actually quite similar to the interfacial species. To that end, mixed quantum-classical molecular dynamics simulations are performed where  $e^-(aq)$  in liquid water and at the interface is described

by two different one-electron pseudopotentials. These simulations suggest that the presence of  $e^-(aq)$  at a liquid-vapor interface is fleeting at ambient conditions, and to experimentally distinguish it from the bulk species using standard spectroscopic techniques would be challenging. Non-equilibrium polarizable continuum models (PCMs), in conjunction with MP2 and DFT methods, are then employed to compute vertical ionization energies in liquid water. However, computing these quantities in anisotropic environments, such as at an interface, is not possible with traditional PCMs. Therefore, a novel methodology is presented for computing vertical ionization energies at a liquid-vapor interface that was developed to incorporate non-equilibrium solvent polarization effects for molecules immersed in arbitrary dielectric environments. This method is applied to  $e^-(aq)$ , alkali metal cations, and halide anions in liquid water, for which vertical ionization energies have been measured, and also to predict these quantities at the interface where some liquid microjet measurements do not yet exist.

This is dedicated to no one person, but to the giants who laid the path before us,  
and to those ensuring it continues.

## Acknowledgments

I would like to express significant gratitude to my advisor, John Herbert, for his patient guidance and support during my graduate career, and demonstrating the meaning of integrity as a person and as a scientist. An enormous amount of gratitude is owed to Ka Un Lao for bridging the continental separation between us with laughter and becoming an extraordinary lifelong friend. I am grateful to the complete set of Herbert group members, spanning both past and present, for insightful conversations and examples of well-written code that were integral to my growth as a graduate student. I would like to thank the faculty and staff members at The Ohio State University who maintain the Department of Chemistry and Biochemistry functioning smoothly and efficiently. Without the generous amount of computing resources from the Ohio Supercomputer Center and the patience of the staff members there, none of this work would have been possible. Finally, in the interest of saving trees, I would like to simultaneously express a deep appreciation for my family and friends. Certainly at least as much as the aforementioned computing resources were necessary for this, so was all your love and support.

## Vita

- 11 August 1987 ..... Born, Canton, OH
- 25 May 2005 ..... Diploma, Jackson High School, Massillon, OH
- 14 June 2009 ..... B.S. *summa cum laude* with honors and research distinction, Chemistry, The Ohio State University, Columbus, OH
- 17 December 2010 ..... M.S., Chemistry, University of California, Berkeley, CA

## Publications

M. P. Coons and J. M. Herbert. *Problems with Polarizable Continuum Calculations that Include Explicit Solvent Molecules, and How to Fix Them.* **in preparation**

M. P. Coons and J. M. Herbert. *A Poisson Equation Solver for Continuum Solvation in Arbitrary Environments: Application to Vertical Ionization Potentials in Liquid Water and at a Liquid-Vapor Interface.* **in preparation**

J. M. Herbert and M. P. Coons. *The Hydrated Electron.* *Annu. Rev. Phys. Chem.*, **68**, 447–472 (2017)

M. P. Coons, Z.- Q. You and J. M. Herbert. *The Hydrated Electron at the Surface of Neat Liquid Water Appears to be Indistinguishable from the Bulk Species.* *J. Am. Chem. Soc.*, **138**, 10879–10886 (2016)

F. Uhlig, J. M. Herbert, M. P. Coons, P. Jungwirth. *Optical Spectroscopy of the Bulk and Interfacial Hydrated Electron from Ab Initio Calculations.* *J. Phys. Chem. A*, **118**, 7507–7515 (2014)

## Fields of Study

Major Field: Chemistry

# Table of Contents

	<b>Page</b>
Abstract . . . . .	ii
Dedication . . . . .	iv
Acknowledgments . . . . .	v
Vita . . . . .	vi
List of Tables . . . . .	x
List of Figures . . . . .	xii
1. Electronic Structure Theory . . . . .	1
1.1 Hartree-Fock Theory . . . . .	1
1.2 Møller-Plesset Second-Order Perturbation Theory . . . . .	9
1.3 Density Functional Theory . . . . .	13
2. The Solvated Electron . . . . .	17
2.1 Liquid Microjet Experiments . . . . .	17
2.2 Modeling $e^-(aq)$ in Liquid and at an Interface . . . . .	19
2.2.1 Mixed Quantum-Classical Simulations . . . . .	19
2.2.2 Equilibrium Solvation . . . . .	23
2.2.3 Non-equilibrium Solvent Effects . . . . .	28
2.3 Electronic Structure Methods . . . . .	35
2.3.1 DFT and MP2 Methods . . . . .	35
2.3.2 Basis Set Convergence . . . . .	41
2.3.3 Non-equilibrium Polarizable Continuum Models . . . . .	42
2.3.4 Non-equilibrium Poisson Equation Solver . . . . .	44
2.4 Results and Discussion . . . . .	49

2.4.1	Simulations of $e^{-}(\text{aq})$ in Liquid and at the Interface . . . . .	49
2.4.2	Vertical Binding Energy of $e^{-}(\text{aq})$ . . . . .	69
2.5	Conclusions . . . . .	73
3.	Continuum Solvation Models . . . . .	77
3.1	The Road to Implicit Solvation . . . . .	77
3.1.1	Generalized Born Model . . . . .	79
3.1.2	Kirkwood-Onsager Model . . . . .	83
3.1.3	The Onsager Self-Consistent Reaction Field Model . . . . .	87
3.2	Isotropic Solvation with Polarizable Continuum Models . . . . .	91
3.2.1	Reference State Formalism . . . . .	94
3.2.2	Non-equilibrium Solvent Effects . . . . .	98
3.3	Anisotropic Solvation with the Poisson Equation Solver . . . . .	104
3.3.1	Revised Reference State Solvation . . . . .	104
3.3.2	State-Specific Non-equilibrium Solvent Effects . . . . .	108
3.4	Computational Details and Methods . . . . .	114
3.4.1	Molecular Dynamics Simulations . . . . .	114
3.4.2	Basis Sets and Electronic Structure Methods . . . . .	115
3.4.3	Molecular Cavities for PCM . . . . .	120
3.4.4	Dielectric Function $\epsilon(\mathbf{r})$ for PEQS . . . . .	125
3.5	Results and Discussion . . . . .	131
3.6	Conclusions . . . . .	137
4.	Numerically Solving Poisson's Equation . . . . .	139
4.1	Finite Difference Method . . . . .	139
4.2	Multigrid Method . . . . .	146
	Bibliography . . . . .	158

## List of Tables

Table	Page
2.1 Average number of water molecules, $\langle N \rangle$ , and optimally-tuned range-separation parameters, $\mu$ , for LRC- $\mu$ BOP. Also shown are the average VEBEs for these QM regions, computed at the LRC- $\mu$ BOP/ and RI-MP2/6-31++G* levels of theory. (Uncertainties represent one standard deviation, averaged over snapshots extracted from the simulations.) In these calculations, the QM region is treated as a gas-phase $(\text{H}_2\text{O})_N^-$ cluster, <i>i.e.</i> , the QM regions are the same as in the calculations presented in Table 2.6 but the latter calculations include PCM boundary conditions. Comparison of these “gas phase” values with the results in Table 2.6 affords the PCM corrections that are listed here, which illustrate the magnitude of the correction for long-range electrostatic effects. . . . .	39
2.2 Average VEBEs computed using two different basis sets, for the interfacial hydrated electron (computed using the DFT/Poisson method) and the bulk hydrated electron (computed using the DFT/PCM method).	41
2.3 Average values for selected observables from one-electron pseudopotential simulations. . . . .	56
2.4 Average values for selected observables from mean-field polarization model simulations with varying numbers of water molecules, $N$ . . . .	65
2.5 Parameters for the Gaussian fits [Eq. 2.34] of the VEBE data in Figure 2.9. . . . .	66
2.6 VEBEs <sup>a</sup> (in eV) computed at the DFT <sup>b</sup> / and RI-MP2 <sup>c</sup> /6-31++G* levels of theory using non-equilibrium solvation models. . . . .	72

3.1	VIPs (in eV) at the RI-MP2 level of theory using the basis sets described in Section 3.4.2 and the non-equilibrium solvation models described in Sections 3.2 and 3.3 for various systems in bulk liquid water. <sup>a</sup> Ref. [1]. <sup>b</sup> Ref. [2]. <sup>c</sup> Ref. [3]. . . . .	135
3.2	VIPs (in eV) at the RI-MP2 level of theory using the basis sets described in Section 3.4.2 and the non-equilibrium PEQS method described in Section 3.3 for various systems at the liquid-vapor interface. The bulk liquid VIPs computed with PEQS from Table 3.1 are reproduced here for comparison. <sup>a</sup> Ref. [9] . . . . .	136
4.1	The finite difference coefficients for the first and second derivatives ( $m = 1$ or $2$ in Eq. 4.19) and the corresponding orders of accuracy. [Note: the order of accuracy is $\mathcal{O}(h^{2n})$ .] . . . . .	145

## List of Figures

Figure	Page
2.1 Illustration of the tuning procedure for LRC- $\mu$ BOP/6-31++G*, demonstrated for three different $(\text{H}_2\text{O})_N^-$ geometries (with $N = 47\text{--}50$ ) that are shown in red, blue, and green and were extracted from an <i>ab initio</i> interfacial $e^-(\text{aq})$ simulation <sup>4</sup> . The solid curves represent $\text{IE}_{\Delta\text{SCF}}(\mu)$ and the dashed curves are $-\epsilon_{\text{SOMO}}(\mu)$ . Where the dashed and solid curves of a particular color intersect is the “optimally tuned” value of $\mu$ for that particular snapshot, and in this case we choose $\mu = 0.18 \text{ a}_0^{-1}$ as the best compromise value for these three snapshots. This value is indicated by a vertical line, and the other vertical lines indicate other values of $\mu$ that have been employed in previous hydrated-electron studies using this functional. These include $\mu = 0.165 \text{ a}_0^{-1}$ (Ref. [5]), $\mu = 0.33 \text{ a}_0^{-1}$ (Ref. [6]), and $\mu = 0.37 \text{ a}_0^{-1}$ (Refs. [7] and [8]). . . . .	38
2.2 Dielectric function $\epsilon(r)$ from Eq. 2.32, as a function of the distance $r$ from the center of the QM region. The parameters of the switching function are $\alpha = 3.9 \text{ \AA}^{-1}$ and $r_{\text{mid}} = 9.375 \text{ \AA}$ , as used in this work for a QM radius of $8.0 \text{ \AA}$ . In this example, $\epsilon(r_{\text{QM}}) = 1.002$ , $\epsilon(r_{\text{mid}}) = 39.695$ , and $\epsilon(r_{\text{cavity}}) = 78.388$ . . . . .	47
2.3 Diagram of the setup for the QM/Poisson calculations at the water/vacuum interface. Within the two shaded regions, the dielectric function is interpolated between $\epsilon = 1.0$ and $\epsilon = 78.39$ , as described in the text. The interfacial region around the GDS uses the switching function in Eq. 2.33, whereas the spherical cavity in bulk water uses the same switching function used for bulk solvation, Eq. 2.32. . . . .	48

2.4 Representative snapshots of a simulation using the mean-field polarization model. In each panel, the opaque blue isocontour of the wave function encapsulates 50% of the one-electron probability distribution  $|\psi(\mathbf{r})|^2$ , while the lighter, translucent isosurface encloses 95%. At  $t = 0$ , where the solvent configuration is taken from an equilibrated simulation of neat liquid water, the electron is localized in a surface trap created by dangling O–H moieties, however the degree of localization ( $r_g \approx 6.25 \text{ \AA}$ ) is far less than for  $e^-(aq)$  in bulk water. After only 0.1 ps, however, the wave function has contracted to  $r_g \approx 3.50 \text{ \AA}$ , as seen in (b), while the VEBE has increased by 1.65 eV. By  $t = 0.5$  ps the wave function has contracted to  $r_g \approx 2.4 \text{ \AA}$  and is comparable in size to the bulk species, yet resides  $\approx 1.0 \text{ \AA}$  above the Gibbs dividing surface. The size of the wave function has changed little by  $t = 10$  ps, but its centroid has moved to  $1.50 \text{ \AA}$  below the Gibbs dividing surface, and water molecules surround the electron. The VEBE has increased by 0.5 eV relative to (c), and is now comparable to the bulk value. Panel (e) shows a snapshot shortly after the electron internalizes, with its centroid situated  $8.74 \text{ \AA}$  below the Gibbs dividing surface and with a VEBE that is similar to the bulk value and essentially unchanged relative to that in (d). By  $t = 40$  ps in (f), the electron centroid has been fluctuating around  $d = -8.50 \text{ \AA}$  for 28 ps with a VEBE that has fluctuated around the bulk value since  $t \approx 10$  ps. . . . . 50

2.5 Time evolution of (a) the VEBE, (b) the radius of gyration, and (c) the distance between the centroid of the  $e^-(aq)$  charge distribution and the Gibbs dividing surface. (The time axis is linear for the first 0.5 ps then switches to a logarithmic scale.) Data are obtained from quantum/classical MD using two different one-electron models, under ambient conditions. . . . . 53

2.6 VEBE as a function of (a) the radius of gyration and (b) the distance between the electron and the Gibbs dividing surface. The data are taken from the trajectories plotted in Figure 2.5, sampled every 20 fs. . . . . 55

2.7 Snapshots from a simulation of an electron initialized delocalized in liquid water using the Turi-Borgis pseudopotential model. In each panel, the opaque blue isocontour of the wave function encapsulates 50% of the one-electron probability distribution  $|\psi(\mathbf{r})|^2$ , while the lighter translucent isosurface encloses 95%. At  $t = 0$ , the electron is initialized as a completely delocalized particle in liquid water, with  $r_g \approx 6.78 \text{ \AA}$ . After 100 fs the wave function has contracted to  $r_g \approx 3.93 \text{ \AA}$ , as seen in (b), while the VEBE has increased by 1.46 eV. By 500 fs, the wave function further contracts to  $r_g \approx 3.12 \text{ \AA}$  and resides  $\approx 1.1 \text{ \AA}$  above the GDS. By  $t = 8.50 \text{ ps}$ , the size of the wave function is comparable to the bulk species with  $r_g = 2.45 \text{ \AA}$ , and its centroid has moved to  $1.50 \text{ \AA}$  below the GDS where water molecules surround the electron, such that the VEBE in (d) is 1.29 eV larger than that in (c), and is similar to the bulk value. Panel (e) shows a snapshot shortly before the electron internalizes, with its centroid situated  $7.75 \text{ \AA}$  below the GDS and with a VEBE that is similar to the bulk value and essentially unchanged relative to that in (d). By  $t = 40 \text{ ps}$  in (f), the electron centroid has been fluctuating around  $d = -13.00 \text{ \AA}$  for 28 ps with a VEBE that has fluctuated around the bulk value since  $t \approx 8.50 \text{ ps}$ . . . . . 61

2.8 Time evolution of (a) the VEBE, (b) the radius of gyration ( $r_g$ ), and (c) the distance ( $d$ ) between the centroid of the  $e^-(\text{aq})$  charge distribution and the Gibbs dividing surface from MD simulations employing a mean-field polarization potential model at ambient temperatures. (The time axis is linear for the first 0.5 ps then switches to a logarithmic scale.) For comparison to the simulations utilizing  $N = 300$  and  $N = 600$  water molecules, the data from representative  $N = 200$  trajectories are reproduced from Figure 2.5. . . . . 62

2.9 Probability distributions for the bulk and interfacial VEBE using (a) the mean-field polarization model, in an  $N = 600$  water box; and (b) the explicit polarization model, in an  $N = 200$  water box, with histogram bin widths of 0.05 eV in each case. Panels (c) and (d) show Gaussian fits to the data in (a) and (b), respectively. (Fitting parameters are listed in Table 2.5.) The interfacial distributions are also shown following a shift in energy equal to the difference in the mean VEBEs of the bulk and interfacial distributions, as reported in Tables 2.3 and 2.4. Differences in the center points of the distributions in (a) versus (b) are partly the result of the effects of finite box size. . . . . 64

2.10	Extrapolation of the average VEBE computed using the mean-field polarization potential, as a function of the inverse unit cell length, $L_x^{-1}$ . Error bars reflect a 95% confidence interval in $\langle \text{VEBE} \rangle$ , which is averaged over trajectories and over time. . . . .	68
2.11	Hypothetical potential energy curves for dissociative electron attachment by an aqueous electron. The molecular anion formed by electron attachment is stabilized by aqueous solvation, relative to its neutral parent molecule. The white band on the right illustrates the window between the vertical electron affinity (VEA) and the adiabatic electron affinity (AEA) in which electron attachment is energetically feasible; the left part of the diagram suggests that the equilibrated species $e^-(\text{aq})$ lies too far below vacuum level for this process to occur. Adapted from Ref. [9]. . . . .	75
3.1	Panels (a) and (c) show the spin density computed using the MOM-SCF procedure for the fluoride-water and neutral water systems, respectively, and Panels (b) and (d) show the spin density computed using the FRAGMO-MOM-SCF procedure for the same systems. In Panels (a) and (b), the fluorine atom is colored pink. In all Panels, the opaque and transparent blue color illustrates 50% and 95% of the spin density, respectively. Panel (a) illustrates the difficulty in separating the fluoride 2p orbital from the orbitals of the surrounding water molecules: removal of an electron from the system results in a collective ionization of both the fluoride and water molecules. In Panel (c), the MOM-SCF procedure for the neutral water system results in ionization of a water molecule close to the solute-solvent boundary, where the effects of hydrogen bonding are neglected. Using the FRAGMO-MOM-SCF procedure fixes this issue, and the desired orbitals are ionized as shown in Panels (b) and (d). . . . .	119
3.2	In Panels (a)–(c), the opaque and transparent blue color illustrates 50% and 95% of the spin density, respectively, while green is employed for the sodium-water system in Panel (d). In Panel (c), the lithium atom is colored pink and in Panel (d), the sodium atom is colored blue. Panel (a) shows the excess spin density of $e^-(\text{aq})$ from a bulk liquid water simulation. Panel (b) illustrates the spin density of the ionized chlorine-water system using a standard SCF approach. Panels (c) and (d) show the spin densities of the lithium- and sodium-water systems, and employing the MOM-SCF approach results in ionization from the desired 1s and 2p orbitals. . . . .	121

3.3	The various molecular cavity surfaces discussed in Section 3.4.3 are illustrated here. A generic, noncovalently bonded solute molecule is represented by the grey regions contained within the vdWS, which is shown in black. The problematic region, where solvent-solvent interactions are double-counted and where solvent dielectric is placed erroneously, is shown as the purple region between the noncovalently bonded regions of the solute. A solvent molecule probe, which is used to generate the SES and SAS cavities, is represented as a green circle. Where the solvent probe makes contact with the solute molecule forms the SES, which is the union of the black (vdWS) and red (re-entrant surface between noncovalently bonded regions of the molecule) surfaces. The SAS shown in blue is effectively a larger vdWS and is the locus of points formed by tracing the centroid of the solvent probe as it “rolls” around the solute molecule. . . . .	124
3.4	A two-dimensional contour plot of Eq. 2.32 applied to bulk $e^-(aq)$ with $\alpha = 4.0 \text{ \AA}^{-1}$ , $r_{QM} = 5.5 \text{ \AA}$ , and $r_{mid} = 6.0 \text{ \AA}$ is illustrated. The contour plot is reflective of the $xz$ -plane through the Cartesian grid origin. These parameters are utilized for all systems investigated in Chapter 3. . . . .	127
3.5	A two-dimensional contour plot of Eq. 3.91 applied to bulk $e^-(aq)$ with $\Delta = 0.265 \text{ \AA}$ is shown. The contour plot is reflective of the $xz$ -plane through the Cartesian grid origin. This parameter is utilized for all systems investigated in Chapter 3 as well as Ref. [10]. . . . .	128

3.6 In Panel (a), a two-dimensional contour plot of Eq. 3.91 applied to bulk  $e^-(aq)$  with  $\Delta = 0.265 \text{ \AA}$  along with a modification to the effective van der Waals radius  $d_\alpha$  for atom  $\alpha$  given by  $d_\alpha \rightarrow d_\alpha + r_{\text{probe}}$  with  $r_{\text{probe}} = 1.4 \text{ \AA}$  is shown. This is a three-dimensional analog of the PCM-SAS cavity, which improves upon the van der Waals cavity, but still has unphysical dielectric in the molecular region of the solute. In panel (b) the two-dimensional contour plot of a hybrid cavity that effectively applies Eq. 3.91 to the outermost atoms of the solute, and then sets  $\epsilon(\mathbf{r}) = 1.0$  within the molecular region of the solute is shown. The hybrid cavity is free of any erroneous solvent dielectric as in the spherical cavity, yet properly conforms to the molecular shape of the molecule as does the van der Waals cavity. These contour plot are reflective of the  $xz$ -plane through the Cartesian grid origin, and the corresponding parameters are utilized for all systems investigated in Chapter 3. . . . . 130

3.7 Panels (a) and (b) show the dielectric environment at the liquid vapor interface employing either the spherical cavity [Panel (a)] or the hybrid cavity [Panel(b)] for the molecular solute, which is an interfacial  $e^-(aq)$  configuration. The location of the GDS is shown as a black dotted line, and the solvent dielectric is smoothly interpolated from the bulk value to vacuum across this region using Eq. 2.33. . . . . 132

4.1 A comparison of the performance for solving Eq. 4.20 to an accuracy of  $10^{-5}$  a.u. for a single water molecule placed at the center of a cubic Cartesian grid of length  $15 \text{ \AA}$  with  $N_{\text{grid}} = 8615125$  using a standard CG routine (green), a fourth-order V-cycle multigrid (red), and a fourth-order W-cycle multigrid (blue). The slow decay of the residual error for the standard CG routine is indicative of  $\lambda > h$ , and using this method requires over 250 iterations. The two multigrid methods achieve convergence more rapidly, with the W-cycle method performing best, and both exhibit a nearly linear decrease of  $\mathbf{r}^h$  with respect to the iteration number that is in stark contrast to the standard CG routine. . . . . 149

- 4.2 An illustration of a second-order multigrid algorithm employed to solve Eq. 4.20 is presented. The input is the known source charge density  $\rho^h$  and the desired output is the electrostatic potential  $\varphi^h$ , both of which are discretized on a target Cartesian grid of resolution  $h$ . Steps 1 and 2 show the formation of the residual error  $r_h$  on the target grid, and then its restriction to a coarsened grid. On the coarse grid, the restricted residual error  $r_H$  is utilized in a CG routine to form a relaxed residual error  $v_H$  in Step 3. The relaxed residual error is then interpolated to the finer target grid to form  $v_h$  in Step 4. In Step 5,  $v_h$  is used to construct a correction to the solution, and this process is repeated until convergence is reached. . . . . 151
- 4.3 The flow diagrams for the fourth-order V-cycle (top) and W-cycle (bottom) multigrid methods are presented. Both methods employ a target grid with the finest resolution  $h$  and three coarse grids of resolutions  $H = 2h$ ,  $H = 4h$ , and  $H = 8h$ . The downward arrows represent restriction of the residual error vector from finer grids to coarser ones where the discretization error is relaxed  $\gamma_1$  times for the target and first two coarse grid levels and relaxed fully on the coarsest grid (denoted by  $\gamma_0$ ). The discretization error is then interpolated from coarser grids to back to finer ones, which is shown as an upward arrow, where it is further relaxed  $\gamma_2$  or  $\gamma_3$  times. After completing the restriction and interpolation operations, convergence of the solution on the target grid is tested, and the V- and W-cycles are repeated until the desired accuracy is reached. Figure adapted from Chapter 3 of Ref [12]. . . . 157

# Chapter 1: Electronic Structure Theory

## 1.1 Hartree-Fock Theory

This Chapter presents the historical and theoretical background for the electronic structure methods that are utilized in this work and are also ubiquitous in other disciplines of science. By the early part of the 20<sup>th</sup> century, the constituent components of atoms had been discovered: the electron by J. J. Thompson in 1897 and the nucleus by E. Rutherford in 1910. Shortly thereafter, the “old” quantum theory was established by N. Bohr and A. Sommerfeld who proposed the existence of orbitals in which the electrons surround atomic nuclei with discrete energy levels. Though this concept was useful in elucidating the relationship between positive nuclear charge, atomic number, and position on the periodic table, extending the notion of electron orbitals to polyatomic atoms proved difficult until the “new” quantum theory was developed. In 1926, E. Schrödinger and W. Heisenberg developed complementary wave and matrix mechanics to describe quantum phenomenon, and the birth of quantum chemistry was the development of the non-relativistic time-independent Schrödinger equation expressed as

$$\hat{H} |\Psi\rangle = E |\Psi\rangle , \tag{1.1}$$

which specifies the energy  $E$  of a particular quantum state  $|\Psi\rangle$  using an atomic or molecular Hamiltonian  $\hat{H}$ . Eq. 1.1 provided an exact solution for the hydrogen atom, the analysis of which agreed with Bohr-Sommerfeld concept of discrete electronic states. Heitler and London applied Eq. 1.1 to diatomic hydrogen in 1927 and provided the first, albeit inaccurate, calculation of a system with a covalent bond. A couple years later in 1929, the relativistic version of Eq. 1.1 was discovered by Paul Dirac, and he famously stated<sup>13</sup>

*The underlying physical laws necessary for the mathematical theory of a large part of physics and the whole of chemistry are thus completely known, and the difficulty is only that the exact application of these laws leads to equations much too complicated to be soluble. It therefore becomes desirable that approximate practical methods of applying quantum mechanics should be developed, which can lead to an explanation of the main features of complex atomic systems without too much computation.*

This revelation made it clear that an exact solution to Eq. 1.1 for complex molecular systems is unobtainable, but endeavors for developing approximate methods would prove useful to the field of chemistry.<sup>13</sup>

The Hartree-Fock (HF) method is one such approximation, and a discussion of this technique is the focus of this section. To proceed, consider a molecule comprised of  $N_{\text{elec}}$  electrons and  $N_{\text{nuc}}$  nuclei: the form of the non-relativistic molecular Hamiltonian  $\hat{H}$  in Eq. 1.1 is specified as

$$\hat{H} = -\frac{1}{2} \sum_{i=1}^{N_{\text{elec}}} \nabla_i^2 - \frac{1}{2} \sum_{\alpha=1}^{N_{\text{nuc}}} \frac{\nabla_{\alpha}^2}{m_{\alpha}} - \sum_{i=1}^{N_{\text{elec}}} \sum_{\alpha=1}^{N_{\text{nuc}}} \frac{Z_{\alpha}}{r_{i\alpha}} + \sum_{i=1}^{N_{\text{elec}}} \sum_{j>i}^{N_{\text{elec}}} \frac{1}{r_{ij}} + \sum_{\alpha=1}^{N_{\text{nuc}}} \sum_{\beta>\alpha}^{N_{\text{nuc}}} \frac{Z_{\alpha}Z_{\beta}}{r_{\alpha\beta}}, \quad (1.2)$$

where atomic units have been utilized ( $\hbar = m_e = 4\pi\epsilon_0 = e = 1$ ). The quantities  $r_{i\alpha} = |\mathbf{r}_i - \mathbf{R}_{\alpha}|$ ,  $r_{ij} = |\mathbf{r}_i - \mathbf{r}_j|$ , and  $r_{\alpha\beta} = |\mathbf{R}_{\alpha} - \mathbf{R}_{\beta}|$  are the electron-nucleus, interelectronic, and internuclear distances, respectively. The first two terms on the right side of Eq. 1.2 are the kinetic energies of the electrons and nuclei, respectively,

and  $m_\alpha$  is the mass of nucleus  $\alpha$ . The third term on the right side is the contribution from the attractive Coulomb interaction of an electron and nucleus of charge  $Z_\alpha$ , and the last two terms are the repulsive electron-electron and nuclear-nuclear Coulomb interactions, respectively. Traditional electronic structure methods employ the Born-Oppenheimer approximation to reduce the complexity of solving Eq. 1.1 for a Hamiltonian given by Eq. 1.2. Within the Born-Oppenheimer approximation, the total molecular Hamiltonian is reduced to the electronic components. The approximation is justified from an argument about the relative masses of electrons and nuclei: since the electrons are effectively three orders of magnitude lighter than the nuclei, they can be viewed as moving in a field generated by fixed nuclei. Therefore the nuclear kinetic energy term is neglected in Eq. 1.2, and the nuclear-nuclear repulsion energy is an additive constant to the electronic energy. In other words, the Born-Oppenheimer approximation neglects any contribution to the electronic structure of a molecule due to nuclear motion. The electronic problem then becomes

$$\hat{H}_{\text{elec}} |\Phi_{\text{elec}}(\mathbf{r}; \mathbf{R})\rangle = E_{\text{elec}}(\mathbf{R}) |\Phi_{\text{elec}}(\mathbf{r}; \mathbf{R})\rangle , \quad (1.3)$$

where the electronic Hamiltonian is

$$\hat{H}_{\text{elec}} = -\frac{1}{2} \sum_{i=1}^{N_{\text{elec}}} \nabla_i^2 - \sum_{i=1}^{N_{\text{elec}}} \sum_{\alpha=1}^{N_{\text{nuc}}} \frac{Z_\alpha}{r_{i\alpha}} + \sum_{i=1}^{N_{\text{elec}}} \sum_{j>i}^{N_{\text{elec}}} \frac{1}{r_{ij}} . \quad (1.4)$$

The electronic wave function  $|\Phi_{\text{elec}}(\mathbf{r}; \mathbf{R})\rangle$  explicitly depends on the electronic coordinates  $\mathbf{r}$  and also parametrically on the coordinates of the fixed nuclei  $\mathbf{R}$ . Furthermore, the electronic energy  $E_{\text{elec}}(\mathbf{R})$  also depends parametrically on  $\mathbf{R}$ : any change in the nuclear positions results in alterations of  $\hat{H}_{\text{elec}}$ ,  $|\Phi_{\text{elec}}(\mathbf{r}; \mathbf{R})\rangle$ , and  $E_{\text{elec}}(\mathbf{R})$ . Finally, the total molecular energy within the Born-Oppenheimer approximation is computed

as the sum of the electronic energy and the nuclear repulsion energy  $E_{\text{nn}}$ .

$$\begin{aligned} E_{\text{HF}} &= E_{\text{elec}} + \sum_{\alpha=1}^{N_{\text{nuc}}} \sum_{\beta>\alpha}^{N_{\text{nuc}}} \frac{Z_{\alpha}Z_{\beta}}{r_{\alpha\beta}} \\ &= E_{\text{elec}} + E_{\text{nn}} . \end{aligned} \quad (1.5)$$

The Hartree-Fock method is couched in terms of a Slater determinant of  $N$  molecular orbitals  $\{\psi_N(\mathbf{r})\}$  to describe the electronic wave function for  $N$  electrons, which is compactly expressed as

$$\begin{aligned} |\Phi_{\text{HF}}\rangle &\equiv \frac{1}{\sqrt{N!}} \begin{vmatrix} \psi_1(\mathbf{r}_1) & \psi_2(\mathbf{r}_1) & \cdots & \psi_N(\mathbf{r}_1) \\ \psi_1(\mathbf{r}_2) & \psi_2(\mathbf{r}_2) & \cdots & \psi_N(\mathbf{r}_2) \\ \vdots & \vdots & \ddots & \vdots \\ \psi_1(\mathbf{r}_N) & \psi_2(\mathbf{r}_N) & \cdots & \psi_N(\mathbf{r}_N) \end{vmatrix} \\ &\equiv |\psi_1(\mathbf{r}_1)\psi_2(\mathbf{r}_2)\cdots\psi_N(\mathbf{r}_N)\rangle , \end{aligned} \quad (1.6)$$

where  $\psi_N(\mathbf{r}_N)$  denotes electron  $N$  occupying molecular orbital  $N$ . A Slater determinant naturally captures the inherent antisymmetric property required for fermionic wave functions: upon permutation of any two electrons in Eq. 1.6, the fermionic wave function acquires a phase factor of  $-1$ . Since the non-relativistic Hamiltonian does not involve electronic spin, only the spatial component of the molecular orbitals, the information of which is entirely encoded in the set  $\{\psi_N(\mathbf{r})\}$ , needs to be considered. The Hartree-Fock energy shown in Eq. 1.5 is then computed as the expectation value of the electronic Hamiltonian with the Slater determinant given by Eq. 1.6:<sup>14</sup>

$$\begin{aligned} E_{\text{HF}}[\Phi_{\text{HF}}] &= \langle \Phi_{\text{HF}} | \hat{H}_{\text{elec}} | \Phi_{\text{HF}} \rangle + E_{\text{nn}} \\ &= \sum_{i=1}^{N_{\text{elec}}} \langle \psi_i | \hat{h} | \psi_i \rangle + \frac{1}{2} \sum_{i=1}^{N_{\text{elec}}} \sum_{j \neq i}^{N_{\text{elec}}} \langle \psi_i \psi_j | | \psi_j \psi_i \rangle + E_{\text{nn}} , \end{aligned} \quad (1.7)$$

where a compact expression for the two-electron contributions is introduced as

$$\langle \psi_i \psi_j | | \psi_j \psi_i \rangle \equiv \langle \psi_i \psi_j | \frac{1}{r_{12}} | \psi_i \psi_j \rangle - \langle \psi_i \psi_j | \frac{1}{r_{12}} | \psi_j \psi_i \rangle . \quad (1.8)$$

The first term on the right side of Eq. 1.7 is referred to as the one-electron core energy and contains the kinetic energy and electron-nuclear Coulomb interaction:

$$\langle \psi_i | \hat{h} | \psi_i \rangle = -\frac{1}{2} \int d\mathbf{r}_1 \psi_i^*(\mathbf{r}_1) \nabla_1^2 \psi_i(\mathbf{r}_1) - \sum_{\alpha=1}^{N_{\text{nuc}}} \int d\mathbf{r}_1 \frac{Z_\alpha}{r_{1\alpha}} |\psi_i(\mathbf{r}_1)|^2. \quad (1.9)$$

The second term reflects the contribution arising from the electron-electron Coulomb repulsion energy in a mean-field way. In other words, it expresses the field that an electron experiences from the other  $N_{\text{elec}} - 1$  electrons in an average manner:

$$\langle \psi_i \psi_j | \frac{1}{r_{12}} | \psi_i \psi_j \rangle \equiv \langle \psi_i | \hat{J}_j | \psi_i \rangle = \int d\mathbf{r}_1 \psi_i^*(\mathbf{r}_1) \hat{J}_j \psi_i(\mathbf{r}_1), \quad (1.10)$$

where the  $\hat{J}_j$  operator is defined as

$$\hat{J}_j \equiv \int d\mathbf{r}_2 \psi_j^*(\mathbf{r}_2) \frac{1}{r_{12}} \psi_j(\mathbf{r}_2). \quad (1.11)$$

The final term on the right side of Eq. 1.7 has no classical analog, and manifests from the antisymmetric property of the Slater determinant wave function. This contribution to the electronic energy is the exchange energy and is expressed as

$$\begin{aligned} \langle \psi_i \psi_j | \frac{1}{r_{12}} | \psi_j \psi_i \rangle &\equiv \langle \psi_i | \hat{K}_j | \psi_i \rangle = \int d\mathbf{r}_1 \psi_i^*(\mathbf{r}_1) \hat{K}_j \psi_i(\mathbf{r}_1) \\ &= \int \int d\mathbf{r}_1 d\mathbf{r}_2 \psi_i^*(\mathbf{r}_1) \psi_j^*(\mathbf{r}_2) \frac{1}{r_{12}} \psi_j(\mathbf{r}_1) \psi_i(\mathbf{r}_2), \end{aligned} \quad (1.12)$$

where the action of  $\hat{K}_j$  on  $\psi_i(\mathbf{r}_1)$  is defined as

$$\begin{aligned} \hat{K}_j \psi_i(\mathbf{r}_1) &\equiv \int d\mathbf{r}_2 \psi_j^*(\mathbf{r}_2) \frac{\hat{P}_{12}}{r_{12}} \psi_j(\mathbf{r}_2) \psi_i(\mathbf{r}_1) \\ &= \int d\mathbf{r}_2 \psi_j^*(\mathbf{r}_2) \frac{1}{r_{12}} \psi_j(\mathbf{r}_1) \psi_i(\mathbf{r}_2), \end{aligned} \quad (1.13)$$

and the permutation operator  $\hat{P}_{12}$  exchanges the electrons in occupied orbitals  $\psi_i(\mathbf{r}_1)$  and  $\psi_j(\mathbf{r}_2)$ . Using these definitions, one can piece together an expression for the orbital energy  $\epsilon_i$  in terms of the Fock operator  $\hat{f}$  and occupied molecular orbital

$\psi_i(\mathbf{r}_i)$ :

$$\begin{aligned}\epsilon_i &= \langle \psi_i | \hat{f} | \psi_i \rangle \equiv \langle \psi_i | \hat{h} + \sum_{j=1}^{N_{\text{elec}}} (\hat{J}_j - \hat{K}_j) | \psi_i \rangle \\ &= \langle \psi_i | \hat{h} | \psi_i \rangle + \sum_{i=1}^{N_{\text{elec}}} \langle \psi_i | \hat{V}_{\text{HF}} | \psi_i \rangle ,\end{aligned}\tag{1.14}$$

where we have introduced the HF operator  $\hat{V}_{\text{HF}} = \sum_{j=1}^{N_{\text{elec}}} (\hat{J}_j - \hat{K}_j)$ . Note that the restriction  $i \neq j$  is removed in the definition of  $\hat{V}_{\text{HF}}$  because the term  $i = j$  is  $\langle \psi_i | (\hat{J}_i - \hat{K}_i) | \psi_i \rangle = 0$ . Therefore, within HF theory, each electron will not experience its own field and is inherently free of self-interaction.

In order to perform an electronic structure calculation, an atom-centered basis set  $\{g_\mu(\mathbf{r})\}$  is chosen to construct the atomic orbitals that describe the electrons for each atom in the desired molecule. The molecular orbitals are then expressed as a linear combination of these atomic orbitals, and such a superposition is given as

$$\psi_i(\mathbf{r}) = \sum_{\mu=1}^{N_{\text{basis}}} C_{\mu i} g_\mu(\mathbf{r}) ,\tag{1.15}$$

where the Latin letters  $\{i, j, k, \dots\}$  denote occupied molecular orbitals, the Greek letters  $\{\mu, \nu, \sigma, \dots\}$  denote atomic orbitals, and  $\{C_{\mu i}\}$  is the set of coefficients that constructs molecular orbital  $\psi_i(\mathbf{r})$  from the basis functions  $g_\mu(\mathbf{r})$ . After the nuclear coordinates  $\mathbf{R}$ , the basis set  $\{g_\mu(\mathbf{r})\}$ , and an initial guess for the set of molecular coefficients are specified, a set of trial molecular orbitals and molecular wave function  $|\Phi_{\text{elec}}\rangle$  is constructed. In order to determine the unknown molecular orbital coefficients  $\{C_{\mu i}\}$ , the electronic energy is considered a functional of the molecular wave function as shown in Eq. 1.7, and  $\{C_{\mu i}\}$  are treated as variational parameters that are optimized to provide the best approximation to the energy and wave function possible for the desired basis set. The Hartree-Fock equations are then obtained by

minimizing the energy expression given in Eq. 1.7 with respect to variations of  $|\Phi_{\text{elec}}\rangle$  while forcing the underlying molecular orbitals to remain orthonormalized:

$$\langle \psi_i | \psi_j \rangle = \delta_{ij} . \quad (1.16)$$

After performing the variational analysis, Eq. 1.14 is the optimal set of integro-differential equations for the Hartree-Fock method.<sup>14</sup>

To derive the Hartree-Fock expressions for practical implementations of the method within electronic structure algorithms, Eq. 1.15 is substituted into Eq. 1.14 resulting in

$$\hat{f} \sum_{\mu=1}^{N_{\text{basis}}} C_{\mu i} g_{\mu}(\mathbf{r}) = \epsilon_i \sum_{\mu=1}^{N_{\text{basis}}} C_{\mu i} g_{\mu}(\mathbf{r}) . \quad (1.17)$$

Multiplying Eq. 1.17 by  $g_{\nu}^*(\mathbf{r})$  and integrating over all space transforms Eq. 1.17 into

$$\sum_{\mu=1}^{N_{\text{basis}}} F_{\nu\mu} C_{\mu i} = \epsilon_i \sum_{\mu=1}^{N_{\text{basis}}} S_{\nu\mu} C_{\mu i} , \quad (1.18)$$

where  $\mathbf{F}$  and  $\mathbf{S}$  are the Fock and overlap matrices with matrix elements

$$F_{\nu\mu} = \int d\mathbf{r} g_{\nu}^*(\mathbf{r}) \hat{f} g_{\mu}(\mathbf{r}) , \quad (1.19)$$

and

$$S_{\nu\mu} = \int d\mathbf{r} g_{\nu}^*(\mathbf{r}) g_{\mu}(\mathbf{r}) , \quad (1.20)$$

respectively. The Hartree-Fock equations in matrix form are expressed as

$$\mathbf{FC} = \mathbf{SC}\epsilon , \quad (1.21)$$

where  $\mathbf{C}$  is the  $N_{\text{basis}} \times N_{\text{basis}}$  matrix containing the coefficients for each molecular orbital expressed as

$$\mathbf{C} = \begin{bmatrix} C_{11} & C_{12} & \cdots & C_{1N_{\text{basis}}} \\ C_{21} & C_{22} & \cdots & C_{2N_{\text{basis}}} \\ \vdots & \vdots & \ddots & \vdots \\ C_{N_{\text{basis}}1} & C_{N_{\text{basis}}2} & \cdots & C_{N_{\text{basis}}N_{\text{basis}}} \end{bmatrix} , \quad (1.22)$$

and  $\epsilon$  is a diagonal matrix that holds the  $N_{\text{basis}}$  molecular orbital eigenvalues  $\epsilon_i$ . Since Eq. 1.21 is nonlinear, an iterative procedure is required to compute  $\mathbf{F}$ ,  $\mathbf{C}$ , and  $\epsilon$ , and such a self-consistent field (SCF) procedure is outlined in detail in Chapter 3 of Ref [14]. By following this procedure, the variational solution of Eq. 1.21 is obtained which yields the molecular orbital eigenvalues as well as the optimal set of molecular orbital coefficients.

This section concludes by pointing out that the Fock matrix elements in Eq. 1.19 are referred to in Chapter 2, and the solvation corrections that are presented in Sections 2.2.2 and 2.2.3 are incorporated into the Hartree-Fock calculations through this matrix. Furthermore, the solvation corrections presented in Sections 3.3.1 and 3.3.2 incorporated into the one-electron matrix,  $\mathbf{h}$ , the matrix elements of which are

$$\begin{aligned} h_{\nu\mu} &= \int d\mathbf{r} g_{\nu}^*(\mathbf{r}) \hat{h} g_{\mu}(\mathbf{r}) \\ &= \frac{1}{2} \int d\mathbf{r} g_{\nu}^*(\mathbf{r}) \nabla^2 g_{\mu}(\mathbf{r}) - \sum_{\alpha=1}^{N_{\text{nuc}}} \int d\mathbf{r} g_{\nu}^*(\mathbf{r}) \frac{Z_{\alpha}}{|\mathbf{r} - \mathbf{R}_{\alpha}|} g_{\mu}(\mathbf{r}) . \end{aligned} \quad (1.23)$$

Note that Eq. 1.23 is the atomic orbital representation of Eq. 1.9, and upon contracting  $\mathbf{h}$  with the one-electron density matrix  $\mathbf{P}$ , the one-electron energy is obtained. The density matrix  $\mathbf{P}$  is constructed using the molecular orbital coefficient matrix:

$$\mathbf{P} = \mathbf{C}\mathbf{C}^{\dagger} \quad (1.24)$$

with matrix elements

$$P_{\mu\nu} = \sum_i^{N_{\text{occ}}} C_{\mu i} C_{i\nu}^* , \quad (1.25)$$

where the summation includes only  $N_{\text{occ}}$  occupied molecular orbitals.

In Chapters 2 and 3, reference is made to the electronic charge density  $\rho_{\text{elec}}(\mathbf{r})$  that is expressed as

$$\begin{aligned} \rho_{\text{elec}}(\mathbf{r}) &= \sum_{i=1}^{N_{\text{elec}}} |\psi_i(\mathbf{r})|^2 \\ &= \sum_{\mu=1}^{N_{\text{basis}}} \sum_{\nu=1}^{N_{\text{basis}}} P_{\mu\nu} g_{\nu}^*(\mathbf{r}) g_{\mu}(\mathbf{r}) . \end{aligned} \tag{1.26}$$

The charge source for the solvation models presented in the coming Chapters arises from the electronic (Eq. 1.26) and nuclear (treated classically as point charges or with Gaussian functions) components of the molecule. Furthermore, modifying the Fock matrix with the solvation correction terms from Chapters 2 and 3 naturally includes the solvent response into the electronic structure calculations and also results in a wave function that is properly polarized by the continuum environment.

## 1.2 Møller-Plesset Second-Order Perturbation Theory

The HF method provided an avenue to more accurately solve the Schrödinger equation in the early 1930s, but was rarely utilized until the advent of computers in the 1950s. Though it has been proven to be a successful methodology for obtaining reasonably accurate energies and molecular orbitals, the HF method is unable to qualitatively describe some physical phenomena. For example, it incorrectly orders the first two ionization potentials of the  $\text{N}_2$  diatomic molecule relative to experiment.<sup>14</sup> More dramatically, the restricted HF method is unable to describe dissociation of diatomic hydrogen into two open shell hydrogen atoms ( $\text{H}_2 \rightarrow 2 \text{H}$ ), and though the unrestricted HF method describes the dissociation correctly in the asymptotic limit, it yields inaccurate potential energy curves.<sup>14</sup> The lack of accuracy exhibited in some cases results from the neglect of electron correlation, which can be thought of as the explicit electron-electron interactions that are ignored when a mean-field approach is

employed to compute the Coulomb energy. The corresponding electron correlation energy  $E_{\text{corr}}$  is defined as the difference between the exact energy  $E_{\text{exact}}$  of the system and the HF energy given by Eq. 1.7:

$$E_{\text{corr}} = E_{\text{exact}} - E_{\text{HF}} . \quad (1.27)$$

This section focuses on a “post-Hartree-Fock” method developed by C. Møller and M. Plesset for improving the HF energy by including electron correlation effects, and is called Møller-Plesset second-order perturbation theory (MP2).

As the name implies, MP2 arises through the application of Rayleigh-Schrödinger perturbation theory (RSPT) for a HF Hamiltonian. From RSPT, the total perturbed Hamiltonian  $\hat{H}$  is the sum of an unperturbed Hamilton  $\hat{H}_0$  and a perturbation  $\hat{V}$ , the strength of which is controlled by a parameter  $\lambda$ :

$$\hat{H} = \hat{H}_0 + \lambda \hat{V} , \quad (1.28)$$

Both the molecular wave function and energy corresponding to the solution of Eq. 1.1 that utilizes Eq. 1.28 for the Hamiltonian can be expressed as an infinite power series in  $\lambda$ :

$$|\Psi\rangle = \sum_m^{\infty} \lambda^m |\Psi^{(m)}\rangle , \quad (1.29)$$

and

$$E = \sum_m^{\infty} \lambda^m E^{(m)} , \quad (1.30)$$

respectively, where  $|\Psi^{(m)}\rangle$  and  $E^{(m)}$  denote the  $m^{\text{th}}$ -order correction terms. Substituting these expressions into Eq. 1.1, the time-independent Schrödinger equation is transformed into

$$\hat{H} \sum_m^{\infty} \lambda^m |\Psi^{(m)}\rangle = \sum_m^{\infty} \lambda^m E^{(m)} \sum_m^{\infty} \lambda^m |\Psi^{(m)}\rangle . \quad (1.31)$$

Expanding the expressions in Eq. 1.31 and equating powers of  $\lambda^m$  yields the appropriate correction terms for both the wave function and energy. The general zeroth-, first-, and second-order RSPT energies are

$$E_0^{(0)} = \langle \Psi_0 | \hat{H}_0 | \Psi_0 \rangle , \quad (1.32)$$

$$E_0^{(1)} = \langle \Psi_0 | \hat{V} | \Psi_0 \rangle , \quad (1.33)$$

and

$$E_0^{(2)} = \sum_{n \neq 0} \frac{|\langle \Psi_n | \hat{V} | \Psi_0 \rangle|^2}{E_0^{(0)} - E_n^{(0)}} , \quad (1.34)$$

respectively.

To apply this to the HF method, we consider the following perturbed Hamiltonian

$$\begin{aligned} \hat{H} &= \hat{H}_{\text{HF}} + \hat{V} \\ &= \sum_{i=1}^{N_{\text{elec}}} [\hat{h}(i) + \hat{V}_{\text{HF}}(i)] + \sum_{i=1}^{N_{\text{elec}}} \sum_{j>i}^{N_{\text{elec}}} \frac{1}{r_{ij}} - \sum_{i=1}^{N_{\text{elec}}} \hat{V}_{\text{HF}}(i) , \end{aligned} \quad (1.35)$$

where the perturbation  $\hat{V}$  is defined as the difference between the exact electron-electron Coulomb interaction and the approximate interaction given by the HF potential energy operator  $\hat{V}_{\text{HF}}$ . The unperturbed molecular wave function  $|\Psi_0\rangle$  is taken to be  $|\Phi_{\text{HF}}\rangle$ , and the zeroth-order energy is evaluated using  $\hat{H}_0 = \hat{H}_{\text{HF}}$  in Eq. 1.32 as<sup>14</sup>

$$E_0^{(0)} = \sum_{i=1}^{N_{\text{elec}}} \epsilon_i = \sum_{i=1}^{N_{\text{elec}}} \left( \langle \psi_i | \hat{h} | \psi_i \rangle \sum_{j=1}^{N_{\text{elec}}} \langle \psi_i \psi_j | | \psi_i \psi_j \rangle \right) , \quad (1.36)$$

where  $\epsilon_i$  is the molecular orbital energy corresponding to  $\psi_i(\mathbf{r})$ . The first-order energy correction is evaluated using Eq. 1.32 to be<sup>14</sup>

$$\begin{aligned} E_0^{(1)} &= \langle \Psi_0 | \sum_{j>i}^{N_{\text{elec}}} \frac{1}{r_{ij}} - \sum_{i=1}^{N_{\text{elec}}} \hat{V}_{\text{HF}}(i) | \Psi_0 \rangle \\ &= -\frac{1}{2} \sum_{i=1}^{N_{\text{elec}}} \sum_{j=1}^{N_{\text{elec}}} \langle \psi_i \psi_j | | \psi_i \psi_j \rangle \end{aligned} \quad (1.37)$$

Note that  $E_0^{(0)}$  from Eq. 1.36 is not equivalent to the HF energy given in Eq. 1.7: the contribution from the two-electron integrals  $\langle \psi_i \psi_j | | \psi_j \psi_i \rangle$  is double counted. Therefore the HF energy is the sum of the zeroth- and first-order energy corrections along with the nuclear-nuclear repulsion energy:

$$E_{\text{HF}} = E_0^{(0)} + E_0^{(1)} + E_{\text{nn}} . \quad (1.38)$$

To evaluate the second-order energy corrections given by Eq. 1.34, all wave functions  $|\Psi_n\rangle$  representing excitations from the ground state wave function must be considered. However, it is shown in Ref [14] that the two-particle nature of the perturbation will couple at most the singly- ( $|\Psi_1\rangle$ ) and doubly-excited ( $|\Psi_2\rangle$ ) states to the ground state. Furthermore, it is also shown in Ref [14] that the coupling between the ground state and singly-excited wave function vanishes ( $\langle \Psi_1 | \hat{V} | \Psi_0 \rangle = 0$ ), and the only non-vanishing term of the type  $\langle \Psi_n | \hat{V} | \Psi_0 \rangle$  results from the doubly-excited wave function. The second-order energy correction is evaluated as<sup>14</sup>

$$E_0^{(2)} = -\frac{1}{4} \sum_{i,j=1}^{N_{\text{occ}}} \sum_{a,b=1}^{N_{\text{virt}}} \frac{|\langle \psi_a \psi_b | | \psi_j \psi_i \rangle|^2}{\epsilon_a + \epsilon_b - \epsilon_i - \epsilon_j} , \quad (1.39)$$

where the Latin letters  $a$  and  $b$  are restricted to  $N_{\text{virt}}$  virtual molecular orbitals  $\psi_a(\mathbf{r})$  and  $\psi_b(\mathbf{r})$ , and the Latin letters  $i$  and  $j$  are restricted to  $N_{\text{occ}}$  occupied molecular orbitals  $\psi_i(\mathbf{r})$  and  $\psi_j(\mathbf{r})$ . The total number of molecular orbitals is equal to the number of atomic orbitals determined by the desired basis set, and therefore  $N_{\text{basis}} = N_{\text{occ}} + N_{\text{virt}}$ .

We conclude this section by mentioning that the investigations in Chapters 2 and 3 employ the HF method to compute the molecular wave function and corresponding energy, and then the second-order energy correction is computed using Eq. 1.39 and augmented to the HF energy upon convergence of the SCF procedure. The total MP2

energy is evaluated as

$$E_{\text{MP2}} = E_{\text{HF}} - \frac{1}{4} \sum_{i,j=1}^{N_{\text{occ}}} \sum_{a,b=1}^{N_{\text{virt}}} \frac{|\langle \psi_a \psi_b | \psi_j \psi_i \rangle|^2}{\epsilon_a + \epsilon_b - \epsilon_i - \epsilon_j}. \quad (1.40)$$

Obtaining the MP2 energy through Eq. 1.40 does not require a modification of the SCF procedure since it is computed as a “one-shot” correction to the HF energy using the optimal set of occupied and virtual molecular orbitals.

### 1.3 Density Functional Theory

As an alternative to the HF and MP2 methods, which utilize the molecular wave function and are aptly named “wave function methods”, density functional theory (DFT) provides a different route for solving Schrödinger’s equation using the electronic density  $\rho_{\text{elec}}(\mathbf{r})$ . Density functional methods appeared contemporaneously with wave function methods, and their origin can be traced to the Thomas-Fermi model, which provides a semiclassical description of many-body systems, and is formulated in terms of functionals that employ an electronic density. The Thomas-Fermi model relies on the assumption that within a small volume of space, the electronic density is uniform, but may vary between different volumes since electrons are nonuniformly distributed in real molecular systems. However, it was not until P. Hohenberg and W. Kohn laid a solid theoretical foundation (the two H-K theorems) that DFT would become the powerful technology it is currently. The first of these theorems demonstrates that the ground state properties of a quantum mechanical system are determined uniquely by the electron density that is a function of only three spatial coordinates, and the second theorem redefines a variational energy functional, such as the one used in the HF method, in terms of the electron density of the many-body quantum system.

DFT is formulated within the Born-Oppenheimer approximation in the same manner as HF theory, and utilizes the electronic Hamiltonian given by Eq. 1.4. The ground state wave function is constructed as a Slater determinant of molecular orbitals as described by Eq. 1.6, and the corresponding  $N_{\text{elec}}$  electron density is defined as

$$\rho_{\text{elec}}(\mathbf{r}) = N_{\text{elec}} \int d\mathbf{r}_2 \cdots \int d\mathbf{r}_N \Phi_{\text{DFT}}^*(\mathbf{r}, \mathbf{r}_2, \cdots \mathbf{r}_N) \Phi_{\text{DFT}}(\mathbf{r}, \mathbf{r}_2, \cdots \mathbf{r}_N). \quad (1.41)$$

A consequence of the H-K theorems is that Eq. 1.41 is reversible: given a variational ground state density  $\rho_{\text{elec}}(\mathbf{r})$ , it is possible to compute  $\Phi_{\text{DFT}}[\rho_{\text{elec}}]$ , which implies that the ground state wave function is a unique functional of the electron density. The ground state energy functional corresponding to  $\rho_{\text{elec}}(\mathbf{r})$  is expressed as

$$E_{\text{DFT}}[\rho_{\text{elec}}] = T_{\text{eff}}[\rho_{\text{elec}}] + E_{\text{en}}[\rho_{\text{elec}}] + E_{\text{ee}}[\rho_{\text{elec}}] + E_{\text{xc}}[\rho_{\text{elec}}], \quad (1.42)$$

where  $T_{\text{eff}}[\rho_{\text{elec}}]$  is the Kohn-Sham kinetic energy functional expressed as

$$T_{\text{eff}}[\rho_{\text{elec}}] = -\frac{1}{2} \sum_{i=1}^{N_{\text{elec}}} \langle \psi_i(\mathbf{r}) | \nabla^2 | \psi_i(\mathbf{r}) \rangle, \quad (1.43)$$

the attractive electron-nuclear interaction energy  $E_{\text{en}}[\rho_{\text{elec}}]$  is

$$E_{\text{en}}[\rho_{\text{elec}}] = - \sum_{\alpha=1}^{N_{\text{nuc}}} Z_{\alpha} \int d\mathbf{r} \frac{\rho_{\text{elec}}(\mathbf{r})}{|\mathbf{r} - \mathbf{R}_{\alpha}|}, \quad (1.44)$$

$E_{\text{ee}}[\rho_{\text{elec}}]$  is the repulsive electron-electron interaction energy, which is qualitatively the same physical quantity as in the HF method, and is evaluated as

$$E_{\text{ee}}[\rho_{\text{elec}}] = \frac{1}{2} \int d\mathbf{r} \int d\mathbf{r}' \frac{\rho_{\text{elec}}(\mathbf{r}) \rho_{\text{elec}}(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}, \quad (1.45)$$

and  $E_{\text{xc}}[\rho_{\text{elec}}]$  accounts for the exchange and correlation energies. Minimizing the energy functional given by Eq. 1.42 with respect to the set of orbitals used to create

$\rho_{\text{elec}}(\mathbf{r})$  results in the Kohn-Sham eigenvalue equation for the orbital  $\psi_i(\mathbf{r})$  and its energy  $\epsilon_i$ :

$$\left[ -\frac{1}{2}\nabla^2 - \sum_{\alpha=1}^{N_{\text{nuc}}} \frac{Z_{\alpha}}{|\mathbf{r} - \mathbf{R}_{\alpha}|} + \int d\mathbf{r}' \frac{\rho_{\text{elec}}(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} + \frac{\delta E_{\text{xc}}[\rho_{\text{elec}}]}{\delta \rho_{\text{elec}}(\mathbf{r})} \right] \psi_i(\mathbf{r}) = \epsilon_i \psi_i(\mathbf{r}). \quad (1.46)$$

Because of the non-interacting nature of the electrons occupying the Kohn-Sham orbitals  $\psi_i(\mathbf{r})$ ,  $\rho_{\text{elec}}(\mathbf{r})$  is constructed using Eq. 1.26 as in the HF method, and this quantity is utilized in the solvation models described in Chapter 2 to obtain the solvent polarization energy as well as in Eqs 1.42-1.46 to compute the ground state energy of the quantum-mechanical solute.

In principle, the DFT energy specified by Eq. 1.42 is exact given the exact exchange-correlation functional, but its form remains elusive, and significant effort is devoted to developing accurate approximations for  $E_{\text{xc}}[\rho_{\text{elec}}]$ . Although HF and DFT methods treat the electronic Coulomb interactions in a qualitatively similar manner, they differ in the description of exchange and correlation effects. For HF, the electron exchange energy arises naturally from the Slater determinant wave function, and as was discussed in Section 1.2, post-HF methods are necessary to incorporate electron correlation effects. Within the DFT methodology, the exchange and correlation effects are captured by  $E_{\text{xc}}[\rho_{\text{elec}}]$ , the form and accuracy of which depends on the chosen functional. A consequence of this for DFT is that the Coulomb interaction of an electron with its own field is not exactly negated by the exchange interaction as in HF, and therefore DFT methods suffer from self-interaction error which tends to elongate chemical bonds or result in poor descriptions of the ground state energy and the manifold of excited or charge-transfer states. As a way of correcting this, “hybrid” or “range-separated” DFT functionals have been developed to include exact HF exchange at long range while using a local or semi-local density functional to describe

the exchange at short range. Such a functional is employed in the investigations of Chapter 2, and the details of its form and performance can be found in Section 2.3.1.

## Chapter 2: The Solvated Electron

### 2.1 Liquid Microjet Experiments

The molecular structure of the vacuum/water interface, including the nature of the dissolved ions that reside there, has attracted attention and controversy for some time.<sup>15-19</sup> An ion that might exist at the interface is the hydrated or aqueous electron,  $e^-(aq)$ , a species long known in the radiation chemistry of water.<sup>20,21</sup> The energetics of  $e^-(aq)$  have attracted considerable attention in the context of DNA radiation damage, where it is known that strand breaks are considerably more common in “wet” than in “dry” DNA,<sup>22-24</sup> suggesting that most DNA damage by ionizing radiation is secondary, with the majority of the energy deposited into water rather than absorbed directly by DNA. This raises the question of whether  $e^-(aq)$  might be an intermediate responsible for such damage, but calculations presented here suggest that  $e^-(aq)$ , whether at the vacuum/water interface or in bulk water, lies too far below vacuum level to induce DNA strand breaks directly.

Motivated by a controversy regarding surface versus cavity electron-binding motifs in gas-phase  $(H_2O)_N^-$  clusters,<sup>8,25-29</sup> along with the recent development of high-pressure liquid microjet photoelectron spectroscopy,<sup>30,31</sup> several groups have reported photoelectron spectra for liquid-phase  $e^-(aq)$ .<sup>9,32-38</sup> These studies concur that the

vertical electron binding energy (VEBE) of  $e^-(aq)$  in bulk water lies in the range 3.3–3.7 eV, with most values around 3.3–3.4 eV,<sup>38</sup> consistent with extrapolations of gas-phase  $(H_2O)_N^-$  cluster photoelectron spectra<sup>39</sup> and detailed theoretical calculations.<sup>40</sup>

In one particular microjet experiment, however, a feature at 1.6 eV is observed, with a lifetime  $\gtrsim 100$  ps, and is attributed to  $e^-(aq)$  bound at the water/vacuum interface.<sup>9</sup> Attempts by others to replicate this result have found only a transient feature with a lifetime of  $\sim 100$  fs,<sup>34,36</sup> even in low kinetic energy experiments that ought to be relatively more sensitive to species solvated near the interface.<sup>36</sup> Nevertheless, there has since arisen much speculation that the putative “interfacial electron” at 1.6 eV might play a role in the radiation chemistry of DNA. Specifically, it is suggested that the much lower VEBE at the interface might provide the proper energetics for dissociative electron attachment to DNA, resulting in single strand breaks.<sup>9,17,41–43</sup>

Although a role in DNA damage for the “pre-solvated electron”, a hot precursor to  $e^-(aq)$ , has been proposed based on other experiments,<sup>21,44</sup> the liquid microjet experiment in Ref. [9] forms the basis of the first claims that the *equilibrated* species  $e^-(aq)$ , as opposed to an excited state, could play a role in radiation damage to DNA. This hypothesis requires that the interfacial species possess an energy  $\leq 2.5$  eV below vacuum level,<sup>41,42</sup> in contrast to the consensus value of  $\geq 3.3$  eV for the binding energy of thermalized  $e^-(aq)$  in bulk water.<sup>9,32–40</sup> Here, we investigate the interfacial hydrated electron with detailed theoretical calculations.

## 2.2 Modeling $e^-(aq)$ in Liquid and at an Interface

This section is comprised of three parts that demonstrate the techniques used to model  $e^-(aq)$  in bulk liquid water and also at a liquid-vapor interface. Section 2.2.1 discusses the one-electron pseudopotential models that were employed to capture the electron-water interactions. Additionally, Section 2.2.1 explains the techniques for performing the classical neat liquid water molecular dynamics simulations as well as the mixed quantum-classical simulations. A methodology that was designed for computing equilibrium solvation free energies for arbitrary dielectric environments is presented in Section 2.2.2, and the non-equilibrium version of this method for computing VIPs is given in Section 2.2.3.

### 2.2.1 Mixed Quantum-Classical Simulations

This section describes the ground state mixed quantum-classical simulations employing  $e^-(aq)$  pseudopotentials that were used to generate the data presented in Section 2.4. The simulations of an excess electron in bulk liquid water and at the liquid water/vapor interface employed two different one-electron pseudopotential models. One of these, which we call the “mean-field polarization model”, was developed by Turi and Borgis<sup>45</sup> and is based upon the non-polarizable SPC water model,<sup>46</sup> but includes a mean-field polarization potential  $\propto r^{-4}$ , where  $r$  is the electron–water distance. The other model, developed by Jacobson and Herbert,<sup>40</sup> is built upon the polarizable AMOEBA water model<sup>47</sup> and includes electron–water polarization explicitly and self-consistently. (Polarization is handled in AMOEBA by means of atom-centered inducible dipoles, and the induced dipoles in turn interact with

the wave function. Schrödinger’s equation for the wave function is iterated to self-consistency alongside the equations for the induced dipoles.<sup>40,48</sup>) These models have been used in myriad previous hydrated electron simulations.<sup>7,8,26,40,45,49–59</sup> Both afford accurate optical spectra for  $e^-(aq)$  in bulk water,<sup>40,45</sup> and in addition the explicit polarization model performs well in reproducing MP2/6-31(1+,3+)G\* benchmarks for  $(H_2O)_N^-$  clusters ( $N = 2–33$ ),<sup>40</sup> including both relative isomer energies and vertical electron binding energies (VEBEs). In the latter case, the mean deviation between the VEBE predicted by the one-electron model and the MP2 benchmark is 0.041 eV for 95 unique cluster structures.<sup>40</sup> With this level of accuracy, the explicit polarization model successfully accounts for trends observed in the gas-phase photoelectron spectra of  $(H_2O)_N^-$  clusters.<sup>8</sup>

Initial structures for the liquid water and liquid-vapor  $e^-(aq)$  simulations were extracted from equilibrated neat liquid water simulations utilizing either the SPC or the AMOEBA water model, as appropriate. These simulations were performed with  $N = 200$  water molecules in a periodic unit cell at  $T = 300$  K, where the unit cell size was set such that the density is  $\rho = 0.997$  g cm<sup>-3</sup>. For a cubic cell, this would correspond to 18.1617 Å on a side, but to simulate the water/vacuum interface we set  $L_x = L_y = 18.1617$  Å in the  $x$  and  $y$  directions, but  $L_z = 5L_x$  in the  $z$  direction, thus creating a periodic “slab” of water that is  $\approx 18$  Å thick in the  $z$  direction and surrounded by vacuum out to  $L_z = 90.8085$  Å. Standard three-dimensional Ewald summation is employed to sum the electrostatic interactions, so that the slab is actually periodically replicated in the  $z$  direction. If the slab of water has a non-zero dipole moment (which, in general, one should expect) then the potential energy must be corrected for the spurious electric fields that are produced

from the dipolar interactions between periodically-replicated slabs. The appropriate correction is derived in Ref. [60] for the case of an infinitely thin slab, and it is noted in Ref. [61] that higher-order terms are required for a slab having finite thickness, unless the spacing between periodic images is three to five times larger than the slab thickness. For this reason we set  $L_x = 5L_z$  and incorporate only the first-order dipolar correction to the Ewald sum.

The neat liquid water simulations described above were equilibrated for 100 ps, and starting structures for the  $e^-(aq)$  simulations were randomly chosen from the final 10 ps of the equilibration run. At  $t = 0$ , an electron is introduced into the system by turning on electron–water interactions using the appropriate pseudopotential. The electronic wave function is represented on a rectangular grid that spans the unit cell in the  $x$  and  $y$  directions and extends into vacuum in the  $z$  direction to a distance equal to the thickness of the slab (*i.e.*, the total span of the grid in the  $z$  direction equals twice the thickness of the slab). The grid spacing is  $\Delta x = \Delta y = \Delta z = 0.95 \text{ \AA}$  in all directions, as in previous work.<sup>40</sup>

For consistency, both the bulk and the interfacial hydrated electron simulations use slab boundary conditions. We do this for two reasons: first, we wish to allow the possibility that an electron in bulk water might spontaneously migrate to the interface; and second, so that even for  $e^-(aq)$  in bulk water we can measure the distance  $d$  between the centroid of the electron’s wave function and the Gibbs dividing surface (GDS), defined as the locus of points in the  $z$  direction where the solvent density has fallen to half its value in bulk solution. This surface is computed on-the-fly at each point  $(x, y)$  and at each time step, and is used to demarcate the liquid/vacuum interface for the purpose of computing  $d$ . Our coordinate system sets  $z = 0$  at the

GDS; this is the coordinate labeled  $d$  in the text, so for a slab whose thickness is  $\approx 18$  Å, the center of the slab corresponds to  $d \approx -9$  Å. [This explains the locus of points near  $d = -9$  Å in Fig 2.6(b).] Whereas trajectories for  $e^-(aq)$  at the water/vacuum interface are initialized from equilibrated neutral water, for which the electron naturally attaches at the interface, all of these trajectories eventually internalize and these equilibrated, internalized trajectories are then used as initial conditions for  $e^-(aq)$  in bulk water.

For each one-electron model, twenty independent trajectories for the interfacial species  $e^-(aq)$  were propagated at  $T = 300$  K using a home-built simulation code, FURRY, that is described in detail in Ref. [48], and has been used extensively for this purpose in other work.<sup>7,8,40,48,55</sup> Ewald summation is used for the  $e^-(aq)$  simulations in the same way that is described above for neutral liquid water, and a Nosé-Hoover thermostat is used to provide temperature control. The interfacial  $e^-(aq)$  simulations were propagated for 80 ps for the mean-field polarization model and for 40 ps in the case of the explicit polarization model in order to obtain statistics. The electron was determined to be internalized when the centroid of its wave function stabilizes around the center of the water slab ( $d \approx -9$  Å), which occurs on a time scale no longer than 25–35 ps. (Snapshots from one such trajectory are shown in Figure 2.4.) Following internalization, each trajectory was propagated for an additional 50 ps to examine whether the electron might return to the surface, but in no case was this observed and the internalized electron fluctuates no farther than 1.0–1.5 Å from the center of the water slab. Following this additional 50 ps of equilibration time, these trajectories were taken to be the initial ( $t = 0$ ) conditions for the simulation of  $e^-(aq)$  in bulk water.

## 2.2.2 Equilibrium Solvation

Two dielectric continuum models are employed to capture the long range electrostatic effects of the solvent, and are designed for use with an all-electron QM treatment of some number of explicit water molecules. The first is the “polarizable continuum model” (PCM)<sup>62,63</sup> that is designed to approximate isotropic (bulk) solvation effects and is discussed in significant detail in Section 3.2.1. The second model is a direct “Poisson equation solver”<sup>10,64,65</sup> (PEQS) which provides a numerical solution to Poisson’s equation, and allows for an anisotropic description of the dielectric environments that is necessary at a liquid-vapor interface. In each case, we use a “non-equilibrium” version of the continuum solvation model,<sup>66,67</sup> as appropriate for describing vertical (rather than an adiabatic) ionization processes. The state-specific non-equilibrium formalism is provided in Sections 3.2.2 and 3.3.2 within the PCM and PEQS frameworks, respectively. This section presents the formal details of the PEQS model for equilibrium solvation and the methodology for incorporating non-equilibrium solvation effects through a perturbative approach. Additionally, two algorithms are outlined for implementing the equilibrium and non-equilibrium PEQS models within an electronic structure code.

The PCMs employed in this work are appropriate for describing  $e^-$  (aq) in isotropic bulk solution, but not at the anisotropic liquid-vapor interface. At the interface, we resort to solving Poisson’s equation,

$$\hat{\nabla} \cdot [\epsilon(\mathbf{r}) \hat{\nabla} \varphi_{\text{tot}}(\mathbf{r})] = -4\pi \rho_{\text{sol}}(\mathbf{r}) , \tag{2.1}$$

for the total solute electrostatic potential  $\varphi_{\text{tot}}(\mathbf{r})$ . The quantity  $\rho_{\text{sol}}(\mathbf{r})$  is the solute charge density, comprised of both the electronic and nuclear components. This density

is embedded in a dielectric medium described by a spatially-inhomogeneous dielectric function,  $\epsilon(\mathbf{r})$ . Solution of Eq. 2.1 affords an electrostatic interaction energy  $E_{\text{elst}}$  expressed as

$$\begin{aligned} E_{\text{elst}} &= \frac{1}{8\pi} \int d\mathbf{r} \epsilon(\mathbf{r}) |\hat{\nabla} \varphi_{\text{tot}}(\mathbf{r})|^2 \\ &= \frac{1}{2} \int d\mathbf{r} \varphi_{\text{tot}}(\mathbf{r}) \rho_{\text{sol}}(\mathbf{r}) \\ &= \frac{1}{2} \int d\mathbf{r} \varphi_{\text{sol}}(\mathbf{r}) \rho_{\text{sol}}(\mathbf{r}) + \frac{1}{2} \int d\mathbf{r} \varphi_{\text{pol}}(\mathbf{r}) \rho_{\text{sol}}(\mathbf{r}) . \end{aligned} \quad (2.2)$$

The first term in the third equality of Eq. 2.2 is the internal self-energy of the solute given by

$$E_{\text{int}} = \frac{1}{2} \int d\mathbf{r} \varphi_{\text{sol}}(\mathbf{r}) \rho_{\text{sol}}(\mathbf{r}) , \quad (2.3)$$

and is equivalent to the explicit inter-particle Coulomb interactions that are already handled by the QM description of the solute. The second term in the third equality of Eq. 2.2 is the electrostatic contribution of the solvation free energy expressed as

$$G_{\text{el}} = \frac{1}{2} \int d\mathbf{r} \varphi_{\text{pol}}(\mathbf{r}) \rho_{\text{sol}}(\mathbf{r}) , \quad (2.4)$$

which is equivalent to the PCM solvation energy, up to discretization errors that can be made quite small,<sup>63,68</sup> and neglecting “volume polarization” arising from the tails of the QM wave function that penetrate beyond the solute cavity.<sup>63,69,70</sup> Volume polarization effects can be significant for anions but are mitigated here by the use of a large number of QM water molecules, such that the diffuse  $e^-(\text{aq})$  wave function has likely decayed to zero well before reaching the QM/continuum interface.

We developed a QM/Poisson approach to non-equilibrium solvation, based on the non-equilibrium polarization formalism described in Refs. [66] and [67]. The difference in the present work is that whereas Refs. [66] and [67] (along with other works on non-equilibrium solvation<sup>71-74</sup>) are couched in terms of a so-called *apparent*

*surface charge* (ASC) PCM,<sup>62,63</sup> in which only the two-dimensional solute/continuum interface (solute cavity surface) must be discretized, here we describe the polarization response to vertical detachment of the excess electron in terms of three-dimensional charge densities. This modification is necessary in order to describe the anisotropic case. The resulting QM/Poisson approach has been implemented in a locally-modified version of the Q-CHEM electronic structure program,<sup>75</sup> following the algorithm in Ref. [10], and is described in more detail below.

Following Ref. [10], the total density on the grid is separated into two parts: the contributions from the solute (electrons and nuclei) and the polarization charges that are induced by the solute-solvent interaction. In detail,

$$\begin{aligned}\rho_{\text{tot}}(\mathbf{r}) &= \rho_{\text{elec}}(\mathbf{r}) + \rho_{\text{nuc}}(\mathbf{r}) + \rho_{\text{pol}}(\mathbf{r}) \\ &= \rho_{\text{sol}}(\mathbf{r}) + \rho_{\text{pol}}(\mathbf{r}) ,\end{aligned}\tag{2.5}$$

where

$$\rho_{\text{sol}}(\mathbf{r}) = \rho_{\text{elec}}(\mathbf{r}) + \rho_{\text{nuc}}(\mathbf{r})\tag{2.6}$$

is the charge density of the solute within the solvent-excluded region. Below we provide the details for computing the solute charge density: the electronic part is taken to be positive, whereas the nuclear part is negative (opposite to the usual convention). The electronic part is expressed as

$$\rho_{\text{elec}}(\mathbf{r}) = \sum_{\mu\nu}^{N_{\text{basis}}} P_{\mu\nu} g_{\mu}(\mathbf{r}) g_{\nu}(\mathbf{r}) ,\tag{2.7}$$

where  $\mathbf{P}$  is the one-electron density matrix,  $g_{\mu}(\mathbf{r})$  is an atom-centered Gaussian basis function, and  $N_{\text{basis}}$  is the number of such functions determined by the choice of basis set. The solute nuclei are treated as classical point charges with strength  $-Z_{\alpha}$  on nucleus  $\alpha$ . Such a description of the nuclei can lead to singularities in grid-based

solutions of Poisson’s equation, and therefore the procedure used in Ref. [76] for classical atom-centered point charges is followed: we add the nuclear charge  $-Z_\alpha$  to the nearest Cartesian grid point but with a value  $-Z_\alpha/(\Delta V)$  that reflects the smearing of this charge over one voxel of the Cartesian grid. Here we take  $\Delta V = \prod_{\beta \in \{x,y,z\}} \Delta\beta$  to be the volume element on the rectangular Cartesian grid with spacings  $\Delta\beta$  for the respective coordinates. Formally, this means that we represent the nuclear contribution to the charge density as

$$\rho_{\text{nuc}}(\mathbf{r}) = - \sum_{\alpha}^{N_{\text{atoms}}} Z_{\alpha} \delta(\mathbf{r} - \mathbf{R}_{\alpha}) , \quad (2.8)$$

where  $\mathbf{R}_{\alpha}$  is the position vector of the Cartesian grid point nearest to nucleus  $\alpha$ , and the volume element  $\Delta V$  is contained within the delta function which has dimensions of inverse volume.

After forming  $\rho_{\text{sol}}(\mathbf{r})$  using Eqs. 2.6–2.8, we solve Eq. 2.1 for  $\varphi_{\text{tot}}(\mathbf{r})$  in the presence of a spatially-varying dielectric. To accomplish this we adapted the procedure outlined for plane-wave electronic structure codes in Refs. [64] and [10] for use in Gaussian basis set codes. The left side of Eq. 2.1 is rewritten to appear as Poisson’s equation for a solute in vacuum:

$$\nabla^2 \varphi_{\text{tot}}(\mathbf{r}) = -4\pi \left[ \rho_{\text{sol}}(\mathbf{r}) + \rho_{\text{pol}}(\mathbf{r}) \right] , \quad (2.9)$$

where the induced polarization charge density,  $\rho_{\text{pol}}(\mathbf{r})$ , has folded within it the effects of the inhomogenous dielectric medium, and takes the form

$$\rho_{\text{pol}}(\mathbf{r}) = \rho_{\text{iter}}(\mathbf{r}) + \left[ \frac{1 - \epsilon(\mathbf{r})}{\epsilon(\mathbf{r})} \right] \rho_{\text{sol}}(\mathbf{r}) . \quad (2.10)$$

The polarization charge density is separated into two components: an iterative charge density,  $\rho_{\text{iter}}(\mathbf{r})$ , and the solute charge density scaled by the factor  $\frac{1-\epsilon(\mathbf{r})}{\epsilon(\mathbf{r})}$ . [Note that

$\rho_{\text{pol}}(\mathbf{r}) = \rho_{\text{iter}}(\mathbf{r})$  at any point  $\mathbf{r}$  for which  $\epsilon(\mathbf{r}) = 1$ .] The iterative charge density is induced by the inhomogeneity of the dielectric function, and acts as an additional source term to the total charge density for Poisson's equation.<sup>10,64</sup> In detail, the iterative charge density is expressed as

$$\rho_{\text{iter}}^{(i)}(\mathbf{r}) = \frac{1}{4\pi} \left[ \hat{\nabla} \ln \epsilon(\mathbf{r}) \right] \cdot \left[ \hat{\nabla} \varphi_{\text{tot}}^{(i)}(\mathbf{r}) \right], \quad (2.11)$$

which explicitly shows the nonlinearity of Eq. 2.9, the solution of which requires an iterative procedure. The details of the conjugate gradient and finite difference methods employed to solve Eq. 2.9 are deferred until Chapter 4.

We now outline a general algorithm for iteratively solving Eq. 2.9 in order to obtain three important quantities necessary for incorporating equilibrium solvation effects into a QM calculation:  $\varphi_{\text{tot}}(\mathbf{r})$ ,  $\varphi_{\text{pol}}(\mathbf{r})$ , and  $\rho_{\text{pol}}(\mathbf{r})$ . Eq. 2.9 is solved for the solute electrostatic potential,  $\varphi_{\text{sol}}(\mathbf{r})$ , which is equivalent to setting  $\rho_{\text{pol}}(\mathbf{r}) = 0$  during the entire iterative procedure. Once  $\varphi_{\text{sol}}(\mathbf{r})$  is obtained, the iterative, polarization, and total charge densities are updated according to Eqs. 2.11 and 2.10 and 2.5, respectively. For the first iteration of the solver routines after obtaining  $\varphi_{\text{sol}}(\mathbf{r})$  in vacuum, corresponding to  $i = 0$  in Eq. 2.11,  $\varphi_{\text{tot}}(\mathbf{r}) = \varphi_{\text{sol}}(\mathbf{r})$ . Subsequently, Eq. 2.9 is solved for the total electrostatic potential using the total charge density at the current iteration, and this in turn is used to update the iterative charge density via Eq. 2.11. Following the procedure outlined in Refs. [10] and [64], a damping procedure is applied to stabilize the updates  $\rho_{\text{iter}}(\mathbf{r})$  in Eq. 2.11. This consists of a linear combination of the densities at iterations  $i$  and  $i + 1$ :

$$\begin{aligned} \rho_{\text{iter}}^{(i+1)}(\mathbf{r}) &= \frac{\eta}{4\pi} \left[ \hat{\nabla} \ln \epsilon(\mathbf{r}) \right] \cdot \left[ \hat{\nabla} \varphi_{\text{tot}}^{(i)}(\mathbf{r}) \right] + (1 - \eta) \rho_{\text{iter}}^{(i)}(\mathbf{r}) \\ &= \eta \rho_{\text{iter}}^{(i+1)} + (1 - \eta) \rho_{\text{iter}}^{(i)}. \end{aligned} \quad (2.12)$$

Both  $\varphi_{\text{tot}}(\mathbf{r})$  and  $\rho_{\text{iter}}(\mathbf{r})$  are considered converged when the Euclidean norm of a residual vector between the current and previous iterations are below a certain threshold. The solvent polarization free energy engendered by the QM solute charge density interaction with the dielectric medium is then evaluated by the expression in Eq. 2.4.

We end this section by discussing the procedure for properly incorporating solvent polarization effects into the QM calculations. Since  $\varphi_{\text{tot}}(\mathbf{r})$  and  $\rho_{\text{tot}}(\mathbf{r})$  are computed self-consistently at each SCF iteration,  $\varphi_{\text{pol}}(\mathbf{r})$  must be added as a correction to the Fock matrix. This correction,  $\Delta\mathbf{F}$  in matrix form, is equal to the functional derivative of the polarization free energy with respect to the one-electron density matrix  $\mathbf{P}$ , and the matrix elements are given by

$$\begin{aligned} \frac{\partial G_{\text{el}}}{P_{\mu\nu}} &= \Delta F_{\mu\nu} \\ &= \int d\mathbf{r} \varphi_{\text{pol}}(\mathbf{r}) g_{\mu}(\mathbf{r}) g_{\nu}(\mathbf{r}) , \end{aligned} \tag{2.13}$$

and the SCF energy is augmented with the quantity  $E_{\text{el}}$  expressed as

$$E_{\text{el}} = \int d\mathbf{r} \varphi_{\text{pol}}(\mathbf{r}) \rho_{\text{sol}}(\mathbf{r}) \tag{2.14}$$

that corresponds to the total solute-solvent interaction without accounting for the work to polarize the dielectric medium. The PEQS method procedure for equilibrium solvation is summarized as Algorithm 1 on Page 29.

### 2.2.3 Non-equilibrium Solvent Effects

To incorporate the solvent polarization response arising from the vertical ionization of an electron, we follow the non-equilibrium solvation formalism presented in Refs. [66] and [67], adapting it from the two-dimensional surface approach that is appropriate for apparent surface charge PCMs to the three-dimensional approach that is required here. Although originally presented in the context of polarizable continuum

---

**Algorithm 1:** PEQS algorithm for equilibrium solvation

---

**Data:** If feasible, precompute the value of all basis functions  $g_\mu(\mathbf{r})$  on the Cartesian grid.

*/\* This requires  $N_{\text{grid}} \times N_{\text{basis}}$  storage, and if this is not feasible then the basis function values can instead be computed on-the-fly at each SCF iteration. \*/*

- 1 **begin** SCF procedure
- 2     Initialize  $\Delta\mathbf{F} \equiv 0$ .
- 3     **repeat**  $n = 1, 2, \dots$  SCF iterations
- 4         Diagonalize the Fock matrix  $\mathbf{F} = \mathbf{F}_0 + \Delta\mathbf{F}$  to obtain the density matrix  $\mathbf{P}^{(n)}$ , where  $\mathbf{F}_0$  is the gas-phase Fock matrix and  $\Delta\mathbf{F}$  is the equilibrium solvation correction.
- 5         Compute  $\rho_{\text{sol}}(\mathbf{r})$ ,  $\rho_{\text{elec}}(\mathbf{r})$ , and  $\rho_{\text{nuc}}(\mathbf{r})$  on the Cartesian grid via Eqs. 2.6–2.8.
- 6         Compute  $\varphi_{\text{sol}}(\mathbf{r})$  via Eq. 2.9 with  $\epsilon(\mathbf{r}) = 1$ .
- 7         Form  $\epsilon(\mathbf{r})$  with  $\epsilon_{\text{solv}}$ .
- 8         *// The form of the dielectric function is discussed in Sections 2.3.4 and 3.4.4.*
- 9         Set  $\rho_{\text{tot}}(\mathbf{r}) = \rho_{\text{sol}}(\mathbf{r})$  and  $\varphi_{\text{tot}}(\mathbf{r}) = \varphi_{\text{sol}}(\mathbf{r})$ .
- 10         Compute  $\rho_{\text{iter}}^{(i=1)}(\mathbf{r})$  via Eq. 2.11 with  $\varphi_{\text{tot}}(\mathbf{r})$  and  $\epsilon(\mathbf{r})$ .
- 11         **repeat**  $i = 1, 2, \dots$  PEQS iterations
- 12             **begin** PEQS routines
- 13                 Compute  $\varphi_{\text{tot}}^{(i)}(\mathbf{r})$  via Eq. 2.9.
- 14                 Update  $\rho_{\text{iter}}^{(i+1)}(\mathbf{r})$  via Eq. 2.12 with  $\epsilon(\mathbf{r})$ ,  $\varphi_{\text{tot}}^{(i)}(\mathbf{r})$ , and  $\rho_{\text{iter}}^{(i)}(\mathbf{r})$ .
- 15                 Update  $\rho_{\text{pol}}(\mathbf{r})$  and  $\rho_{\text{tot}}(\mathbf{r})$  via Eqs. 2.10 and 2.5.
- 16             **until**  $\|\rho_{\text{iter}}^{(i+1)}(\mathbf{r}) - \rho_{\text{iter}}^{(i)}(\mathbf{r})\| < T_{\text{solver}}$
- 17             Compute  $\varphi_{\text{pol}}(\mathbf{r}) = \varphi_{\text{tot}}(\mathbf{r}) - \varphi_{\text{sol}}(\mathbf{r})$ .
- 18             Generate  $\Delta\mathbf{F}$  via Eq. 2.13 with  $\varphi_{\text{pol}}(\mathbf{r})$ .
- 19             Compute  $E_{\text{el}}$  via Eq. 2.14 and add it to the SCF energy.
- 20             **until** *DIIS error*  $< T_{\text{SCF}}$
- 21 Compute the equilibrium solvation free energy  $G_{\text{el}}$  via Eq. 2.4.

---

solvation effects on vertical *excitation* energies, it is noted in Refs. [66] and [67] that the formalism is equally valid for vertical *ionization*. The non-equilibrium effects for the vertical ionization potentials (VIPs) reported in Chapter 2 are computed through a perturbative approach, whereas the VIPs in Chapter 3 employ a more rigorous state-specific technique. An in-depth discussion of the origin of non-equilibrium effects for vertical excitation and ionization processes as well as the state-specific methodology within the context of PCMs and the PEQS method is reserved for Sections 3.2.2 and 3.3.2. Here we present the theoretical details for the perturbative approach that was adjusted to work within the PEQS method described in Section 2.2.2.

Within the perturbative scheme, the effective Hamiltonian for either the equilibrium reference state ( $i = 0$ ) or the non-equilibrium ionized state ( $i = 1$ ) is given by<sup>66</sup>

$$\hat{H}_i = \hat{H}_i^{\text{vac}} + \hat{V}_0 + \lambda \left( \hat{V}_i^{\text{fast}} - \hat{V}_0^{\text{fast}} \right), \quad (2.15)$$

where  $\hat{H}_i^{\text{vac}}$  is the gas-phase Hamiltonian for state  $i$  (either the Fock operator for HF or Kohn-Sham operator for DFT),  $\hat{V}_i$  is the reaction field operator that generates  $\varphi_{\text{pol},i}(\mathbf{r})$  for state  $i$ . The “fast” superscript designates that only the fast components of the solvent response are considered; see discussions in Section 3.2.2. The quantity  $\lambda$  is the usual perturbation parameter. For the reference state, in which  $i = 0$ , the effective Hamiltonian becomes

$$\hat{H}_0 = \hat{H}_0^{\text{vac}} + \hat{V}_0, \quad (2.16)$$

and the total free energy of the reference state, now denoted as  $G_0^{\text{el}}$ , is computed following the equilibrium solvation methodology described in Section 2.2.2. For the

case  $i = 1$  corresponding to the ionized state, the effective Hamiltonian becomes

$$\hat{H}_1 = \hat{H}_0^{\text{vac}} + \hat{V}_0 + \lambda (\hat{V}_1^{\text{fast}} - \hat{V}_0^{\text{fast}}) , \quad (2.17)$$

and the difference between the fast components of the ionized and reference state polarization responses is treated as a perturbation. It is important to note that the reference state reaction field operator,  $\hat{V}_0$ , appears in the effective Hamiltonian for the ionized state, and therefore two separate SCF calculations must be performed: one for the reference state following Algorithm 1 to generate and store in core memory  $G_0^{\text{el}}$  and  $\varphi_{\text{pol},0}(\mathbf{r})$ , and a second one to compute  $G_1^{\text{el}}$  for the ionized state, the details of which are described below.

After the performing the SCF calculation for the reference state, the second calculation begins for the ionized state. Algorithm 1 outlined in Section 2.2.2 requires a few adjustments for the non-equilibrium case, the first of which is a modification of Eq. 2.9 to compute the fast components of the solvent polarization response. The Marcus partitioning (MP) scheme<sup>67,77–80</sup> is employed to separate the polarization response into fast and slow components, and a more detailed discussion this procedure is provided in Sections 3.2.2 and 3.3.2. The following equations describe the MP scheme for the reference state, which splits the total polarization charge density into slow and fast components:

$$\rho_{\text{pol},0}(\mathbf{r}) = \rho_{\text{pol},0}^{\text{slow}}(\mathbf{r}) + \rho_{\text{pol},0}^{\text{fast}}(\mathbf{r}) \quad (2.18)$$

$$\rho_{\text{pol},0}^{\text{slow}}(\mathbf{r}) = \left( \frac{\epsilon_{\text{solv}} - \epsilon_{\text{opt}}}{\epsilon_{\text{solv}} - 1} \right) \rho_{\text{pol},0}(\mathbf{r}) , \quad (2.19)$$

where  $\epsilon_{\text{solv}}$  and  $\epsilon_{\text{opt}}$  are the equilibrium and optical solvent dielectric constants, respectively. The optical solvent dielectric is generally computed as  $\epsilon_{\text{opt}} = n^2$ , where  $n$  is the solvent index of refraction. The requisite fast and slow components of the

reference state polarization potentials are found using Eqs. 2.18 and 2.19, and then by exploiting the canonical definition of the electrostatic potential,

$$\varphi_{\text{pol},0}^{\text{slow/fast}}(\mathbf{r}) = \int d\mathbf{r}' \frac{\rho_{\text{pol},0}^{\text{slow/fast}}(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}, \quad (2.20)$$

the fast and slow components of the solvent response are obtained.

In order to obtain the fast components of the solvent response for the ionized state, Poisson's equation is modified according to

$$\nabla^2 \varphi_{\text{tot},1}(\mathbf{r}) = -4\pi [\rho_{\text{sol},1}(\mathbf{r}) + \rho_{\text{pol},0}^{\text{slow}}(\mathbf{r}) + \rho_{\text{pol},1}^{\text{fast}}(\mathbf{r})], \quad (2.21)$$

where  $\varphi_{\text{tot},1}(\mathbf{r})$ ,  $\rho_{\text{sol},1}(\mathbf{r})$ , and  $\rho_{\text{pol},1}^{\text{fast}}(\mathbf{r})$  are the total electrostatic potential, solute charge density, and fast component polarization charge density of the ionized state, and  $\rho_{\text{pol},0}^{\text{slow}}(\mathbf{r})$  is the slow component polarization charge density of the reference state; see discussions in Sections 3.2.2 and 3.3.2 for the origin of Eq. 2.21. In a similar manner as  $\rho_{\text{sol}}(\mathbf{r})$  served as the source of the reference state solvent polarization response in Eq. 2.9, the source for the ionized state solvent polarization response is  $\rho_{\text{sol},1}(\mathbf{r}) + \rho_{\text{pol},0}^{\text{slow}}(\mathbf{r})$ , where  $\rho_{\text{pol},1}^{\text{fast}}(\mathbf{r})$  takes the form

$$\rho_{\text{pol},1}^{\text{fast}}(\mathbf{r}) = \rho_{\text{iter},1}(\mathbf{r}) + \left[ \frac{1 - \epsilon(\mathbf{r})}{\epsilon(\mathbf{r})} \right] [\rho_{\text{sol},1}(\mathbf{r}) + \rho_{\text{pol},0}^{\text{slow}}(\mathbf{r})], \quad (2.22)$$

and  $\rho_{\text{iter},1}(\mathbf{r})$  is computed and updated by Eqs. 2.11 and 2.12. [Note: the function  $\epsilon(\mathbf{r})$  must employ the optical solvent dielectric when using Eqs. 2.11 and 2.12 for the ionized state.] The total charge density for the ionized state is then computed as

$$\rho_{\text{tot},1}(\mathbf{r}) = \rho_{\text{sol},1}(\mathbf{r}) + \rho_{\text{pol},0}^{\text{slow}}(\mathbf{r}) + \rho_{\text{pol},1}^{\text{fast}}(\mathbf{r}). \quad (2.23)$$

After converging  $\varphi_{\text{tot},1}(\mathbf{r})$  and  $\rho_{\text{pol},1}^{\text{fast}}(\mathbf{r})$ , the fast component of the solvent polarization response is computed as

$$\varphi_{\text{pol},1}^{\text{fast}}(\mathbf{r}) = \varphi_{\text{tot},1}(\mathbf{r}) - \varphi_{\text{sol},1}(\mathbf{r}) - \varphi_{\text{pol},0}^{\text{slow}}(\mathbf{r}) \quad (2.24)$$

The non-equilibrium solvation free energy for the ionized state is

$$G_1^{\text{el}} = E_1^{\text{int}} + W_0 + W_1^{(1)} + W_{0,1} , \quad (2.25)$$

where  $E_1^{\text{int}}$  is the self-energy of the ionized state and  $W_0$ ,  $W_1^{(1)}$ , and  $W_{0,1}$  are the free energy contributions from the reference state reaction field, the first-order perturbation correction, and a correction for the Coulomb interaction between the slow and fast polarization charges that arises due to the MP scheme, respectively. In detail, the latter three contributions are given by the following expressions:

$$W_0 = \frac{1}{2} \int d\mathbf{r} \varphi_{\text{pol},0}(\mathbf{r}) \rho_{\text{sol},0}(\mathbf{r}) , \quad (2.26)$$

$$W_1^{(1)} = \frac{1}{2} \int d\mathbf{r} \varphi_{\text{sol},1}(\mathbf{r}) \left[ \rho_{\text{pol},1}^{\text{fast}}(\mathbf{r}) - \rho_{\text{pol},0}^{\text{fast}}(\mathbf{r}) \right] , \quad (2.27)$$

and

$$W_{0,1} = \frac{1}{2} \int d\mathbf{r} \varphi_{\text{pol},0}^{\text{slow}}(\mathbf{r}) \left[ \rho_{\text{pol},1}^{\text{fast}}(\mathbf{r}) - \rho_{\text{pol},0}^{\text{fast}}(\mathbf{r}) \right] . \quad (2.28)$$

Eq. 2.28 describes a physical situation in which the “earlier” slow component of the reference state solvent response affects the “current” fast response for the ionized state;<sup>67,72,77,81</sup> see discussion in Section 3.2.2 for details regarding the concept of fictitious polarization time. Finally, the Fock matrix correction for the ionized state is identical to Eq. 2.13 because the reference state reaction field is used in Eq. 2.17. The perturbative correction terms  $W_1^{(1)}$  and  $W_{0,1}$  only contribute to the ionized state solvation free energy through Eq. 2.25 and do not appear as a correction to the Fock matrix. An outline of the non-equilibrium PEQS procedure is presented as Algorithm 2 on Page 34.

---

**Algorithm 2:** Non-equilibrium PEQS procedure using the MP scheme

---

```

1 begin Reference state procedure
2   ┌ Proceed with Algorithm 1 to compute the relevant energies, electrostatic
   └ potentials, and charge densities and save to disk.
   Input:  $G_0^{\text{el}}$ ,  $\varphi_{\text{pol},0}(\mathbf{r})$ ,  $\rho_{\text{pol},0}(\mathbf{r})$ , and  $\rho_{\text{sol},0}(\mathbf{r})$ .
3 begin Ionized state procedure
4   Initialize  $\Delta\mathbf{F} \equiv 0$ .
5   repeat  $n = 1, 2, \dots$  SCF iterations
6     Diagonalize the Fock matrix  $\mathbf{F} = \mathbf{F}_0 + \Delta\mathbf{F}$  to obtain the density matrix
        $\mathbf{P}^{(n)}$ , where  $\mathbf{F}_0$  is the gas-phase Fock matrix and  $\Delta\mathbf{F}$  is the
       non-equilibrium solvation correction.
7     Compute  $\rho_{\text{sol},1}(\mathbf{r})$ ,  $\rho_{\text{elec},1}(\mathbf{r})$ , and  $\rho_{\text{nuc},1}(\mathbf{r})$  via Eqs. 2.6–2.8.
8     Compute  $\varphi_{\text{sol},1}(\mathbf{r})$  via Eq. 2.9 with  $\epsilon(\mathbf{r}) = 1$ .
9     Form  $\epsilon(\mathbf{r})$  with  $\epsilon_{\text{opt}}$ .
10    Compute  $\rho_{\text{pol},0}^{\text{slow}}(\mathbf{r})$  and  $\varphi_{\text{pol},0}^{\text{slow}}(\mathbf{r})$  via Eqs. 2.18 and 2.20.
11    Set  $\rho_{\text{tot},1}(\mathbf{r}) = \rho_{\text{sol},1}(\mathbf{r}) + \rho_{\text{pol},0}^{\text{slow}}(\mathbf{r})$  and  $\varphi_{\text{tot},1}(\mathbf{r}) = \varphi_{\text{sol},1}(\mathbf{r}) + \varphi_{\text{pol},0}^{\text{slow}}(\mathbf{r})$ .
12    Compute  $\rho_{\text{iter},1}^{(i=1)}(\mathbf{r})$  via Eq. 2.11 with  $\epsilon_{\text{opt}}(\mathbf{r})$  and  $\varphi_{\text{tot},1}(\mathbf{r})$ .
13    repeat  $i = 1, 2, \dots$  PEQS iterations
14      begin PEQS routines
15        ┌ Compute  $\varphi_{\text{tot},1}^{(i)}(\mathbf{r})$  via Eq. 2.21.
16        └ Update  $\rho_{\text{iter},1}^{(i+1)}(\mathbf{r})$  via Eq. 2.12 with  $\epsilon_{\text{opt}}(\mathbf{r})$ ,  $\varphi_{\text{tot},1}^{(i)}$ , and  $\rho_{\text{iter},1}^{(i)}$ .
17        Update  $\rho_{\text{pol},1}^{\text{fast}}(\mathbf{r})$  and  $\rho_{\text{tot},1}(\mathbf{r})$  via Eqs. 2.22 and 2.23.
18      until  $\|\rho_{\text{iter},1}^{(i+1)}(\mathbf{r}) - \rho_{\text{iter},1}^{(i)}(\mathbf{r})\| < T_{\text{solver}}$ 
19      Generate  $\Delta\mathbf{F}$  via Eq. 2.13 with  $\varphi_{\text{pol},0}(\mathbf{r})$ .
20      Compute  $\varphi_{\text{pol},1}^{\text{fast}}(\mathbf{r})$  via Eq. 2.24.
21      Compute  $E_{\text{el}}$  and add it to the SCF energy.
22    until DIIS error  $< T_{\text{SCF}}$ 
23 Compute the non-equilibrium polarization free energy  $G_1^{\text{el}}$  via Eq. 2.25.
24 Compute  $\text{VEBE} = G_1^{\text{el}} - G_0^{\text{el}}$ 

```

---

## 2.3 Electronic Structure Methods

In Section 2.2.1 from above, the procedure for performing molecular dynamics simulations of  $e^-(aq)$  utilizing the one-electron pseudopotentials was explained. A discussion of the computational methods required for performing the *ab initio* calculations is provided here. Long-range-corrected DFT (LRC-DFT) and RI-MP2 methods are used for the QM description of  $e^-(aq)$ , and the process of obtaining the optimal range-separation parameter for the LRC-DFT functional is the subject of Section 2.3.1. Convergence tests of the LRC-DFT and RI-MP2 results are performed with respect to the choice basis set in Section 2.3.2. Finally, Sections 2.3.3 and 2.3.4 conclude this section with a discussion of the relevant details for the non-equilibrium solvation calculations.

### 2.3.1 DFT and MP2 Methods

DFT calculations reported in this work use a long-range-corrected (LRC) version of the “BOP” density functional. The BOP functional is comprised of Becke’s generalized-gradient exchange functional (B88),<sup>82</sup> along with the “one-parameter progressive” (OP) correlation functional,<sup>83</sup> which is similar to the more familiar Lee-Yang-Parr (LYP) correlation functional.<sup>84</sup> The resulting exchange-correlation functional is

$$E_{xc}^{\text{LRC-}\mu\text{BOP}} = E_c^{\text{OP}} + E_x^{\mu\text{B88,SR}} + E_x^{\text{HF,LR}}, \quad (2.29)$$

where “SR” and “LR” express that only the short-range or long-range parts of the Coulomb operator are employed when evaluating these components of the energy. The functional in Eq. 2.29 has been implemented<sup>85,86</sup> in the Q-CHEM electronic structure program,<sup>75</sup> where it is called LRC- $\mu$ BOP.<sup>87</sup> This functional has been shown

to afford accurate VEBEs for  $(\text{H}_2\text{O})_N^-$  clusters as compared to correlated wave function benchmarks.<sup>6,48</sup>

The range-separation parameter,  $\mu$ , dictates the partition between what is considered short-range versus long-range in the Coulomb operator. (Roughly speaking, this partition occurs on a length scale  $\sim \mu^{-1}$ .) Baer and co-workers<sup>88</sup> suggest a non-empirical “tuning” procedure to determine  $\mu$ , according to the criterion

$$\text{IE}_{\Delta\text{SCF}}(\mu) = -\epsilon_{\text{SOMO}}(\mu) . \quad (2.30)$$

Here,  $\epsilon_{\text{SOMO}}$  represents the self-consistent field (SCF) orbital energy for the singly-occupied molecular orbital (SOMO) and  $\text{IE}_{\Delta\text{SCF}}$  is the ionization energy determined from a “ $\Delta\text{SCF}$ ” approach, *i.e.*,

$$\text{IE}_{\Delta\text{SCF}}(\mu) = E_{\text{neutral}}(\mu) - E_{\text{anion}}(\mu) . \quad (2.31)$$

(The neutral and anion energies are computed at the geometry of the anion.) The condition in Eq. 2.30 is a rigorous one for the exact Kohn-Sham functional, and because the ionization energy (IE) controls the long-range decay of the SCF wave function, this tuning procedure ensures an exchange-correlation potential with proper long-range behavior.<sup>89</sup>

Unfortunately, the optimally-tuned value of  $\mu$  is known to exhibit a strong dependence on system size,<sup>5,87,90,91</sup> even for chemically homologous systems. In previous QM/MM calculations of the bulk and interfacial hydrated electron, we have shown that  $\mu$  should be tuned separately for each distinct size of the QM region.<sup>5</sup> This leads to an electronic absorption spectrum for  $e^-(\text{aq})$  in bulk water that is in excellent agreement with the experimental spectrum, whereas values of  $\mu$  that are tuned for

small clusters afford a significantly shifted spectrum, as do global hybrid functionals such as B3LYP.<sup>5,57</sup>

We use QM regions that range in size from a radius of 5.5–8.0 Å around the centroid of the spin density, and we tune  $\mu$  separately for each different size. Figure 2.1 demonstrates the tuning procedure for three snapshots containing 47–50 QM water molecules, corresponding to a QM radius of 7.5 Å, extracted from an *ab initio* simulation of interfacial  $e^-(aq)$ <sup>4</sup>. Each of the three snapshots is separated in time by 4 ps. As  $\mu$  increases (and thus Hartree-Fock exchange is introduced on increasingly short length scales), the SOMO is increasingly destabilized. This is the result of eliminating self-interaction error associated with the unpaired electron, which is primarily a long-range phenomena for this particular system since the unpaired electron exists largely outside of the valence-electron regions of the water molecules. The IE, in contrast, is less sensitive to changes in  $\mu$ . In order to satisfy the criterion of Eq. 2.30 in an average way across all three snapshots in Figure 2.1, we set  $\mu = 0.18 a_0^{-1}$ . The value of  $\mu$  used for each QM size for both the bulk and interfacial species was determined independently using this procedure. These “tuned” values of  $\mu$  are listed in Table 2.1.

In a recent study of the optical spectrum of  $e^-(aq)$  using time-dependent DFT (TD-DFT),<sup>5</sup> we showed that this spectrum is sensitive to the value of  $\mu$ , and that the spectrum is significantly blue-shifted (as compared to experiment) for  $\mu \gtrsim 0.3 a_0^{-1}$ . This is despite the fact that values of  $\mu > 0.3 a_0^{-1}$  afford accurate VEBEs for small, gas-phase  $(H_2O)_N^-$  clusters.<sup>6,48</sup> These discrepancies in the optical spectrum (with respect to experiment) are not reconciled by increasing the size of the QM region.<sup>5</sup> In Ref. [5], a range-separation parameter  $\mu = 0.165 a_0^{-1}$  was ultimately chosen based on the tuning procedure described above (see Table 2.1). We emphasize that this

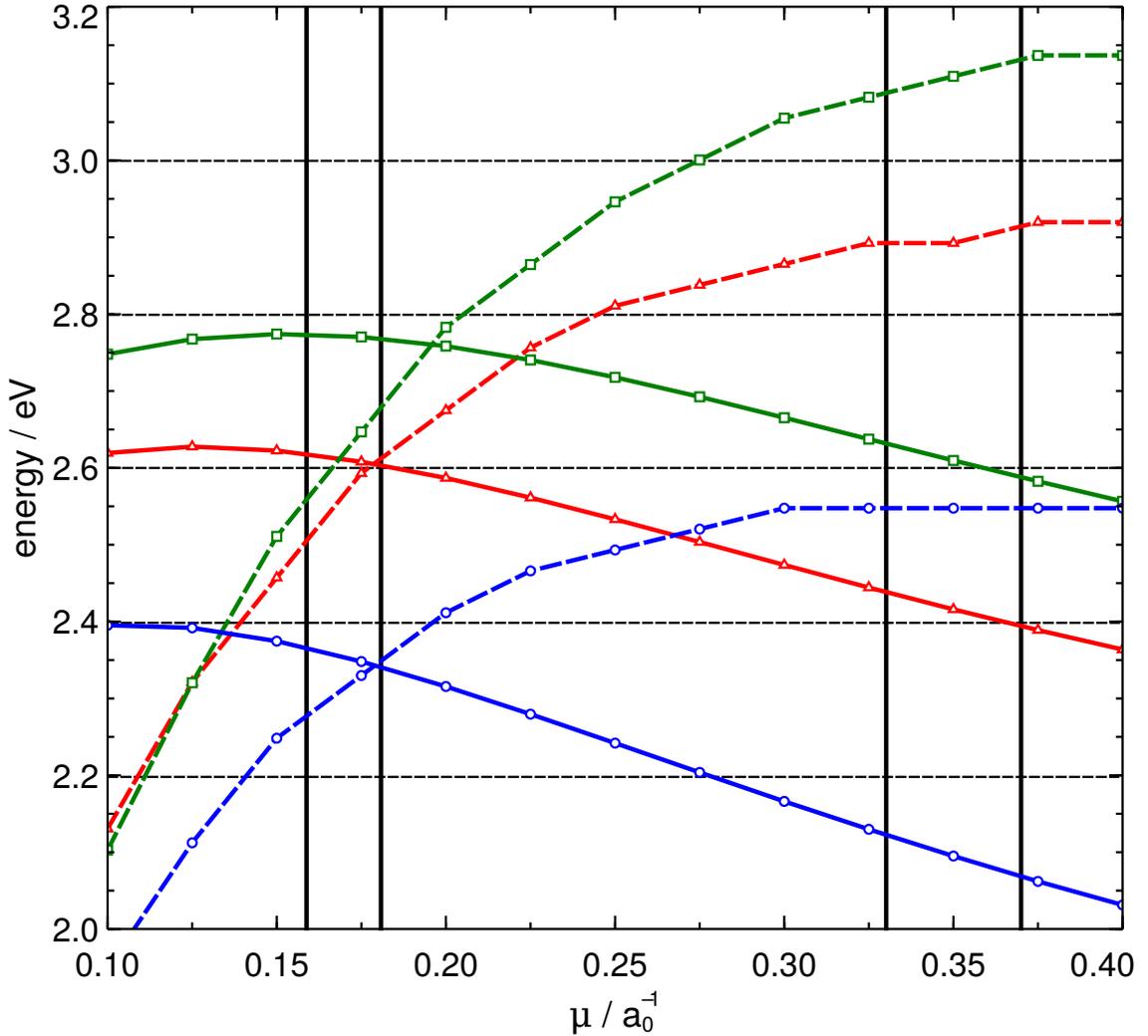


Figure 2.1: Illustration of the tuning procedure for LRC- $\mu$ BOP/6-31++G\*, demonstrated for three different  $(\text{H}_2\text{O})_N^-$  geometries (with  $N = 47-50$ ) that are shown in red, blue, and green and were extracted from an *ab initio* interfacial  $e^-$ (aq) simulation<sup>4</sup>. The solid curves represent  $\text{IE}_{\Delta\text{SCF}}(\mu)$  and the dashed curves are  $-\epsilon_{\text{SOMO}}(\mu)$ . Where the dashed and solid curves of a particular color intersect is the “optimally tuned” value of  $\mu$  for that particular snapshot, and in this case we choose  $\mu = 0.18 \text{ a}_0^{-1}$  as the best compromise value for these three snapshots. This value is indicated by a vertical line, and the other vertical lines indicate other values of  $\mu$  that have been employed in previous hydrated-electron studies using this functional. These include  $\mu = 0.165 \text{ a}_0^{-1}$  (Ref. [5]),  $\mu = 0.33 \text{ a}_0^{-1}$  (Ref. [6]), and  $\mu = 0.37 \text{ a}_0^{-1}$  (Refs. [7] and [8]).

System	QM region radius/Å	$\langle N \rangle$	$\mu / a_0^{-1}$	VEBE / eV		PCM Correction / eV		Poisson Correction/ eV	
				DFT	RI-MP2	DFT	RI-MP2	DFT	RI-MP2
interface	5.5	21	0.205	$2.05 \pm 0.38$	$1.96 \pm 0.37$	–	–	1.05	1.11
interface	6.0	26	0.195	$2.17 \pm 0.39$	$2.05 \pm 0.38$	–	–	1.01	0.97
interface	6.5	32	0.190	$2.22 \pm 0.37$	$2.11 \pm 0.34$	–	–	0.94	0.95
interface	7.0	39	0.185	$2.30 \pm 0.33$	$2.20 \pm 0.33$	–	–	0.87	0.88
interface	7.5	47	0.180	$2.37 \pm 0.34$	$2.26 \pm 0.35$	–	–	0.77	0.83
interface	8.0	58	0.170	$2.45 \pm 0.33$	$2.35 \pm 0.33$	–	–	0.74	0.75
bulk	5.5	29	0.195	$2.12 \pm 0.29$	$2.04 \pm 0.31$	1.40	1.40	1.21	1.14
bulk	6.0	37	0.180	$2.31 \pm 0.32$	$2.22 \pm 0.34$	1.30	1.20	1.05	0.99
bulk	6.5	47	0.175	$2.36 \pm 0.31$	$2.29 \pm 0.37$	1.21	1.08	1.01	0.88
bulk	7.0	58	0.170	$2.40 \pm 0.32$	$2.33 \pm 0.33$	1.13	1.08	0.97	0.86
bulk	7.5	71	0.165	$2.49 \pm 0.35$	$2.39 \pm 0.32$	1.10	1.03	0.86	0.81
bulk	8.0	86	0.160	$2.58 \pm 0.33$	$2.51 \pm 0.31$	0.95	0.94	0.80	0.71

Table 2.1: Average number of water molecules,  $\langle N \rangle$ , and optimally-tuned range-separation parameters,  $\mu$ , for LRC- $\mu$ BOP. Also shown are the average VEBEs for these QM regions, computed at the LRC- $\mu$ BOP/ and RI-MP2/6-31++G\* levels of theory. (Uncertainties represent one standard deviation, averaged over snapshots extracted from the simulations.) In these calculations, the QM region is treated as a gas-phase  $(\text{H}_2\text{O})_N^-$  cluster, *i.e.*, the QM regions are the same as in the calculations presented in Table 2.6 but the latter calculations include PCM boundary conditions. Comparison of these “gas phase” values with the results in Table 2.6 affords the PCM corrections that are listed here, which illustrate the magnitude of the correction for long-range electrostatic effects.

tuning procedure needs to be performed separately at each unique QM system size, and we have done so here. Evident from Table 2.1 is a small but clear anti-correlation between the number of QM water molecules and the optimally-tuned value of  $\mu$ .

In addition to listing the optimally-tuned values of  $\mu$ , Table 2.1 provides the average VEBE computed for the gas-phase QM region, *i.e.*, a negatively-charged water cluster with vacuum boundary conditions. As in Table 2.6, these calculations are performed at both the LRC- $\mu$ BOP/6-31++G\* level of theory and the resolution-of-identity<sup>92,93</sup> (RI)-MP2/6-31++G\* level. For each snapshot (extracted from the QM/MM simulations reported in Refs. [11] and [4]), the Turi-Borgis pseudopotential model<sup>45</sup> was employed to estimate the centroid of the spin density, and then all water molecules contained within a sphere centered at this point were included in the QM region for the DFT or RI-MP2 calculations. We vary the radius of this sphere to watch the convergence of the VEBE with respect to the size of the QM region. The average number of H<sub>2</sub>O molecules contained in this sphere is shown as  $\langle N \rangle$  in Table 2.1.

Note that the only difference between the calculations reported in Table 2.1 and those in Table 2.6 is that in the latter case, continuum boundary conditions are applied in order to obtain quantitative agreement with experiment, whereas in Table 2.1 the QM region is treated as a gas-phase (H<sub>2</sub>O)<sub>N</sub><sup>-</sup> cluster. Comparison of the two data sets allows us to assess the quantitative importance of the boundary conditions. We find that the non-equilibrium solvation model increases the VEBE by  $\approx 0.7$ – $1.1$  eV for the interfacial species and by  $\approx 0.7$ – $1.4$  eV for e<sup>-</sup>(aq) in bulk water. Notably, the VEBE converges more rapidly, as a function of the size of the QM region, for calculations that use continuum boundary conditions (Table 2.6) than it does for calculations with

Table 2.2: Average VEBEs computed using two different basis sets, for the interfacial hydrated electron (computed using the DFT/Poisson method) and the bulk hydrated electron (computed using the DFT/PCM method).

System	QM radius/ Å	VEBE / eV	
		6-31++G*	6-31(1+,3+)G*
interface	5.5	3.15 ± 0.41	3.18 ± 0.43
interface	7.5	3.14 ± 0.41	3.16 ± 0.44
bulk	5.5	3.52 ± 0.31	3.55 ± 0.33
bulk	7.5	3.59 ± 0.35	3.62 ± 0.30

vacuum boundary conditions (Table 2.1), suggesting that the dielectric continuum is providing the desired long-range electrostatic effect of the bulk solvent.

### 2.3.2 Basis Set Convergence

We used the 6-31++G\* basis set to compute VEBEs in Tables 2.1 and 2.6, as in previous simulations of the optical spectrum of  $e^-(aq)$ .<sup>5,7,57</sup> The choice of basis set has been carefully examined in previous work<sup>5,57,89,94-96</sup> (see Ref. [57] in particular), with the conclusion that in a condensed-phase aqueous environment, convergence of the VEBE is rapid and requires only a single set of diffuse basis functions, despite the fact that the electron occupies an excluded volume in the structure of liquid water. Additional diffuse shells have a very limited effect.<sup>57</sup>

Nevertheless, to test basis-set effects in the present context we computed VEBEs using the 6-31(1+,3+)G\* basis set,<sup>95</sup> which includes two additional diffuse shells as compared to 6-31++G\*, with exponents of  $\zeta_1 = 0.2700058 a_0^{-2}$  and  $\zeta_2 = 0.0845 a_0^{-2}$  that are reduced by successive factors of 3.2 relative to the most diffuse exponent in the 6-31++G\* basis set. Results are listed in Table 2.2. Only minor deviations of

$\approx 0.03$  eV are observed between the two basis sets, and we thus regard the 6-31++G\* basis set as being sufficiently diffuse.

The DFT and RI-MP2 calculations described above were performed on snapshots extracted from the QM/MM simulations reported in Refs. [11] and [4]. Although we use large QM regions in these single-point calculations—up to a radius of 8.0 Å, which encompasses  $\approx 90$  QM water molecules—longer-range polarization effects must still be incorporated in order to obtain a VEBE that is in quantitative agreement with the consensus value for  $e^-(aq)$  in bulk water. This is evident in the fact that the VEBE changes by  $\sim 1$  eV between the  $N = 600$  simulation cell and the infinite-dilution limit; see Figure 2.10. (The difference is smaller, but still not negligible, when a polarizable pseudopotential model is employed.<sup>40</sup>) This speaks to the importance of very long-range Coulomb effects on the VEBE.

### 2.3.3 Non-equilibrium Polarizable Continuum Models

The specific version of PCM employed here for the bulk species  $e^-(aq)$  is the “integral equation formalism”, IEF-PCM.<sup>63,68,97–99</sup> The *equilibrium* version of IEF-PCM affords aqueous solvation free energies within a few kcal/mol of experimental values for small to medium-size molecules and ions, if non-electrostatic corrections are applied,<sup>100–102</sup> but accurate solvatochromatic shifts can be obtained without such corrections, using the non-equilibrium version of the model.<sup>66</sup> The non-equilibrium model is designed to describe electronic polarization of the solvent upon vertical excitation or ionization of the solute.<sup>71–74</sup> In other words, while the ionization process may be vertical (in the sense that the vibrational and orientational degrees of freedom of the solvent molecules are not allowed to relax or respond in any way), electronic

polarization of the solvent should take place on the same time scale as excitation or ionization and ought to be considered, even in a continuum treatment. This effect is represented by re-polarizing the continuum solvent in the excited state of the solute using the solvent’s *optical* (rather than static) dielectric constant,  $\epsilon_\infty = n^2$ , where  $n$  is the solvent’s index of refraction. The total polarization response of the solvent is partitioned into a “fast” component (due to the electrons and described using  $\epsilon_\infty = 1.8$  for water) and a “slow” component (due to nuclear reorganization and described using  $\epsilon_{\text{static}} = 78.39$  for water). Details of the PCM and its implementation in Q-CHEM<sup>75</sup> can be found in Sections 3.2.1 and 3.2.2 as well as Refs. [66] and [67].

The predictions of PCM-type methods are sensitive to the details of how the “solute cavity” (*i.e.*, the boundary between the atomistic QM region and the continuum solvent) is constructed. Most often this cavity is taken to be a union of atom-centered spheres, whose radii may be considered to be parameters of the model. The calculations presented here, however, include a large number of explicit water molecules, obviating the need for such crude boundary conditions. These are suspect anyway for a species like  $e^-(\text{aq})$  where one electron is clearly outside of the van der Waals radii of the atoms and would therefore directly inhabit in continuum region. Moreover, in the presence of explicit solvent molecules, the aforementioned “van der Waals” cavity construction places dielectric medium between the explicit solvent molecules, which should not be there since Coulomb interactions are treated explicitly in the QM region. A more detailed investigation of this problem is presented in Sections 3.4.3 and 3.4.4.

As an alternative, we use a QM/continuum interface consisting of single sphere whose radius is selected to be 2.75 Å larger than the radius of the QM region. (The

QM radii are reported in Table 2.6.) This cavity is discretized using a Lebedev integration grid with 5,294 points. The 2.75 Å buffer reflects the estimated diameter of a water molecule, according to effective ionic radii for the isoelectronic ions  $\text{O}^{2-}$ ,  $\text{OH}^-$ , and  $\text{H}_3\text{O}^+$ .<sup>103</sup> In addition, our selection criterion for placing  $\text{H}_2\text{O}$  molecules in the QM region is to do so if any one of the three atoms lies within the specified QM radius, and as such there are cases where a hydrogen nucleus ends up as much as 2.75 Å farther away from the origin than the proscribed radius of the QM region. In any case, the predicted VEBEs are relatively insensitive to small changes in the size of the spherical cavity provided that it encompasses the entire QM region. (When QM nuclei extend beyond the cavity, non-sensical results are obtained or in some cases the SCF + PCM procedure fails to converge at all.)

### 2.3.4 Non-equilibrium Poisson Equation Solver

This section outlines the specific computational details needed to obtain non-equilibrium solvation free energies using the algorithms described in Sections 2.2.2 and 2.2.3 as well as specifying the form of the dielectric function  $\epsilon(\mathbf{r})$  for bulk liquid and liquid-vapor interfaces. To solve Eqs. 2.9 and 2.21, we discretize the QM electron density  $\rho_{\text{elec}}(\mathbf{r})$  on a  $25 \text{ \AA} \times 25 \text{ \AA} \times 25 \text{ \AA}$  Cartesian grid with a grid spacing of  $\Delta x = 0.22 \text{ \AA}$  in each direction, which closely follows the details of both the classical electrostatics calculations reported in Ref. [76] (although the grid used here is slightly more conservative, reflecting the somewhat more complicated topology of the electrostatic potential generated by a QM charge distribution<sup>104</sup>) as well as those of the quantum electrostatics calculations reported in Refs. [10] and [64]. The Turi-Borgis pseudopotential model<sup>45</sup> is used to estimate the location of the centroid of the spin

density, for reasons of computational expediency, and this point is selected as the origin for the grid. Solution of Eq. 2.9 is then accomplished using a finite-difference, pre-conditioned conjugate gradient algorithm adapted from Ref. [10]. The iterative charge density, and therefore the polarization charge density, is updated until the Euclidean norm of the residual vector between iterations falls below a threshold of  $T_{\text{solver}} = 10^{-5}$  a.u. We take  $\eta = 0.6$  in Eq. 2.12. The SCF procedure is considered converged when the *DIIS* error falls below a threshold of  $T_{\text{SCF}} = 10^{-5}$  a.u.

In order to describe anisotropic dielectric environments, we employ a dielectric function  $\epsilon(\mathbf{r})$  that can assume a different value at each grid point in three-dimensional space. For testing and comparison purposes we have also implemented this approach for *isotropic* solvation, which we describe first. For  $e^-(\text{aq})$  in bulk water, we imagine a spherical cavity whose radius is 2.75 Å larger than the radius of the QM region ( $r_{\text{QM}}$ ), as in the PCM calculations described in Section 2.3.3. Across this boundary, we interpolate  $\epsilon(\mathbf{r})$  between the values  $\epsilon_{\text{vac}} = 1.0$  inside the cavity, where matter is described atomistically, and  $\epsilon_{\text{static}} = 78.39$  outside the cavity. (This smooth change in the dielectric aids with convergence of the finite-difference Poisson solver.<sup>105,106</sup>) Interpolation is accomplished as a function of the radial distance  $r$  from the center of the grid using a hyperbolic tangent switching function. In detail, the  $r$ -dependent dielectric function that we use is

$$\epsilon(r) = \frac{1}{2} \left\{ \epsilon_{\text{static}} + \epsilon_{\text{vac}} + (\epsilon_{\text{static}} - \epsilon_{\text{vac}}) \tanh[\alpha(r - r_{\text{mid}})] \right\}. \quad (2.32)$$

The parameter  $\alpha = 3.9 \text{ \AA}^{-1}$  controls the length scale of the switching process and  $r_{\text{mid}}$  is the midpoint of the interpolation. We take  $r_{\text{mid}} = r_{\text{QM}} + 1.375 \text{ \AA}$ , so that the dielectric assumes the value  $\epsilon = (\epsilon_{\text{vac}} + \epsilon_{\text{static}})/2$  halfway between  $r_{\text{QM}}$  and  $r_{\text{cavity}} = r_{\text{QM}} + 2.75 \text{ \AA}$ . (See Figure 2.2 for an example.) Numerical results are not strongly

dependent on the details of this smoothing procedure, and the calculations reported in Table 2.6 for  $e^-(aq)$  in bulk water demonstrate the extent to which IEF-PCM results agree with this QM/Poisson procedure.

The QM/Poisson procedure for the interfacial species is similar. We first use the Turi-Borgis model to estimate the center point of the electronic wave function and thus locate the center of the grid. Across a spherical cavity of radius  $r_{\text{cavity}} = r_{\text{QM}} + 2.75 \text{ \AA}$  surrounding the Turi-Borgis centroid, we smoothly interpolate  $\epsilon(r)$  as indicated in Eq. 2.32. However, we also compute the instantaneous distance  $d$  between the centroid of the electron and the GDS, for each snapshot from the simulations. Since the vapor phase should have a dielectric of  $\epsilon_{\text{vac}} = 1$ , we use a similar hyperbolic tangent function to interpolate  $\epsilon(z)$  across the GDS:

$$\epsilon(z) = \frac{1}{2} \left[ \epsilon_{\text{static}} + \epsilon_{\text{vac}} + (\epsilon_{\text{vac}} - \epsilon_{\text{static}}) \tanh(\alpha|z - d|) \right]. \quad (2.33)$$

Here,  $z = 0$  defines the middle of the slab, so Eq. 2.33 corresponds to a midpoint of  $z = d$  for the switching function at the water/vapor interface. A diagram of this interfacial setup appears in Figure 2.3.

It should be noted that employing Eqs. 2.32 and 2.33 to create spatially-varying dielectric functions for the multiplicative factor in Eq. 2.19 is unnecessary; the component of the smoothing functions containing the spatial dependence algebraically cancels out, and thus the ratio  $\chi_{\text{slow}}/\chi_{\text{e}}$  is constant across the grid. The spatially-varying nature of the dielectric is carried solely by the static dielectric function used to compute the polarization charge in Eq. 2.10 and iterative charge densities in Eqs. 2.11 and 2.12.

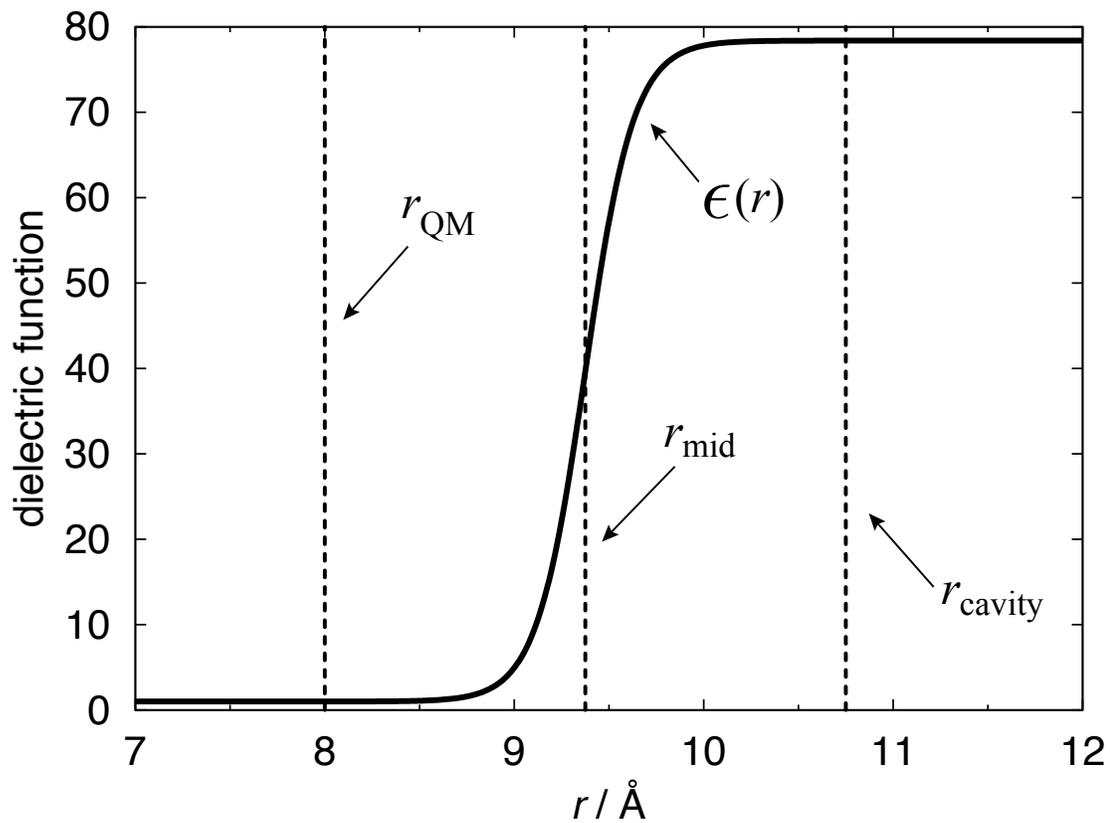


Figure 2.2: Dielectric function  $\epsilon(r)$  from Eq. 2.32, as a function of the distance  $r$  from the center of the QM region. The parameters of the switching function are  $\alpha = 3.9 \text{ \AA}^{-1}$  and  $r_{\text{mid}} = 9.375 \text{ \AA}$ , as used in this work for a QM radius of  $8.0 \text{ \AA}$ . In this example,  $\epsilon(r_{\text{QM}}) = 1.002$ ,  $\epsilon(r_{\text{mid}}) = 39.695$ , and  $\epsilon(r_{\text{cavity}}) = 78.388$ .

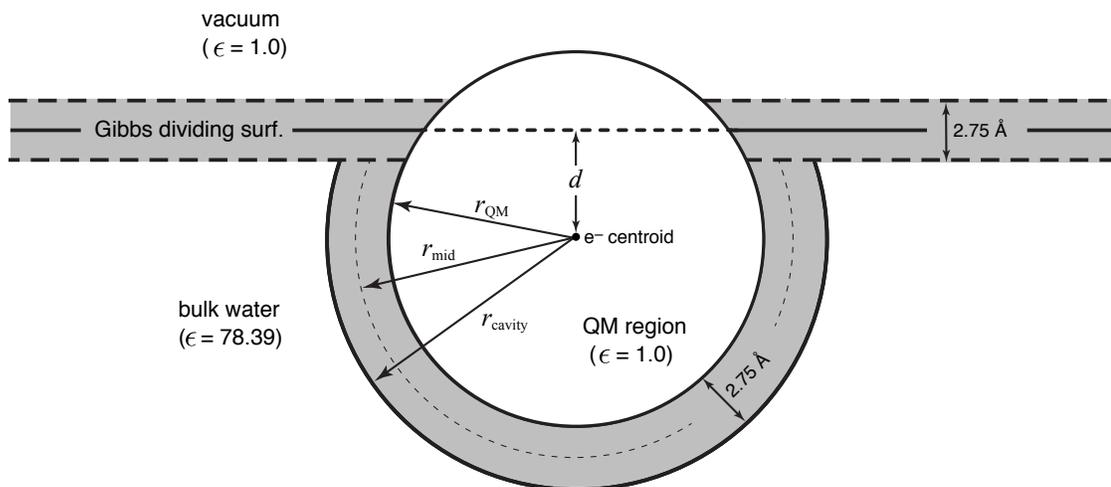


Figure 2.3: Diagram of the setup for the QM/Poisson calculations at the water/vacuum interface. Within the two shaded regions, the dielectric function is interpolated between  $\epsilon = 1.0$  and  $\epsilon = 78.39$ , as described in the text. The interfacial region around the GDS uses the switching function in Eq. 2.33, whereas the spherical cavity in bulk water uses the same switching function used for bulk solvation, Eq. 2.32.

## 2.4 Results and Discussion

In Section 2.4.1, we present the results from the mixed quantum-classical molecular dynamics simulations of  $e^-(aq)$  in liquid water and at the liquid-vapor interface and provide a discussion of these results in the context of recent liquid microjet experiments. The VEBE of bulk  $e^-(aq)$  computed with IEF-PCM and the PEQS methods are provided in Section 2.4.2 as well as the VEBE of interfacial  $e^-(aq)$  with the PEQS method. A discussion of the performance of the PEQS method relative to PCM for the bulk case is also included.

### 2.4.1 Simulations of $e^-(aq)$ in Liquid and at the Interface

There is a long history of mixed quantum/classical molecular dynamics (MD) studies of  $e^-(aq)$  using one-electron pseudopotential models.<sup>107–109</sup> We report results from simulations employing two variants of such a model: a fully polarizable model that treats electron–water polarization explicitly and self-consistently,<sup>40</sup> and a simpler model<sup>45</sup> wherein polarization is described by a pairwise-additive  $r^{-4}$  term in the interaction potential, which is tantamount to a mean-field description of electron–water polarization.<sup>107</sup> Both models successfully reproduce the steady-state absorption spectrum of  $e^-(aq)$  as well as the radius of gyration

$$r_g = \left\langle (\mathbf{r} - \langle \mathbf{r} \rangle) \cdot (\mathbf{r} - \langle \mathbf{r} \rangle) \right\rangle^{1/2}$$

that is inferred from a moment analysis of this spectrum,<sup>110</sup> and both provide a qualitatively correct value for the rather large diffusion coefficient of  $e^-(aq)$ .<sup>40,57</sup> The explicit-polarization model also reproduces the aforementioned “consensus” VEBE, along with a variety of *ab initio* benchmarks for  $(H_2O)_N^-$  cluster anions.<sup>40</sup>

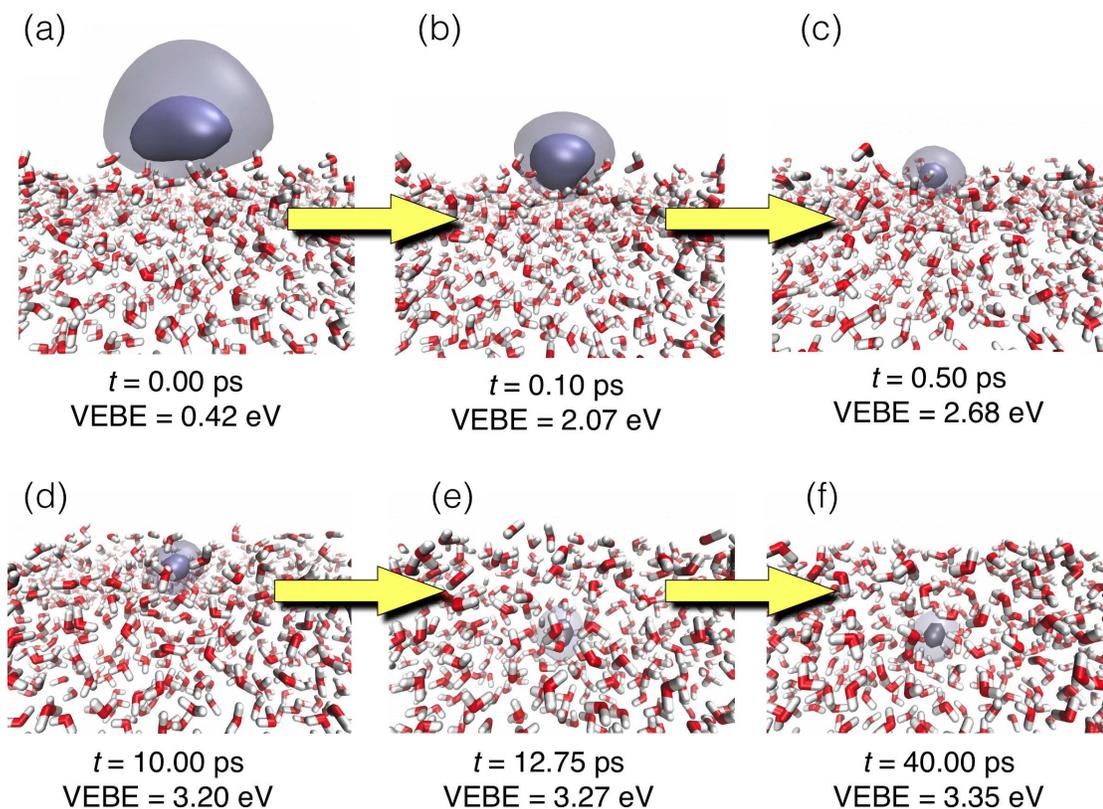


Figure 2.4: Representative snapshots of a simulation using the mean-field polarization model. In each panel, the opaque blue isocontour of the wave function encapsulates 50% of the one-electron probability distribution  $|\psi(\mathbf{r})|^2$ , while the lighter, translucent isosurface encloses 95%. At  $t = 0$ , where the solvent configuration is taken from an equilibrated simulation of neat liquid water, the electron is localized in a surface trap created by dangling O–H moieties, however the degree of localization ( $r_g \approx 6.25$  Å) is far less than for  $e^-(aq)$  in bulk water. After only 0.1 ps, however, the wave function has contracted to  $r_g \approx 3.50$  Å, as seen in (b), while the VEBE has increased by 1.65 eV. By  $t = 0.5$  ps the wave function has contracted to  $r_g \approx 2.4$  Å and is comparable in size to the bulk species, yet resides  $\approx 1.0$  Å above the Gibbs dividing surface. The size of the wave function has changed little by  $t = 10$  ps, but its centroid has moved to 1.50 Å below the Gibbs dividing surface, and water molecules surround the electron. The VEBE has increased by 0.5 eV relative to (c), and is now comparable to the bulk value. Panel (e) shows a snapshot shortly after the electron internalizes, with its centroid situated 8.74 Å below the Gibbs dividing surface and with a VEBE that is similar to the bulk value and essentially unchanged relative to that in (d). By  $t = 40$  ps in (f), the electron centroid has been fluctuating around  $d = -8.50$  Å for 28 ps with a VEBE that has fluctuated around the bulk value since  $t \approx 10$  ps.

Our simulations begin from neat, equilibrated liquid water at  $T = 300$  K and  $\rho = 0.997$  g/cm<sup>3</sup>, in a periodic simulation cell containing 200 water molecules. Convergence tests with respect to the size of the simulation cell are discussed in detail below. At time  $t = 0$  we introduce the extra electron simply by turning on the electron–water interaction potential, and from this point the energy and forces are obtained from the ground-state eigenvalue of the one-electron model Hamiltonian. This does not constitute a direct simulation of electron injection into liquid water, except possibly for an electron with zero kinetic energy, but does allow us to examine the interface  $\rightarrow$  bulk internalization process because at  $t = 0$  the solvent configuration is representative of neat liquid water, and dangling O–H moieties at the water/vacuum interface create a number of shallow potential energy “traps” for the electron to inhabit. As such, the ground-state wave function is fairly localized (at the interface) at  $t = 0$ , as shown in Figure 2.4(a). Subsequent panels in Figure 2.4 show how the wave function changes as it evolves in time.

The very early-time dynamics of  $e^-(aq)$  at the interface suggest that within even the first 100 fs of solvation dynamics, the electron is pulled farther below vacuum level than the reported 1.6 eV for the putative surface-bound electron.<sup>9,41</sup> Although not yet as compact as the equilibrated species, the electron’s wave function is certainly localized, whereas previous discussion has assumed that an electron  $\sim 1.6$  eV below vacuum level would be delocalized.<sup>24</sup> That assumption is based on a value  $V_0 = -1.5$  eV for the band gap of liquid water,<sup>24</sup> meaning that the conduction band sits 1.5 eV below vacuum level. Extrapolations of cluster ion data, however, suggest instead that  $V_0 \sim 0$ .<sup>111</sup> Moreover, on the surface of amorphous solid water, where the librational dynamics that ultimately solvate the interfacial electron are frozen

out, the surface-bound electron is stable for *minutes*, and is thought to localize in a Bjerrum defect.<sup>112</sup> The measured VEBE of  $e^-(aq)$  on the surface of amorphous solid water is 1.4 eV,<sup>113</sup> consistent with values intermediate between those in Figure 2.4(a) and (b).

Figure 2.5 depicts the time evolution of the VEBE, the radius of gyration, and the distance  $d$  between the centroid of the excess electron wave function and the Gibbs dividing surface, which is defined as the locus of points in the  $z$  direction for which the solvent density has fallen to half its value in bulk solution, and is used to demarcate the liquid/vapor interface. Although initially quite diffuse and weakly bound, the interfacial electron undergoes an ultrafast hydration process that begins to stabilize and localize the diffuse charge within 25 fs, via rotation of O–H moieties into the charge cloud. This process increases the VEBE by  $\sim 1$  eV. A deuterium isotope effect observed in the early-time signal (35–80 fs) in pump/probe experiments has previously been taken as evidence that librational motion of the waters is responsible for the first stages of electron localization,<sup>114,115</sup> and we observe this directly in our simulations. An additional 1.0–1.5 eV of stabilization occurs in the next 125 fs, and by  $t = 0.5$ – $1.0$  ps the VEBE of the interfacial species is essentially indistinguishable from that of  $e^-(aq)$  in bulk water. Concurrently, the size of the interfacial charge collapses from  $r_g \sim 6$ – $8$  Å to  $r_g \sim 2$ – $3$  Å as shown in Figure 2.5(b). The electron localization timescale that we observe directly in bulk water is consistent with that inferred from time-resolved terahertz spectroscopy.<sup>116</sup>

Twenty trajectories were examined for both the bulk and interfacial electron, for each pseudopotential model and for periodic simulation cells containing 200, 300, and 600 water molecules, and in every single case the interfacial electron undergoes a

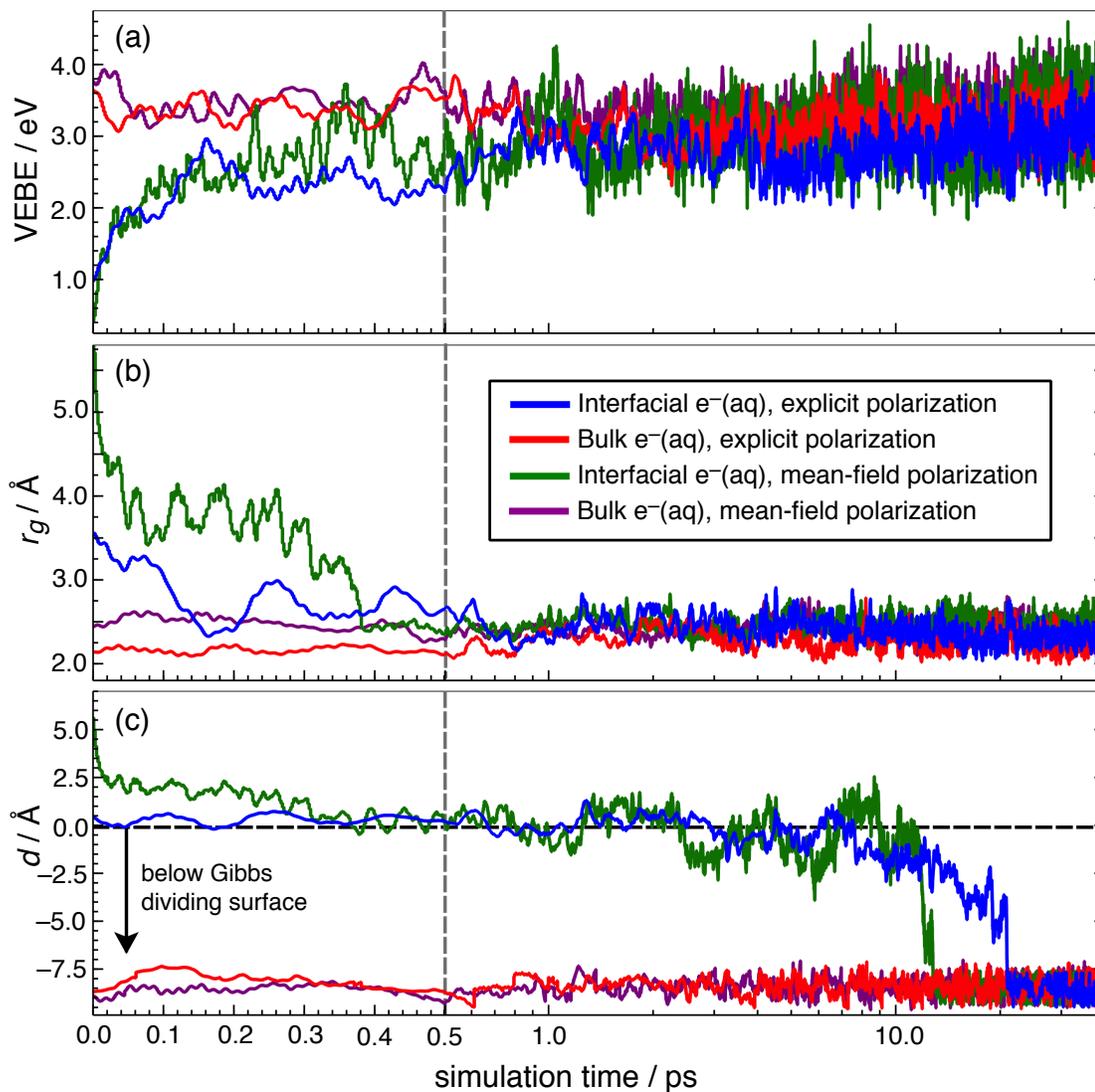


Figure 2.5: Time evolution of (a) the VEBE, (b) the radius of gyration, and (c) the distance between the centroid of the  $e^-(aq)$  charge distribution and the Gibbs dividing surface. (The time axis is linear for the first 0.5 ps then switches to a logarithmic scale.) Data are obtained from quantum/classical MD using two different one-electron models, under ambient conditions.

transition into bulk water on a timescale of  $\approx 10\text{--}25$  ps for the mean-field polarization model and  $\approx 25\text{--}35$  ps for the explicitly polarizable model [see Figure 2.5(c)]. This is consistent with previous simulations of relaxation dynamics of  $e^-(\text{aq})$  at the water/vacuum interface at ambient temperature,<sup>52</sup> and also with recent work in which the mean-field polarization model was used to compute the potential of mean force for dragging an electron through the interface.<sup>58,59</sup> In the latter simulations, the free energy profile for moving the electron from the interface to a depth of  $9\text{--}10$  Å is either strictly downhill,<sup>58</sup> or else exhibits a barrier to internalization of  $< 2k_B T$  (at  $T = 300$  K),<sup>59</sup> depending upon how the interface is defined. Electron–water polarization is the only component of the interaction energy that drives the electron inward from the interface,<sup>58</sup> and the mean-field  $r^{-4}$  polarization potential is known to overestimate the polarization energy relative to explicit, many-body treatments of electron–solvent polarization.<sup>117,118</sup> One might therefore expect that an explicitly-polarizable model would afford a slightly longer timescale for internalization as compared to a mean-field polarization model, which is in fact what we observe. In any case, the potential of mean force is found to be completely flat for  $d < -1$  nm,<sup>58,59</sup> so the fact that we do not observe bulk-solvated electrons spontaneously return to the interface can be explained based simply on the fact that the volume of phase space consistent with  $d < -1$  nm is far less than that within 1 nm of the interface.

Figure 2.6 is a scatter plot examining the correlation between the VEBE and the quantities  $r_g$  and  $d$ . Anti-correlation between VEBE and  $r_g$  [Figure 2.6(a)] is observed for all of the simulations, regardless of initial conditions or which one-electron model is used. This has also been observed in previous simulations of both  $e^-(\text{aq})$ <sup>4,11,45,52</sup> as well as finite  $(\text{H}_2\text{O})_N^-$  clusters.<sup>8,49</sup> In clusters, the correlation between the VEBE and

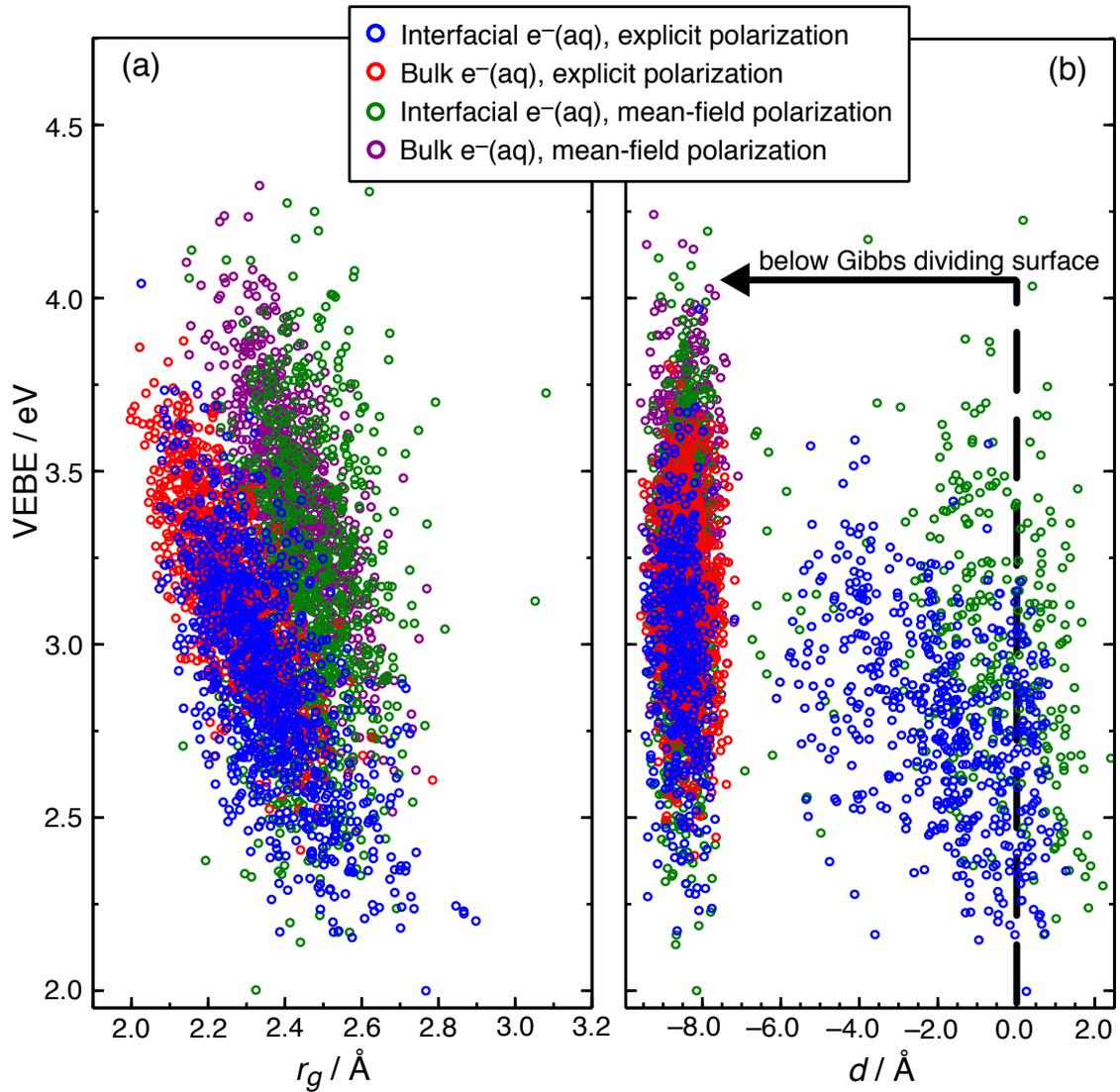


Figure 2.6: VEBE as a function of (a) the radius of gyration and (b) the distance between the electron and the Gibbs dividing surface. The data are taken from the trajectories plotted in Figure 2.5, sampled every 20 fs.

Table 2.3: Average values for selected observables from one-electron pseudopotential simulations.

Polarization Method	System	$\langle \text{VEBE} \rangle$ (eV)	$\langle r_g \rangle$ (Å)	$\langle d \rangle$ (Å)
explicit	interface <sup>a</sup>	3.10	2.27	-2.05
explicit	bulk	3.30	2.21	-8.54
mean-field	interface <sup>a</sup>	3.13	2.55	-1.76
mean-field	bulk	3.39	2.43	-8.46

<sup>a</sup>Averages exclude data points after the electron internalizes.

$r_g$  is observed for both surface-bound isomers (the cluster analogs of the interfacial hydrated electron) as well as “internalized” isomers (analogs of the electron in bulk water). A similar correlation is found between  $r_g$  and the electronic absorption maximum, in both clusters and for  $e^-(aq)$  in bulk water, although the clusters span a much wider range of  $r_g$  values and make it apparent that both the VEBE and the absorption maximum vary as  $r_g^{-2}$ , a dependence that can be derived from a particle-in-a-cavity model for bulk  $e^-(aq)$ , or from a Rydberg atom model for  $(H_2O)_N^-$ .<sup>8</sup> On the other hand, the VEBE data span a range of  $\sim 1$  eV (see Figure 2.6), so while these toy models successfully predict the qualitative  $r_g^{-2}$  trend that is somewhat evident in the data in Figure 2.6(a), and more obvious in the cluster data,<sup>8</sup> they cannot quantitatively predict the VEBE. For this, detailed atomistic calculations are required.

In contrast to the correlation with  $r_g$ , the VEBE appears to be *uncorrelated* with the distance  $d$  from the water/vacuum interface, as has been seen in previous simulations as well.<sup>4,52</sup> Averages for these observables are reported in Table 2.4, and are essentially unchanged even if half of the trajectories are discarded. It should be noted

that there are important finite-size corrections to the VEBE, which is especially sensitive to long-range Coulomb effects,<sup>40,119</sup> and these are addressed in Section 2.3.1. as well as by all-electron quantum chemistry calculations with continuum boundary conditions that are described below. What is most important to the present discussion is that the interfacial and bulk VEBEs differ only by about 0.2 eV. Full probability distributions for the bulk and interfacial VEBEs (Figure 2.9) are consistent with a shift of 0.2 eV, and a broadening of  $< 0.2$  eV, in the VEBE distribution at the interface relative to that in bulk water.

For the bulk species, values of  $d$  span a narrow range around  $-8.5$  Å that reflects about half the thickness of the periodic slab used in the simulations. (This is true in larger simulation cells as well; in a 300 H<sub>2</sub>O cell with a slab thickness of 20.8 Å, the bulk species fluctuates around  $d = -8.9$  Å and in a 600 H<sub>2</sub>O cell with a thickness of 26.2 Å it fluctuates around  $d = -13.0$  Å, as shown in Figure 2.5.) Evidence of the prompt disappearance of the interfacial species can be seen in Figure 2.6(b), where trajectories initialized at  $d = 0$  migrate to  $d \approx -8.5$  Å within tens of picoseconds. Once it passes into bulk solution, the electron VEBE fluctuates between 2.5–3.5 eV, which is roughly the same range observed for the original, interfacial state, for which  $d$  ranges from  $-5$  Å to  $+1$  Å. Evidently, the VEBE is controlled by the size of the electron ( $r_g$ ) but is uncorrelated with its distance from the interface.

This, too, is similar to what is observed in finite clusters.<sup>8</sup> Whereas when the unpaired electron is far from the water cluster, it is dipole-bound and exhibits a VEBE  $< 0.5$  eV, as the centroid of its wave function approaches the cluster surface, the distribution of VEBEs broadens and approaches values similar to the case of a fully solvated electron. Similarly, the present calculations show that even for an interfacial

electron that is 1–2 Å above the surface of liquid water, the VEBE spans a range of values comparable to what is observed for  $e^-(aq)$  in bulk water [see Figure 2.6(b)].

Simulations reported here and elsewhere<sup>52</sup> suggest that the interfacial species may not be stable beyond a timescale of  $\sim 30$  ps, and perhaps more importantly, that measurement of the VEBE by photoelectron spectroscopy may be unable to differentiate between  $e^-(aq)$  in bulk solution and  $e^-(aq)$  at the water/vacuum interface. (It is possible that measurement of photoelectron angular distributions might be able to make such a distinction; such a technique has recently been used in an attempt to elucidate electron binding motifs in sodium-doped water clusters.<sup>120</sup>) Photoelectron spectroscopy likely probes the compactness of the solvated electron’s spin density, rather than its position relative to the interface, much like terahertz spectroscopy of  $e^-(aq)$  in bulk water.<sup>116</sup> Recently, we also concluded that the the steady-state absorption spectrum of  $e^-(aq)$  that, historically speaking, is the primary observable used to monitor production and annihilation of this species,<sup>121–125</sup> also fails to discriminate  $e^-(aq)$  in bulk water from that at the interface.<sup>5</sup> As in the photoelectron case, absorption spectroscopy is primarily a probe of the electron’s compactness, regardless of binding motif.<sup>8</sup>

Although  $e^-(aq)$  has been detected in experiments using various surface-sensitive spectroscopies,<sup>126,127</sup> those experiments have been interpreted to imply that the species responsible for the “interfacial” signal might reside anywhere within the first 1–2 nm of the liquid,<sup>126</sup> and that its lifetime in the vicinity of the interface is  $< 100$  ps.<sup>127</sup> Even at a depth of 1 nm, our calculations suggest that the properties of the nominally “interfacial” electron are actually indistinguishable from those of the equilibrated species in bulk water, owing to the presence of sufficiently many solvation

shells to converge  $r_g$ . In this interpretation, surface-sensitive spectroscopy simply detects whatever bulk-like electrons happen to diffuse close enough to the surface to break spatial inversion symmetry and thus become observable. This is consistent with an electron attenuation length of  $\sim 5$  nm at 3 eV of electron kinetic energy,<sup>128</sup> meaning that photoemission from any putative surface-bound species should always be accompanied by a strong signal from the bulk species.<sup>31,38</sup>

Additional testing was performed to examine the dependence of our results on the number of water molecules, choice of simulation grid and periodic unit cell, and initial conditions. To this end we performed twenty simulations each with  $N = 300$  and  $N = 600$  water molecules in the unit cell, using the mean-field Turi-Borgis polarization model at  $T = 300$  K. The density was kept the same as the smaller simulations discussed above ( $\rho = 0.997$  g cm<sup>-3</sup>), which in the present cases means  $L_x = L_y = 20.7961$  Å for  $N = 300$  and  $L_x = L_y = 26.2105$  Å for  $N = 600$ , with  $L_z = 5L_x$  in all cases. The simulation grid representing the electronic wave function was chosen in the same manner as above, with  $\Delta x = \Delta y = \Delta z = 0.95$  Å. To make the calculations tractable (because the dimension of the Hamiltonian is equal to the total number of grid points in three dimensions), we used the same simulation grid for  $N = 600$  that we did for  $N = 300$ , so in the former case the grid does not extend quite to the edges of the simulation cell. (Since the electron localizes rapidly in each trajectory, this is not a serious limitation.)

As with the  $N = 200$  simulations described above, initial structures were extracted from an equilibrated simulation of neat liquid water (SPC water model) with either  $N = 300$  or 600 molecules in the unit cell. Trajectories were propagated for 80 ps, and as in the smaller unit cells the electron internalizes into the bulk in every single

trajectory. Simulations were propagated for an additional 50 ps to equilibrate the internalized  $e^-(aq)$  for use in bulk simulations, as discussed above, and the initial structures were randomly selected from the last 10 ps of these equilibration trajectories. The bulk simulations were also propagated for 80 ps.

It should be noted that the initialization procedure used in these simulations does not represent a simulation of electron injection into water, except possibly that of an electron with zero kinetic energy. Rather, we simulate adiabatic dynamics on the ground-state Born-Oppenheimer potential energy surface generated by the one-electron model Hamiltonian. At  $t = 0$ , the electron is found wherever the potential energy is low, and that is consistently found to be a surface trap state formed by dangling O–H moieties at the water/vacuum interface, as shown for example in Figure 2.4(a). We have verified that different initial guesses for the wave function at  $t = 0$  lead to the same solution, demonstrating that our eigensolver (the details of which are described in Ref. [40]) is robust. For example, Figure 2.7 is analogous to Figure 2.4 but for a unit cell containing 600 water molecules, and with an initial guess wave function that is completely delocalized. Very similar localization and internalization dynamics can be seen in this trajectory as compared to that in Figure 2.4.

Figure 2.8 shows representative trajectories for all three unit cell sizes, with the  $N = 200$  trajectories being the same ones that are plotted in Figure 2.5. Despite being initialized as a delocalized particle below the GDS, the interfacial  $e^-(aq)$  immediately appears as a diffuse charge cloud [ $r_g \approx 5\text{--}6 \text{ \AA}$ , see Figure 2.8(b)] located at  $d \approx 3\text{--}6 \text{ \AA}$  above the liquid surface as illustrated in Figure 2.7(a). This behavior is easily understood. Unlike in previous simulations of  $e^-(aq)$  in bulk water that used an isotropic unit cell (*i.e.*, no water/vacuum interface), and for which the electron takes

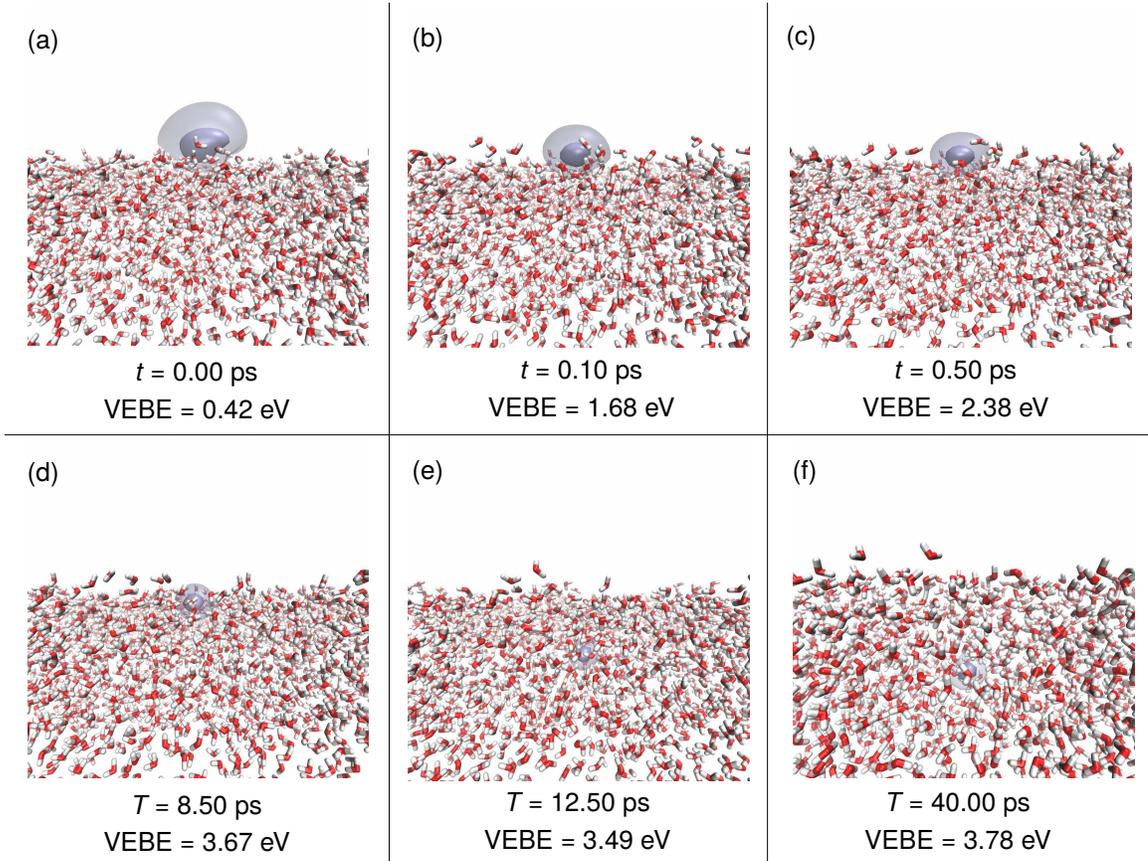


Figure 2.7: Snapshots from a simulation of an electron initialized delocalized in liquid water using the Turi-Borgis pseudopotential model. In each panel, the opaque blue isocontour of the wave function encapsulates 50% of the one-electron probability distribution  $|\psi(\mathbf{r})|^2$ , while the lighter translucent isosurface encloses 95%. At  $t = 0$ , the electron is initialized as a completely delocalized particle in liquid water, with  $r_g \approx 6.78 \text{ \AA}$ . After 100 fs the wave function has contracted to  $r_g \approx 3.93 \text{ \AA}$ , as seen in (b), while the VEBE has increased by 1.46 eV. By 500 fs, the wave function further contracts to  $r_g \approx 3.12 \text{ \AA}$  and resides  $\approx 1.1 \text{ \AA}$  above the GDS. By  $t = 8.50 \text{ ps}$ , the size of the wave function is comparable to the bulk species with  $r_g = 2.45 \text{ \AA}$ , and its centroid has moved to  $1.50 \text{ \AA}$  below the GDS where water molecules surround the electron, such that the VEBE in (d) is 1.29 eV larger than that in (c), and is similar to the bulk value. Panel (e) shows a snapshot shortly before the electron internalizes, with its centroid situated  $7.75 \text{ \AA}$  below the GDS and with a VEBE that is similar to the bulk value and essentially unchanged relative to that in (d). By  $t = 40 \text{ ps}$  in (f), the electron centroid has been fluctuating around  $d = -13.00 \text{ \AA}$  for 28 ps with a VEBE that has fluctuated around the bulk value since  $t \approx 8.50 \text{ ps}$ .

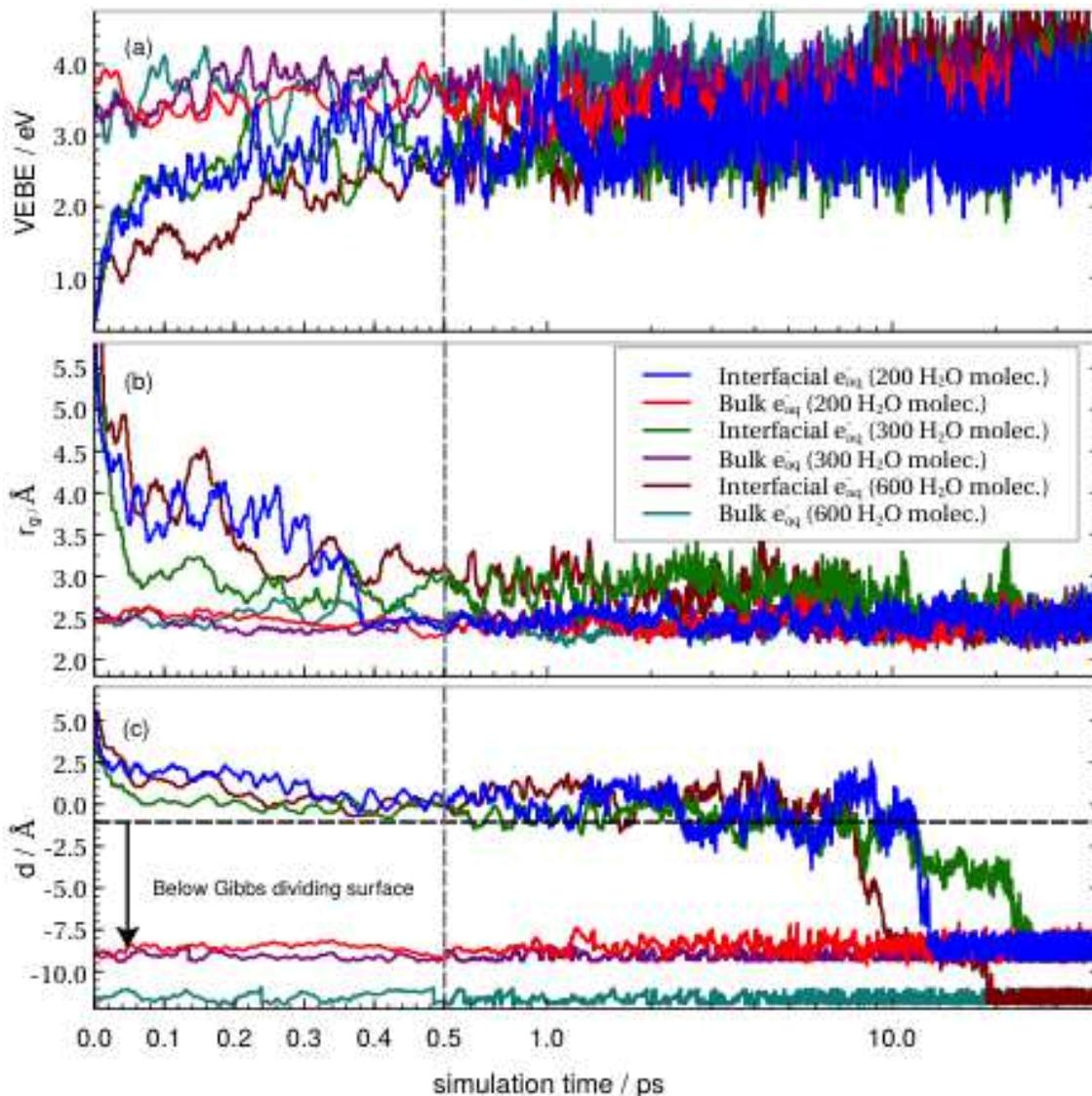


Figure 2.8: Time evolution of (a) the VEBE, (b) the radius of gyration ( $r_g$ ), and (c) the distance ( $d$ ) between the centroid of the  $e^-(aq)$  charge distribution and the Gibbs dividing surface from MD simulations employing a mean-field polarization potential model at ambient temperatures. (The time axis is linear for the first 0.5 ps then switches to a logarithmic scale.) For comparison to the simulations utilizing  $N = 300$  and  $N = 600$  water molecules, the data from representative  $N = 200$  trajectories are reproduced from Figure 2.5.

$\sim 1$  ps to carve out and localize in an excluded volume,<sup>40</sup> in the present simulations there exist, already at  $t = 0$ , shallow potential energy “traps” around dangling O–H moieties at the interface. The ground-state wave function at  $t = 0$  occupies in these traps, one of which is apparently deep enough relative to the others such that the electron localizes in it, as shown in Figure 2.7(a).

Essentially instantaneous electron localization is observed in every trajectory at every box size, as is the rapid increase in VEBE during the first 0.5 ps, as exemplified by panels (b) and (c) of Figures 2.4 and 2.7. The internalization process is spontaneous on a timescale of 20–25 ps regardless of the size of the simulation cell or the initial conditions used to introduce the electron at  $t = 0$ , and upon internalization the electron fluctuates around  $d \approx 9\text{--}10$  Å for  $N = 200$  and  $N = 300$  simulations and  $d \approx 13$  Å for  $N = 600$ , which reflects half the thickness of the simulated water slab. (The  $N = 200$  and  $N = 300$  simulation cells are sufficiently similar in size, at  $L_x = 18.1617$  and  $20.7961$  Å, respectively, that the 1.0–1.5 Å fluctuations about the midpoint blur the distinction and the electron fluctuates around similar values of  $d$  in both cases.) The good agreement that we see amongst all three simulation cell sizes is consistent with the fact that the potential of mean force for pulling an electron through the interface, which was computed in Refs. [58] and [59] using the same Turis-Borgis pseudopotential that is used here, flattens out at  $d \approx 9.5$  Å, just slightly more than half the width of the slab in our smallest simulation cell.

Table 2.4 reports averages for the quantities  $r_g$ , VEBE, and  $d$ , comparing different simulation cell sizes from all twenty simulations. For each given box size, values of the VEBE and radius of gyration are rather similar, regardless of whether the electron was initialized in the bulk or at the interface, although the average VEBE for the

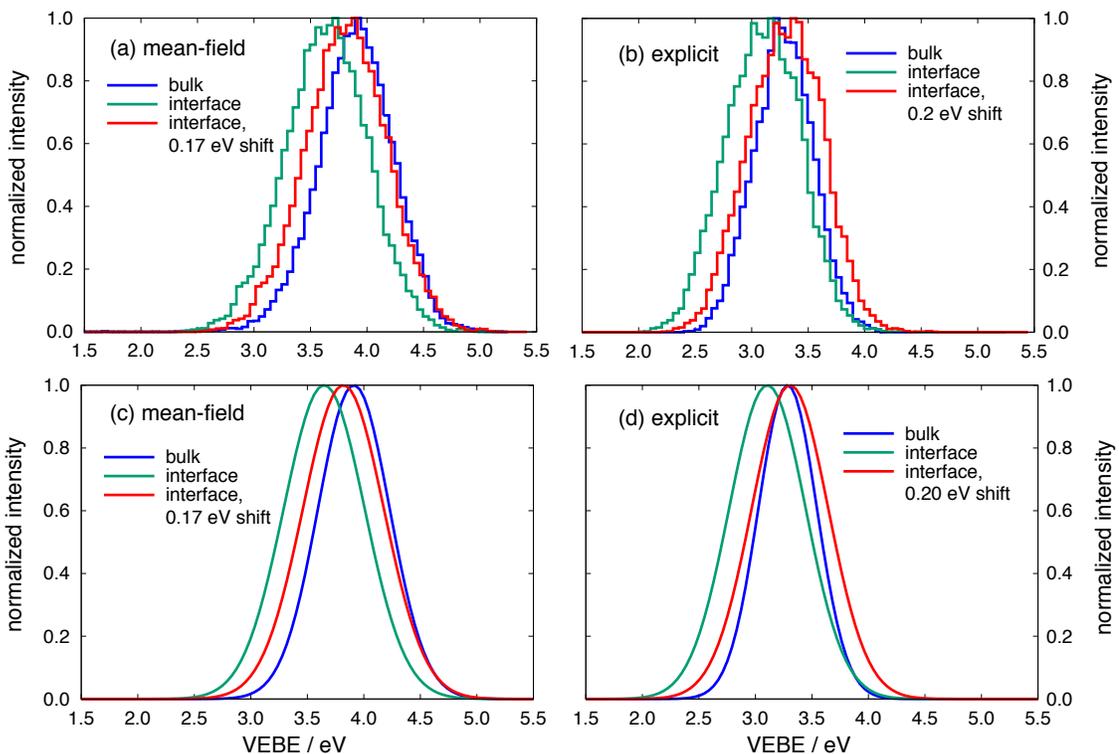


Figure 2.9: Probability distributions for the bulk and interfacial VEBE using (a) the mean-field polarization model, in an  $N = 600$  water box; and (b) the explicit polarization model, in an  $N = 200$  water box, with histogram bin widths of  $0.05$  eV in each case. Panels (c) and (d) show Gaussian fits to the data in (a) and (b), respectively. (Fitting parameters are listed in Table 2.5.) The interfacial distributions are also shown following a shift in energy equal to the difference in the mean VEBEs of the bulk and interfacial distributions, as reported in Tables 2.4 and 2.5. Differences in the center points of the distributions in (a) versus (b) are partly the result of the effects of finite box size.

Table 2.4: Average values for selected observables from mean-field polarization model simulations with varying numbers of water molecules,  $N$ .

$N$	System	$\langle \text{VEBE} \rangle$ (eV)	$\langle r_g \rangle$ (Å)	$\langle d \rangle$ (Å)
200	interface <sup>a</sup>	3.13	2.55	-1.76
200	bulk	3.39	2.43	-8.46
300	interface <sup>a</sup>	3.34	2.49	-1.97
300	bulk	3.57	2.40	-8.78
600	interface <sup>a</sup>	3.67	2.44	-1.66
600	bulk	3.84	2.39	-13.01

<sup>a</sup>Averages exclude data points after the electron internalizes.

interfacial species is about 0.2 eV smaller than that of the bulk species. As a test of convergence, we can throw out half of the trajectories and recompute the averages, and we find that they change hardly at all. Average VEBEs change by only about 0.05 eV and  $\langle r_g \rangle$  by about 0.05 Å, for both the bulk and interfacial species. The average distance  $\langle d \rangle$  to the GDS is essentially unchanged for the bulk species and changes by  $< 0.1$  Å for the interfacial species. Full probability distributions for the bulk and interfacial VEBE are depicted in Figure 2.9 and are fit quite well to Gaussian probability distributions,

$$g(E) = \exp\left(\frac{-(E - \langle E \rangle)^2}{2\sigma^2}\right), \quad (2.34)$$

with fitting parameters given in Table 2.5. Although the distributions of interfacial VEBEs are 0.1–0.2 eV wider than those obtained in bulk water, we find that even

Table 2.5: Parameters for the Gaussian fits [Eq. 2.34] of the VEBE data in Figure 2.9.

Polarization	System	Parameter (eV)		
Method		$\langle E \rangle$	$\sigma$	FWHM <sup>b</sup>
explicit	interface	3.1060	0.3392	0.7987
explicit	bulk <sup>a</sup>	3.2856	0.2596	0.6113
mean-field	interface	3.6489	0.3583	0.8437
mean-field	bulk <sup>a</sup>	3.9115	0.3175	0.7476

<sup>a</sup>Averages exclude data points after the electron internalizes.

<sup>b</sup>Full width at half maximum of the Gaussian distribution.

in terms of the raw data both distributions look quite similar if the interfacial distribution is simply shifted 0.2 eV higher in energy, which is the difference in the mean VEBE at the interface relative to that in bulk water.

Very recently, Casey *et al.*,<sup>59</sup> also reported distributions of bulk and interfacial VEBEs based on simulations employing the mean-field Turi-Borgis pseudopotential, but based on an alternative definition of the instantaneous liquid interface.<sup>129</sup> (In our simulations, the Gibbs dividing surface is also updated at each time step.) The simulations and sampling protocol reported in Ref. [59] afford a shift approaching 0.5 eV between the interfacial and bulk VEBE distributions, as opposed to the 0.2 eV shift reported here. The reasons for this discrepancy are unclear, although there is certainly a tail in our interfacial VEBE distributions that is shifted much more than 0.5 eV from the mean value in bulk water, so it could simply be that the definition of the liquid surface that is used in Ref. [59] excludes more of the water molecules than does the Gibbs dividing surface, giving it a preference for identifying as “interfacial” simulation snapshots where the electron is more weakly solvated, as compared to those snapshots that the Gibbs surface identifies as interfacial. The 0.2 eV shift that

we obtain using both pseudopotential models agrees quantitatively with the shift obtained from DFT and RI-MP2 calculations based on QM/MM trajectories; see Table 2.6.

The averages in Table 2.4 do show that the VEBE increases with box size, due to the importance of long-range Coulomb interactions in the ionization process, as we have reported previously.<sup>40</sup> As in previous work,<sup>40</sup> we can correct for this by extrapolating  $\langle \text{VEBE} \rangle$  as a function of the inverse unit cell length,  $L_x^{-1}$ , and such an extrapolation is shown in Figure 2.10. Remarkably, both the interfacial and bulk VEBE extrapolate essentially to the same value in the infinite-dilution limit: 4.87 eV for the interfacial species versus 4.85 eV for the bulk species. The extrapolated bulk value is also consistent with a previous extrapolation of 4.79 eV for the bulk species, using the same Turis-Borgis pseudopotential.<sup>40</sup> When the VEBEs for large  $(\text{H}_2\text{O})_N^-$  clusters are extrapolated to  $N \rightarrow \infty$  using the Turis-Borgis pseudopotential, a value of 4.4 eV is obtained for the interior (cavity) states, versus 3.9 eV for surface-bound electrons.<sup>119</sup>

Each of these extrapolated values for the bulk VEBE is significantly larger than the accepted range of experimental values, 3.3–3.7 eV.<sup>9,32,33,36,39</sup> In Ref. [40], the infinite-dilution extrapolation was performed using the explicit polarization model (in addition to the mean-field Turi-Borgis model), and it was found that the electronic reorganization energy associated with relaxing the inducible dipole moments on the water molecules is significant, and reduces the VEBE by 1.37 eV in the limit  $L \rightarrow \infty$ . A very similar correction to the non-polarizable Turi-Borgis VEBE, 1.3 eV, was obtained using a continuum model,<sup>40</sup> which inspires some confidence that this is indeed the electronic (clamped-nuclei) reorganization energy associated with vertical

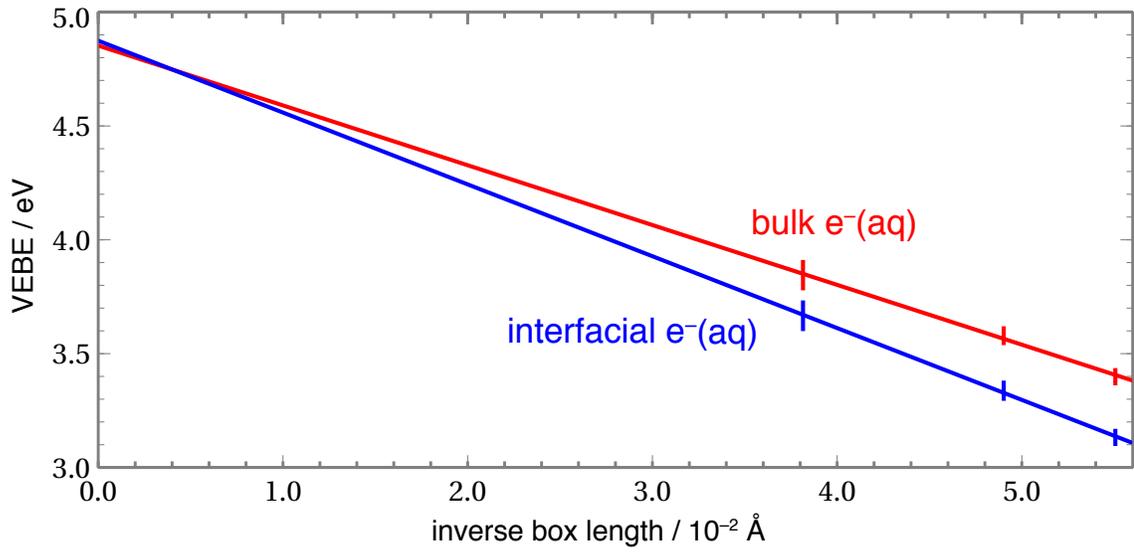


Figure 2.10: Extrapolation of the average VEBE computed using the mean-field polarization potential, as a function of the inverse unit cell length,  $L_x^{-1}$ . Error bars reflect a 95% confidence interval in  $\langle \text{VEBE} \rangle$ , which is averaged over trajectories and over time.

ionization of  $e^-(aq)$ . A correction of 1.3 eV applied to the extrapolated Turis-Borgis VEBE from Figure 2.10 affords a corrected VEBE of 3.6 eV, within the experimental range.

## 2.4.2 Vertical Binding Energy of $e^-(aq)$

To complement the one-electron pseudopotential simulations discussed above, we also computed the VEBE at the level of DFT and RI-MP2 including a quantum-mechanical (QM) description of a significant number of surrounding water molecules. DFT calculations employed a long-range-corrected functional (LRC- $\mu$ BOP) that has been shown to provide accurate VEBEs for  $(H_2O)_N^-$  clusters<sup>6,48</sup> as well as an accurate absorption spectrum for  $e^-(aq)$  in bulk water.<sup>5</sup> The range-separation parameter,  $\mu$ , is tuned in a non-empirical way to satisfy the ionization energy (IE) condition<sup>88</sup> shown in Eq. 2.30, and must be adjusted when the size of the QM region changes.<sup>5</sup> (Consult Section 2.3.1 for details and for tuned values of  $\mu$ .) RI-MP2 calculations, in contrast, have no such adjustable parameter and have been shown to provide VEBEs within  $\sim 0.05$  eV of CCSD(T) values for strongly-bound  $(H_2O)_N^-$  clusters.<sup>94,130</sup>

The *ab initio* calculations reported here were performed on snapshots extracted from mixed quantum/classical (QM/MM) simulations of both bulk and interfacial  $e^-(aq)$ ,<sup>4,11</sup> which we have previously used to obtain quantitative agreement with the electronic absorption spectrum of the bulk species.<sup>5</sup> From these snapshots, we extract QM regions ranging in radius from 5.5–8.0 Å around the centroid of the spin density  $(\rho_\alpha - \rho_\beta)$ . For the bulk species, this corresponds to  $\approx 30$  H<sub>2</sub>O molecules for the smallest QM regions and  $\approx 90$  for the largest, whereas for the interfacial species the corresponding numbers are  $\approx 20$  and  $\approx 60$  H<sub>2</sub>O molecules, respectively.

To incorporate longer-range solvent effects for  $e^-(aq)$  in bulk water, we employ a non-equilibrium polarizable continuum model (PCM) that uses the solvent optical dielectric constant to model electronic polarization upon vertical ionization.<sup>66,67</sup> PCM calculations in quantum chemistry usually describe the QM/continuum boundary by means of a union of atom-centered van der Waals spheres,<sup>63</sup> but this makes little sense in the present context given the highly delocalized nature of the solute (an electron), which is not assignable to any particular atom. In addition, with a significant number of solvent molecules considered explicitly within the QM region, the van der Waals cavity construction would allow the interstitial excess electron to inhabit the dielectric medium, to unknown effect. To avoid these problems, we define the boundary between atomistic and continuum regions to consist of a single spherical surface, centered at the centroid of the spin density and whose radius extends 2.75 Å beyond the farthest atom in the QM region. (Consult Section 2.3.3 for a justification of this choice.)

This spherical-boundary approach is not possible for the interfacial species, where the dielectric medium is spatially anisotropic. In this case, we directly solve Poisson’s equation expressed in Eq. 2.1 using an algorithm adapted from Ref. [10], for a spatially-inhomogeneous dielectric function  $\epsilon(\mathbf{r})$  that takes the value  $\epsilon = 78$  for  $d < 0$  and  $\epsilon = 1$  for  $d > 0$ . The QM charge density  $\rho(\mathbf{r})$  is discretized onto a three-dimensional grid and we solve for the electrostatic potential  $\varphi(\mathbf{r})$  at each self-consistent field iteration, whereupon the (equilibrium) solvation free energy is given by Eq. 2.4. The converged electrostatic potential  $\varphi(\mathbf{r})$  can then be used in the non-equilibrium reaction-field method of Refs. [66] and [67], in order to compute the change in solvation free energy upon vertical ionization. This represents a model in which electronic polarization is considered following ionization, but vibrational and

orientational polarization is frozen, since the ionization process is vertical in terms of the nuclear coordinates.

VEBEs computed using these non-equilibrium QM/PCM and QM/Poisson methods are listed in Table 2.6. For the bulk species, DFT/PCM and MP2/PCM calculations both afford VEBEs of 3.4–3.6 eV, in agreement with the experimental consensus of 3.3–3.7 eV,<sup>9,32–39</sup> and with prior QM/MM simulations.<sup>11</sup> The computed VEBEs converge rapidly with respect to the size of the QM region, and even the smallest QM region (comprising approximately two solvation shells around the spin density) affords a VEBE within 0.1 eV of the converged result. This is consistent with observations in previous calculations that the spin density extends outward through two solvation shells, but not three.<sup>40,96,108</sup> Continuum boundary conditions contribute 0.7–1.4 eV to the VEBE (see Table 2.1), and are therefore indispensable for obtaining agreement with experiment. Simulations using the explicitly-polarizable one-electron model also suggest a polarization response of  $\approx 1.4$  eV for vertical ionization of  $e^-(aq)$ .<sup>40</sup>

For the bulk species, where the dielectric function is isotropic, solution of Poisson’s equation should afford very similar results to the PCM, up to discretization errors and the fact that the PCM provides only an approximate treatment of the volume polarization that arises from the tail of the wave function that penetrates beyond the continuum boundary.<sup>63</sup> Comparing Poisson- and PCM-based VEBEs for the bulk species (Table 2.6), we find that the former are consistently 0.2 eV smaller but still within the experimentally measured range of bulk  $e^-(aq)$ . This favorable comparison validates our implementation of the non-equilibrium Poisson solver that is described in Sections 2.2.2 and 2.3.3.

Table 2.6: VEBEs<sup>a</sup> (in eV) computed at the DFT<sup>b</sup>/ and RI-MP2<sup>c</sup>/6-31++G\* levels of theory using non-equilibrium solvation models.

QM radius (Å)	Bulk e <sup>-</sup> (aq)				Interfacial e <sup>-</sup> (aq)	
	DFT	RI-MP2	DFT	RI-MP2	DFT	RI-MP2
	PCM	PCM	Poisson	Poisson	Poisson	Poisson
5.5	3.52 ± 0.31	3.44 ± 0.35	3.33 ± 0.31	3.18 ± 0.32	3.15 ± 0.41	3.07 ± 0.46
6.0	3.61 ± 0.35	3.42 ± 0.34	3.36 ± 0.32	3.21 ± 0.34	3.18 ± 0.43	3.02 ± 0.40
6.5	3.57 ± 0.38	3.37 ± 0.30	3.40 ± 0.33	3.17 ± 0.34	3.16 ± 0.42	3.06 ± 0.46
7.0	3.53 ± 0.35	3.41 ± 0.31	3.37 ± 0.34	3.19 ± 0.32	3.17 ± 0.41	3.08 ± 0.44
7.5	3.59 ± 0.35	3.42 ± 0.32	3.35 ± 0.31	3.20 ± 0.33	3.14 ± 0.41	3.09 ± 0.43
8.0	3.54 ± 0.32	3.45 ± 0.33	3.39 ± 0.33	3.22 ± 0.35	3.19 ± 0.44	3.10 ± 0.47

<sup>a</sup>Using structures extracted from QM/MM simulations reported in Refs. [11] and [4]. Uncertainties represent one standard deviation. <sup>b</sup>LRC- $\mu$ BOP with  $\mu$  tuned individually at each QM size so that  $\epsilon_{\text{SOMO}} = -\text{IE}$ . <sup>c</sup>Using the resolution-of-identity approximation.

Only the QM/Poisson approach is available for the interfacial species due to the anisotropic nature of the dielectric boundary in this case. The QM/Poisson approach affords an interfacial VEBE that is consistently 0.1–0.2 eV smaller than the bulk value computed with the same algorithm, but this difference is well within statistical fluctuations. Furthermore the VEBEs computed with the QM/Poisson methodology are in agreement with QM/MM simulations of the interfacial species under periodic boundary conditions,<sup>4</sup> as well as mixed quantum/classical MD simulations of the relaxation dynamics of an excess electron at the vacuum/water interface that employ a one-electron pseudopotential and a much larger periodic simulation cell.<sup>52</sup> Given the quantitative agreement between these various simulation methods, and the fact that they all agree with experimental results for  $e^-(aq)$  in bulk water, our calculations for the interfacial species appear to exclude the possibility of an interfacial solvated electron whose VEBE is significantly smaller than that of  $e^-(aq)$  in bulk water.

## 2.5 Conclusions

The picture that emerges from this work is one in which librational modes contribute to ultrafast ( $< 1$  ps) electron localization, both in bulk water and at the water/vacuum interface. Simulations indicate that an electron spawned at the interface of neat liquid water is unstable and migrates into bulk water within  $\sim 30$  ps. Even during the short time that the electron samples a truly interfacial environment, however, we find that its energy (measured relative to vacuum level) is essentially indistinguishable from that of equilibrated aqueous electron in bulk water. *Ab initio* calculations of the VEBE for the bulk species afford values of 3.4–3.6 eV that agree quantitatively with the experimental consensus of 3.3–3.7 eV, but also suggest that

the VEBE is not significantly different for the short-lived interfacial species. As such, these calculations lend no support to the idea of an electron at the vacuum/water interface with a VEBE that is significantly smaller than the bulk value.

Our results for both the energy and lifetime of the interfacial species are at odds with a recent experimental report,<sup>9</sup> in which a long-lived ( $\gtrsim 100$  ps) signal at 1.6 eV in a liquid microjet photoelectron spectrum was assigned to an electron bound at the water/vacuum interface. Based on electron attachment energies computed for temporary anion resonances in nucleotide monophosphates,<sup>131,132</sup> it was suggested in subsequent review articles<sup>17,41–43</sup> that a hydrated electron situated  $\leq 2.5$  eV below vacuum level is in the right range to induce DNA single strand breaks, via dissociative electron attachment. An electron whose VEBE is  $\approx 1.6$  eV would thus fall into this range, but one whose VEBE is  $\geq 3.3$  eV would not.

Recent QM and QM/MM calculations of double-stranded DNA<sup>133</sup> and individual nucleobases<sup>134</sup> in aqueous solution shown that the adiabatic electron affinities (EAs) of the DNA bases increase dramatically upon solvation, although the effect saturates after about two solvation shells and values in the range 0.75–1.20 eV are obtained.<sup>134</sup> Both the vertical and adiabatic EAs are relevant in the discussion of electron attachment to DNA, but in aqueous solution one expects the latter to be larger,<sup>9</sup> hence these values set an upper bound on the VEBE that would make the interfacial hydrated electron relevant to dissociative electron attachment reactions in DNA. Our calculations suggest that the energetics of an electron solvated at the surface of neat liquid water lie well outside of this bound, with both the bulk *and* the interfacial species residing too far below vacuum level. The energetics of this scenario are illustrated schematically in Figure 2.11.

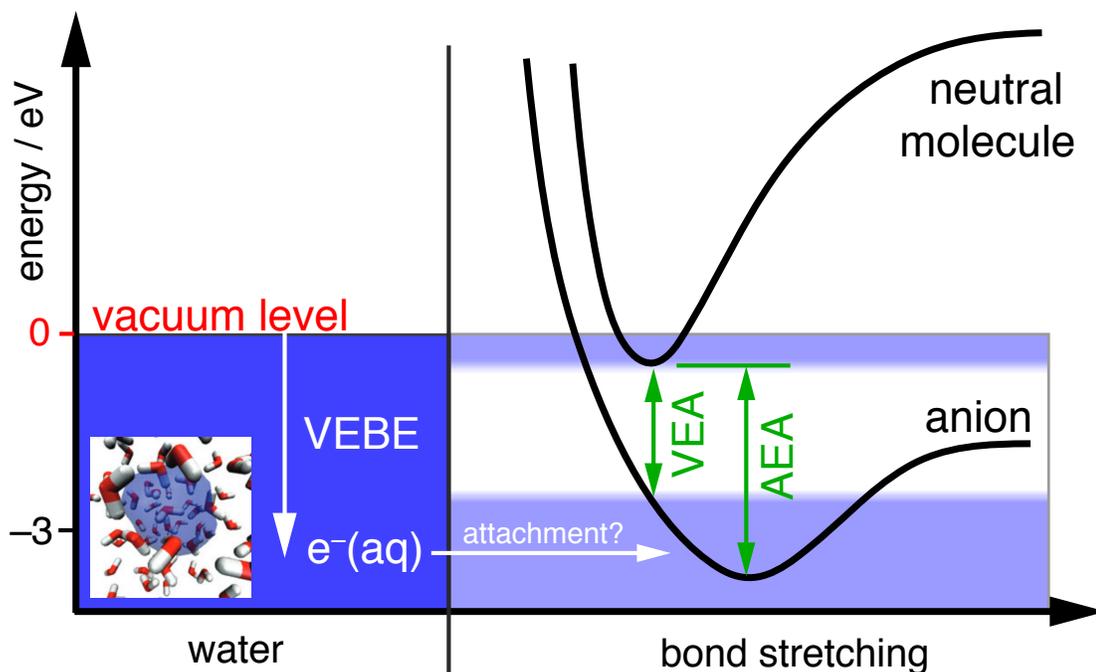


Figure 2.11: Hypothetical potential energy curves for dissociative electron attachment by an aqueous electron. The molecular anion formed by electron attachment is stabilized by aqueous solvation, relative to its neutral parent molecule. The white band on the right illustrates the window between the vertical electron affinity (VEA) and the adiabatic electron affinity (AEA) in which electron attachment is energetically feasible; the left part of the diagram suggests that the equilibrated species  $e^{-}(\text{aq})$  lies too far below vacuum level for this process to occur. Adapted from Ref. [9].

Notably, the only other microjet experiment to report a feature at 1.6 eV for  $e^-(aq)$  observed this signal to decay in  $\lesssim 100$  fs.<sup>36</sup> That timescale is consistent with the excited-state lifetime of  $e^-(aq)$ ,<sup>114,115,135,136</sup> and thus it is suggested that the 1.6 eV feature might simply be a short-lived excited state of  $e^-(aq)$ .<sup>36</sup> Given that the most probable  $s \rightarrow p$  excitation energy of the equilibrated species in bulk water is 1.7 eV,<sup>39</sup> a VEBE of 1.6 eV for an excited state is consistent with a ground-state VEBE of 3.3 eV.

Results presented here appear to close the door on the notion that an electron at the surface of neat liquid water is responsible for the 1.6 eV binding energy that is reported in Ref. [9], but leave open the question of the chemical identity of the species responsible for this feature. Its energetics are uncannily similar to the  $s \rightarrow p$  excited state of  $e^-(aq)$ ,<sup>5</sup> whether in bulk water or at the interface, but the  $\gtrsim 100$  ps lifetime for this signal that is reported in Ref. [9] is inconsistent with the  $< 1$  ps excited-state lifetime that is reported elsewhere.<sup>114,115,135,136</sup> There is speculation that this signal might be attributable to a Rydberg anion supported by  $H_3O^+$  at the interface, where the hydronium ion might occur naturally or might have been created photochemically as a side product of two-photon ionization of water.<sup>43</sup> Such speculation may suggest new calculations aimed to interpret experimental measurements that seek to understand the nature of the water/vacuum interface.

## Chapter 3: Continuum Solvation Models

### 3.1 The Road to Implicit Solvation

As technology continues to push computing capabilities to astounding limits, quantum mechanical (QM) methodologies can investigate increasingly more elaborate chemical phenomenon and are becoming ubiquitous in other disciplines of science. Perhaps one of the most useful utilities of quantum chemistry is its ability to treat large molecular systems embedded in complex environments, which makes it possible to, for example, simulate biologically relevant processes or examine the properties of new materials. This computational feat is made possible by the development of solvation models that treat the environment classically. The classical description of the solvent is either made explicitly through the use of molecular mechanics (MM) or implicitly using continuum models. Although the computational cost for a large MM representation of the solvent is negligible relative to the expense of a molecular QM calculation for the solute, an exhaustive amount of configurational sampling must be performed to observe the desired chemical phenomenon such as protein folding in aqueous solution. An implicit treatment of the solvent environment with a uniform dielectric media obviates the need for extensive sampling by accounting for thermal

fluctuations of the solvent in an average way. The tradeoff for this computational efficiency is the loss of important solute-solvent interactions such as hydrogen-bonding which can be captured only with explicit solvent molecules. However, it is possible to layer these methodologies by including a smaller MM region (e.g., one or two solvation shells of water molecules) around the QM solute molecule of interest, and then immersing the QM/MM system in a dielectric continuum. The rest of this introductory section will focus on the origin of continuum solvation models, beginning with Max Born’s model for solvating a point charge, continuing with a discussion of John Kirkwood and Lars Onsager’s seminal work on solvating a point dipole in polar media, and ending with a survey of modern methods.

Before proceeding it should be pointed out that to obtain chemically accurate solvation free energies with continuum models, it is necessary to incorporate interactions that go beyond electrostatic contributions. However, these nonelectrostatic terms, which account for cavitation energies, attractive van der Waals interactions, repulsive solute-solvent steric interactions, and thermal motions of the nuclei, will not be considered any further, and the solvation free energies discussed below reflect only the electrostatic contributions. The underlying physical basis for continuum solvation models is relatively straightforward: upon immersion of a solute in a dielectric medium, the solute charge distribution polarizes the surrounding dielectric medium. This interaction results in a polarization potential that is a self-consistent solution to Poisson’s equation, which directly influences the classical continuum and indirectly affects the quantum charge distribution through Schrödinger’s equation. The interaction of the resulting polarization potential,  $\varphi_{\text{pol}}(\mathbf{r})$ , with the solute charge density,  $\rho_{\text{sol}}(\mathbf{r})$ , gives rise to the polarization free energy,  $G_{\text{pol}}$ , and the electrostatic

contribution to the total free energy of the solvated solute molecule,  $G_{\text{el}}$ , which is relative to a reference system where the energy is taken to be zero, is expressed as

$$\begin{aligned} G_{\text{el}} &= E_{\text{int}} + \frac{1}{2} \int d\mathbf{r} \varphi_{\text{pol}}(\mathbf{r}) \rho_{\text{sol}}(\mathbf{r}) \\ &= E_{\text{int}} + G_{\text{pol}} , \end{aligned} \tag{3.1}$$

where  $E_{\text{int}}$  is the internal energy of the solute and the factor of one-half in  $G_{\text{pol}}$  accounts for the work associated with polarizing the dielectric medium.

### 3.1.1 Generalized Born Model

By way of introducing the fundamental concepts of continuum solvation, we begin with a simplistic model<sup>137</sup> proposed by Max Born in 1920. The model sets forth to approximate  $G_{\text{pol}}$  for solvating a point charge of strength  $Ze$  located in the center of solvent-excluded sphere of radius  $a$  within a dielectric medium. We define  $\Delta G_{\text{Born}}$  in a slightly different manner than  $G_{\text{el}}$  by making reference to a system composed of a pure liquid at equilibrium and an isolated solute molecule in vacuum:

$$\begin{aligned} \Delta G_{\text{Born}} &= G_{\text{el}} - E_{\text{vac}} \\ &= E_{\text{int}} - E_{\text{vac}} + G_{\text{pol}} , \end{aligned} \tag{3.2}$$

where  $E_{\text{vac}}$  is the self-energy of a point charge in vacuum. The terms  $E_{\text{int}}$  in Eq. 3.1 and  $E_{\text{vac}}$  in Eq. 3.2 are not necessarily the same quantity: in the context of QM calculations,  $E_{\text{int}}$  is the eigenvalue of a gas-phase Hamiltonian in which the eigenvectors have been polarized by the dielectric medium. Therefore,  $E_{\text{int}}$  is not truly the gas-phase internal energy of the solute molecule because its underlying wave function has been polarized. (A more in-depth discussion of continuum model/QM calculations is presented later). However, for the simplistic model we are considering here,  $E_{\text{int}} = E_{\text{vac}}$  and  $\Delta G_{\text{Born}} = G_{\text{pol}}$

The theoretical origin of the Born model is Poisson's equation,

$$\hat{\nabla} \cdot [\epsilon(\mathbf{r}) \hat{\nabla} \varphi_{\text{tot}}(\mathbf{r})] = -4\pi \rho_{\text{sol}}(\mathbf{r}) , \quad (3.3)$$

which provides the framework for computing the total electrostatic potential,  $\varphi_{\text{tot}}$ , for a given solute charge density interacting with a continuum dielectric described by  $\epsilon(\mathbf{r})$ . Once  $\varphi_{\text{tot}}$  is obtained from the solution of Eq. 3.3, the electrostatic interaction energy is computed as

$$\begin{aligned} G_{\text{el}} &= \frac{1}{8\pi} \int d\mathbf{r} \mathbf{E}(\mathbf{r}) \cdot \mathbf{D}(\mathbf{r}) \\ &= \frac{1}{8\pi} \int d\mathbf{r} \epsilon(\mathbf{r}) |\hat{\nabla} \varphi_{\text{tot}}(\mathbf{r})|^2 , \end{aligned} \quad (3.4)$$

where  $\mathbf{E}(\mathbf{r}) = -\hat{\nabla} \varphi_{\text{tot}}(\mathbf{r})$  is the total electric field and  $\mathbf{D}(\mathbf{r})$  is the displacement field composed of  $\mathbf{E}(\mathbf{r})$  and the polarization field,  $\mathbf{P}(\mathbf{r})$ :

$$\begin{aligned} \mathbf{D}(\mathbf{r}) &= \mathbf{E}(\mathbf{r}) + 4\pi \mathbf{P}(\mathbf{r}) \\ &= \epsilon(\mathbf{r}) \mathbf{E}(\mathbf{r}) . \end{aligned} \quad (3.5)$$

Furthermore, we show that  $G_{\text{el}}$  defined by Eq. 3.4 is identical to the first definition given in Eq. 3.1 by integrating the second equality by parts and applying Eq. 3.3 to yield

$$G_{\text{el}} = \frac{1}{2} \int d\mathbf{r} \varphi_{\text{sol}}(\mathbf{r}) \rho_{\text{sol}}(\mathbf{r}) + \frac{1}{2} \int d\mathbf{r} \varphi_{\text{pol}}(\mathbf{r}) \rho_{\text{sol}}(\mathbf{r}) , \quad (3.6)$$

where  $\varphi_{\text{tot}}(\mathbf{r})$  is decomposed into the solute molecule electrostatic potential,  $\varphi_{\text{sol}}(\mathbf{r})$ , and the polarization potential,  $\varphi_{\text{pol}}(\mathbf{r})$ , generated from the solute interaction the dielectric medium. By comparing Eqs. 3.1 and 3.6, it is immediately obvious that the second term on the right-side of Eq. 3.6 is  $G_{\text{pol}}$  and the first term is an expression for the internal energy of the solute,  $E_{\text{int}}$ .

The simplicity of the Born model lies in the assumption of a uniform dielectric for both the solvating environment and in vacuum, and therefore  $\epsilon(\mathbf{r}) \rightarrow \epsilon_{\text{r}}$  where  $\epsilon_{\text{r}}$

either takes the value of  $\epsilon_{\text{solv}}$  for the solvent or 1 in vacuum. Therefore, Eq. 3.3 can be simplified to a second-order differential equation:

$$\nabla^2 \varphi_{\text{tot}}(\mathbf{r}) = -\frac{4\pi Ze\delta(\mathbf{r})}{\epsilon_0 \epsilon_r}, \quad (3.7)$$

where the  $Ze\delta(\mathbf{r})$  on the right side of Eq. 3.7 describes the point charge centered at the origin.  $\Delta G_{\text{Born}}$  is then computed as the difference between electrostatic energies given by Eq. 3.4 using the respective dielectric constants and the second equality in Eq. 3.5 to define the displacement field. Putting all these ingredients together yields

$$\Delta G_{\text{Born}} = \frac{\epsilon_{\text{solv}}}{8\pi} \int d\mathbf{r} |\mathbf{E}_{\text{el}}(\mathbf{r})|^2 - \frac{1}{8\pi} \int d\mathbf{r} |\mathbf{E}_{\text{vac}}(\mathbf{r})|^2, \quad (3.8)$$

and the exact electric fields are derived from Eq. 3.3 with the respective dielectric constants:

$$\mathbf{E}_{\text{el}}(\mathbf{r}) = \begin{cases} \frac{Ze}{\epsilon_0} \frac{\mathbf{r}}{r^3} & \text{for } r < a \\ \frac{Ze}{\epsilon_{\text{solv}}} \frac{\mathbf{r}}{r^3} & \text{for } r \geq a \end{cases} \quad (3.9)$$

and

$$\mathbf{E}_{\text{vac}}(\mathbf{r}) = \frac{Ze}{\epsilon_0} \frac{\mathbf{r}}{r^3}, \quad (3.10)$$

where  $\epsilon_0$  is the permittivity of free space and the point charge is located at the origin which defines the vector  $\mathbf{r}$ . With the electric fields in hand, the Born model solvation energy in atomic units is given by

$$\Delta G_{\text{Born}} = -\frac{Z^2}{2a} \left( \frac{\epsilon_{\text{solv}} - 1}{\epsilon_{\text{solv}}} \right) \quad (3.11)$$

(Note that  $4\pi\epsilon_0 = 1$  and  $e = 1$  in atomic units.) Physically, the quantity in Eq. 3.11 illustrates two important features that manifest in a continuum description of solvation: the interaction of charges with a dielectric medium must lower the total energy of the system and the magnitude of the polarization free energy is inversely

proportional with respect to the size of the cavity in which the charge resides. The generalized Born model (GBM), which extends this model to a set of  $N$  interacting charged spheres, required significant effort,<sup>138</sup> and an effective Coulomb operator  $r_{\text{eff}}^{-1}$  was derived to account for the interactions among the charged spheres as well as the surrounding dielectric medium. The resulting expression for the polarization free energy in this case is given by

$$\Delta G_{\text{Born}} = - \left( \frac{\epsilon_{\text{solv}} - 1}{\epsilon_{\text{solv}}} \right) \sum_{ij}^N \frac{Z_i Z_j}{2r_{\text{eff}}}, \quad (3.12)$$

where  $r_{\text{eff}}$  takes the form

$$r_{\text{eff}} = \sqrt{r_{ij}^2 + a_{ij}^2 e^{-D_{ij}}}. \quad (3.13)$$

In Eq. 3.13,  $r_{ij} = |\mathbf{r}_i - \mathbf{r}_j|$  is the distance between two point charges  $Z_i$  and  $Z_j$ ,  $a_{ij}$  is the geometric mean of two Born radii  $a_i$  and  $a_j$ , and the exponent  $D_{ij} = r_{ij}^2/2a_{ij}^2$  is necessary to smoothly transition  $r_{\text{eff}}$  to the usual Coulomb operator when two charged spheres are merged together to a dampened Coulomb operator for spheres that are separated by large distances.<sup>62</sup> The accuracy of the GBM depends sensitively on the underlying description of the Born radii, which for a large molecular system, define the solute cavity that separates it from the surrounding solvent. Although a few of definitions have been proposed<sup>139,140</sup> for determining the Born radii, perhaps the most useful parameterization for use in large protein systems arises by minimizing the differences between Born radii computed using Eq. 3.12 and those computed by finite differencing the Poisson-Boltzmann equation.<sup>141</sup> It is important to note that issues arising in the determination of such a solute-solvent boundaries are ubiquitous in continuum solvation models and will be a recurring theme in the upcoming sections. There is no unique methodology for defining the solute cavity encompassing

the molecular region, however there are certain physical features that should be captured. We reserve a discussion of this for Sections 3.4.3 and 3.4.4 where we explore the nature of the solute cavity in the context of modern continuum models by examining the successes and pitfalls of common definitions and then proposing techniques to improve them. This concludes our brief tour of the Born model and its variations where we have introduced the fundamental concepts of continuum solvation. We now move onto a discourse of the Kirkwood-Onsager solvation model that laid the theoretical foundation on which more sophisticated models were built.

### 3.1.2 Kirkwood-Onsager Model

The discussion of the previous section revolved around the concept of solvating a monopole in a solvent-excluded region characterized by a sphere of radius  $a$ . To extend this idea further, the natural progression is to consider a point dipole which is meant to model a molecule with no net electric charge but possesses a nonzero dipole moment,  $\boldsymbol{\mu} = \mu_0 \mathbf{u}_r + \alpha \mathbf{F}$ , where  $\mu_0$  is its permanent dipole moment,  $\mathbf{u}_r$  is the axis along which its dipole is oriented,  $\alpha$  is the polarizability of the molecule, and  $\mathbf{F}$  is an external electric field. In Lars Onsager's seminal work<sup>142</sup> describing the interaction of a point dipole with a polarizable medium in terms of a total electrostatic potential, he set out to improve the theory of dipoles first explored by Peter Debye. Debye's dipole theory qualitatively captures the dielectric properties in dielectric liquid, and posits that equality exists between an *internal field* that polarizes a molecule in a dielectric medium and an *external field* representative of  $(4\pi/3)$  times the electric moment induced per unit volume of the liquid.<sup>142</sup> Out of Debye's analysis arose the idea that the force from a polarized dielectric medium acting on an electrically asymmetric

molecule is directly related to this *internal field*, which leads to a formula for the dielectric medium of a substance depending on  $\boldsymbol{\mu}$ . Since real molecules possess a positive polarizability, interaction with a polarizing medium will act to enhance its own electrical asymmetry, and an assumption of a static dipole moment is therefore problematic. In fact, this “self-enhancement” is through a *reaction field*, and in Ref. [142] the fundamental concepts of a molecular cavity and this *reaction field* are introduced. It is on these foundations that all modern continuum solvation models are built. As will be seen in the following section, the continuum models employ a reaction field operator as a perturbative potential in the electronic Hamiltonian, and it is through this perturbation that environmental effects are captured in *ab initio* calculations.

A couple years prior to Onsager’s work, John Kirkwood formulated<sup>143</sup> the interaction energy for a classical charge distribution arranged in a solvent-excluded sphere of radius  $a$  in terms of a multipole expansion (MPE). Such a distribution of classical point charges is represented as  $\rho_{\text{sol}}(\mathbf{r}') = e \sum_m^N Z_m \delta(\mathbf{r}' - \mathbf{r}_m)$ . The electrostatic potential arising from both the solute charge distribution (first term) and the solvent reaction potential (second term) are derived from Poisson’s equation with appropriate boundary conditions, and is expressed as<sup>144</sup>

$$\varphi_{\text{tot}}(r) = \sum_k^{\infty} \frac{1}{r^{k+1}} \sum_{l=-k}^k M_k^l Y_k^l(\theta, \phi) + \sum_k^{\infty} c_k r^{k+1} \sum_{l=-k}^k M_k^l Y_k^l(\theta, \phi), \quad (3.14)$$

where  $M_k^l$  are known as the multipole moments defined as

$$M_k^l = \int_0^{\infty} d\mathbf{r}' \int_0^{\pi} d\theta \sin(\theta) \int_0^{2\pi} d\phi Y_k^l(\theta, \phi) |\mathbf{r}'|^{2+k} \rho_{\text{sol}}(\mathbf{r}') \quad (3.15)$$

and  $Y_k^l(\theta, \phi)$  are the spherical harmonic functions. The reaction potential coefficients,  $c_k$ , contain the effect of the solvent dielectric and take the form

$$c_k = -\frac{(k+1)(\epsilon_{\text{solv}} - 1)}{k + (k+1)\epsilon_{\text{solv}}} \frac{1}{a^{2k+1}}. \quad (3.16)$$

The free energy of solvation is one-half the solute-solvent interaction energy, which is expressed elegantly in the compact form<sup>62</sup>

$$\Delta G_{MPE} = \frac{1}{2} \sum_k \sum_l M_k^l R_k^l. \quad (3.17)$$

where  $R_k^l$  is the reaction field component corresponding to multipole moment  $M_k^l$ , which can be written as<sup>62</sup>

$$R_k^l = \sum_{k'} \sum_{l'} f_{kk'}^{ll'} M_{k'}^{l'}, \quad (3.18)$$

and  $f_{kk'}^{ll'}$  are the generalized reaction field coefficients.

Returning now to Onsager's work for computing the interaction of a point dipole with a dielectric medium, the results of Kirkwood's multipole analysis (Eqs. 3.14 and 3.15) can be applied to derive the necessary electrostatic potentials, where  $k = 1$  is appropriate for the dipole terms, and in polar coordinates the expressions are<sup>142</sup>

$$\mathbf{E}_{\text{el}}(\mathbf{r}, \theta) = \begin{cases} -\left(\frac{2\mu}{r^3} + D\mu\right) \frac{\cos(\theta)}{\epsilon_0} \mathbf{u}_{\mathbf{r}} + \left(D\mu - \frac{\mu}{r^3}\right) \frac{\sin(\theta)}{\epsilon_0} \mathbf{u}_{\theta} & \text{for } r < a \\ -\frac{2\mu\mu^*}{r^3} \frac{\cos(\theta)}{\epsilon_{\text{solv}}} \mathbf{u}_{\mathbf{r}} - \frac{\mu\mu^*}{r^3} \frac{\sin(\theta)}{\epsilon_{\text{solv}}} \mathbf{u}_{\theta} & \text{for } r \geq a \end{cases} \quad (3.19)$$

and

$$\mathbf{E}_{\text{vac}}(\mathbf{r}, \theta) = -\frac{2\mu \cos(\theta)}{r^3} \frac{1}{\epsilon_0} \mathbf{u}_{\mathbf{r}} - \frac{\mu \sin(\theta)}{r^3} \frac{1}{\epsilon_0} \mathbf{u}_{\theta}, \quad (3.20)$$

where  $\mathbf{u}_{\mathbf{r}}$  and  $\mathbf{u}_{\theta}$  are unit vectors along the dipole moment and rotation axes, respectively. The pinnacle of Onsager's analysis is the arrival at the expressions for  $\mu^*$  and  $D$ , the former representing the strength of the external force exerted on a charge in the dielectric medium and the latter measures the strength of the reaction field generated

by the electric field acting on the dipole as a result of its own electric displacements **D**.<sup>142</sup> Onsager showed that

$$\mu^* = \left( \frac{3\epsilon_{\text{solv}}}{2\epsilon_{\text{solv}} + 1} \right) \quad (3.21)$$

and

$$D = \frac{2(\epsilon_{\text{solv}} - 1)}{2\epsilon_{\text{solv}} + 1} \frac{1}{a^3}, \quad (3.22)$$

which is equivalent to Eq. 3.16 with  $k = 1$ . The free energy of solvation for the Onsager model in the absence of an external electric field,  $\Delta G_{\text{Onsager}}$ , is computed as before for the Born model with the electric fields given by Eqs. 3.19 and 3.20:

$$\Delta G_{\text{Onsager}} = -\frac{D\mu_0^2}{2}, \quad (3.23)$$

which is equivalent to Eq. 3.17 for the case of a point dipole in a sphere of radius  $a$ .  $\Delta G_{\text{Onsager}}$  in Eq. 3.23 can be made arbitrarily accurate by adjusting the size of the solvent-excluded sphere. However, there are a variety of procedures for estimating the size of the sphere, the most common of which is the relationship

$$a = \left( \frac{3M_v}{4\pi N_A} \right)^{1/3}, \quad (3.24)$$

where  $M_v$  is the molar volume of the solute and  $N_A$  is Avogadro's number. Another way to estimate  $a$  is to set this value equal to the maximum distance between the solute center of mass and solute atoms, where the distances include the atomic van der Waals radii. A second alternative to Eq. 3.24 is to set  $2a$  equal to the largest solute-solvent internuclear distance, again with the atomic van der Waals radii included.

We conclude this section with a discussion of transitioning the classical description of these models to one that incorporates quantum effects. The charge distributions considered were arbitrary arrangements of point charges in a solvent-excluded sphere,

but Kirkwood<sup>143</sup> posits that the results are general such that a continuous charge distribution can be used. Furthermore, these continuous charge distributions can be described with the aid of quantum mechanics that employ the molecular wave function for the electrons, while the nuclear components are satisfactorily described as point charges. This was done for the electronic components of the MPE, where the  $M_k^l$  terms were recast in terms of one-electron integrals using atom-centered basis functions  $g_\mu(\mathbf{r})$  and the corresponding one-electron density matrix  $\mathbf{P}$ , as shown here:<sup>62,145</sup>

$$M_k^l(I) = \sum_{\mu \in I} \sum_{\nu \in I} P_{\mu\nu} \langle g_\mu(\mathbf{r}) | \hat{M}_k^l | g_\nu(\mathbf{r}) \rangle . \quad (3.25)$$

In Eq. 3.25, the multipole moments are computed for each nucleus  $I$  and therefore only the basis functions centered on each nucleus are considered. The reaction field specified by Eq. 3.18 can then be added to the molecular Hamiltonian, and a self-consistent procedure is performed to simultaneously update the electronic wave function which in turn produces a new reaction field. This describes a quantum mechanical implementation of the Kirkwood-Onsager solvation model. In the final introductory section below, the computational details of a self-consistent reaction field (SCRF) technique that exploits Onsager’s dipole reaction field to describe solvent environmental effects is outlined, and its performance is discussed.

### 3.1.3 The Onsager Self-Consistent Reaction Field Model

For the remaining part of this work, reference will be made to an effective Hamiltonian that is defined explicitly here:

$$\hat{H}_{\text{eff}} = \hat{H}_0 + \hat{V} , \quad (3.26)$$

where  $\hat{H}_0$  is the molecular Hamiltonian that governs the internal energy of the solute and  $\hat{V}$  is the reaction field operator that acts on the solute charge distribution to yield

the electrostatic effects from the surrounding dielectric environment. The functional form of  $\hat{H}_0$  depends on the chosen *ab initio* method such as HF, MP2, or DFT (see Chapter 1). Likewise the form  $\hat{V}$  takes depends on the desired description of the solute-solvent interaction (e.g. the multipole moment operator  $\hat{M}_k^l$  of Eq. 3.25). In the original implementation of the Onsager-SCRF, both HF and MP2 methods are employed for the QM description of the solute molecules, and  $\hat{H}_0$  is the Hartree-Fock operator represented as the Fock matrix  $\mathbf{F}$ . The reaction field operator couples the solute and solvent together through the molecular dipole moment and takes the form  $\hat{V} = -\hat{\boldsymbol{\mu}} \cdot \vec{\mathbf{R}}$ , and  $\vec{\mathbf{R}} = D\vec{\boldsymbol{\mu}}$ . The strength of the Onsager dipole reaction field  $\vec{\mathbf{R}}$  is controlled through  $D$  (see Eq. 3.22 and surrounding discussion) and  $\vec{\boldsymbol{\mu}}$  is the molecular dipole moment comprised of the electronic and nuclear components defined as<sup>146</sup>

$$\vec{\boldsymbol{\mu}} = \left( \frac{Q_{\text{sol}} + N_{\text{elec}}}{N_{\text{elec}}} \right) \langle \Psi | \hat{\boldsymbol{\mu}}_{\text{elec}} | \Psi \rangle + \sum_i^{N_{\text{atom}}} q_i \mathbf{r}_i . \quad (3.27)$$

In Eq. 3.27,  $Q_{\text{sol}}$  is the total charge of the molecule,  $N_{\text{elec}}$  is the total number of electrons,  $|\Psi\rangle$  is the molecular wave function,  $\hat{\boldsymbol{\mu}}_{\text{elec}}$  is the electronic dipole moment operator expressed as

$$\hat{\boldsymbol{\mu}}_{\text{elec}} = \hat{x}\hat{\mathbf{x}} + \hat{y}\hat{\mathbf{y}} + \hat{z}\hat{\mathbf{z}} . \quad (3.28)$$

The second term on the right side of Eq. 3.27 is the classical dipole moment operator where each  $q_i$  is a nuclear charge and  $\mathbf{r}$  is the vector of nuclear positions referenced to an arbitrary origin that is generally taken to be the charge distribution centroid.

Sections 3.1.1 and 3.1.2 have dealt with deriving polarization free energies employing the principles of classical electrostatics. In this section we have introduced an effective QM Hamiltonian that quantizes the Onsager dipole reaction field. As discussed in Chapter 1, within the context of HF, MP2, and DFT calculations, the

Hamiltonians represented by Fock matrices are nonlinear and require a self-consistent procedure for the simultaneous solution of the molecular wave function (i.e. the optimized molecular orbitals) and corresponding energy. Therefore to explicitly incorporate the Onsager dipole reaction field into the effective Hamiltonian and self-consistent field equations, the Fock matrix elements are modified according to<sup>146</sup>:

$$F_{\mu\nu} = F_{\mu\nu}^0 - \vec{\mathbf{R}} \cdot \langle g_{\mu}(\mathbf{r}) | \hat{\boldsymbol{\mu}} | g_{\nu}(\mathbf{r}) \rangle , \quad (3.29)$$

where  $F_{\mu\nu}^0$  corresponds to the Fock matrix for the solute in the absence of solvent and  $g_{\mu}(\mathbf{r})$  is an atom-centered basis function. The total free energy corresponding to the Fock matrix in Eq. 3.29 is expressed as<sup>146</sup>

$$G_{\text{el}} = \mathbf{F}^0 \cdot \mathbf{P} - \frac{1}{2} \vec{\boldsymbol{\mu}} \cdot \vec{\mathbf{R}} , \quad (3.30)$$

where the first term on the right side is the solute internal energy,  $E_{\text{int}}$ , that is computed as the trace of the unperturbed solute Fock matrix with the one-electron density matrix,  $\mathbf{P}$ , and the second term is the polarization free energy. For neutral molecules,  $Q_{\text{sol}} = 0$  and there is no monopole contribution to the solvation energy, and therefor only the dipole moment of the molecule will lead to solvation (granted that it is nonzero). However, for nonzero  $Q_{\text{sol}}$ , it is trivial to incorporate the Born model monopole term by adding Eq. 3.11 (with  $Z = Q_{\text{sol}}$ ) to Eq. 3.30.

In Refs. [146] and [147], Eqs. 3.29 and 3.30, along with the corresponding gradients, were implemented to investigate solvent effects on various molecular properties, rotational barriers, charge distributions, and the equilibria between conformational isomers of different ionic and neutral dipolar organic molecules. The sensitive behavior of the Onsager model of solvation with respect to the radius  $a$  of the solvent-excluded sphere is highlighted by some computed molecular properties

of 1,2-dichloroethane. For example, the gauche–trans rotational barrier for 1,2-dichloroethane in liquid water at the HF and MP2 levels of theory can be made arbitrarily negative for small values of the cavity, for which the dipole moment is made larger due to the inverse relationship of the solvation free energy and cavity radius, or can be made arbitrarily positive for large values. Using an alternative definition for the cavity radius that computes  $a$  as the largest internuclear distance with the van der Waals radii included, both HF and MP2 predict relatively accurate barrier heights in different solvents with low and high dielectrics. Furthermore, for ionic molecules that include both the monopole and dipole terms, it was found that high dielectric solvents have a dramatic effect on the molecular structure.

We conclude Section 3.1 by noting the success of Onsager’s reaction field theory for incorporating solvation effects in QM methods. With the theoretical foundation for implicit solvation laid, we move onto modern methods. In Section 3.2 we explore isotropic polarizable continuum models (PCMs) for equilibrium and non-equilibrium solvation processes. We then show how these methodologies are applied to compute vertical ionization energies for  $e^-(aq)$ , alkali metal cations, and halide ions in liquid water. In Section 3.3 we present a recently developed method for computing equilibrium solvation free energies for an arbitrary description of the dielectric medium, and then discuss a method to incorporate of non-equilibrium solvent effects. This state-of-the-art technique is then applied to the aforementioned systems with an isotropic description of the dielectric for validation against experimental measurements and PCM calculations, and also with an anisotropic dielectric in order to predict vertical ionization energies at the liquid-vapor interface.

## 3.2 Isotropic Solvation with Polarizable Continuum Models

This section discusses the formulation of modern PCMs within the context of apparent surface charge (ASC) techniques, where the methodology to compute  $\varphi_{\text{tot}}(\mathbf{r})$  through Poisson's equation in Eq. 3.3 is combined with electronic structures methods employing Eq. 3.26 to compute the solute charge distribution in the presence of an induced polarization potential,  $\varphi_{\text{pol}}(\mathbf{r})$ . In Section 3.1.1 it was shown that the polarization free energy is computed as a free energy difference between an ion in a solvent-excluded sphere in dielectric medium and the ion in vacuum, and this free energy depends sensitively on the sphere radius,  $a$ . Modern PCMs describe the solvent-excluded cavity with complex shapes and sizes, and employ a variety of sophisticated algorithms to accomplish this. A more thorough discussion of this topic is presented in Section 3.4.3, but for the time being we will denote a complex molecular cavity by  $\mathbf{M}$ . Similar to the Born solvation model, modern PCMs make use of a simplistic, discontinuous form of  $\epsilon(\mathbf{r})$  that mathematically manifests as<sup>148</sup>

$$\epsilon(\mathbf{r}) = \begin{cases} 1 & \text{for } \mathbf{r} \in \mathbf{M} \\ \epsilon_{\text{solv}} & \text{for } \mathbf{r} \notin \mathbf{M} \end{cases} \quad (3.31)$$

This assumption for the dielectric function reduces the complexity of Poisson's such that two equations for  $\varphi_{\text{tot}}(\mathbf{r})$  must be solved simultaneously, and are expressed as

$$\varphi_{\text{tot}}(\mathbf{r}) \rightarrow \begin{cases} \nabla^2 \varphi_{\text{tot}}(\mathbf{r}) = -4\pi\rho_{\text{sol}}(\mathbf{r}) & \text{for } r \in \mathbf{M} \\ -\epsilon_{\text{solv}} \nabla^2 \varphi_{\text{tot}}(\mathbf{r}) = 0 & \text{for } r \notin \mathbf{M} , \end{cases} \quad (3.32)$$

where we explicitly note that the total electrostatic potential is comprised of the solute and solvent-induced polarization components:

$$\varphi_{\text{tot}}(\mathbf{r}) = \varphi_{\text{sol}}(\mathbf{r}) + \varphi_{\text{pol}}(\mathbf{r}) \quad (3.33)$$

The potential and its electric field must be continuous across the surface,  $\Gamma$ , of  $\mathbf{M}$ , and therefore<sup>62</sup>

$$[\varphi_{\text{tot}}(\mathbf{r} \in \Gamma)] = \varphi_{\text{tot}}^{\mathbf{r} \in \mathbf{M}}(\mathbf{r} \in \Gamma) - \varphi_{\text{tot}}^{\mathbf{r} \notin \mathbf{M}}(\mathbf{r} \in \Gamma) = 0 \quad (3.34)$$

$$[\partial\varphi_{\text{tot}}(\mathbf{r} \in \Gamma)] = \left( \frac{\partial\varphi_{\text{tot}}}{\partial\mathbf{n}} \right) \Big|_{\mathbf{r} \in \mathbf{M}} - \epsilon_{\text{solv}} \left( \frac{\partial\varphi_{\text{tot}}}{\partial\mathbf{n}} \right) \Big|_{\mathbf{r} \notin \mathbf{M}} = 0, \quad (3.35)$$

where it is noted that the discontinuity of the electric field involves only its normal components,  $\mathbf{n}$ . Taken together, Eqs. 3.31–3.35 lay the groundwork for ASC methods. In such methods, the jump conditions given in Eqs. 3.34 and 3.35 limit the solution of Poisson’s equation to  $\Gamma$ , and this leads to an unequivocal definition of  $\varphi_{\text{pol}}(\mathbf{r})$  expressed as

$$\varphi_{\text{pol}}(\mathbf{r}) = \int_{\Gamma} d\mathbf{s} \frac{\sigma_{\text{pol}}(\mathbf{s})}{|\mathbf{r} - \mathbf{s}|}, \quad (3.36)$$

where the integral is over the molecular cavity surface  $\Gamma$ , the coordinates of which are denoted by  $\mathbf{s}$ , and  $\sigma_{\text{pol}}(\mathbf{s})$  are the ASCs induced by  $\rho_{\text{sol}}(\mathbf{r})$  at  $\Gamma$ . The solute charge density  $\rho_{\text{sol}}(\mathbf{r})$  can be expressed as the sum of its electronic and nuclear contributions:

$$\begin{aligned} \rho_{\text{sol}}(\mathbf{r}) &= \rho_{\text{elec}}(\mathbf{r}) + \rho_{\text{nuc}}(\mathbf{r}) \\ &= |\Psi(\mathbf{r})|^2 + \sum_{\alpha}^{N_{\text{atom}}} Z_{\alpha} \delta(\mathbf{r} - \mathbf{R}_{\alpha}), \end{aligned} \quad (3.37)$$

where the  $|\Psi(\mathbf{r})|^2$  is the modulus-square of the electronic wave function and the nuclear component is represented by classical point charges. The linearity of Poisson’s equation allows the reaction potential operator in Eq. 3.26 that generates  $\sigma_{\text{pol}}(\mathbf{s})$  to be split into its electronic and nuclear components as

$$\begin{aligned} \hat{V} &= \hat{V}_{\text{elec}} + \int_{\Gamma} d\mathbf{s} \frac{\sigma_{\text{pol}}^{\text{nuc}}(\mathbf{s})}{|\mathbf{r} - \mathbf{s}|} \\ &= \hat{V}_{\text{elec}} + \varphi_{\text{pol}}^{\text{nuc}}(\mathbf{r}), \end{aligned} \quad (3.38)$$

where  $\sigma_{\text{pol}}^{\text{nuc}}(\mathbf{s})$  are the ASCs generated at  $\Gamma$  arising from  $\rho_{\text{nuc}}(\mathbf{r})$ , and  $\varphi_{\text{pol}}^{\text{nuc}}(\mathbf{r})$  is the corresponding nuclear contribution to the polarization potential. The polarization

free energy,  $G_{\text{el}}$ , is then computed as

$$\begin{aligned}
G_{\text{el}} &= \frac{1}{2} \langle \Psi | \hat{V}_{\text{elec}} | \Psi \rangle + \frac{1}{2} \int_{\mathbf{V}} d\mathbf{r} \rho_{\text{nuc}}(\mathbf{r}) \varphi_{\text{pol}}^{\text{nuc}}(\mathbf{r}) \\
&= \frac{1}{2} \int_{\mathbf{V}} d\mathbf{r} \rho_{\text{elec}}(\mathbf{r}) \varphi_{\text{pol}}^{\text{elec}}(\mathbf{r}) + \frac{1}{2} \int_{\mathbf{V}} d\mathbf{r} \rho_{\text{nuc}}(\mathbf{r}) \varphi_{\text{pol}}^{\text{nuc}}(\mathbf{r}) \\
&= \frac{1}{2} \int_{\mathbf{V}} d\mathbf{r} \rho_{\text{sol}}(\mathbf{r}) \varphi_{\text{pol}}(\mathbf{r}),
\end{aligned} \tag{3.39}$$

where the last equality is equivalent to the expression for  $G_{\text{el}}$  in Eq. 3.1 and it is explicitly noted that the integrals are over the volume  $\mathbf{V}$  in  $\mathbb{R}^3$ . Furthermore, the expressions in Eq. 3.39 can be recast in terms of surface integrals, which significantly reduces the computational burden from integrals over  $\mathbb{R}^3$  to  $\mathbb{R}^2$ .<sup>67</sup>

$$\begin{aligned}
G_{\text{el}} &= \frac{1}{2} \int_{\mathbf{V}} d\mathbf{r} \rho_{\text{sol}}(\mathbf{r}) \int_{\Gamma} ds \frac{\sigma_{\text{pol}}(\mathbf{s})}{|\mathbf{r} - \mathbf{s}|} \\
&= \frac{1}{2} \int_{\Gamma} ds \sigma_{\text{pol}}(\mathbf{s}) \int_{\mathbf{V}} d\mathbf{r} \frac{\rho_{\text{sol}}(\mathbf{r})}{|\mathbf{r} - \mathbf{s}|} \\
&= \frac{1}{2} \int_{\Gamma} ds \sigma_{\text{pol}}(\mathbf{s}) \varphi_{\text{sol}}^{\rho}(\mathbf{s}),
\end{aligned} \tag{3.40}$$

where the notation  $\varphi_{\text{sol}}^{\rho}(\mathbf{s})$  is introduced to signify the electrostatic on  $\Gamma$  generated from the solute charge distribution, which is mathematically expressed as

$$\varphi_{\text{sol}}^{\rho}(\mathbf{s}) = \int_{\mathbf{V}} d\mathbf{r} \frac{\rho_{\text{sol}}(\mathbf{r})}{|\mathbf{r} - \mathbf{s}|}. \tag{3.41}$$

We conclude this section by noting that, up to this point in the discussion, no approximations have been made, and Eq. 3.40 is formally the exact solution to the boundary conditions specified by Eqs. 3.31–3.35. In practice, Eqs. 3.36 and 3.40 are used for all flavors of PCM, and the computational problem is reduced to solving for the ASCs on the surrounding molecular cavity. To accomplish this, the molecular cavity surface is discretized into a number of finite elements using an algorithm such as GEOPOL.<sup>149–151</sup> The discretized elements are called tesserae and the GEOPOL algorithm constructs them small enough such that  $\sigma_{\text{pol}}(\mathbf{s})$  is nearly constant within

each. This discretization process transforms the integral expression in Eq. 3.36 into a summation over the surface charge elements  $\sigma_{\text{pol}}(\mathbf{s}_i)$ :

$$\begin{aligned}\varphi_{\text{pol}}(\mathbf{r}) &\simeq \sum_i^{N_{\text{tess}}} \frac{\sigma_{\text{pol}}(\mathbf{s}_i) A_i}{|\mathbf{r} - \mathbf{s}_i|} \\ &\simeq \sum_i^{N_{\text{tess}}} \frac{q_{\text{pol}}(\mathbf{s}_i)}{|\mathbf{r} - \mathbf{s}_i|},\end{aligned}\tag{3.42}$$

where  $\{\mathbf{s}_i\}$  is the set of ASC coordinates for the  $N_{\text{tess}}$  is the number of tesserae needed to discretize the surface. In the next section, the details for computing  $\sigma_{\text{pol}}(\mathbf{s}_i)$  are outlined for various types of PCMs, some of which are employed for the calculations presented in this Chapter.

### 3.2.1 Reference State Formalism

With the foundation of ASC techniques laid out in the previous section, we are now in a position to formally introduce the methodologies for determining  $\sigma_{\text{pol}}(\mathbf{s}_i)$ . All PCM calculations reported in this work employ IEF-PCM<sup>152–154</sup> for computing the ASCs on a molecular cavity surface. Furthermore, the formalisms to follow assume that the solute and solvent are in thermal equilibrium for a given temperature and pressure, and that the solvent is described fully by a homogenous and isotropic dielectric constant. At the beginning of Chapter 3, it was noted in Eq. 3.5 that the displacement field,  $\mathbf{D}(\mathbf{r})$ , can be expressed in two unique ways, and equating these expressions to eliminate  $\mathbf{D}(\mathbf{r})$  allows one to solve for the polarization field in terms of the electric field:

$$\mathbf{P}(\mathbf{r}) = \left( \frac{\epsilon_{\text{solv}} - 1}{4\pi} \right) \mathbf{E}(\mathbf{r}),\tag{3.43}$$

where we have used an isotropic assumption for the solvent dielectric: that is  $\epsilon(\mathbf{r}) = \epsilon_{\text{solv}}$ . The predecessor to CPCM and IEF-PCM, called DPCM<sup>155</sup>, computes  $\sigma_{\text{pol}}(\mathbf{s})$

directly from Eq. 3.43 by enforcing continuity of  $\mathbf{P}(\mathbf{r})$  across the molecular cavity surface. Such a process results in

$$\sigma_{\text{pol}}(\mathbf{s}) = \left( \frac{\epsilon_{\text{solv}} - 1}{4\pi\epsilon_{\text{solv}}} \right) \frac{\partial}{\partial \mathbf{n}} [\varphi_{\text{solv}}^{\rho}(\mathbf{s}) + \varphi_{\text{pol}}(\mathbf{s})] , \quad (3.44)$$

where  $\mathbf{n}$  is the vector normal to the molecular cavity surface, pointing into the dielectric medium, and a new quantity  $\varphi_{\text{pol}}(\mathbf{s})$  has been introduced that is representative of a surface potential generated from  $\sigma_{\text{pol}}(\mathbf{s})$ . The quantity  $\varphi_{\text{pol}}(\mathbf{s})$  can be generated from the action of the surface integral operator  $\hat{S}$  on  $\sigma_{\text{pol}}(\mathbf{s})$  as

$$\hat{S}\sigma_{\text{pol}}(\mathbf{s}) = \int_{\Gamma} ds' \frac{\sigma_{\text{pol}}(\mathbf{s}')}{|\mathbf{s} - \mathbf{s}'|} = \varphi_{\text{pol}}(\mathbf{s}) . \quad (3.45)$$

Furthermore, the expression given by Eq. 3.44 can be recast in terms of an integral operator  $\hat{D}^*$  in matrix form as<sup>148</sup>

$$\left[ 2\pi \left( \frac{\epsilon_{\text{solv}} + 1}{\epsilon_{\text{solv}} - 1} \right) \mathbf{I} - \mathbf{D}^* \right] \sigma_{\text{pol}}(\mathbf{s}) = \frac{\partial}{\partial \mathbf{n}} \varphi_{\text{solv}}^{\rho}(\mathbf{s}) , \quad (3.46)$$

where the action of  $\hat{D}^*$  on  $\sigma_{\text{pol}}(\mathbf{s})$  is expressed as<sup>148</sup>

$$\hat{D}^* \sigma_{\text{pol}}(\mathbf{s}) = \int_{\Gamma} ds' \left[ \frac{\partial}{\partial \mathbf{n}} \frac{1}{|\mathbf{s} - \mathbf{s}'|} \right] \sigma_{\text{pol}}(\mathbf{s}') , \quad (3.47)$$

and  $\mathbf{I}$  is the identity matrix. Taken together, Eqs. 3.46 and 3.47 specify a method for computing the D-PCM ASCs. The CPCM method takes a more simplistic approach to this by equating  $\varphi_{\text{pol}}(\mathbf{s})$  to  $\varphi_{\text{solv}}(\mathbf{s})$  that has been scaled a factor  $f(\epsilon_{\text{solv}})$  as<sup>148</sup>

$$\hat{S}\sigma_{\text{pol}}(\mathbf{s}) = -f(\epsilon_{\text{solv}})\varphi_{\text{solv}}(\mathbf{s}) , \quad (3.48)$$

where

$$f(\epsilon_{\text{solv}}) = \frac{\epsilon_{\text{solv}} - 1}{\epsilon_{\text{solv}}} . \quad (3.49)$$

The IEF-PCM methodology is expressed by the following integral equation<sup>148</sup>

$$\left[ 2\pi \left( \frac{\epsilon_{\text{solv}} + 1}{\epsilon_{\text{solv}} - 1} \right) \mathbf{I} - \mathbf{D} \right] \sigma_{\text{pol}}(\mathbf{s}) = [2\pi \mathbf{I} - \mathbf{D}] \varphi_{\text{sol}}^{\rho}(\mathbf{s}) , \quad (3.50)$$

where  $\mathbf{D}$  is the adjoint of the integral operator  $\hat{D}^*$  in matrix form. As noted in Refs. [97], [98], and [70], IEF-PCM overcomes a technical issue of neglecting the solute electronic density escaping the solvent-excluded molecular cavity. The ideal molecular cavity should contain the entirety of the solute charge distribution. The simplistic DPCM formulation presented above suffers from errors in neglecting electronic charge penetration into the solvent, and it has been shown that IEF-PCM correctly accounts for this implicitly and therefore yields the correct reaction potential within the molecular cavity,<sup>148</sup> which is important in systems like  $e^-(\text{aq})$  where the excess charge density can be diffuse.

Regardless of which PCM method is chosen, once the molecular cavity is tessellated and the numerical mesh for evaluating the quantities in Eqs. 3.46, 3.48, and 3.50 is made concrete, those equations can be reduced to the following linear matrix equation:

$$\mathbf{K}\mathbf{q} = -\mathbf{R}\boldsymbol{\varphi} , \quad (3.51)$$

where  $\mathbf{K}$  and  $\mathbf{R}$  are known matrices that correspond to the respective operators in the aforementioned equations, and  $\mathbf{q}$  and  $\boldsymbol{\varphi}$  are column vectors containing the unknown discretized values of  $q_{\text{pol}}(\mathbf{s}_i)$  and the known values of  $\varphi_{\text{sol}}^{\rho}(\mathbf{s}_i)$ , respectively. Note that the form of ASCs on the left side of Eq. 3.51 has changed from  $\boldsymbol{\sigma}$  to  $\mathbf{q}$  reflecting the specification of the set of tesserae surface area elements  $\{\mathbf{A}\}_{i \in N_{\text{tess}}}$ . The unknown ASCs in Eq. 3.52,  $\mathbf{q}$ , are then computed as

$$\begin{aligned} \mathbf{q} &= -\mathbf{K}^{-1}\mathbf{R} \boldsymbol{\varphi} \\ &= \mathbf{Q} \boldsymbol{\varphi} , \end{aligned} \quad (3.52)$$

which is generally accomplished for small- or medium-sized systems through direct matrix inversion, the scaling of which is  $\mathcal{O}(N_{\text{tess}}^3)$  in time and  $\mathcal{O}(N_{\text{tess}}^2)$  in memory, or for large systems through an iterative conjugate gradient method that scales as  $\mathcal{O}(N_{\text{tess}}^2)$  in time and  $\mathcal{O}(N_{\text{tess}})$  in memory. The notation  $\mathbf{K}^{-1}\mathbf{R} = \mathbf{Q}$  is introduced to represent the polarization weights<sup>156</sup>. In practice  $\mathbf{Q}$  is symmetrized for computational efficiency, and  $\mathbf{q}$  is recast as

$$\begin{aligned}\tilde{\mathbf{q}} &= \left( \frac{\mathbf{Q} + \mathbf{Q}^T}{2} \right) \boldsymbol{\varphi} \\ &= \mathbf{Q}_s \boldsymbol{\varphi} .\end{aligned}\tag{3.53}$$

This section concludes in a similar fashion as Section 3.1.3 with a discussion for incorporating  $\tilde{\mathbf{q}}$  into modern electronic structure methods. Referring back to the effective Hamiltonian in Eq. 3.26, we now recast the reaction field operator in terms of the ASCs. Recognizing that the reaction field operator  $\hat{V}$  generates the polarization potential in Eq. 3.42, the continuous expression of the operator in Eq 3.38 is rewritten in terms of the discrete surface elements as

$$\hat{V} = \sum_i^{N_{\text{tess}}} q_{\text{pol}}(\mathbf{s}_i) \hat{V}(\mathbf{s}_i) ,\tag{3.54}$$

After computing  $\mathbf{q}$  from Eq. 3.53 the solvation free energy is expressed as

$$\begin{aligned}G_{\text{el}} &= \langle \Psi | \hat{H}_0 | \Psi \rangle + \frac{1}{2} \tilde{\mathbf{q}} \cdot \boldsymbol{\varphi} \\ &= E_{\text{int}} + G_{\text{pol}} ,\end{aligned}\tag{3.55}$$

where the second expression in the first or second equality is  $G_{\text{pol}}$  within the PCM framework. To include the effects of the solvent reaction field operator  $\hat{V}$  into electronic structure calculations, the Fock matrix must be modified similar to Eq. 3.29. Formally, the correction that must be added to the isolated solute molecular Fock matrix is the functional derivative of  $G_{\text{el}}$  with respect to the electronic density matrix,

expressed as<sup>148</sup>

$$\begin{aligned}
 F_{\mu\nu} &= \frac{\partial G_{\text{el}}}{\partial P_{\mu\nu}} = \frac{\partial E_{\text{int}}}{\partial P_{\mu\nu}} + \frac{1}{2} \frac{\partial \tilde{\mathbf{q}} \cdot \boldsymbol{\varphi}}{\partial P_{\mu\nu}} \\
 &= F_{\mu\nu}^0 + (\tilde{\mathbf{q}} \cdot \boldsymbol{\varphi})_{\mu\nu} ,
 \end{aligned}
 \tag{3.56}$$

where the matrix elements  $(\tilde{\mathbf{q}} \cdot \boldsymbol{\varphi})_{\mu\nu}$  are added as a correction to the one-electron matrix as follows:

$$(\tilde{\mathbf{q}} \cdot \boldsymbol{\varphi})_{\mu\nu} = \int_{\Gamma} ds \tilde{q}(\mathbf{s}) \int_{\mathbf{V}} d\mathbf{r} \frac{g_{\mu}(\mathbf{r}) g_{\nu}(\mathbf{r})}{|\mathbf{r} - \mathbf{s}|} .
 \tag{3.57}$$

Eqs. 3.53, 3.55, and 3.56 specify the ingredients needed for a fully variational and SCF procedure (see Chapter 1) for a discussion of incorporating solvation effects).

### 3.2.2 Non-equilibrium Solvent Effects

In Section 2.2.3 a methodology for incorporating non-equilibrium solvent polarization effects through a perturbative approach was outlined. In this section a more robust state-specific formalism<sup>66,67</sup> is discussed for use with the PCM methods of Section 3.2.1. Furthermore, the state-specific formalism presented here will also serve as the foundation for the non-equilibrium PEQS method discussed in Section 3.3.2. Before proceeding with the state-specific formalism, a brief introduction to the concept of non-equilibrium solvation within a classical picture is provided.

The fundamental principle for non-equilibrium solvation is the idea that the solvent polarization field, described by Eq. 3.43, can be partitioned into fast and slow components. For the solvation models discussed up to this point, such a partitioning of  $\mathbf{P}(\mathbf{r})$  is unnecessary because it is assumed the electric field due to the solute charge distribution is in equilibrium with the solvent polarization field. However, for vertical excitation or ionization processes of the solute molecule, there is an abrupt change in

its electronic charge density and electric field, and an equilibrium approach for solvation effects is no longer sufficient, which necessitates the inclusion of non-equilibrium effects. Non-equilibrium solvation effects arise from the fast components of the polarization field,  $\mathbf{P}_{\text{fast}}(\mathbf{r})$ , and the details depend on which of the two commonly used partition schemes proposed by Marcus<sup>77,78</sup> and Pekar<sup>157</sup> is employed.

Although the formalisms presented in this Chapter are time-independent, the concept of “polarization time” will be introduced to facilitate the discussion of the partitioning schemes. Within the Marcus partitioning scheme (MP), upon a rapid disruption of the solute electronic charge density,  $\mathbf{P}_{\text{fast}}(\mathbf{r})$  adjusts simultaneously to the new electric field generated from the solute molecule, while the slow components,  $\mathbf{P}_{\text{slow}}(\mathbf{r})$ , remain in equilibrium with the ground state or un-ionized solute charge distribution. Mathematically,  $\mathbf{P}_{\text{fast}}(\mathbf{r})$  at the current polarization time  $\tau_{\text{cur}}$  can be expressed as<sup>67</sup>

$$\begin{aligned} \mathbf{P}_{\text{fast}}^{\text{MP}}(\mathbf{r}; \tau_{\text{cur}}) &= \left( \frac{\epsilon_{\text{opt}} - 1}{4\pi} \right) \mathbf{E} [\rho_{\text{sol}}(\mathbf{r}; \tau_{\text{cur}}), \mathbf{P}_{\text{fast}}(\mathbf{r}; \tau_{\text{cur}}), \mathbf{P}_{\text{slow}}(\mathbf{r}; \tau_{\text{past}})] \\ &= \chi_{\text{fast}} \mathbf{E} [\rho_{\text{sol}}(\mathbf{r}; \tau_{\text{cur}}), \mathbf{P}_{\text{fast}}(\mathbf{r}; \tau_{\text{cur}}), \mathbf{P}_{\text{slow}}(\mathbf{r}; \tau_{\text{past}})] \end{aligned} \quad (3.58)$$

where  $\epsilon_{\text{opt}}$  is the solvent optical dielectric and the notation in brackets signifies that the electric field arises from the vertically excited or ionized solute charge density at  $\tau_{\text{cur}}$  and is affected by the slow components of the early polarization at time  $\tau_{\text{past}}$ . Generally  $\epsilon_{\text{opt}} = n^2$ , where  $n$  is the solvent index of refraction, and takes on values in the range of 1.5 – 3.0.<sup>67</sup> Similarly,  $\mathbf{P}_{\text{slow}}(\mathbf{r})$  is given by the following expression<sup>67</sup>

$$\begin{aligned} \mathbf{P}_{\text{slow}}^{\text{MP}}(\mathbf{r}; \tau_{\text{cur}}) &= \mathbf{P}_{\text{total}}(\mathbf{r}; \tau_{\text{cur}}) - \mathbf{P}_{\text{fast}}(\mathbf{r}; \tau_{\text{cur}}) \\ &= \left( \frac{\epsilon_{\text{solv}} - \epsilon_{\text{opt}}}{4\pi} \right) \mathbf{E} [\rho_{\text{sol}}(\mathbf{r}; \tau_{\text{cur}}), \mathbf{P}_{\text{fast}}(\mathbf{r}; \tau_{\text{cur}}), \mathbf{P}_{\text{slow}}(\mathbf{r}; \tau_{\text{past}})] \\ &= \chi_{\text{slow}} \mathbf{E} [\rho_{\text{sol}}(\mathbf{r}; \tau_{\text{cur}}), \mathbf{P}_{\text{fast}}(\mathbf{r}; \tau_{\text{cur}}), \mathbf{P}_{\text{slow}}(\mathbf{r}; \tau_{\text{past}})] \end{aligned} \quad (3.59)$$

The Pekar partition (PP) differs from the MP scheme in that the polarization at  $\tau_{\text{cur}}$  is not affected from earlier polarization at  $\tau_{\text{past}}$ . Therefore, within the PP scheme, the

solvent polarization depends only on the solute charge density at  $\tau_{\text{cur}}$  as follows:<sup>67</sup>

$$\mathbf{P}_{\text{fast}}^{\text{PP}}(\mathbf{r}; \tau_{\text{cur}}) = \chi_{\text{fast}} \mathbf{E} [\rho_{\text{sol}}(\mathbf{r}; \tau_{\text{cur}}), \mathbf{P}_{\text{fast}}(\mathbf{r}; \tau_{\text{cur}})] , \quad (3.60)$$

and the difference between the total polarization and Eq. 3.60 yields the the slow polarization field:

$$\mathbf{P}_{\text{slow}}^{\text{PP}}(\mathbf{r}; \tau_{\text{cur}}) = \mathbf{P}_{\text{total}}(\mathbf{r}; \tau_{\text{cur}}) - \mathbf{P}_{\text{fast}}(\mathbf{r}; \tau_{\text{cur}}) . \quad (3.61)$$

The manner in which the “polarization time” manifests is made concrete by examining the equations above:  $\mathbf{P}_{\text{slow}}^{\text{MP}}(\mathbf{r}; \tau_{\text{cur}})$  employs both the equilibrium and optical solvent dielectric constants, whereas  $\mathbf{P}_{\text{slow}}^{\text{PP}}(\mathbf{r}; \tau_{\text{cur}})$  makes use of only the latter. Although the methods describe the same physical phenomenon and therefore should yield identical non-equilibrium solvation free energies within the regime for which the solvent may be described by a linear dielectric, it was shown in Ref. [67] that differences in computed vertical excitations arise as a result of the numerical implementation of certain PCMs such as IEF-PCM, although these differences were quite small. However the subtlety in which the polarization components are separated does impose an additional computational burden for the PP scheme. As will be shown in the upcoming discussions of the state-specific methodology, the non-equilibrium reaction field operator in the solute Hamiltonian requires a contribution from the slow components of the ground state or un-ionized polarization field. For the MP scheme, one needs only compute the equilibrium polarization field as described in Section 3.2.1 and multiply by a constant factor in order to obtain the slow components. However, for the PP scheme, a consequence of Eq. 3.61 is that a computation for both  $\mathbf{P}_{\text{total}}(\mathbf{r}; \tau_{\text{cur}})$  and  $\mathbf{P}_{\text{fast}}(\mathbf{r}; \tau_{\text{cur}})$  is required to obtain  $\mathbf{P}_{\text{slow}}^{\text{PP}}(\mathbf{r}; \tau_{\text{cur}})$ , which effectively doubles the total CPU time.

The details for incorporating non-equilibrium solvation effects into an electronic structure calculation for the ionization potential are now presented, the origin of which is a state-specific form of Schrödinger’s equation:

$$\hat{H}_i^{\text{SS}} |\Psi_i\rangle = E_i^{\text{SS}} |\Psi_i\rangle \quad (3.62)$$

where the state-specific Hamiltonian takes the form

$$\hat{H}_i^{\text{SS}} = \hat{H}_i^{\text{vac}} + \hat{V}_0^{\text{slow}} + \hat{V}_i^{\text{fast}} . \quad (3.63)$$

For consistency with the perturbative approach presented in Section 2.2.3, the subscript notation is reused. To reiterate,  $i = 0$  reflects an energy, electrostatic potential, or charge density arising from the equilibrium reference state, whereas  $i = 1$  is for the ionized state.  $\hat{H}_i^{\text{vac}}$  is the isolated solute molecular Hamiltonian in the absence of solute-solvent interactions, which corresponds to the Hartree-Fock or Kohn-Sham operators.  $\hat{V}_i$  is the total reaction field operator for either the reference or ionized solute state that generates the solvent polarization potential through Eq. 3.36 using either  $\sigma(\mathbf{s})_{\text{pol},0}$  (reference state ASCs) or  $\sigma(\mathbf{s})_{\text{pol},1}$  (ionized state ASCs). The superscripts “slow” and “fast” designate which component of the ASCs are used to construct the response potential through

$$\varphi_{\text{pol},i}^{\text{slow/fast}}(\mathbf{r}) = \int_{\Gamma} d\mathbf{s} \frac{\sigma_{\text{pol},i}^{\text{slow/fast}}(\mathbf{s})}{|\mathbf{r} - \mathbf{s}|} . \quad (3.64)$$

For the case that  $i = 0$ , the state-specific Hamiltonian in Eq. 3.63 becomes

$$\begin{aligned} \hat{H}_0^{\text{SS}} &= \hat{H}_0^{\text{vac}} + \hat{V}_0^{\text{slow}} + \hat{V}_0^{\text{fast}} \\ &= \hat{H}_0^{\text{vac}} + \hat{V}_0 , \end{aligned} \quad (3.65)$$

which is identical to Eq. 3.26, and an equilibrium PCM methodology is used to compute the reference state solvation free energy,  $G_{\text{el}}$ , via Eq. 3.55, the notation of

which is changed to  $G_0^{\text{el}}$  for clarity in this discussion. For the ionized case with  $i = 1$ , the state-specific Hamiltonian in Eq. 3.63 takes the form

$$\hat{H}_1^{\text{SS}} = \hat{H}_1^{\text{vac}} + \hat{V}_0^{\text{slow}} + \hat{V}_1^{\text{fast}} , \quad (3.66)$$

where the non-equilibrium effects of the vertical ionization process manifest in the representation of the total reaction field operator, which contains contributions from the slow components of the reference state and the fast components of the ionized state.

In order to compute a non-equilibrium solvation free energy from Eq. 3.66, the MP or PP scheme is utilized within the PCM framework to generate the appropriate reaction field operators. For the reference state, the symmetrized operator  $\mathbf{Q}_s$  in Eq. 3.53 is employed on the vector  $\boldsymbol{\varphi}$  to compute the vector of ASCs,  $\tilde{\mathbf{q}}$ , which determines the equilibrium solvent response. A similar approach is taken for the non-equilibrium state, and  $\mathbf{Q}_s$  is constructed with the optical dielectric constant  $\epsilon_{\text{opt}}$  and is denoted as  $\mathbf{Q}_s^{\text{eopt}}$ . The resulting expressions for the fast components of the ionized ASCs  $\tilde{\mathbf{q}}_i^{\text{fast}}$  for the MP and PP schemes are<sup>67,158–160</sup>

$$\tilde{\mathbf{q}}_i^{\text{fast}} = \begin{cases} \mathbf{Q}_s^{\text{eopt}} \left( \boldsymbol{\varphi}^{\sigma_{\text{sol},i}} + \boldsymbol{\varphi}^{\sigma_{\text{pol},0}^{\text{slow}}} \right) & \text{for MP} \\ \mathbf{Q}_s^{\text{eopt}} \boldsymbol{\varphi}^{\sigma_{\text{sol},i}} & \text{for PP} . \end{cases} \quad (3.67)$$

The difference between the partitioning schemes arises from the different electrostatic potentials on the right side of Eq. 3.67: for the Marcus partition,  $\tilde{\mathbf{q}}_1^{\text{fast}}$  is generated from the sum of the electrostatic potential due to the ionized solute charge distribution,  $\boldsymbol{\varphi}^{\sigma_{\text{sol},1}}$ , and the slow component of the polarization reaction field from the reference state,  $\boldsymbol{\varphi}^{\sigma_{\text{pol},0}^{\text{slow}}}$ , whereas for the Pekar partition, only  $\boldsymbol{\varphi}^{\sigma_{\text{sol},1}}$  is relevant. The slow components of the reference state ASCs are defined as

$$\tilde{\mathbf{q}}_0^{\text{slow}} = \begin{cases} \left( \frac{\chi_{\text{slow}}}{\chi_{\text{fast}} + \chi_{\text{slow}}} \right) \tilde{\mathbf{q}}_0 & \text{for MP} \\ \tilde{\mathbf{q}}_0 - \tilde{\mathbf{q}}_0^{\text{fast}} & \text{for PP} . \end{cases} \quad (3.68)$$

For the Marcus partitioning scheme, generation of  $\tilde{\mathbf{q}}_0^{\text{slow}}$  is trivial: once  $\tilde{\mathbf{q}}_0$  is obtained from the reference state calculation, the fast and slow components are immediately available after multiplying by the appropriate factors. The Pekar partitioning scheme requires two separate PCM calculations: first for  $\tilde{\mathbf{q}}_0$  with  $\epsilon_{\text{solv}}$  and the second for  $\tilde{\mathbf{q}}_0^{\text{fast}}$  using Eq. 3.67 with  $i = 0$  and  $\epsilon_{\text{opt}}$ .

With the various components of the solvent polarization field now defined, the non-equilibrium solvation free energy,  $G_1^{\text{el}}$ , is computed as<sup>67</sup>

$$G_1^{\text{el}} = \begin{cases} E_1^{\text{int}} + \frac{1}{2} \tilde{\mathbf{q}}_0^{\text{slow}} \cdot \boldsymbol{\varphi}_0 + \frac{1}{2} \tilde{\mathbf{q}}_1^{\text{fast}} \cdot \boldsymbol{\varphi}_1 + W_{0,1} & \text{for MP} \\ E_1^{\text{int}} + \frac{1}{2} \tilde{\mathbf{q}}_0^{\text{slow}} \cdot \boldsymbol{\varphi}_0 + \frac{1}{2} \tilde{\mathbf{q}}_1^{\text{fast}} \cdot \boldsymbol{\varphi}_1 & \text{for PP} , \end{cases} \quad (3.69)$$

where the additional term  $W_{0,1}$  present for the MP scheme accounts for the effect on the fast polarization components at  $\tau_{\text{cur}}$  from the slow components of the polarization at  $\tau_{\text{past}}$ , and is given by<sup>62,67,70</sup>

$$W_{0,1} = \frac{1}{2} \boldsymbol{\varphi}_{\text{pol},0}^{\text{slow}} \cdot (\tilde{\mathbf{q}}_1^{\text{fast}} - \tilde{\mathbf{q}}_0^{\text{fast}}) \quad (3.70)$$

This section concludes with the expression for the vertical ionization potential (VIP), computed as the difference in solvation free energies between the ionized and reference states at the reference state molecular geometry, that is employed for all non-equilibrium PCM calculations reported in this work:

$$\text{VIP} = \begin{cases} (E_1^{\text{int}} - E_0^{\text{int}}) + \frac{1}{2} \tilde{\mathbf{q}}_0^{\text{fast}} \cdot \boldsymbol{\varphi}_0 + \frac{1}{2} \tilde{\mathbf{q}}_1^{\text{fast}} \cdot \boldsymbol{\varphi}_1 + W_{0,1} & \text{for MP} \\ (E_1^{\text{int}} - E_0^{\text{int}}) + \frac{1}{2} \tilde{\mathbf{q}}_0^{\text{fast}} \cdot \boldsymbol{\varphi}_0 + \frac{1}{2} \tilde{\mathbf{q}}_1^{\text{fast}} \cdot \boldsymbol{\varphi}_1 & \text{for PP} , \end{cases} \quad (3.71)$$

where the term in parentheses represents the difference in internal energies of the ionized and reference solute states. Finally to incorporate the non-equilibrium effects into the QM calculation, the Fock matrix is modified according to

$$F_{\mu\nu} = F_{\mu\nu}^1 + \left( \tilde{\mathbf{q}}_0^{\text{slow}} \cdot \boldsymbol{\varphi}_0 \right)_{\mu\nu} + \left( \tilde{\mathbf{q}}_1^{\text{fast}} \cdot \boldsymbol{\varphi}_1 \right)_{\mu\nu} , \quad (3.72)$$

regardless of the chosen partitioning scheme, and  $F_{\mu\nu}^1$  is the Fock matrix of the gas-phase ionized solute molecule.

### 3.3 Anisotropic Solvation with the Poisson Equation Solver

In Section 2.2.2 the PEQS method for computing the equilibrium solvent response of immersing a molecule in an arbitrary and spatially-varying dielectric function,  $\epsilon(\mathbf{r})$ , was presented for use in electronic structure calculations, and a form of  $\epsilon(\mathbf{r})$  for bulk liquid and the liquid-vapor interface was discussed in Section 2.3.4. In this section we focus on improvements made to the PEQS methodology for incorporating equilibrium (Section 3.3.1) and non-equilibrium (Section 3.3.2) solvation effects. A more detailed explanation of the polarization partitioning scheme that was deferred in Chapter 2 (see Eqs. 2.21–2.24) is also provided in Section 3.3.2, and this discussion ties together the concepts introduced in Section 3.2.2 regarding non-equilibrium solvation with PCMs. Alternative definitions of  $\epsilon(\mathbf{r})$  will be presented in Section 3.4.4.

#### 3.3.1 Revised Reference State Solvation

Although the fundamental concept of the PEQS method remains unchanged (i.e., Eq. 2.9 is still used for obtaining  $\varphi_{\text{tot}}(\mathbf{r})$  and  $\rho_{\text{pol}}(\mathbf{r})$  through an iterative procedure), modifying the manner in which the electronic and nuclear electrostatic potentials and charge densities are computed improves the efficiency of Algorithm 1. Obtaining  $\varphi_{\text{sol}}(\mathbf{r})$  is necessary for isolating  $\varphi_{\text{pol}}(\mathbf{r})$  after computing  $\varphi_{\text{tot}}(\mathbf{r})$ , but the methodology presented in Section 2.2.2 requires a separate solution of Poisson’s equation with  $\epsilon(\mathbf{r}) = 1$  for this, which effectively doubles the computational cost. This is avoided by directly computing the exact electrostatic potentials and charge densities (up to

discretization errors on the Cartesian grid) for the electronic and nuclear components, and then summing them together to form  $\varphi_{\text{sol}}(\mathbf{r})$  and  $\rho_{\text{sol}}(\mathbf{r})$ , as shown below.

For electronic structure methods utilizing atom-centered Gaussian basis functions  $g_{\mu}(\mathbf{r})$ , the electronic electrostatic potential is expressed as

$$\begin{aligned}\varphi_{\text{elec}}(\mathbf{r}) &= \sum_{\mu\nu}^{N_{\text{basis}}} P_{\mu\nu} \int d\mathbf{r}' \frac{g_{\mu}(\mathbf{r}')g_{\nu}(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \\ &= \sum_{\mu\nu}^{N_{\text{basis}}} P_{\mu\nu} \varphi_{\mu\nu}(\mathbf{r})\end{aligned}\tag{3.73}$$

where  $\mathbf{P}$  is the one-electron density matrix,  $N_{\text{basis}}$  is the number of basis functions for the desired basis set, and  $\varphi_{\mu\nu}(\mathbf{r})$  is introduced to denote the electrostatic potential in the atomic orbital representation. The electronic charge density is computed as

$$\rho_{\text{elec}}(\mathbf{r}) = -\frac{1}{4\pi} \nabla^2 \varphi_{\text{elec}}(\mathbf{r}),\tag{3.74}$$

where the Laplacian operator is discretized on a Cartesian grid using a high-order finite difference scheme. A formal discussion of finite difference methods for obtaining numerical solutions to Poisson's equation, and the errors associated with them, is reserved for Section 4.1. The solute nuclei are treated as classical point charges approximated by Gaussian functions of finite width  $\sigma$ :

$$\rho_{\text{nuc}}(\mathbf{r}) = -\sum_{\alpha}^{N_{\text{atom}}} \frac{Z_{\alpha}}{(2\pi\sigma)^{3/2}} \exp\left(-\frac{|\mathbf{r} - \mathbf{R}_{\alpha}|^2}{2\sigma^2}\right),\tag{3.75}$$

where  $-Z_{\alpha}$  and  $\mathbf{R}_{\alpha}$  are the charge and position vector of nucleus  $\alpha$  and  $N_{\text{atom}}$  is the number of nuclei in the solute molecule. The exact nuclear electrostatic potential corresponding to  $\rho_{\text{nuc}}(\mathbf{r})$  in Eq. 3.75 is

$$\varphi_{\text{nuc}}(\mathbf{r}) = -\sum_{\alpha}^{N_{\text{atom}}} \frac{Z_{\alpha}}{|\mathbf{r} - \mathbf{R}_{\alpha}|} \text{erf}\left(\frac{|\mathbf{r} - \mathbf{R}_{\alpha}|}{\sqrt{2}\sigma}\right).\tag{3.76}$$

In the event solute nucleus  $\alpha$  lies on top of a Cartesian grid point, such that  $\mathbf{r} = \mathbf{R}_\alpha$ , the contribution to the nuclear electrostatic potential is

$$\varphi_{\text{nuc}}(\mathbf{r} = \mathbf{R}_\alpha) = -\sqrt{\frac{2}{\pi}} \frac{Q_{\text{sol}}}{\sigma}, \quad (3.77)$$

where  $Q_{\text{sol}} = \sum_{\alpha}^{N_{\text{atom}}} Z_{\alpha}$  is the total nuclear charge of the solute. Thus, using Eq. 3.76 for  $\varphi_{\text{nuc}}(\mathbf{r})$  alleviates singularities that arise with a point charge description of the nuclei on Cartesian grids. Having computed the components of  $\varphi_{\text{sol}}(\mathbf{r})$  and  $\rho_{\text{sol}}(\mathbf{r})$ , which bypasses the need to solve Poisson’s equation with  $\epsilon(\mathbf{r}) = 1$ , the iterative procedure outlined in Section 2.2.2 is followed for obtaining  $\varphi_{\text{tot}}(\mathbf{r})$ ,  $\varphi_{\text{pol}}(\mathbf{r})$ , and  $\rho_{\text{pol}}(\mathbf{r})$ , and the equilibrium solvation free energy,  $G_{\text{el}}$ , is computed using Eq. 2.4. One final modification is made for including the solvent response into electronic structure calculations: instead of augmenting the Fock matrix with the correction term given by Eq. 2.13, Eq. 2.20 is exploited so that  $\rho_{\text{pol}}(\mathbf{r})$  can be included as a correction to the one-electron matrix  $\mathbf{h}$  instead of using  $\varphi_{\text{pol}}(\mathbf{r})$  as a correction to the Fock matrix  $\mathbf{F}$ . The correction term  $\Delta\mathbf{h}$  has matrix elements

$$\Delta h_{\mu\nu} = \int d\mathbf{r} \rho_{\text{pol}}(\mathbf{r}) \varphi_{\mu\nu}(\mathbf{r}). \quad (3.78)$$

Obtaining  $\varphi_{\mu\nu}(\mathbf{r})$  requires evaluation of one-electron integrals (see Eq 3.73), and this is accomplished by utilizing the OpenMP parallelized integrals-engine within the Q-CHEM software package.<sup>75</sup> The definitions of  $\rho_{\text{elec}}(\mathbf{r})$  in Eq. 2.7 and  $\Delta F_{\mu\nu}$  in Eq. 2.13 require evaluation of  $N_{\text{basis}}$  basis functions on  $N_{\text{grid}}$  Cartesian grid points at *each* SCF iteration because it is generally not feasible to store these values for production-quality grids requiring  $\mathcal{O}(10^6)$  points. Although the new methodology is subject to the same computational cost, a significant amount of overhead is eliminated by utilizing the optimized routines available in Q-CHEM for the one-electron integrals as

opposed to using a previous implementation that serially evaluates the Fock matrix correction. Thus, by computing  $\varphi_{\text{elec}}(\mathbf{r})$  with Eq. 3.73 and then  $\rho_{\text{elec}}(\mathbf{r})$  with a finite difference method as well as correcting  $\mathbf{h}$  according to Eq. 3.78, a speedup of the PEQS routines are observed at each SCF cycle. The improved PEQS method for equilibrium solvation discussed in this section is presented as Algorithm 3 on Page 107.

---

**Algorithm 3:** Improved PEQS algorithm for equilibrium solvation

---

```

1 begin SCF procedure
2   Initialize  $\Delta\mathbf{h} \equiv 0$ .
3   repeat  $n = 1, 2, \dots$  SCF iterations
4     Diagonalize the Fock matrix  $\mathbf{F} = \mathbf{F}_0 + \Delta\mathbf{h}$  to obtain the density matrix
        $\mathbf{P}^{(n)}$ , where  $\mathbf{F}_0$  is the gas-phase Fock matrix and  $\Delta\mathbf{h}$  is the equilibrium
       solvation correction to the one-electron matrix  $\mathbf{h}$ .
5     Compute  $\varphi_{\text{elec}}(\mathbf{r})$ ,  $\rho_{\text{elec}}(\mathbf{r})$ ,  $\rho_{\text{nuc}}(\mathbf{r})$ , and  $\varphi_{\text{nuc}}(\mathbf{r})$  on the Cartesian grid via
       Eqs. 3.73–3.76.
6     Form  $\rho_{\text{sol}}(\mathbf{r}) = \rho_{\text{elec}}(\mathbf{r}) + \rho_{\text{nuc}}(\mathbf{r})$  and  $\varphi_{\text{sol}}(\mathbf{r}) = \varphi_{\text{elec}}(\mathbf{r}) + \varphi_{\text{nuc}}(\mathbf{r})$ .
7     Form  $\epsilon(\mathbf{r})$  with  $\epsilon_{\text{solv}}$ .
8     // The form of the dielectric function is discussed in
       Sections 2.3.4 and 3.4.4.
9     Set  $\rho_{\text{tot}}(\mathbf{r}) = \rho_{\text{sol}}(\mathbf{r})$  and  $\varphi_{\text{tot}}(\mathbf{r}) = \varphi_{\text{sol}}(\mathbf{r})$ .
10    Compute  $\rho_{\text{iter}}^{(i=1)}(\mathbf{r})$  via Eq. 2.11 with  $\varphi_{\text{tot}}(\mathbf{r})$  and  $\epsilon(\mathbf{r})$ .
11    repeat  $i = 1, 2, \dots$  PEQS iterations
12      begin PEQS routines
13        Compute  $\varphi_{\text{tot}}^{(i)}(\mathbf{r})$  via Eq. 2.9.
14        Update  $\rho_{\text{iter}}^{(i+1)}(\mathbf{r})$  via Eq. 2.12 with  $\epsilon(\mathbf{r})$ ,  $\varphi_{\text{tot}}^{(i)}(\mathbf{r})$ , and  $\rho_{\text{iter}}^{(i)}(\mathbf{r})$ .
15        Update  $\rho_{\text{pol}}(\mathbf{r})$  and  $\rho_{\text{tot}}(\mathbf{r})$  via Eqs. 2.10 and 2.5.
16      until  $\|\rho_{\text{iter}}^{(i+1)}(\mathbf{r}) - \rho_{\text{iter}}^{(i)}(\mathbf{r})\| < T_{\text{solver}}$ 
17      Compute  $\varphi_{\text{pol}}(\mathbf{r}) = \varphi_{\text{tot}}(\mathbf{r}) - \varphi_{\text{sol}}(\mathbf{r})$ .
18      If performing a non-equilibrium solvation calculation with the MP
        scheme, compute  $\rho_{\text{pol},0}^{\text{slow}}(\mathbf{r})$  and  $\varphi_{\text{pol},0}^{\text{slow}}(\mathbf{r})$  via Eqs. 3.79 and 2.20.
19      Generate  $\Delta\mathbf{h}$  via Eq. 3.78 with  $\varphi_{\text{pol}}(\mathbf{r})$ .
20      Compute  $E_{\text{el}}$  via Eq. 2.14 and add it to the SCF energy.
21    until DIIS error  $< T_{\text{SCF}}$ 
22 Compute the equilibrium solvation free energy  $G_0^{\text{el}}$  via Eq. 2.4.

```

---

### 3.3.2 State-Specific Non-equilibrium Solvent Effects

A perturbative approach for the PEQS method to incorporate non-equilibrium effects into the solvation free energies computed for vertically ionized systems is presented in Section 2.2.3. This is tantamount to “freezing” the reference state reaction field such that, upon vertical ionization, the slow and fast inertial components of the reference polarization field are unable to adjust to the ionized charge distribution. Therefore the ionized solute-solvent interaction is governed by this “frozen” reference state reaction field. In this section a novel state-specific PEQS methodology for non-equilibrium solvation is adapted from the technique in Section 3.2.2 for PCMs. Whereas the PCM approach is derived for two-dimensional ASCs, the PEQS approach is applicable for arbitrary three-dimensional boundary conditions. Furthermore, the details for adapting the MP and PP schemes to the PEQS method are elucidated here.

The state-specific Hamiltonian in Eqs. 3.62 and 3.63 is the theoretical foundation for the state-specific PEQS method, and the notation is reused to preserve consistency. For  $i = 0$ , the state-specific and perturbative PEQS methods employ an identical effective Hamiltonian, and the improved procedure given in Algorithm 3 is performed to obtain  $G_0^{\text{el}}$  and the necessary reference state potentials and charge densities. The state-specific and perturbative PEQS methods differ when  $i = 1$  for the ionized state, and the effective Hamiltonian for the state-specific PEQS method takes the form as expressed in Eq. 3.66, where it is necessary to partition the solvent polarization response into fast and slow components utilizing either the MP or PP schemes.

For the slow component of the reference state polarization response field, Eq. 3.68 is used:

$$\varphi_{\text{pol},0}^{\text{slow}}(\mathbf{r}) = \begin{cases} \left(\frac{\chi_{\text{slow}}}{\chi_{\text{fast}} + \chi_{\text{slow}}}\right) \varphi_{\text{pol},0}(\mathbf{r}) & \text{for MP} \\ \varphi_{\text{pol},0}(\mathbf{r}) - \varphi_{\text{pol},0}^{\text{fast}}(\mathbf{r}) & \text{for PP} . \end{cases} \quad (3.79)$$

To obtain  $\varphi_{\text{pol},0}^{\text{slow}}(\mathbf{r})$  from the polarization potential within the MP scheme, it is only necessary to multiply by the constant factor  $\left(\frac{\chi_{\text{slow}}}{\chi_{\text{fast}} + \chi_{\text{slow}}}\right)$ . For the PP scheme, Eq. 2.9 must be solved twice: once using  $\epsilon_{\text{solv}}$  in  $\epsilon(\mathbf{r})$ , and the second time using  $\epsilon_{\text{opt}}$ . The former will yield  $\varphi_{\text{pol},0}(\mathbf{r})$  using the methods from Section 3.3.1, whereas the latter will result in a modified version of Eq. 2.9 that provides the fast components of both the *total* reference electrostatic potential and *polarization* charge density:

$$\hat{\nabla}^2 \varphi_{\text{tot},0}^{\text{fast}}(\mathbf{r}) = -4\pi \left[ \rho_{\text{sol},0}(\mathbf{r}) + \rho_{\text{pol},0}^{\text{fast}}(\mathbf{r}) \right] , \quad (3.80)$$

and the fast component of the reference state *polarization* charge density is obtained by

$$\rho_{\text{pol},0}^{\text{fast}}(\mathbf{r}) = \rho_{\text{iter},0}(\mathbf{r}) + \left[ \frac{1 - \epsilon_{\text{opt}}(\mathbf{r})}{\epsilon_{\text{opt}}(\mathbf{r})} \right] \rho_{\text{sol},0}(\mathbf{r}) , \quad (3.81)$$

where  $\rho_{\text{iter},0}(\mathbf{r})$  is computed using Eq. 2.11 with  $\epsilon_{\text{opt}}(\mathbf{r})$  and  $\varphi_{\text{tot},0}^{\text{fast}}(\mathbf{r})$ . Finally the fast component of the reference state *polarization* response is obtained by

$$\varphi_{\text{pol},0}^{\text{fast}}(\mathbf{r}) = \varphi_{\text{tot},0}^{\text{fast}}(\mathbf{r}) - \varphi_{\text{sol},0}(\mathbf{r}) , \quad (3.82)$$

which then allows  $\varphi_{\text{pol},0}^{\text{slow}}(\mathbf{r})$  to be computed using Eq. 3.79.

For the fast components of the ionized state polarization field, Eq. 3.67 that describes the partitioning schemes for PCMs must be modified for the PEQS method. The MP scheme uses a combined potential arising from the ionized solute electrostatic potential and the slow contribution of the solvent reaction field to generate the fast ASCs using the polarization weight operator that employs  $\epsilon_{\text{opt}}$ . For the PP scheme,

the same polarization weight operator is utilized, but the potential that generates the fast ASCs arises only from the the ionized solute electrostatic potential. To capture the physical nature of these partitioning schemes for the PEQS method, Poisson's equation is modified as

$$\hat{\nabla} \cdot [\epsilon_{\text{opt}}(\mathbf{r}) \hat{\nabla} \varphi_{\text{tot},1}^{\text{fast}}(\mathbf{r})] = \begin{cases} -4\pi [\rho_{\text{sol},1}(\mathbf{r}) + \rho_{\text{pol},0}^{\text{slow}}(\mathbf{r})] & \text{for MP} \\ -4\pi \rho_{\text{sol},1}(\mathbf{r}) & \text{for PP} , \end{cases} \quad (3.83)$$

where  $\varphi_{\text{tot},1}^{\text{fast}}(\mathbf{r})$  is the fast component of the *total* electrostatic potential of the ionized state. The difference between these partitioning schemes manifests in the description of the charge density source on the right side of Eq. 3.83. For both partitioning schemes, however, the polarization response is generated by the interaction of their respective sources with the fast component of the solvent dielectric that is encoded in  $\epsilon_{\text{opt}}$ .

To apply the procedures of the PEQS method introduced in Section 2.2.3, it is useful to rewrite Eq. 3.83 as

$$\hat{\nabla}^2 \varphi_{\text{tot},1}^{\text{fast}}(\mathbf{r}) = \begin{cases} -4\pi [\rho_{\text{sol},1}(\mathbf{r}) + \rho_{\text{pol},0}^{\text{slow}}(\mathbf{r}) + \rho_{\text{pol},1}^{\text{fast}}(\mathbf{r})] & \text{for MP} \\ -4\pi [\rho_{\text{sol},1}(\mathbf{r}) + \rho_{\text{pol},1}^{\text{fast}}(\mathbf{r})] & \text{for PP} , \end{cases} \quad (3.84)$$

where  $\rho_{\text{pol},1}^{\text{fast}}(\mathbf{r})$  and  $\varphi_{\text{tot},1}^{\text{fast}}(\mathbf{r})$  are iteratively computed self-consistently. For the respective partitioning schemes,  $\rho_{\text{pol},1}^{\text{fast}}(\mathbf{r})$  takes the form

$$\rho_{\text{pol},1}^{\text{fast}}(\mathbf{r}) = \begin{cases} \rho_{\text{iter},1}(\mathbf{r}) + \left[ \frac{1-\epsilon_{\text{opt}}(\mathbf{r})}{\epsilon_{\text{opt}}(\mathbf{r})} \right] [\rho_{\text{sol},1}(\mathbf{r}) + \rho_{\text{pol},0}^{\text{slow}}(\mathbf{r})] & \text{for MP} \\ \rho_{\text{iter},1}(\mathbf{r}) + \left[ \frac{1-\epsilon_{\text{opt}}(\mathbf{r})}{\epsilon_{\text{opt}}(\mathbf{r})} \right] \rho_{\text{sol},1}(\mathbf{r}) & \text{for PP} , \end{cases} \quad (3.85)$$

where  $\rho_{\text{iter},1}(\mathbf{r})$  is computed and updated using Eqs. 2.11 and 2.12 with  $\epsilon_{\text{opt}}(\mathbf{r})$  and  $\varphi_{\text{tot},1}^{\text{fast}}(\mathbf{r})$ . The corresponding fast polarization response for the ionized state is then computed as

$$\varphi_{\text{pol},1}^{\text{fast}}(\mathbf{r}) = \begin{cases} \varphi_{\text{tot},1}^{\text{fast}}(\mathbf{r}) - \varphi_{\text{sol},1}(\mathbf{r}) - \varphi_{\text{pol},0}^{\text{slow}}(\mathbf{r}) & \text{for MP} \\ \varphi_{\text{tot},1}^{\text{fast}}(\mathbf{r}) - \varphi_{\text{sol},1}(\mathbf{r}) & \text{for PP} . \end{cases} \quad (3.86)$$

Utilizing these quantities, the non-equilibrium solvation free energy of the ionized state,  $G_1^{\text{el}}$ , for the two partitioning schemes is computed as

$$G_1^{\text{el}} = \begin{cases} E_1^{\text{int}} + \frac{1}{2} \int d\mathbf{r} \varphi_{\text{pol},0}^{\text{slow}}(\mathbf{r}) \rho_{\text{sol},0}(\mathbf{r}) + \frac{1}{2} \int d\mathbf{r} \varphi_{\text{pol},1}^{\text{fast}}(\mathbf{r}) \rho_{\text{sol},1}(\mathbf{r}) + W_{0,1} & \text{for MP} \\ E_1^{\text{int}} + \frac{1}{2} \int d\mathbf{r} \varphi_{\text{pol},0}^{\text{slow}}(\mathbf{r}) \rho_{\text{sol},0}(\mathbf{r}) + \frac{1}{2} \int d\mathbf{r} \varphi_{\text{pol},1}^{\text{fast}}(\mathbf{r}) \rho_{\text{sol},1}(\mathbf{r}) & \text{for PP} . \end{cases} \quad (3.87)$$

The correction term arising from the MP scheme,  $W_{0,1}$ , represents the same physical quantity described by Eq. 3.70, and is computed identically for the perturbative and state-specific PEQS methods by Eq. 2.28. The non-equilibrium VIPs reported for the state-specific PEQS method are computed as  $\text{VIP} = G_{\text{el}}^1 - G_{\text{el}}^0$  shown explicitly here:

$$\text{VIP} = \begin{cases} (E_1^{\text{int}} - E_0^{\text{int}}) + \frac{1}{2} \int d\mathbf{r} \varphi_{\text{pol},0}^{\text{fast}}(\mathbf{r}) \rho_{\text{sol},0}(\mathbf{r}) + \frac{1}{2} \int d\mathbf{r} \varphi_{\text{pol},1}^{\text{fast}}(\mathbf{r}) \rho_{\text{sol},1}(\mathbf{r}) + W_{0,1} & \text{for MP} \\ (E_1^{\text{int}} - E_0^{\text{int}}) + \frac{1}{2} \int d\mathbf{r} \varphi_{\text{pol},0}^{\text{fast}}(\mathbf{r}) \rho_{\text{sol},0}(\mathbf{r}) + \frac{1}{2} \int d\mathbf{r} \varphi_{\text{pol},1}^{\text{fast}}(\mathbf{r}) \rho_{\text{sol},1}(\mathbf{r}) & \text{for PP} . \end{cases} \quad (3.88)$$

Finally, to include the non-equilibrium solvent response in the electronic structure calculations, the matrix elements of the corresponding correction term made to the one-electron matrix  $\mathbf{h}$  are

$$\Delta h_{\mu\nu} = \int d\mathbf{r} \rho_{\text{pol},0}^{\text{slow}}(\mathbf{r}) \varphi_{\mu\nu,0}(\mathbf{r}) + \int d\mathbf{r} \rho_{\text{pol},1}^{\text{fast}}(\mathbf{r}) \varphi_{\mu\nu,1}(\mathbf{r}) , \quad (3.89)$$

and the SCF energy is augmented with the quantity  $E_{\text{el},1}$  expressed as

$$E_{\text{el},1} = \int d\mathbf{r} \varphi_{\text{pol},0}^{\text{slow}}(\mathbf{r}) \rho_{\text{sol},0}(\mathbf{r}) + \int d\mathbf{r} \varphi_{\text{pol},1}^{\text{fast}}(\mathbf{r}) \rho_{\text{sol},1}(\mathbf{r}) \quad (3.90)$$

that corresponds to the total non-equilibrium solute-solvent interaction without accounting for the polarization work of the dielectric medium. A modified algorithm showing the changes that must be made to the equilibrium solvation algorithm in order to compute the required reference state quantities for the non-equilibrium PP scheme is presented in Algorithm 4 on Page 112. Then, the improved state-specific PEQS method for non-equilibrium solvation with both partitioning schemes is presented as Algorithm 5 on Page 113.

---

**Algorithm 4:** Improved PEQS algorithm for equilibrium solvation with the PP scheme

---

```

1 begin SCF procedure
2   Initialize  $\Delta\mathbf{h} \equiv 0$ .
3   repeat  $n = 1, 2, \dots$  SCF iterations
4     Diagonalize the Fock matrix  $\mathbf{F} = \mathbf{F}_0 + \Delta\mathbf{h}$  to obtain the density matrix
        $\mathbf{P}^{(n)}$ , where  $\mathbf{F}_0$  is the gas-phase Fock matrix and  $\Delta\mathbf{h}$  is the equilibrium
       solvation correction to the one-electron matrix  $\mathbf{h}$ .
5     Compute  $\varphi_{\text{elec},0}(\mathbf{r})$ ,  $\rho_{\text{elec},0}(\mathbf{r})$ ,  $\rho_{\text{nuc},0}(\mathbf{r})$ , and  $\varphi_{\text{nuc},0}(\mathbf{r})$  via Eqs. 3.73–3.76.
6     Form  $\rho_{\text{sol},0}(\mathbf{r}) = \rho_{\text{elec},0}(\mathbf{r}) + \rho_{\text{nuc},0}(\mathbf{r})$  and  $\varphi_{\text{sol},0}(\mathbf{r}) = \varphi_{\text{elec},0}(\mathbf{r}) + \varphi_{\text{nuc},0}(\mathbf{r})$ .
7     Form  $\epsilon(\mathbf{r})$  with  $\epsilon_{\text{solv}}$ .
8     // The form of the dielectric function is discussed in
       Sections 2.3.4 and 3.4.4.
9     Set  $\rho_{\text{tot},0}(\mathbf{r}) = \rho_{\text{sol},0}(\mathbf{r})$  and  $\varphi_{\text{tot},0}(\mathbf{r}) = \varphi_{\text{sol},0}(\mathbf{r})$ .
10    Compute  $\rho_{\text{iter},0}^{(i=1)}(\mathbf{r})$  via Eq. 2.11 with  $\varphi_{\text{tot},0}(\mathbf{r})$  and  $\epsilon(\mathbf{r})$ .
11    repeat  $i = 1, 2, \dots$  PEQS iterations
12      begin PEQS routines
13        | Compute  $\varphi_{\text{tot},0}^{(i)}(\mathbf{r})$  via Eq. 2.9.
14        | Update  $\rho_{\text{iter},0}^{(i+1)}(\mathbf{r})$  via Eq. 2.12 with  $\epsilon(\mathbf{r})$ ,  $\varphi_{\text{tot},0}^{(i)}(\mathbf{r})$ , and  $\rho_{\text{iter},0}^{(i)}(\mathbf{r})$ .
15        | Update  $\rho_{\text{pol},0}(\mathbf{r})$  and  $\rho_{\text{tot},0}(\mathbf{r})$  via Eqs. 2.10 and 2.5.
16      until  $\|\rho_{\text{iter},0}^{(i+1)}(\mathbf{r}) - \rho_{\text{iter},0}^{(i)}(\mathbf{r})\| < T_{\text{solver}}$ 
17      Reform  $\epsilon(\mathbf{r})$  with  $\epsilon_{\text{opt}}$ .
18      Reset  $\rho_{\text{tot},0}(\mathbf{r}) = \rho_{\text{sol},0}(\mathbf{r})$  and  $\varphi_{\text{tot},0}(\mathbf{r}) = \varphi_{\text{sol},0}(\mathbf{r})$ .
19      Compute  $\rho_{\text{iter},0}^{(i=1)}(\mathbf{r})$  via Eq. 2.11 with  $\varphi_{\text{tot},0}(\mathbf{r})$  and  $\epsilon_{\text{opt}}(\mathbf{r})$ .
20      repeat  $i = 1, 2, \dots$  PEQS iterations
21        begin PEQS routines
22          | Compute  $\varphi_{\text{tot},0}^{\text{fast},(i)}(\mathbf{r})$  via Eq. 3.80.
23          | Update  $\rho_{\text{iter},0}^{(i+1)}(\mathbf{r})$  via Eq. 2.12 with  $\epsilon_{\text{opt}}(\mathbf{r})$ ,  $\varphi_{\text{tot},0}^{\text{fast},(i)}(\mathbf{r})$ , and  $\rho_{\text{iter},0}^{(i)}(\mathbf{r})$ .
24          | Update  $\rho_{\text{pol},0}^{\text{fast}}(\mathbf{r})$  via Eq. 3.81.
25        until  $\|\rho_{\text{iter},0}^{(i+1)}(\mathbf{r}) - \rho_{\text{iter},0}^{(i)}(\mathbf{r})\| < T_{\text{solver}}$ 
26        Compute  $\varphi_{\text{pol},0}(\mathbf{r}) = \varphi_{\text{tot},0}(\mathbf{r}) - \varphi_{\text{sol},0}(\mathbf{r})$ .
27        Compute  $\varphi_{\text{pol},0}^{\text{fast}}(\mathbf{r})$  via Eq. 3.82 and  $\varphi_{\text{pol},0}^{\text{fast}}(\mathbf{r})$  via Eq. 3.79.
28        Generate  $\Delta\mathbf{h}$  via Eq. 3.78 with  $\varphi_{\text{pol},0}(\mathbf{r})$ .
29        Compute  $E_{\text{el}}$  via Eq. 2.14 and add it to the SCF energy.
30      until  $DIIS \text{ error} < T_{\text{SCF}}$ 
31 Compute the equilibrium solvation free energy  $G_0^{\text{el}}$  via Eq. 2.4.

```

---

---

**Algorithm 5:** Improved non-equilibrium PEQS procedure using the MP and PP schemes

---

```

1 begin Reference state procedure
2   | Proceed with Algorithms 3 or 4 to compute the energies, electrostatic
   | potentials, and charge densities for the MP or PP schemes, respectively,
   | and save to disk.
   Input:  $G_0^{\text{el}}$ ,  $\varphi_{\text{pol},0}^{\text{fast/slow}}(\mathbf{r})$ ,  $\rho_{\text{pol},0}^{\text{fast/slow}}(\mathbf{r})$ , and  $\rho_{\text{sol},0}(\mathbf{r})$ .
3 begin Ionized state procedure
4   | Initialize  $\Delta\mathbf{h} \equiv 0$ .
5   repeat  $n = 1, 2, \dots$  SCF iterations
6     | Diagonalize the Fock matrix  $\mathbf{F} = \mathbf{F}_0 + \Delta\mathbf{h}$  to obtain the density matrix
       |  $\mathbf{P}^{(n)}$ , where  $\mathbf{F}_0$  is the gas-phase Fock matrix and  $\Delta\mathbf{h}$  is the
       | non-equilibrium solvation correction.
7     | Compute  $\varphi_{\text{elec},1}(\mathbf{r})$ ,  $\rho_{\text{elec},1}(\mathbf{r})$ ,  $\rho_{\text{nuc},1}(\mathbf{r})$ , and  $\varphi_{\text{nuc},1}(\mathbf{r})$  via Eqs. 3.73–3.76.
8     | Form  $\rho_{\text{sol},1}(\mathbf{r}) = \rho_{\text{elec},1}(\mathbf{r}) + \rho_{\text{nuc},1}(\mathbf{r})$  and  $\varphi_{\text{sol},1}(\mathbf{r}) = \varphi_{\text{elec},1}(\mathbf{r}) + \varphi_{\text{nuc},1}(\mathbf{r})$ .
9     | Form  $\epsilon(\mathbf{r})$  with  $\epsilon_{\text{opt}}$ .
10    | If (MP), set  $\rho_{\text{tot},1}(\mathbf{r}) = \rho_{\text{sol},1}(\mathbf{r}) + \rho_{\text{pol},0}^{\text{slow}}(\mathbf{r})$  and
       |  $\varphi_{\text{tot},1}(\mathbf{r}) = \varphi_{\text{sol},1}(\mathbf{r}) + \varphi_{\text{pol},0}^{\text{slow}}(\mathbf{r})$ .
11    | Else if (PP), set  $\rho_{\text{tot},1}(\mathbf{r}) = \rho_{\text{sol},1}(\mathbf{r})$  and  $\varphi_{\text{tot},1}(\mathbf{r}) = \varphi_{\text{sol},1}(\mathbf{r})$ .
12    | Compute  $\rho_{\text{iter},1}^{(i=1)}(\mathbf{r})$  via Eq. 2.11 with  $\epsilon_{\text{opt}}(\mathbf{r})$  and  $\varphi_{\text{tot},1}(\mathbf{r})$ .
13    | repeat  $i = 1, 2, \dots$  PEQS iterations
14      | begin PEQS routines
15        |  $\varphi_{\text{tot},1}^{\text{fast},(i)}(\mathbf{r})$  via Eq. 3.84.
16        | Update  $\rho_{\text{iter},1}^{(i+1)}(\mathbf{r})$  via Eq. 2.12 with  $\epsilon_{\text{opt}}(\mathbf{r})$ ,  $\varphi_{\text{tot},1}^{\text{fast},(i)}(\mathbf{r})$ , and  $\rho_{\text{iter},1}^{(i)}(\mathbf{r})$ .
17        |  $\rho_{\text{pol},1}^{\text{fast}}(\mathbf{r})$  Update and via Eq. 3.85.
18        | until  $\|\rho_{\text{iter},1}^{(i+1)}(\mathbf{r}) - \rho_{\text{iter},1}^{(i)}(\mathbf{r})\| < T_{\text{solver}}$ 
19        | Compute  $\varphi_{\text{pol},1}^{\text{fast}}(\mathbf{r})$  via Eq. 3.86.
20        | Generate  $\Delta\mathbf{h}$  via Eq. 3.89.
21        | Compute  $E_{\text{el},1}$  via Eq. 3.90 and add it to the SCF energy.
22    | until DIIS error  $< T_{\text{SCF}}$ 
23  | Compute the non-equilibrium polarization interaction energy  $G_1^{\text{el}}$  via Eq. 3.87.
24  | Compute VIP via Eq. 3.88.

```

---

## 3.4 Computational Details and Methods

In this section, the computational details for the methods discussed in the previous sections are presented. Section 3.4.1 outlines the molecular dynamics simulations that were performed to generate the classical nuclear configurations for the alkali metal cation and anion halide systems that are investigated in this Chapter. The electronic structure methodology and choice of basis sets employed for the quantum-mechanical description of these systems is discussed in Section 3.4.2. Sections 3.4.3 and 3.4.4 discuss the details for computing the VIPs within the non-equilibrium PCM and PEQS frameworks, respectively, and emphasis is placed on discussing the various forms of the molecular cavity used in PCM and the functional forms of the dielectric function  $\epsilon(\mathbf{r})$  for the PEQS method.

### 3.4.1 Molecular Dynamics Simulations

Simulations of neat liquid water were performed with 222 water molecules in a cubic unit cell of side 18.8 Å (corresponding to a density of 0.9995 g/cm<sup>3</sup>) that was replicated periodically in three dimensions at 300 K. The AMOEBA<sup>47</sup> polarizable force field within the TINKER molecular modeling package<sup>161</sup> was employed to describe the ions and water molecules and perform the molecular dynamics simulations, and electrostatic interactions were computed using standard Ewald summation with an interaction cutoff of 9.4 Å. The ion polarizabilities for Li<sup>+</sup>, Na<sup>+</sup>, F<sup>-</sup>, and Cl<sup>-</sup> are 0.0280, 0.1200, 1.350, and 4.000 Å<sup>3</sup>, respectively, within this forcefield.<sup>47</sup> The neat liquid water simulations were equilibrated for 1 ns, and the last 500 ps of the equilibration were extracted for further use. For neat liquid-vapor simulations, the last configuration from the bulk water equilibration was extracted and the simulation box

was extended to 90.0 Å in the  $z$ -direction (18.8 Å × 18.8 Å × 90.0 Å). The resulting water “slab” was equilibrated for an additional 1 ns at 300 K. The last configuration from either of these equilibrations was used to generate the initial configuration for the alkali metal and halide simulations. For the simulations in liquid water, the water molecule nearest the center of the cubic unit cell was located and replaced with Li<sup>+</sup>, Na<sup>+</sup>, F<sup>-</sup>, or Cl<sup>-</sup>, whereas the water molecule nearest the interface was replaced for the liquid-vapor simulations. The location of the interface is characterized by the Gibbs dividing surface (GDS), and the procedure for determining this quantity is discussed below in Section 3.4.4. After insertion of the ion, the simulations were equilibrated for 250 ps at 300 K followed by a 500 ps production run where the configurations were stored every 10 ps for further use in the electronic structure/continuum model calculations. The configurations for e<sup>-</sup>(aq) in liquid water and at the interface are identical to the ones used in Chapter 2, and were extracted from the simulations of Refs. [11] and [4].

### 3.4.2 Basis Sets and Electronic Structure Methods

All electronic structure computations presented in this Chapter employ the resolution-of-identity<sup>92,93</sup> (RI)-MP2 method, and the results of Section 3.5 utilize the following basis sets:

1. The alkali metal cations Li<sup>+</sup> and Na<sup>+</sup> are treated with the cc-pVQZ and auxiliary rimp2-cc-pVQZ basis sets and the 6-311+G\* and auxiliary rimp2-cc-pVTZ basis sets are employed for the explicit solvent water molecules.

2.  $e^-(aq)$ , the halide anions  $F^-$  and  $Cl^-$ , and the explicit solvent water molecules in these systems are treated with the 6-311++G\* and auxiliary rimp2-aug-cc-pVTZ basis sets.
3. The neutral liquid water systems are treated with the 6-311+G\* and auxiliary rimp2-aug-cc-pVTZ basis sets.

[In Ref. [1], it was noted that a very diffuse sp even-tempered set of basis functions were added for  $Cl^-$  to properly describe the CTTS levels in aqueous solution. However, this affected the computed ionization potentials of  $Cl^-$  by  $\approx 7$  meV, and was therefore omitted for this work.] The details of all basis sets listed above were obtained from Ref. [162] and references therein.

Furthermore, it was anticipated that the magnitude of the computed ionization potentials of  $Li^+$  and  $Na^+$  would be far greater than for liquid water, and additional considerations for performing the SCF procedure are required. The measured ionization potentials for the 1s state of  $Li^+$  and the 2p state of  $Na^+$  are  $60.4 \pm 0.07$  eV<sup>1</sup> and  $35.4 \pm 0.04$  eV<sup>1</sup>, respectively, whereas ionization from the 1b<sub>1</sub> orbital of liquid water produces a broad absorption band centered at 11.23 eV<sup>2</sup>. Since the cation-water systems considered here contain nearly two solvation shells of explicit water molecules as part of the QM solute, and the cation ionization states lie much higher in energy compared to the lowest ionization of liquid water, a standard SCF procedure cannot be used. The variational nature of the HF method that underlies RI-MP2 would provide the lowest energy ionization state for the cation-water system, which is removal of an electron from the surrounding water molecules. We are interested in computing the ionization potentials for the cations, and therefore we employ the maximum overlap method<sup>163</sup> (MOM) to circumvent this issue. As discussed in Sections 3.2 and 3.3,

two separate SCF procedures must be performed for the non-equilibrium solvation models: the first for the reference state and a second for the ionized state. After completing the reference state SCF procedure, the 1s orbital for  $\text{Li}^+$  or the 2p orbital for  $\text{Na}^+$  is identified. Before the ionized state calculation begins, an electron is explicitly removed from the 1s or 2p reference state orbital of  $\text{Li}^+$  or  $\text{Na}^+$ , respectively, and this set of “ionized” reference state orbitals is saved to disk. Starting with the first SCF cycle for the ionized state, the MOM procedure maximizes the overlap at each cycle between the computed orbitals for the ionized state and the set of “ionized” reference state orbitals that were saved from the previous calculation. Using this procedure results in ionization from the desired 1s and 2p orbitals of  $\text{Li}^+$  or  $\text{Na}^+$ , respectively, and the spin density  $[\rho_{\text{spin}}(\mathbf{r}) = \rho_{\alpha}(\mathbf{r}) - \rho_{\beta}(\mathbf{r})]$  is shown for each is shown in Fig 3.2(c) and (d).

The success of this MOM-SCF procedure for the cation-water systems relies on the fact that the 1s and 2p orbitals for  $\text{Li}^+$  and  $\text{Na}^+$  are energetically well-separated from the molecular orbitals of the surrounding water molecules. This is no longer the case for  $\text{F}^-$ , for which the 2p ionization potential has recently been measured<sup>2</sup> to be 11.58 eV. Since the relative energies of the 2p orbitals of  $\text{F}^-$  and 1b<sub>1</sub> orbital of water are similar, significant mixing occurs and it is impossible to identify them explicitly. Using the MOM-SCF procedure just described results in an undesired collective ionization of fluoride and multiple water molecules in the QM region, and the spin density of the resulting system is shown in Fig 3.1(a). Furthermore, we are also interested in obtaining the ionization potential of a single solvated water molecule, where the solvent is comprised of other quantum-mechanical water molecules and the classical continuum. Again, using the MOM-SCF procedure previously described

results in ionization of a water molecule close to the continuum boundary as shown in Fig 3.1(c). This is undesirable because it neglects explicit hydrogen bonds present for a more centrally located water molecule. In both cases, the computed ionization potentials differ from experiment by nearly 1.5 eV.

These issues are addressed by utilizing a similar approach that exploits a different method for generating the initial guess of the molecular wave function. Instead of using a set of “ionized” reference state orbitals as for the cation-water systems, a FRAGMO<sup>75</sup> procedure is utilized. For the fluoride-water and single water molecule ionizations, all ions and molecules are treated as individual fragments for which a complete SCF procedure is performed. In doing this, the fluoride anion and a single water molecule can be explicitly ionized, while the surrounding water molecules remained unaffected, neutral, and closed-shell. A superposition of the resulting fragment molecular orbitals creates a wave function for the total molecular system that reflects an ionized fluoride atom or single water molecule at the center of the QM solute region, and this superposition is used as the initial guess in the ionized state SCF procedure. MOM-SCF is then employed to ensure the overlap between the computed molecular orbitals at each SCF cycle and the FRAGMO initial guess wave function is maximized. The success of this FRAGMO-MOM-SCF method is shown in Fig 3.1, where the resulting spin densities reflect ionization from the desired 2p orbital on fluoride [Panel (b)] and the 1b<sub>1</sub> orbital on a central water molecule [Panel (d)] that is far from the continuum boundary and properly participating in hydrogen bonding with the surrounding quantum-mechanical solvent water molecules. We end this section by noting that the SCF procedure for e<sup>-</sup>(aq) and Cl<sup>-</sup> required no modifications

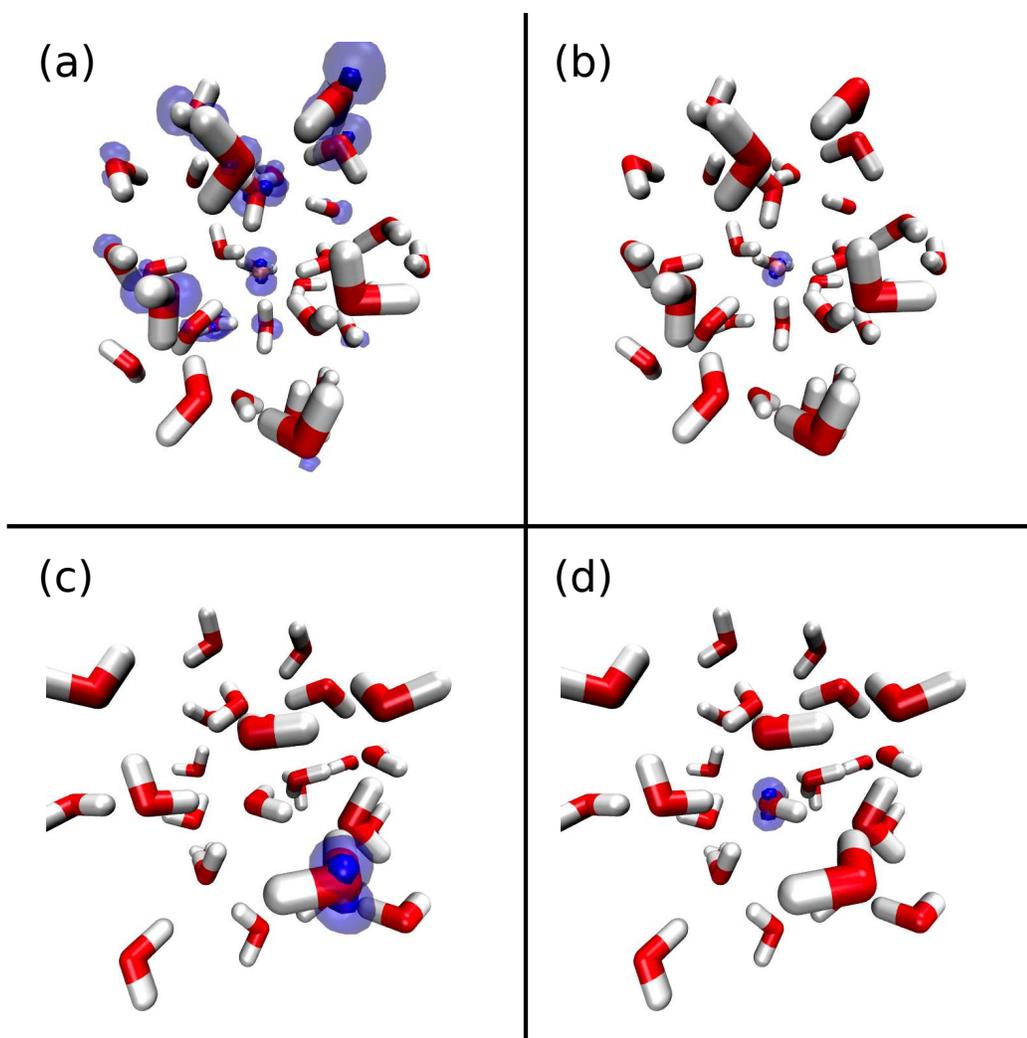


Figure 3.1: Panels (a) and (c) show the spin density computed using the MOM-SCF procedure for the fluoride-water and neutral water systems, respectively, and Panels (b) and (d) show the spin density computed using the FRAGMO-MOM-SCF procedure for the same systems. In Panels (a) and (b), the fluorine atom is colored pink. In all Panels, the opaque and transparent blue color illustrates 50% and 95% of the spin density, respectively. Panel (a) illustrates the difficulty in separating the fluoride 2p orbital from the orbitals of the surrounding water molecules: removal of an electron from the system results in a collective ionization of both the fluoride and water molecules. In Panel (c), the MOM-SCF procedure for the neutral water system results in ionization of a water molecule close to the solute-solvent boundary, where the effects of hydrogen bonding are neglected. Using the FRAGMO-MOM-SCF procedure fixes this issue, and the desired orbitals are ionized as shown in Panels (b) and (d).

because the ionization potentials for these species lie below and are energetically well-separated from that of liquid water. The excess spin density of  $e^-(aq)$  is shown in Fig 3.2(a) and the spin density of the ionized chlorine-water system using a standard SCF approach is shown in Fig 3.2(b).

### 3.4.3 Molecular Cavities for PCM

The simulations for the alkali metal cations and halide anions described in Section 3.4.1 provide the configurations that were utilized for all the calculations reported in Section 3.5, except for the  $e^-(aq)$  configurations that were extracted from Refs. [4] and [11]. For each of the systems investigated, the ion or excess electron and a certain number of explicit water molecules are treated quantum-mechanically with RI-MP2 as described in Section 3.4.2. The number of QM water molecules is determined by carving a sphere of radius 5.5 Å around the ion or excess electron centroid, and all water molecules within this sphere are treated explicitly. This roughly corresponds to two solvation shells of water molecules that are treated quantum-mechanically. For all of these configurations, the origin of the system is placed at its center-of-mass.

The last detail that needs specified for the non-equilibrium IEF-PCM technique employed to compute the VIPs presented in Section 3.5 is the form of the solute molecular cavity, which is the topic of this section. In addition to introducing the concept of a reaction field, L. Onsager explicates the properties of the solute cavity and discusses its physical meaning:<sup>142</sup> the molecular cavity should contain within it as much of the solute charge distribution as possible, be devoid of solvent, and conform as closely to the molecular shape as possible. Ensuring the solute charge distribution is entirely within the cavity becomes difficult to achieve for quantum mechanical

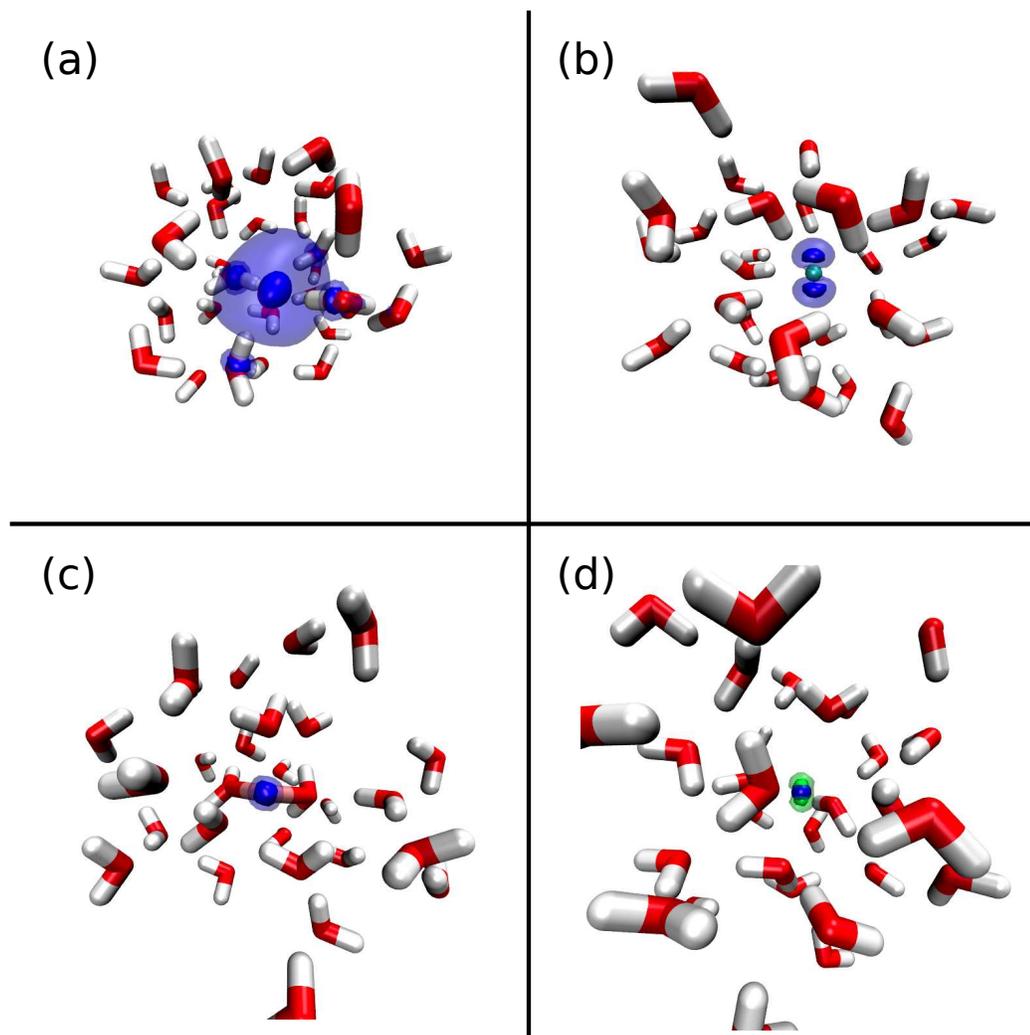


Figure 3.2: In Panels (a)–(c), the opaque and transparent blue color illustrates 50% and 95% of the spin density, respectively, while green is employed for the sodium-water system in Panel (d). In Panel (c), the lithium atom is colored pink and in Panel (d), the sodium atom is colored blue. Panel (a) shows the excess spin density of  $e^{-}(\text{aq})$  from a bulk liquid water simulation. Panel (b) illustrates the spin density of the ionized chlorine-water system using a standard SCF approach. Panels (c) and (d) show the spin densities of the lithium- and sodium-water systems, and employing the MOM-SCF approach results in ionization from the desired 1s and 2p orbitals.

densities because these, in principle, asymptotically decay to zero. As was discussed in Section 3.2.1, simpler PCMs neglect “charge penetration” effects where electron density leaks into the surrounding dielectric, but the IEF-PCM method employed in this work implicitly accounts for this.<sup>70,97,98</sup> Furthermore, molecular cavities that are significantly different than the actual molecular shape may suffer from deformed charge densities and unphysical values for various molecular properties.<sup>62</sup>

The simplest of these cavities is a solvent-excluded sphere that encompasses the solute molecule, the radius of which is taken to be 2.025 Å larger than the corresponding QM radius. In other words, since the molecular solute region is 5.5 Å, the size of the spherical cavity is 7.525 Å, which is 0.5 Å smaller than the cavities used in Section 2.3.3. Though the spherical cavity is a crude model that does not adequately reproduce the shape of the solute, it ensures that solvent dielectric is properly absent within the molecular region that is treated quantum-mechanically. The second type of molecular cavity employed in this investigation is derived by placing spheres on each of the atoms comprising the solute molecule, the size of which is dictated by their respective van der Waals radius. We utilize the Bondi definition for the atomic van der Waals radii, and each radius is multiplied by a constant factor of 1.2. The resulting surface that arises from the collection of these atomic spheres is called the van der Waals surface (vdWS), and Figure 3.3 illustrates this as a black surface surrounding two noncovalently bonded solute molecules. The grey shaded areas within the vdWS are representative of the solvent-excluded regions for this particular solute. The vdWS in conjunction with the Bondi definition is the most commonly used molecular cavity as it provides a suitable description of a surface that conforms well to the molecule and only needs calculated once at the beginning of the calculation

once the nuclear positions are specified. However, for pathological cases like  $e^-(aq)$ , where the system possesses a diffuse excess electron that is not associated with a particular solute atom or molecule, a van der Waals description of the molecular cavity becomes ill-defined. Furthermore, Figure 3.3 shows a “problematic region” (purple shaded area) where a hypothetical solvent molecule is too large to fit between the noncovalently bonded molecular regions, yet solvent dielectric will be placed there anyway because the vdWS prescription has left a gap between these disjointed parts of the solute.

In an effort to address this problem exhibited by  $e^-(aq)$ , two variants of the vdWS are examined. To derive both of these variations, consider a solvent molecule represented by a single “probe” sphere with a volume equal to its van der Waals volume. The unscaled vdWS (i.e., one using unscaled Bondi radii for the atomic spheres) is first constructed for the molecule as described above. The spherical solvent probe is then “rolled” along the vdWS, and the volume enclosed by this process results in the Solvent-Excluded Surface (SES) depicted in Figure 3.3. The points where the solvent probe, which is shown as a green circle in Figure 3.3, is unable to penetrate into the molecular region are called the re-entry points, and these points are shown as a red “re-entrant” surface in the figure. The union of the black and red surfaces forms the SES, and by doing this, the problematic region is encompassed by the re-entrant surfaces, and solvent dielectric will be properly eliminated. Additionally, the solvent probe will smooth out cusps that arise at the intersections of atomic spheres used to generate the vdWS. The second variation of an unscaled vdWS is also created by rolling a spherical solvent probe along the vdWS, and the surface traced out by the center of the probe forms the Solvent-Accessible-Surface (SAS), which is shown as

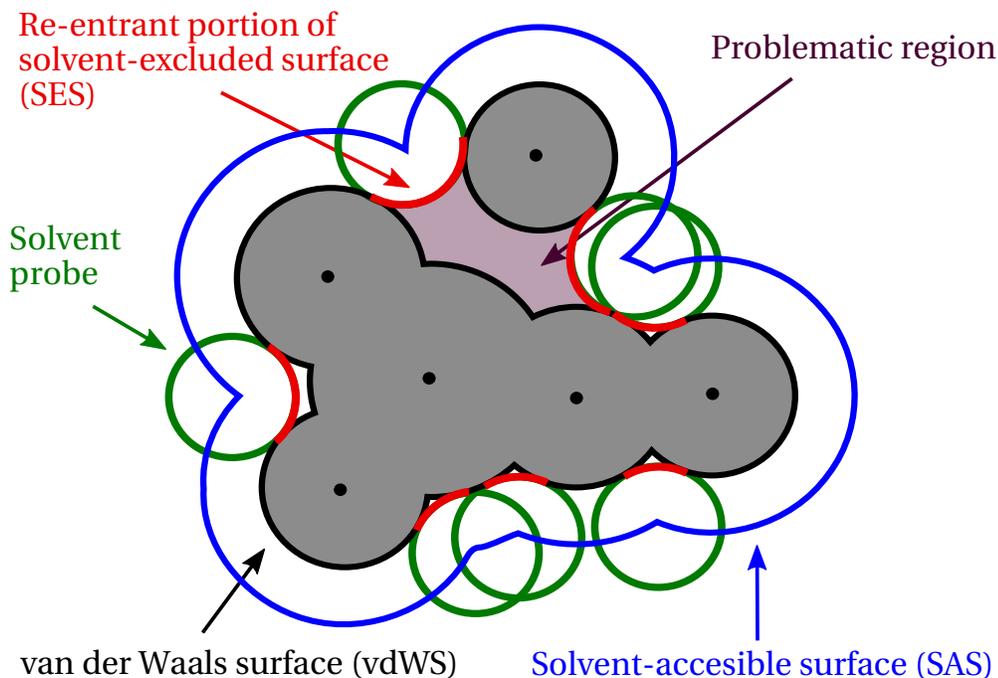


Figure 3.3: The various molecular cavity surfaces discussed in Section 3.4.3 are illustrated here. A generic, noncovalently bonded solute molecule is represented by the grey regions contained within the vdWS, which is shown in black. The problematic region, where solvent-solvent interactions are double-counted and where solvent dielectric is placed erroneously, is shown as the purple region between the noncovalently bonded regions of the solute. A solvent molecule probe, which is used to generate the SES and SAS cavities, is represented as a green circle. Where the solvent probe makes contact with the solute molecule forms the SES, which is the union of the black (vdWS) and red (re-entrant surface between noncovalently bonded regions of the molecule) surfaces. The SAS shown in blue is effectively a larger vdWS and is the locus of points formed by tracing the centroid of the solvent probe as it “rolls” around the solute molecule.

a blue surface in Figure 3.3. In other words, the SAS is the surface enclosing the volume where the probe center cannot enter the solute region.<sup>62</sup>

### 3.4.4 Dielectric Function $\epsilon(\mathbf{r})$ for PEQS

This section outlines the specific computational details needed to obtain non-equilibrium solvation free energies using the improved algorithms described in Sections 3.3.1 and 3.3.2, and discusses the development of an improved description of the dielectric function in liquid and at the liquid-vapor interface. To solve Eqs. 2.9, 3.80, and 3.84, we employ a three-dimensional, cubic Cartesian grid of side length of 25 Å and a grid spacing of  $\Delta x = 0.24$  Å, and its origin is placed at the center-of-mass of the quantum-mechanical solute. The solution of Poisson’s equation is obtained using a finite difference scheme with a fourth order W-cycle multigrid conjugate gradient algorithm, which is described in Chapter 4. The requisite charge densities and electrostatic potentials are then discretized on the Cartesian grid, and the iterative procedures are performed to compute the solvent polarization response. As in Section 2.3.4, the iterative charge density is updated until the Euclidean norm of the residual vector between iterations falls below a threshold of  $T_{\text{solver}} = 10^{-5}$  a.u., we use  $\eta = 0.6$  in Eq. 2.12, and when the *DIIS* error falls below a threshold of  $T_{\text{SCF}} = 10^{-5}$  a.u., the SCF procedure is considered converged.

Ultimately, we chose to employ two models for describing the spatially-varying dielectric function, the first of which is the spherical cavity that is described in Section 2.3.4. Its functional form is shown in Eq. 2.32, and an example is illustrated in Figure 2.2. The parameter  $\alpha$  controls the length scale of the switching process of  $\epsilon_{\text{vac}}$  to  $\epsilon_{\text{solv}}$ , and we take  $\alpha = 4.0 \text{ \AA}^{-1}$  which corresponds to a interpolation length of 1 Å.

As an example, Eq. 2.32 is applied to a bulk  $e^-$ (aq) configuration using the parameters described above, and a two-dimensional contour plot is provided in Figure 3.4. The contour plot is reflective of the  $xz$ -plane through the Cartesian grid origin.

The second functional form of  $\epsilon(\mathbf{r})$  was developed while investigating the various solute-solvent surfaces discussed in Section 3.4.3. For traditional PCM calculations, the default solute cavity is treated as a rigid, two-dimensional surface constructed from a union of atom-centered spheres, the size of which are taken to be the atomic van der Waals radii. At the surface of such solute cavities, there is an abrupt and discontinuous change in the value of the dielectric which assumes the value of  $\epsilon_{\text{vac}}$  inside the cavity and  $\epsilon_{\text{solv}}$  outside. A continuous and differentiable three-dimensional analog of this model is adapted from Ref. [10], which employs a set of spherically symmetric error functions centered on the solute nuclei, the coordinates of which are specified as  $\mathbf{R}_\alpha$ :

$$\epsilon(\mathbf{r}) = (\epsilon_{\text{solv}} - 1) \left\{ \prod_{\alpha}^{N_{\text{atom}}} h(d_\alpha, \Delta; |\mathbf{r} - \mathbf{R}_\alpha|) \right\}, \quad (3.91)$$

where

$$h(d_\alpha, \Delta; |\mathbf{r} - \mathbf{R}_\alpha|) = \frac{1}{2} \left[ 1 + \text{erf} \left( \frac{|\mathbf{r} - \mathbf{R}_\alpha| - d_\alpha}{\Delta} \right) \right], \quad (3.92)$$

and the van der Waals radius of atom  $\alpha$  is given by  $d_\alpha$ . Eq. 3.91 smoothly interpolates the dielectric from vacuum to solvent ( $\epsilon_{\text{solv}}$ ) over a length scale of  $\approx 4\Delta$ . Following Ref. [10], we take  $\Delta = 0.265 \text{ \AA}$ , which smoothly interpolates the dielectric for each atom-centered sphere over a length scale of  $\approx 1.0 \text{ \AA}$ . Using the same bulk  $e^-$ (aq) configuration for which the example in Figure 3.4 utilizes, a two-dimensional contour plot of the  $xz$ -plane is shown in Fig 3.5. Although Eq. 3.91 performs well for creating a cavity that correctly conforms to the molecular shape of the solute, Figure 3.5 shows that it suffers from placing unphysical solvent dielectric in the interstices of the solute

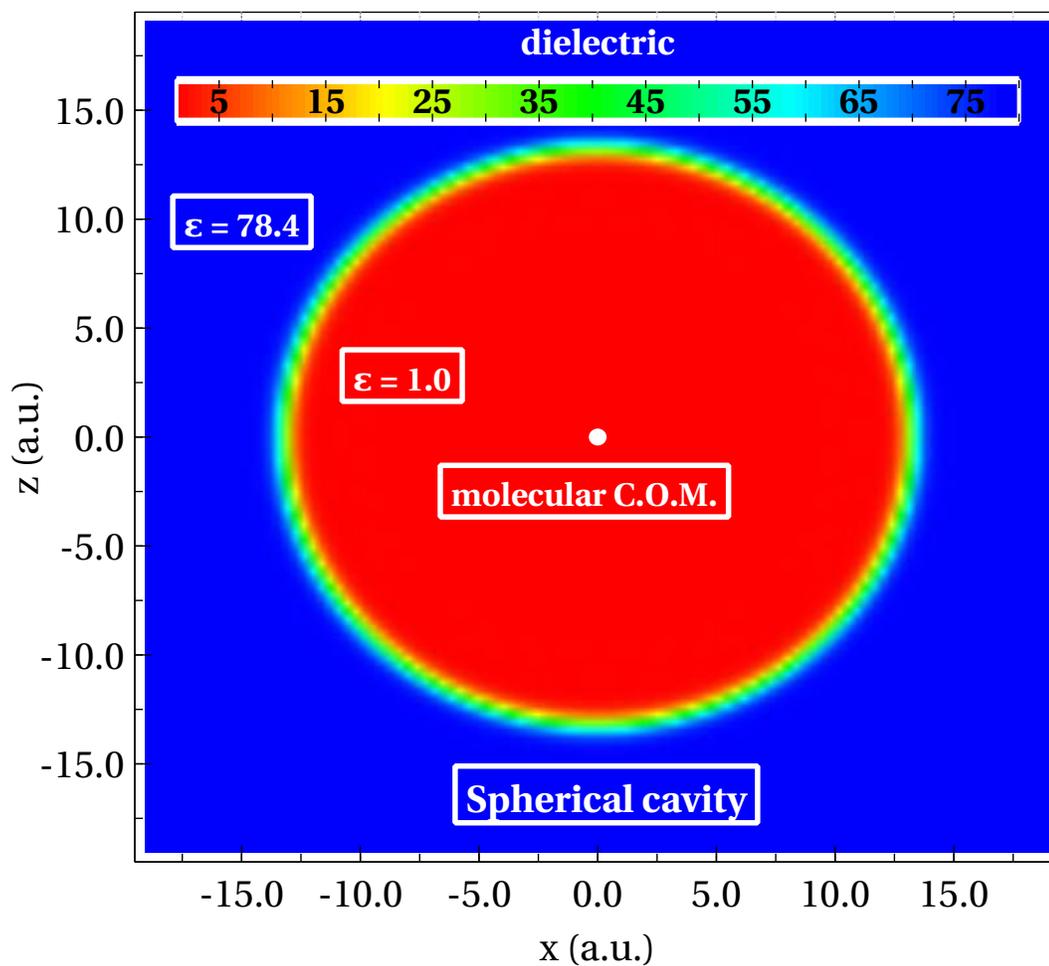


Figure 3.4: A two-dimensional contour plot of Eq. 2.32 applied to bulk  $e^-(aq)$  with  $\alpha = 4.0 \text{ \AA}^{-1}$ ,  $r_{QM} = 5.5 \text{ \AA}$ , and  $r_{mid} = 6.0 \text{ \AA}$  is illustrated. The contour plot is reflective of the  $xz$ -plane through the Cartesian grid origin. These parameters are utilized for all systems investigated in Chapter 3.

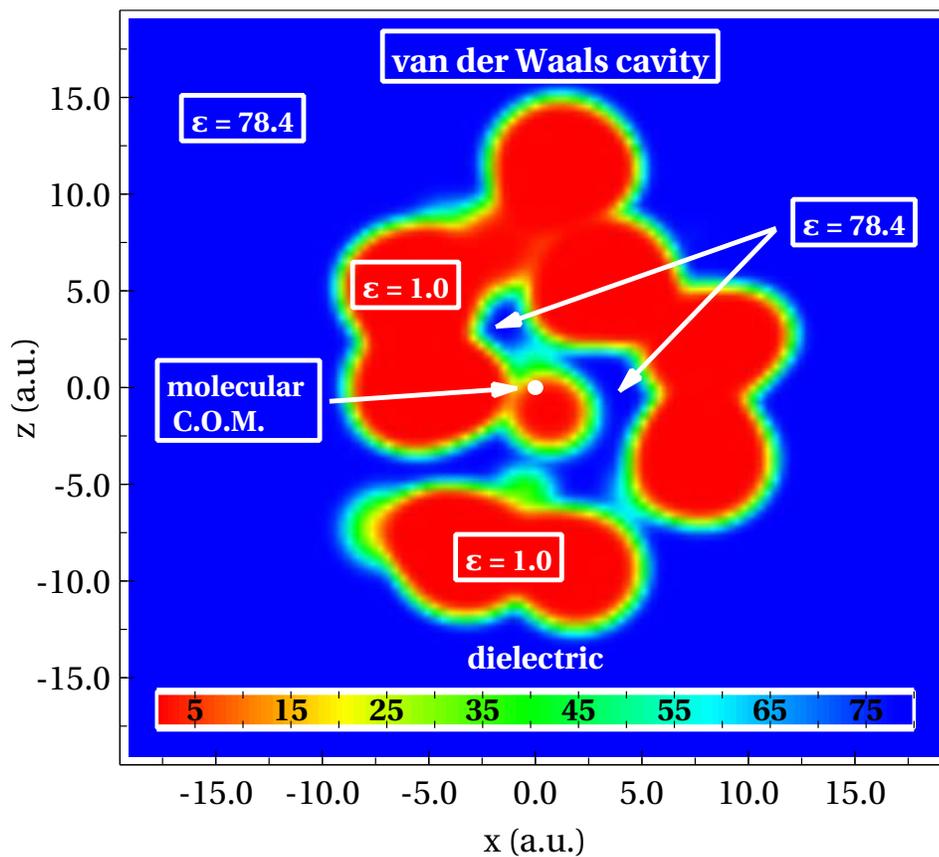


Figure 3.5: A two-dimensional contour plot of Eq. 3.91 applied to bulk  $e^-(aq)$  with  $\Delta = 0.265 \text{ \AA}$  is shown. The contour plot is reflective of the  $xz$ -plane through the Cartesian grid origin. This parameter is utilized for all systems investigated in Chapter 3 as well as Ref. [10].

molecule region. For the large molecular clusters being investigated in this work, where there are a significant number of explicit solvent water molecules included in the QM region, this erroneous solvent dielectric results in an overcounting of the electrostatic interactions that are handled by the inter-particle Coulomb interactions of the electronic Hamiltonian. This will overpolarize the solute wave function and result in unphysical and unmeaningful results.

Therefore in an effort to remedy this issue, we adapt the approach taken in PCM to create the SAS, which effectively uses a larger van der Waals radius for each solute atom. To accomplish this, Eqs. 3.91 and 3.92 are employed as for the van der Waals cavity, except with the modification  $d_\alpha \rightarrow d_\alpha + r_{\text{probe}}$ , where  $r_{\text{probe}}$  is the radius of a fictitious solvent “probe”, which is taken to be 1.4 Å for water. Such a modification is shown in Figure 3.6(a), and comparing this to Figure 3.5 shows that the three-dimensional analog of the PCM-SAS cavity improves the description of the dielectric function, but there are still regions within the molecular solute region where the dielectric assumes nonphysical values (show as  $\epsilon > 1.0$  in the plot). The “hybrid cavity” that is developed for the PEQS method and shown in Figure 3.6(b) was created to overcome the shortcomings of the spherical and van der Waals cavities discussed above, yet retain some of their desired features. The spherical cavity results in a form of the dielectric that properly takes the value of vacuum in the molecular region of the solute molecule, but fails to adequately conform to its molecular shape; the converse of this is true for the van der Waals cavity. Therefore the hybrid model is created by applying Eqs. 3.91 and 3.92 (with  $\Delta = 0.265$  Å) to the outermost atoms of the solute, and then manually setting  $\epsilon(\mathbf{r}) = 1.0$  in the internal molecular regions of the solute. As seen in Figure 3.6(b), the hybrid cavity provides a satisfactory

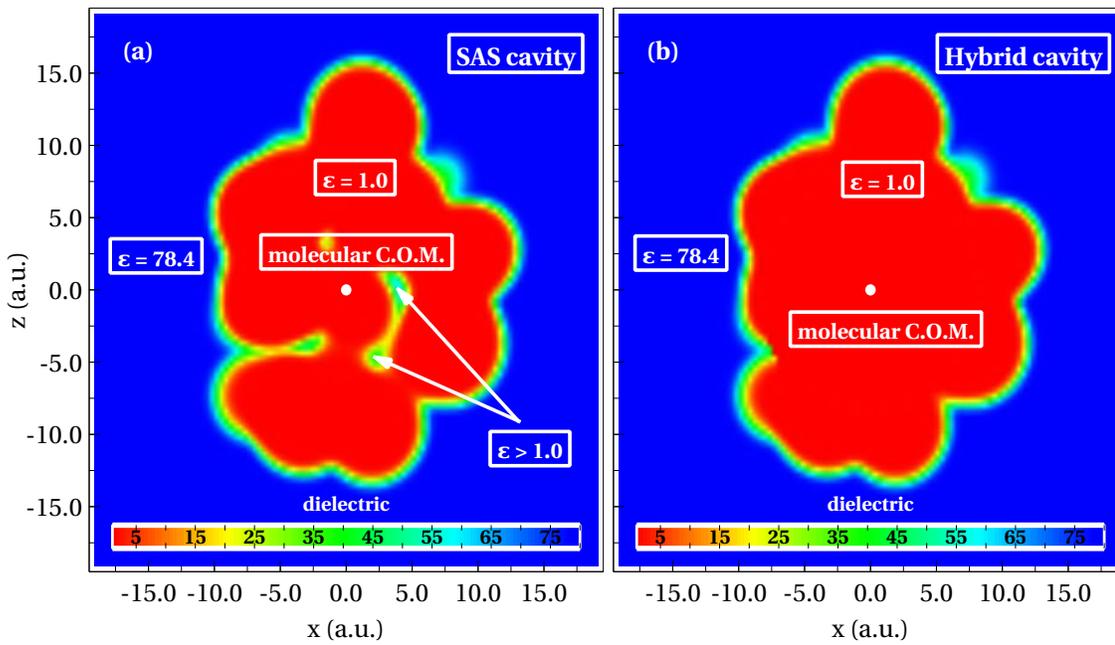


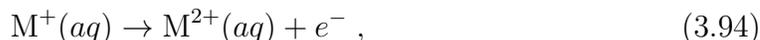
Figure 3.6: In Panel (a), a two-dimensional contour plot of Eq. 3.91 applied to bulk  $e^{-}(\text{aq})$  with  $\Delta = 0.265 \text{ \AA}$  along with a modification to the effective van der Waals radius  $d_{\alpha}$  for atom  $\alpha$  given by  $d_{\alpha} \rightarrow d_{\alpha} + r_{\text{probe}}$  with  $r_{\text{probe}} = 1.4 \text{ \AA}$  is shown. This is a three-dimensional analog of the PCM-SAS cavity, which improves upon the van der Waals cavity, but still has unphysical dielectric in the molecular region of the solute. In panel (b) the two-dimensional contour plot of a hybrid cavity that effectively applies Eq. 3.91 to the outermost atoms of the solute, and then sets  $\epsilon(\mathbf{r}) = 1.0$  within the molecular region of the solute is shown. The hybrid cavity is free of any erroneous solvent dielectric as in the spherical cavity, yet properly conforms to the molecular shape of the molecule as does the van der Waals cavity. These contour plot are reflective of the  $xz$ -plane through the Cartesian grid origin, and the corresponding parameters are utilized for all systems investigated in Chapter 3.

description of the dielectric environment, and this cavity, along with the spherical cavity, is used with the PEQS method method for computing VIPs in bulk liquid water that are reported in Section 3.5.

To create the interfacial dielectric environments, a similar approach from Section 2.3.4 is followed. The solvent dielectric is smoothly interpolated in the  $z$ -direction across the GDS using the hyperbolic tangent switching function given by Eq. 2.33 (see Figure 2.3). This procedure is performed for an interfacial  $e^-(aq)$  configuration using the spherical and hybrid cavities, and a two-dimensional contour plot of the  $xz$ -plane is illustrated in Figure 3.7 as an example.

### 3.5 Results and Discussion

The VIPs for  $e^-(aq)$ , alkali metal cations, halide anions, and a water molecule in liquid water and at the liquid-vapor interface utilizing the solvation models presented in Sections 3.2 and 3.3 in conjunction with the computational details from Section 3.4 are presented in this section. Recent liquid microjet measurements of these species are reported in Refs. [1], [2], and [3], and the ionization potentials computed with non-equilibrium the PEQS and PCM methods are discussed in context of those data. The ionization potentials of alkali metal cations and halide anions were measured directly by photoelectron spectroscopy of aqueous salt solutions with 100 eV photon energies that can access both outer- and inner-shell electrons.<sup>1,2</sup> The  $e^-(aq)$ , metal cation ( $M^+$ ), anion ( $A^-$ ), and water molecule photoionization processes are expressed by the following equations:



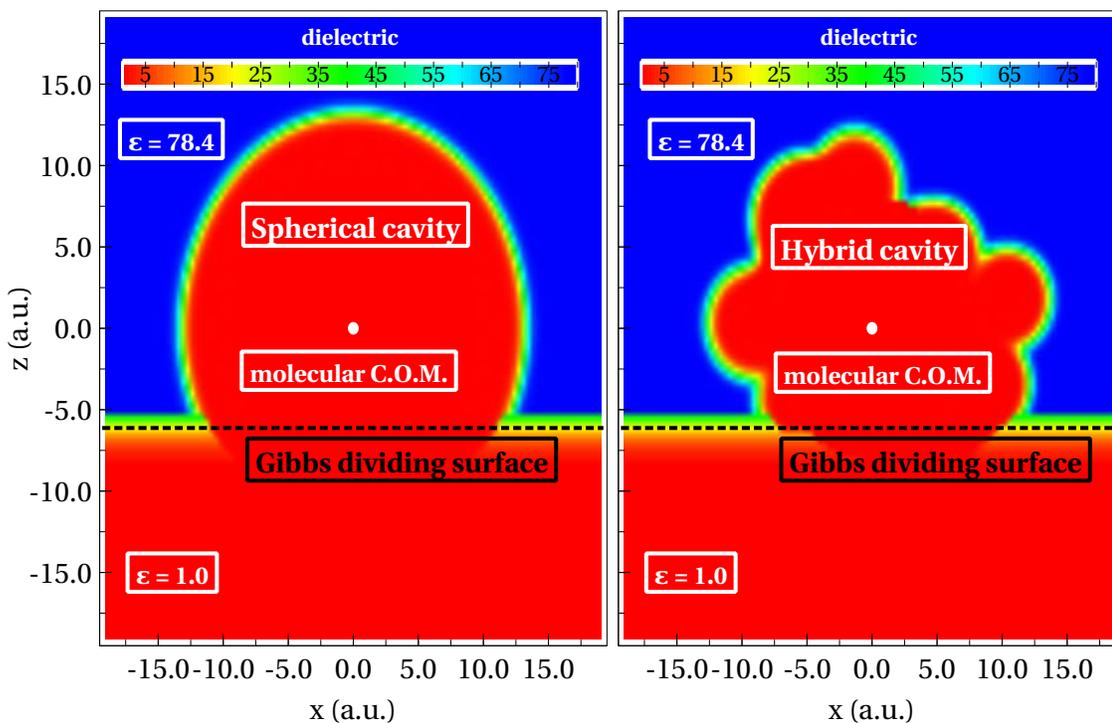
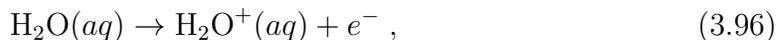


Figure 3.7: Panels (a) and (b) show the dielectric environment at the liquid vapor interface employing either the spherical cavity [Panel (a)] or the hybrid cavity [Panel(b)] for the molecular solute, which is an interfacial  $e^-(aq)$  configuration. The location of the GDS is shown as a black dotted line, and the solvent dielectric is smoothly interpolated from the bulk value to vacuum across this region using Eq. 2.33.



and



respectively. Figures 1 and 2 of Ref. [1] illustrate the photoemission spectra of liquid water,  $\text{Li}^+(aq)$ ,  $\text{Na}^+(aq)$ , and  $\text{Cl}^-(aq)$ , the latter three of which result from 3 m solutions of  $\text{NaCl}(aq)$  and  $\text{Li}(aq)$ . The reported ionization energies obtained from these spectra are 11.16 eV,  $60.4 \pm 0.07$  eV,  $35.4 \pm 0.04$  eV, and  $9.6 \pm 0.07$  eV for the  $1b_1$  orbital of liquid water, the  $1s$  orbital of  $\text{Li}^+(aq)$ , the  $2p$  orbital of  $\text{Na}^+(aq)$ , and the  $3p$  orbital  $\text{Cl}^-(aq)$ , respectively. In Ref. [2], the vertical ionization potential of  $\text{F}^-(aq)$  was remeasured after its value was erroneously assigned to the leading edge of the  $1b_1$  spectral feature arising from liquid water, and is reported to be 11.58 eV.

In Ref. [1], calculations were performed at the MP2 and CCSD(T) levels of theory in conjunction with an equilibrium PCM approach for incorporating long-range solvent effects to estimate the ionization potentials for the metal cation and anion halide systems. For those PCM calculations, a single ion is placed in a spherical cavity, the size of which is determined self-consistently through the electronic density at each SCF iteration, and the corresponding equilibrium solvation free energy is obtained upon convergence. Similarly for the ionized state, a separate, equilibrium PCM calculation is performed, and the difference between the ionized and unionized free energies is compared with the liquid microjet photoelectron measurements. Additionally, Eq. 3.11 is utilized to compute the classical Born solvation free energy for each state with the appropriate values of  $Z$  and  $a$ , and the free energy difference yields a second estimate to the reported ionization potentials. Following such a “ $\Delta$ SCF”, where the ionized and unionized states are allowed to fully relax in the

presence of the dielectric medium, or a “ $\Delta$ Born” approach provides an adiabatic ionization energy, whereas the experiments are probing a vertical ionization energy. The adiabatic ionization energies using PCM-CCSD(T), PCM-MP2, and the Born model are 62.23 eV, 62.13 eV, and 60.95 eV for  $\text{Li}^+(\text{aq})$  and 35.45 eV, 35.50 eV, and 35.70 eV for  $\text{Na}^+(\text{aq})$ .<sup>1</sup> The MP2 values are converged to within 0.1 eV of the CCSD(T) value for the cationic systems, which suggests that MP2 is sufficiently capturing electron correlation effects. Despite the agreement between CCSD(T) and MP2, however, the computed adiabatic ionization energies for  $\text{Li}^+(\text{aq})$  with these methods are  $\approx 2$  eV too large relative to experiment, whereas the Born model provides reasonable agreement to within 0.55 eV. In contrast, equilibrium PCM and the classical Born model provide ionization potentials in excellent agreement with experiment, differing by at most 0.3 eV for  $\text{Na}^+(\text{aq})$ .<sup>1</sup> For the anionic systems, the PCM-CCSD(T), PCM-MP2, and the Born model adiabatic ionization potentials are 7.8 eV, 8.09 eV, and 8.3 eV for  $\text{F}^-(\text{aq})$  and 6.85 eV, 6.97 eV, and 7.3 eV for  $\text{Cl}^-(\text{aq})$ , which deviate from experiment by  $\approx 3 - 4$  eV for  $\text{F}^-(\text{aq})$  and  $\approx 2.5 - 3$  eV for  $\text{Cl}^-(\text{aq})$ .<sup>1,2</sup> The poor performance of the adiabatic ionization energies for the anions suggests that non-equilibrium solvent polarization effects are particularly important for these systems, and that an adiabatic treatment of the ionization is not appropriate for these systems.

The results of Table 3.1 are a remarkable improvement compared to the data presented in Refs. [1] and [2]. However, it is interesting to note that the adiabatic ionization energies of the cations from Ref. [1] were more accurate than for the anions. Upon vertical ionization of an equilibrated cation  $\text{M}^+$ , the nuclear polarization of the surrounding water molecules will not have to adjust much to accommodate the dicationic species  $\text{M}^{2+}$  since they have already preferentially oriented around the  $\text{M}^+$

Species	Expt. / eV	PEQS / eV		PCM / eV	
		Hybrid cavity	Spherical cavity	SAS cavity	
Li <sup>+</sup>	60.40 ± 0.07 <sup>a</sup>	60.43 ± 0.55	61.27 ± 0.50	61.13 ± 0.44	
Na <sup>+</sup>	35.40 ± 0.04 <sup>a</sup>	35.99 ± 0.73	36.35 ± 0.73	36.53 ± 0.50	
H <sub>2</sub> O	11.23 <sup>b</sup>	11.28 ± 0.52	11.51 ± 0.52	11.41 ± 0.45	
e <sup>-</sup>	3.7 ± 0.1	3.51 ± 0.29	3.15 ± 0.32	3.16 ± 0.29	
F <sup>-</sup>	11.58 <sup>b</sup>	11.45 ± 0.40	11.19 ± 0.40	11.27 ± 0.36	
Cl <sup>-</sup>	9.60 ± 0.07 <sup>a</sup>	9.31 ± 0.37	9.23 ± 0.41	9.13 ± 0.41	

Table 3.1: VIPs (in eV) at the RI-MP2 level of theory using the basis sets described in Section 3.4.2 and the non-equilibrium solvation models described in Sections 3.2 and 3.3 for various systems in bulk liquid water. <sup>a</sup> Ref. [1]. <sup>b</sup> Ref. [2]. <sup>c</sup> Ref. [3].

species. A possible explanation for the VIP of Li<sup>+</sup> being too large by  $\approx 2$  eV in Ref. [1] is the inability of the employed adiabatic PCM approach to account for changes in electronic polarization. The electric field from the ionized M<sup>2+</sup> species is significantly larger than for the M<sup>+</sup> species, and therefore a non-equilibrium treatment of the electronic polarization is most important.<sup>1</sup> Table 3.1 shows that a state-specific, non-equilibrium treatment of *both* the nuclear and electronic polarizations provides VIPs for Li<sup>+</sup> and Na<sup>+</sup> to be within a few tenths of an eV relative to experiment. Furthermore, inclusion of two solvation shells of water molecules for these systems, which is not done in Ref. [1], captures most of the short-range polarization effects at a high level of theory.

For the halide anions, where the adiabatic solvation techniques of Ref. [1] provide VIPs that are too small by  $> 2.5$  eV, one can expect that nuclear polarization effects will be more significant. Prior to ionization of an equilibrated species A<sup>-</sup>, the hydrogen atoms of the surrounding water molecules will be oriented towards the

Species	Hybrid Cavity	
	Bulk	Interface
Li <sup>+</sup>	60.43 ± 0.55	60.99 ± 0.50
Na <sup>+</sup>	35.99 ± 0.73	36.66 ± 0.69
H <sub>2</sub> O	11.28 ± 0.52	11.54 ± 0.55
e <sup>-</sup> (1.6 eV <sup>a</sup> )	3.51 ± 0.29	3.34 ± 0.44
F <sup>-</sup>	11.45 ± 0.40	11.07 ± 0.42
Cl <sup>-</sup>	9.31 ± 0.37	9.05 ± 0.42

Table 3.2: VIPs (in eV) at the RI-MP2 level of theory using the basis sets described in Section 3.4.2 and the non-equilibrium PEQS method described in Section 3.3 for various systems at the liquid-vapor interface. The bulk liquid VIPs computed with PEQS from Table 3.1 are reproduced here for comparison. <sup>a</sup> Ref. [9]

anion. Upon ionization to generate the species A, the oxygen atoms will preferentially orient towards the neutral species, which will require significant reorientation of the solvation shells.<sup>1</sup> In contrast to the cationic systems, ionization of the anion will not result in a strong change in the electronic polarization from the surrounding solvent, and including two solvation shells of explicit water molecules should be sufficient to capture the short-range electronic polarization effects. Table 3.1 illustrates that the non-equilibrium PCM and PEQS methods are more accurately describing these nuclear polarization effects, and provide VIPs that are within 0.3 – 0.5 eV of experiment.

In addition to the bulk liquid VIPs presented above, which served as a metric for verifying the robustness of the PEQS method against experimental measurements and the well-established PCM methodology, the non-equilibrium PEQS technique was employed to compute the VIPs of these species at a liquid-vapor interface. Such calculations are not possible with standard PCMs due to the constraint of isotropic

boundary conditions (see Section 3.2). Table 3.2 contains the interfacial VIPs using the non-equilibrium PEQS method with the hybrid form of  $\epsilon(\mathbf{r})$  for the solute cavity as well as employing a smoothing function to interpolate the dielectric from its value in liquid water to vacuum across the GDS (see Section 3.4.4). Since experimental values for interfacial VIPs of these species have not yet been measured, except for  $e^-(aq)$  which is shown in parentheses and is taken from Ref. [9], the bulk liquid VIPs computed using PEQS in Table 3.1 are reproduced in Table 3.2 for comparison. The interfacial VIPs are consistently smaller for anionic species by 0.2 – 0.5 eV and larger for the cationic species and water molecule by 0.4 – 0.6 eV. At ambient conditions, the interfacial species likely does not linger around the GDS for long, and since two solvation shells of explicit water molecules are also included in the calculations, the short-range solvation structure is similar to that of the bulk species. Furthermore, the VIP of  $e^-(aq)$  in liquid and at the liquid-vapor were also reported in Section 2.4.2 using a perturbative approach for the non-equilibrium solvation models. The VIP of  $e^-(aq)$  using the improved, state-specific methodology presented in Section 3.3.2 are slightly more accurate relative to the previous perturbative approach and are in agreement with previous *ab initio* results.<sup>4,11</sup>

### 3.6 Conclusions

We present the details for a newly developed continuum solvation model that includes non-equilibrium solvent polarization effects and also employs an arbitrary description of the dielectric environment. The details for constructing dielectric functions  $\epsilon(\mathbf{r})$  representing bulk liquid and liquid-vapor interfaces are also discussed.

To examine the effectiveness of this model, the VIPs of alkali metal cations, anion halides,  $e^-(aq)$ , and neutral liquid water were computed with PEQS and also with non-equilibrium PCM. These systems were chosen because experimental and computational endeavors have previously measured and calculated the VIPs of these species with liquid microjet photoelectron spectroscopy and an adiabatic approach with equilibrium PCM. For these species in liquid water, PEQS provides VIPs that are in reasonable agreement with both non-equilibrium PCM and experimental measurements, providing confidence that the method is robust. Furthermore, previously calculated VIPs employing equilibrium PCM neglected important polarization effects that lead to errors as large as 4 eV in some cases. For our calculations with non-equilibrium PEQS and PCM, the electronic and nuclear polarization effects from the solvent are treated on an equal footing, which is necessary for a proper description of the vertical ionization process, and we reduce the discrepancy between experimental and calculated VIPs to 0.2 – 1.0 eV. We then use PEQS in conjunction with an interfacial description of the dielectric to make predictions for the VIPs at the liquid-vapor interface. This is not possible with standard PCMs because of the isotropic nature of the methodology.

## Chapter 4: Numerically Solving Poisson’s Equation

### 4.1 Finite Difference Method

In Chapters 2 and 3, a variety of methods for incorporating solvent effects into electronic structure calculations was discussed. For the PEQS method, the charge densities and electrostatic potentials are discretized on a three-dimensional Cartesian grid, and  $\mathcal{O}(10^6)$  grid points are required to accurately represent these quantities. For PCMs, the two-dimensional surface of the solute cavity, which separates the quantum-mechanical solute particles from the classical continuum solvent, is discretized into tesserae where the ASCs are computed. The required number of tesserae to achieve satisfactory accuracy, which depends on the solute identity and type of cavitation scheme, is orders of magnitude smaller by comparison. Therefore, to solve linear equations such as Eq. 3.52, it is commonplace for PCM methods to obtain the ASCs  $\mathbf{q}$  by direct inversion of the polarization weight matrix  $\mathbf{Q}$ , the cost of which is  $\mathcal{O}(N_{\text{tess}}^3)$  and  $\mathcal{O}(N_{\text{tess}}^2)$  in computational time and memory, respectively. However, this approach is prohibitively expensive for the PEQS method, where the memory cost alone is  $> 7$  TB for storing the matrix representation of the discretized Laplacian operator, and it is necessary to use relaxation techniques such as an iterative conjugate gradient (CG) procedure. This Chapter focuses on a finite difference discretization scheme for

Poisson's equation on large, rectangular Cartesian grids and also on improving the efficiency of obtaining a numerical solution by exploiting a multigrid method.

Partial differential equations are ubiquitous in science and engineering, and many techniques have been developed to provide accurate solutions to them. The analytical solution of these equations is usually available only in asymptotic limits, and therefore numerical approaches are required for complex problems at realistic length and time scales. To begin our discussion, consider the generalized three-dimensional version of Poisson's equation recast in terms of a polarization charge density that was introduced in Chapter 2:

$$\begin{aligned}\nabla^2\varphi_{\text{tot}}(\mathbf{r}) &= -4\pi \left[ \frac{\rho_{\text{sol}}(\mathbf{r})}{\epsilon(\mathbf{r})} + \frac{\hat{\nabla} \ln \epsilon(\mathbf{r}) \cdot \hat{\nabla} \varphi_{\text{tot}}(\mathbf{r})}{4\pi} \right] \\ &= -4\pi [\rho_{\text{sol}}(\mathbf{r}) + \rho_{\text{pol}}(\mathbf{r})] .\end{aligned}\tag{4.1}$$

For the discussions of this Chapter, Eq. 4.1 is rewritten in its simplest form as

$$\nabla^2\varphi(\mathbf{r}) = \rho(\mathbf{r}) ,\tag{4.2}$$

where the factor of  $-4\pi$  is included in  $\rho(\mathbf{r})$ , the subscript notation signifying the origin of the particular quantity (i.e. sol, pol, or tot) used previously is suppressed for clarity, and it is understood that the unknown quantity on the left side of Eq. 4.2 is  $\varphi_{\text{tot}}(\mathbf{r})$ , the source of which is  $\rho_{\text{tot}}(\mathbf{r}) = \rho_{\text{sol}}(\mathbf{r}) + \rho_{\text{pol}}(\mathbf{r})$ . The Laplacian operator in three-dimensional Cartesian coordinates is expanded as

$$\begin{aligned}\frac{\partial^2\varphi}{\partial x^2} + \frac{\partial^2\varphi}{\partial y^2} + \frac{\partial^2\varphi}{\partial z^2} &= \rho(x, y, z) && \text{for } (x, y, z) \in \Omega \\ \varphi(x, y, z) &= 0 && \text{for } (x, y, z) \in \delta\Omega .\end{aligned}\tag{4.3}$$

The domain  $\Omega$  is defined as the region

$$\Omega = \left\{ [-L_x/2 < x < L_x/2], [-L_y/2 < y < L_y/2], [-L_z/2 < z < L_z/2] \right\}\tag{4.4}$$

of a rectangular grid centered at the origin  $O = (0, 0, 0)$  with side lengths  $L_x$ ,  $L_y$ , and  $L_z$ . The second equality of Eq. 4.3 enforces the Dirichlet boundary condition that  $\phi(\mathbf{r}) \rightarrow 0$  at the domain boundary surface  $\delta\Omega$ .

The rectangular grids utilized for the PEQS method can be thought of as three infinite sets of evenly-spaced parallel lines orthogonal to each other, and the grid points are defined as the intersections of these lines over the domain  $\Omega$ . The process of discretization is to specify  $\{L_x, L_y, L_z\}$  and the corresponding set of spacings  $\{h_x, h_y, h_z\}$  between the respective parallel lines. In doing this, the complexity of the problem is reduced, and it is transformed into a set of equations corresponding to a set of unknowns. Additionally, by mapping the continuous function  $\varphi(x, y, z)$  that contains an infinite amount of information onto a domain of finite size, it is only necessary to solve the problem on a finite number of grid points. For a uniform rectangular grid with side lengths  $\{L_x, L_y, L_z\}$  containing  $\{N_x, N_y, N_z\}$  grid points, the spacings between the grid points in the three Cartesian coordinate directions are given by  $h_\alpha = L_\alpha / (N_\alpha - 1)$  for  $\alpha \in \{x, y, z\}$ . The Cartesian coordinates are mapped onto the grid coordinates through  $x_i = -L_x/2 + ih_x$  where  $i = 0, \dots, (N_x - 1)$  indexes the  $x$ -component of the grid point. The mappings are similar for the  $y$ - and  $z$ -components using the indices  $j$  and  $k$ , respectively. The value of the unknown function  $\varphi$  at the grid point  $(x_i, y_j, z_k)$  is denoted as  $\varphi_{i,j,k} = \varphi(x_i, y_j, z_k)$ .

To discretize the function derivatives on the grid required for the PEQS method, a Taylor series expansion is utilized. The power of such an expansion lies in the ability to relate a value of  $\varphi_{i,j,k}$  or its derivatives at a particular point  $\{i, j, k\}$  to values at neighboring grid points. Using a Taylor series expansion, advancing one grid point

forward or backward in the  $x$ -direction is expressed as

$$\varphi_{i\pm 1,j,k} = \varphi_{i,j,k} \pm \sum_{n=0}^{\infty} \frac{h_x^{(2n+1)}}{(2n+1)!} \left. \frac{\partial^{(2n+1)} \varphi}{\partial x^{(2n+1)}} \right|_{i,j,k} + \sum_{n=0}^{\infty} \frac{h_x^{(2n+2)}}{(2n+2)!} \left. \frac{\partial^{(2n+2)} \varphi}{\partial x^{(2n+2)}} \right|_{i,j,k}, \quad (4.5)$$

and it is then possible to construct the first derivative of the electrostatic potential with respect to perturbations in the  $x$ -coordinate:

$$\begin{aligned} \left. \frac{\partial \varphi}{\partial x} \right|_{i,j,k} &= \frac{\pm (\varphi_{i\pm 1,j,k} - \varphi_{i,j,k})}{h_x} - \sum_{n=1}^{\infty} \frac{h_x^{(2n)}}{(2n+1)!} \left. \frac{\partial^{(2n+1)} \varphi}{\partial x^{(2n+1)}} \right|_{i,j,k} \\ &\mp \sum_{n=0}^{\infty} \frac{h_x^{(2n+1)}}{(2n+2)!} \left. \frac{\partial^{(2n+2)} \varphi}{\partial x^{(2n+2)}} \right|_{i,j,k} \end{aligned} \quad (4.6)$$

with similar expressions for perturbations in the other coordinates. The first term in parentheses on the right side of Eq. 4.6 is either a forward  $[+(\varphi_{i+1,j,k} - \varphi_{i,j,k})]$  or backward  $[-(\varphi_{i-1,j,k} - \varphi_{i,j,k})]$  finite difference approximation to the first derivative, while the remaining terms quantify the truncation error,  $T_{i,j,k}$ .<sup>12</sup>

$$\left. \frac{\partial \varphi}{\partial x} \right|_{i,j,k} = \frac{\pm (\varphi_{i\pm 1,j,k} - \varphi_{i,j,k})}{h_x} + T_{i,j,k}, \quad (4.7)$$

where

$$T_{i,j,k} = - \sum_{n=1}^{\infty} \frac{h_x^{(2n)}}{(2n+1)!} \left. \frac{\partial^{(2n+1)} \varphi}{\partial x^{(2n+1)}} \right|_{i,j,k} \mp \sum_{n=0}^{\infty} \frac{h_x^{(2n+1)}}{(2n+2)!} \left. \frac{\partial^{(2n+2)} \varphi}{\partial x^{(2n+2)}} \right|_{i,j,k}. \quad (4.8)$$

Formally, in the limit that  $h_x \rightarrow 0$ ,  $T_{i,j,k} \rightarrow 0$  and Eq. 4.6 is the exact *continuous* derivative of  $\varphi_{i,j,k}$ . Unfortunately, it is impossible to compute the infinite sums required for an exact treatment of the derivatives within any discretization scheme, so it necessary to truncate the summations in Eq. 4.6.

The order of accuracy for the finite difference approximations is determined by the power of  $h_x$  in the leading term of  $T_{i,j,k}$ . Both the forward and backward approximations in Eq. 4.6 exhibit  $\mathcal{O}(h_x)$  error, or, in words, exhibit errors that are first-order with respect to the grid spacing  $h_x$ . However, it is possible to systematically improve

the discretization error by combining terms from the Taylor series expansion and eliminating terms of lower order. For example, the first derivative of  $\varphi_{i,j,k}$  can also be expressed as

$$\frac{\partial \varphi}{\partial x} \Big|_{i,j,k} = \frac{(\varphi_{i+1,j,k} - \varphi_{i-1,j,k})}{2h_x} + T_{i,j,k}, \quad (4.9)$$

which is known as the central difference approximation, and the truncation order becomes

$$T_{i,j,k} = - \sum_{n=1}^{\infty} h_x^{(2n)} \frac{\partial^{(2n+1)} \varphi}{\partial x^{(2n+1)}} \Big|_{i,j,k} - \sum_{n=1}^{\infty} h_x^{(2n+1)} \frac{\partial^{(2n+2)} \varphi}{\partial x^{(2n+2)}} \Big|_{i,j,k}. \quad (4.10)$$

The leading term in Eq. 4.10 contains a factor of  $h_x^2$ , and the central difference approximation to the first derivative displays  $\mathcal{O}(h_x^2)$  error. A similar analysis can be performed for the Laplacian operator that is required for solving Poisson's equation in Eq. 4.3. Focusing on the second derivative of the electrostatic potential with respect to  $x$ , the central difference finite difference expression is

$$\frac{\partial^2 \varphi}{\partial x^2} \Big|_{i,j,k} = \frac{(\varphi_{i+1,j,k} - 2\varphi_{i,j,k} + \varphi_{i-1,j,k})}{h_x^2} + T_{i,j,k}, \quad (4.11)$$

where the  $\mathcal{O}(h_x^2)$  error truncation error is

$$T_{i,j,k} = -2 \sum_{n=1}^{\infty} \frac{h_x^{(2n)}}{(2n+2)!} \frac{\partial^{(2n+2)} \varphi}{\partial x^{(2n+2)}} \Big|_{i,j,k}. \quad (4.12)$$

The three-dimensional Laplacian operator in Eq. 4.3 is rewritten in a discretized form as

$$\begin{aligned} \nabla^2 \varphi_{i,j,k} &= \frac{\varphi_{i+1,j,k} - 2\varphi_{i,j,k} + \varphi_{i-1,j,k}}{h_x^2} + \frac{\varphi_{i,j+1,k} - 2\varphi_{i,j,k} + \varphi_{i,j-1,k}}{h_y^2} \\ &+ \frac{\varphi_{i,j,k+1} - 2\varphi_{i,j,k} + \varphi_{i,j,k-1}}{h_z^2} + T_{i,j,k}, \end{aligned} \quad (4.13)$$

where

$$T_{i,j,k} = -2 \sum_{\alpha \in \{x,y,z\}} \sum_{n=1}^{\infty} \frac{h_\alpha^{(2n)}}{(2n+2)!} \frac{\partial^{(2n+2)} \varphi}{\partial \alpha^{(2n+2)}} \Big|_{i,j,k}, \quad (4.14)$$

and the error order is  $\max[\mathcal{O}(h_x^2), \mathcal{O}(h_y^2), \mathcal{O}(h_z^2)]$ . Using Eq. 4.13 in Eq. 4.3 and omitting the truncation error, the second-order discretized Poisson equation becomes

$$\begin{aligned} & \frac{\varphi_{i+1,j,k} - 2\varphi_{i,j,k} + \varphi_{i-1,j,k}}{h_x^2} + \frac{\varphi_{i,j+1,k} - 2\varphi_{i,j,k} + \varphi_{i,j-1,k}}{h_y^2} \\ & + \frac{\varphi_{i,j,k+1} - 2\varphi_{i,j,k} + \varphi_{i,j,k-1}}{h_z^2} = \rho_{i,j,k} \end{aligned} \quad (4.15)$$

over the domain

$$\Omega = \{[1 \leq i \leq (N_x - 2)], [1 \leq j \leq (N_y - 2)], [1 \leq k \leq (N_z - 2)]\} \quad (4.16)$$

and  $\varphi_{i,j,k} = 0$  on the boundary surface

$$\delta\Omega \equiv \{i = 0 \wedge j = 0 \wedge k = 0 \wedge i = (N_x - 1) \wedge j = (N_y - 1) \wedge k = (N_z - 1)\}. \quad (4.17)$$

For the case  $\{h_x, h_y, h_z\} \rightarrow 0$ , Eq. 4.15 approaches the continuous form of Poisson's equation that provides the exact solution  $\varphi_{i,j,k}^{\text{exact}}$ . For  $\{h_x, h_y, h_z\} > 0$ , the discretization error  $v_{i,j,k}^h$  is defined as the difference between  $\varphi_{i,j,k}^{\text{exact}}$  and the finite difference approximation  $\varphi_{i,j,k}^h$ .<sup>12</sup>

$$v_{i,j,k}^h = \varphi_{i,j,k}^{\text{exact}} - \varphi_{i,j,k}^h. \quad (4.18)$$

In general, the discretization and truncations errors are of the same order, but they are inherently different quantities. The quantity in Eq. 4.18 is reflective of error incurred from a numerical solution of Poisson's equation, whereas the truncation error arises from the choice of finite difference discretization scheme and is independent of the partial differential equation being solved.

In order to achieve the best possible accuracy for the numerical method, it is essential to employ finite difference schemes that exhibit high-order errors (e.g.  $\mathcal{O}(h^m)$  for large  $m$ ). The central difference approximation given in Eqs. 4.9 and 4.13 are

$m$	$n$	$c_{m,-n}$				$c_{m,0}$	$c_{m,n}$			
1	1	-	-	-	-1/2	0	1/2	-	-	-
	2	-	-	1/12	-2/3	0	2/3	-1/12	-	-
	3	-	-1/60	3/20	-3/4	0	3/4	-3/20	1/60	-
	4	1/280	-4/105	1/5	-4/5	0	4/5	-1/5	4/105	-1/280
2	2	-	-	-	1	-2	1	-	-	-
	2	-	-	-1/12	4/3	-5/2	4/3	-1/12	-	-
	3	-	1/90	-3/20	3/2	-49/18	3/2	-3/20	1/90	-
	4	-1/560	8/315	-1/5	8/5	-205/72	8/5	-1/5	8/315	-1/560

Table 4.1: The finite difference coefficients for the first and second derivatives ( $m = 1$  or  $2$  in Eq. 4.19) and the corresponding orders of accuracy. [Note: the order of accuracy is  $\mathcal{O}(h^{2n})$ .]

accurate up to  $\mathcal{O}(h^2)$ , but higher-order methods are desired. Considering differentiation along the  $x$ -coordinate as an example, with similar equations for  $y$ - and  $z$ -derivatives, the general  $m^{\text{th}}$ -order finite difference derivative expression is

$$\left. \frac{\partial^m \varphi}{\partial x^m} \right|_{i,j,k} = \sum_{-n}^n c_{m,n} \frac{\varphi_{i+n,j,k}}{h_x^m}, \quad (4.19)$$

where  $c_{m,n}$  are the finite difference coefficients (see Table 4.1) for the  $m^{\text{th}}$ -order derivative with  $\mathcal{O}(h_x^{2n})$  accuracy. For the PEQS implementation within the QCHEM software package,<sup>75</sup> eighth-order accurate finite difference schemes for the gradient and Laplacian operators were utilized. In addition to providing chemically accurate free energies with  $\mathcal{O}(10^6)$  grid points and high-order finite difference schemes, it is desirable that the method be computationally affordable. This is a challenging task given the enormous number of grid points required to sufficiently capture the topological features of the electronic density. Therefore, to make the PEQS method a useful computational tool, an efficient numerical technique must be employed. To this end,

a multigrid method is adapted for use within the PEQS algorithms to accelerate convergence of the CG routines used to solve Eq. 4.3, and that is focus of the following section.

## 4.2 Multigrid Method

The process of discretizing Poisson's equation given by Eq. 4.2 with a high-order finite difference scheme was presented in the previous section, but it is left to discuss the manner in which a numerical solution is obtained. By recasting Poisson's equation as Eq. 4.2, the numerical problem is reduced to a set of linear equations expressed in matrix-vector form as

$$\mathcal{L}^h \boldsymbol{\varphi}^h = \boldsymbol{\rho}^h, \quad (4.20)$$

where  $\mathcal{L}^h$  is the  $N_{\text{grid}} \times N_{\text{grid}}$  matrix representation of the discretized Laplacian operator and the vectors  $\boldsymbol{\varphi}^h$  and  $\boldsymbol{\rho}^h$  contain the  $N_{\text{grid}}$  discretized values of  $\varphi_{i,j,k}^h$ , and  $-4\pi\rho_{i,j,k}^h$ , respectively, subject to the boundary conditions  $\Omega$  and  $\delta\Omega$ . The simplest approach to solve Eq. 4.20 is to invert  $\mathcal{L}^h$  and apply  $(\mathcal{L}^h)^{-1}$  to  $\boldsymbol{\rho}^h$ , but it was noted at the beginning of this Chapter that such a route is not practical due to excessive computational overhead.

As an alternative to a direct matrix inversion approach, a CG routine is implemented to numerically solve Eq. 4.20. The CG computational time and memory storage requirements scale as  $\mathcal{O}(N_{\text{grid}}^2)$  and  $\mathcal{O}(N_{\text{grid}})$ , respectively, which is an order of magnitude improvement over a direct solver procedure. Furthermore, for the CG routines, the matrix  $\mathcal{L}^h$  is never explicitly formed, and only its action on the vector  $\boldsymbol{\varphi}^h$  is required. For symmetric and positive-semidefinite matrices such as  $\mathcal{L}^h$ , the CG routine is guaranteed to converge in  $N_{\text{grid}}$  iterations, which is somewhat unsettling

given its magnitude. Furthermore, in Chapter 2 of Ref [12], a Fourier analysis of the discretization error shows that it contains a spectrum of wavelengths  $\lambda$ , the magnitudes of which are either short, comparable, or long relative to the grid resolution  $h$ . The CG routine efficiently eliminates discretization error where  $\lambda \simeq h$ , but struggles with components of the error where  $\lambda > h$ . Thus, iterative techniques such as CG can effectively smooth out the short wavelength discretization errors, but they do not perform well for efficiently obtaining a fully converged solution due to the long wavelength components. To illustrate this, the CG routine was employed to compute  $\varphi^h$  in Eq. 4.20 to an accuracy of  $10^{-5}$  a.u. for a single water molecule placed at the center of a 15 Å cubic Cartesian grid with resolution  $h = 0.074$  Å, which corresponds to  $N_{\text{grid}} = 8615125$ . Figure 4.1 shows the Euclidean norm of the residual error vector for the electrostatic potential, the expression of which is

$$\mathbf{r}^h = \boldsymbol{\rho}^h - \mathcal{L}^h \boldsymbol{\varphi}^h, \quad (4.21)$$

as a function of iteration number. The “V-cycle” and “W-cycle” methods will be addressed later, but for now consider the green points representing a CG routine that does not utilize a multigrid method. There is a rapid drop in  $\mathbf{r}^h$  during the first few iterations of the CG method, but over 250 iterations were necessary to achieve convergence. Furthermore, the inability of the CG routine to eliminate the long wavelength error components manifests as a broad and slowly decaying shoulder feature present between  $\approx 10$ -110 iterations.

Before discussing the details of multigrid methods, it is useful to revisit the discretization error  $v_{i,j,k}^h$  that was introduced in Eq. 4.18. Both the residual error vector  $\mathbf{r}^h$  and the discretization error vector  $\mathbf{v}^h$  are utilized in the multigrid algorithms, and the relationship between them is described below. Since the charge density  $\rho(\mathbf{r})$  is

known to arbitrary accuracy, it is free of discretization error upon formation of  $\rho^h$ . Supposing that Eq. 4.20 can be solved exactly for the electrostatic potential, then  $\rho^h = \mathcal{L}^h \varphi^{\text{exact}}$  and  $\varphi^h = \varphi^{\text{exact}}$ , which implies  $\mathbf{r}^h = 0$ . Using these expressions, Eq. 4.21 can be written in terms of  $\mathbf{v}^h$ , when the exact solution is not known, as follows<sup>12</sup>

$$\begin{aligned}
 \mathbf{r}^h &= \mathcal{L}^h \varphi^{\text{exact}} - \mathcal{L}^h \varphi^h \\
 &= \mathcal{L}^h \left[ \varphi^{\text{exact}} - \varphi^h \right] \\
 &= \mathcal{L}^h \mathbf{v}^h,
 \end{aligned}
 \tag{4.22}$$

where we have used Eq. 4.18 to obtain the last equality. Despite never actually acquiring  $\varphi^{\text{exact}}$  for real applications where  $h > 0$ , Eq. 4.22 provides an avenue for computing  $\mathbf{v}^h$  by solving Poisson's equation with the residual error vector replacing the charge density that appears on the right side of Eq. 4.20. Although the importance of  $\mathbf{v}^h$  may not be clear at the moment, its role in the multigrid method will be discussed shortly.

The multigrid method seeks to obviate undesired computational effort spent eliminating long wavelength error components that results in the slow convergence exhibited by the CG routine in Figure 4.1. For this discussion all grids are assumed to be cubic, although generalization to rectangular grids is straightforward. A solution to Eq. 4.20 is desired on a rectangular grid with a fine resolution, and this will be referred to as the target grid with resolution  $h$ . The central idea behind the multigrid method is to relax the iterative solution on the target grid, where the computational cost is highest, only a handful of times in order to ensure that the short wavelength error components are eliminated. Then,  $\mathbf{r}^h$  from the target grid is *restricted* to a coarsened grid with half as many grid points along each Cartesian coordinate, the

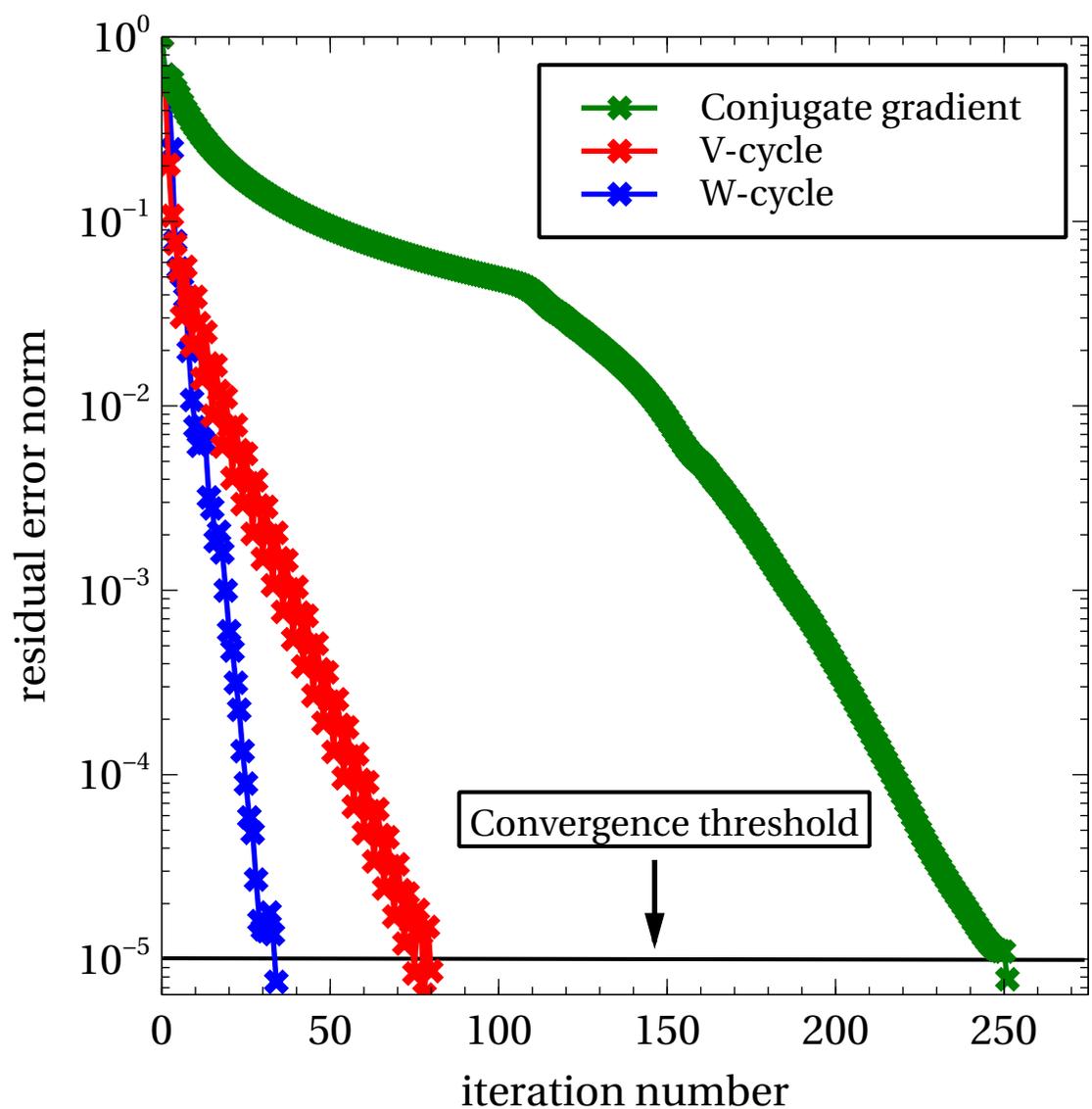


Figure 4.1: A comparison of the performance for solving Eq. 4.20 to an accuracy of  $10^{-5}$  a.u. for a single water molecule placed at the center of a cubic Cartesian grid of length  $15 \text{ \AA}$  with  $N_{\text{grid}} = 8615125$  using a standard CG routine (green), a fourth-order V-cycle multigrid (red), and a fourth-order W-cycle multigrid (blue). The slow decay of the residual error for the standard CG routine is indicative of  $\lambda > h$ , and using this method requires over 250 iterations. The two multigrid methods achieve convergence more rapidly, with the W-cycle method performing best, and both exhibit a nearly linear decrease of  $\mathbf{r}^h$  with respect to the iteration number that is in stark contrast to the standard CG routine.

resolution of which will be denoted  $H = 2h$ , to relax the residual error and eliminate the long wavelength components. For three-dimensions, reducing the number of grid points by half for each coordinate lessens  $N_{\text{grid}}$  by a factor of  $\approx 8$ , and this translates to a speedup of  $\approx 64$  for the computational time on the coarsened grid. After relaxing  $\mathbf{r}^h$  on the coarsened grid, it can either be restricted to an even coarser grid or it can be *interpolated* back to the target grid where it will be used to update the solution  $\varphi^h$ . The simplest implementation of a scheme following this procedure is called a second-order multigrid, and a schematic of it is shown in Figure 4.2 with an explanation of each step in the following paragraph. The multigrid order refers to the number of grids with different resolutions that are employed for the procedure: a second-order multigrid method uses a target grid of resolution  $h$  and one coarse grid of resolution  $H$ .

Both the CG routine and multigrid method start with Step 1 of Figure 4.2: an initial guess for  $\varphi^h$  is specified, and then  $\mathbf{r}^h$  is computed. For the CG algorithm,  $\mathbf{r}^h$  is used to correct  $\varphi^h$  on the target grid, which in turn is used to create a new vector of residuals, and this process is repeated until the desired accuracy is obtained. The resulting solution  $\varphi^h$  is considered fully relaxed on the target grid with error components reduced below a certain threshold value. For the multigrid method,  $\varphi^h$  is relaxed on the target grid a small number of times (see discussion of fourth-order methods below), which is in contrast to fully relaxing the solution, and this unrelaxed solution is used in Eq. 4.21 to compute  $\mathbf{r}^h$ . The residual error vector is then restricted from the target grid to a coarsened one resulting in  $\mathbf{r}^H$ , and Step 2 of Figure 4.2 illustrates this process. The restricted residual error  $\mathbf{r}^H$  is expressed as

$$\mathbf{r}^H = \mathbf{I}_h^H \mathbf{r}^h, \quad (4.23)$$

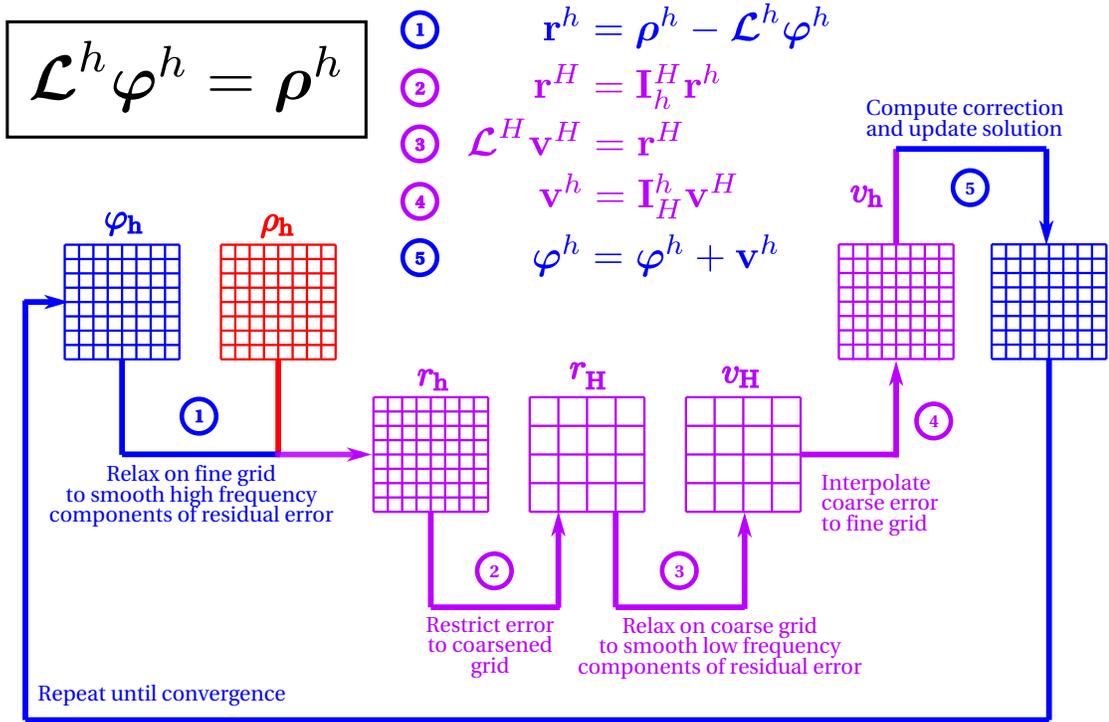


Figure 4.2: An illustration of a second-order multigrid algorithm employed to solve Eq. 4.20 is presented. The input is the known source charge density  $\rho^h$  and the desired output is the electrostatic potential  $\varphi^h$ , both of which are discretized on a target Cartesian grid of resolution  $h$ . Steps 1 and 2 show the formation of the residual error  $r_h$  on the target grid, and then its restriction to a coarsened grid. On the coarse grid, the restricted residual error  $r_H$  is utilized in a CG routine to form a relaxed residual error  $v_H$  in Step 3. The relaxed residual error is then interpolated to the finer target grid to form  $v_h$  in Step 4. In Step 5,  $v_h$  is used to construct a correction to the solution, and this process is repeated until convergence is reached.

where  $\mathbf{I}_h^H$  is the matrix representation of the restriction operator. However,  $\mathbf{I}_h^H$  is not formed in practice, and only its action on  $\mathbf{r}^h$  to generate  $\mathbf{r}^H$  is relevant:

$$\begin{aligned}
r_{I,J,K}^H &= \frac{r_{i,j,k}^h}{8} \\
&+ \frac{1}{16} \left( r_{i+1,j,k}^h + r_{i-1,j,k}^h + r_{i,j+1,k}^h + r_{i,j-1,k}^h + r_{i,j,k+1}^h + r_{i,j,k-1}^h \right) \\
&+ \frac{1}{32} \left( r_{i+1,j+1,k}^h + r_{i+1,j-1,k}^h + r_{i-1,j+1,k}^h + r_{i-1,j-1,k}^h \right) \\
&+ \frac{1}{32} \left( r_{i+1,j,k+1}^h + r_{i+1,j,k-1}^h + r_{i-1,j,k+1}^h + r_{i-1,j,k-1}^h \right) \\
&+ \frac{1}{32} \left( r_{i,j+1,k+1}^h + r_{i,j+1,k-1}^h + r_{i,j-1,k+1}^h + r_{i,j-1,k-1}^h \right) \\
&+ \frac{1}{64} \left( r_{i+1,j+1,k+1}^h + r_{i+1,j+1,k-1}^h + r_{i+1,j-1,k+1}^h + r_{i+1,j-1,k-1}^h \right) \\
&+ \frac{1}{64} \left( r_{i-1,j+1,k+1}^h + r_{i-1,j+1,k-1}^h + r_{i-1,j-1,k+1}^h + r_{i-1,j-1,k-1}^h \right) ,
\end{aligned} \tag{4.24}$$

where the notation  $\{I, J, K\}$  is introduced to denote that a different mapping scheme for the coarse grid coordinates is required. Recall from Section 4.1 that the target grid coordinates were mapped to the grid points through  $x_i = -L_x/2 + ih_x$  for  $i = 0, \dots, (N_x - 1)$  and  $h_x = L_x/(N_x - 1)$ . For the coarse grid, there are half as many grid points for each coordinate, and therefore the mapping scheme becomes  $x_I = -L_x/2 + IH_x$  for  $I = 0, \dots, (N_x - 1)/2$  and  $H_x = 2h_x$ . The target grid points  $\{i, j, k\}$  are mapped to the coarse grid points  $\{I, J, K\}$  through  $\{i, j, k\} = \{2I, 2J, 2K\}$ . Eq. 4.24 is valid for three dimensions and shows that a particular value of the restricted error on the coarse grid is an average of the values from all surrounding grid points of the target grid with a weight determined by its proximity to coarse grid point  $r_{I,J,K}^H$ .

In Step 3, after restricting the residual error vector to form  $\mathbf{r}^H$ , an analog of Eq. 4.22 is solved for the discretization error on the coarse grid  $\mathbf{v}^H$  :

$$\mathcal{L}^H \mathbf{v}^H = \mathbf{r}^H . \tag{4.25}$$

The discretization error measures the deviation between the exact solution and the finite difference approximation, and is only indirectly obtainable by solving Eq. 4.22 or 4.25. By fully relaxing  $\mathbf{v}^H$  on the coarse grid, its problematic long wavelength components are reduced, and since the input quantity  $\mathbf{r}^H$  of Eq. 4.25 has already relaxed on the target grid, the error components of  $\mathbf{v}^H$  will be eliminated on multiple length scales. Thus,  $\mathbf{v}^H$  can be used to create a better correction for  $\varphi^h$  as opposed to using  $\mathbf{r}^h$  on the target grid as in the standard CG routine. However,  $\mathbf{v}^H$  cannot be used directly to correct  $\varphi^h$  because it is the coarse grid approximation to  $\mathbf{v}^h$ . Therefore, it is necessary to interpolate  $\mathbf{v}^H$  to the target grid, the process of which is effectively the inverse of restriction, and this is demonstrated as Step 4 of Figure 4.2. Interpolating the coarse grid discretization error to form  $\mathbf{v}^h$  on the target grid is accomplished by applying the matrix representation of the interpolation operator  $\mathbf{I}_H^h$  to  $\mathbf{v}^H$ :

$$\mathbf{v}^h = \mathbf{I}_H^h \mathbf{v}^H . \quad (4.26)$$

The action of the interpolation operator that maps  $\mathbf{I}_H^h$  to the target grid is expressed with the following set of equations:

$$\begin{aligned}
v_{i,j,k}^h &= v_{I,J,K}^H \\
v_{i+1,j,k}^h &= \frac{1}{2} \left( v_{I,J,K}^H + v_{I+1,J,K}^H \right) \\
v_{i,j+1,k}^h &= \frac{1}{2} \left( v_{I,J,K}^H + v_{I,J+1,K}^H \right) \\
v_{i,j,k+1}^h &= \frac{1}{2} \left( v_{I,J,K}^H + v_{I,J,K+1}^H \right) \\
v_{i+1,j+1,k}^h &= \frac{1}{4} \left( v_{I,J,K}^H + v_{I+1,J,K}^H + v_{I,J+1,K}^H + v_{I+1,J+1,K}^H \right) \\
v_{i+1,j,k+1}^h &= \frac{1}{4} \left( v_{I,J,K}^H + v_{I+1,J,K}^H + v_{I,J,K+1}^H + v_{I+1,J,K+1}^H \right) \\
v_{i,j+1,k+1}^h &= \frac{1}{4} \left( v_{I,J,K}^H + v_{I,J+1,K}^H + v_{I,J,K+1}^H + v_{I,J+1,K+1}^H \right) \\
v_{i+1,j+1,k+1}^h &= \frac{1}{8} \left( v_{I,J,K}^H + v_{I+1,J,K}^H + v_{I,J+1,K}^H + v_{I,J,K+1}^H \right) \\
&\quad + \frac{1}{8} \left( v_{I+1,J+1,K}^H + v_{I+1,J,K+1}^H + v_{I,J+1,K+1}^H + v_{I+1,J+1,K+1}^H \right) .
\end{aligned} \tag{4.27}$$

The final step of the second-order method is to correct the target grid solution with  $\mathbf{v}^h$ :

$$\boldsymbol{\varphi}^h = \boldsymbol{\varphi}^h + \mathbf{v}^h , \tag{4.28}$$

which is shown as Step 5 of Figure 4.2. Convergence is tested by computing  $\mathbf{r}^h$  using Eq. 4.21 with the updated solution  $\boldsymbol{\varphi}^h$  from Eq. 4.28, and then comparing the value of its Euclidean norm to a chosen threshold value. If convergence is achieved,  $\boldsymbol{\varphi}^h$  is the fully relaxed solution to Poisson's equation given by Eq. 4.20; otherwise the multigrid process continues with Step 2 as described above.

The reported results in Chapter 3 from production-quality calculations employed two modifications of the method presented above. Figure 4.3 illustrates these modifications as flow diagrams for the V-cycle and W-cycle multigrid implementations of the PEQS method and introduces four parameters  $\gamma_0$ ,  $\gamma_1$ ,  $\gamma_2$ , and  $\gamma_3$  that control the

number of CG iterations spent at a particular grid level. Both the V-cycle and W-cycle methods use a fourth-order scheme that employs four Cartesian grids: a target grid possessing the finest resolution  $h$  and three levels of coarse grids with increasing resolutions  $H = 2h$ ,  $H = 4h$ , and  $H = 8h$ . The quantity  $\gamma_0$  signifies that the solution on the coarsest grid level is fully relaxed with the CG routine, and the values for the other parameters are taken from Ref [12] as  $\gamma_1 = 2$ ,  $\gamma_2 = 3$ , and  $\gamma_3 = \gamma_1 + \gamma_2$ . In doing this, the solution  $\varphi^h$  on the target grid, where the computational cost is largest, is relaxed a total of  $\gamma_1 + \gamma_2$  times for each V-cycle or W-cycle. Furthermore, a fourth-order scheme will eliminate error components on a variety of wavelengths that should, in principle, produce better corrections at each grid level and achieve convergence on the target grid more rapidly (i.e. as few iterations on the grid of finest resolution as possible).

The fourth-order V-cycle method presented in Figure 4.3 proceeds similarly to the second-order method described above. At the target grid level, the residual error  $\mathbf{r}^h$  corresponding to an unrelaxed solution  $\varphi^h$  after  $\gamma_1$  CG iterations is restricted, which is signified by a downward arrow in Figure 4.3, to the first coarse grid level with resolution  $H = 2h$ , denoted as  $\mathbf{r}^{H=2h}$ . Using  $\mathbf{r}^{H=2h}$  in Eq. 4.25,  $\mathbf{v}^{H=2h}$  is relaxed  $\gamma_1$  times and is stored in memory. This process is repeated twice more:  $\mathbf{r}^{H=2h}$  is further restricted to the second and third coarse grid levels forming  $\mathbf{r}^{H=4h}$  and  $\mathbf{r}^{H=8h}$ , and at each of those coarse grid levels, the solution is relaxed  $\gamma_1$  and  $\gamma_0$  times, respectively. The upward arrows in Figure 4.3 signify interpolation of the discretization error from a coarser grid to a finer one using Eq. 4.26 (i.e.  $\mathbf{v}^{H=8h} \rightarrow \mathbf{v}^{H=4h}$ ,  $\mathbf{v}^{H=4h} \rightarrow \mathbf{v}^{H=2h}$ , and  $\mathbf{v}^{H=2h} \rightarrow \mathbf{v}^h$ ) that is then used to update the corresponding solution at a particular grid level using Eq. 4.28. After forming the appropriate correction at each grid level,

the solution is further relaxed  $\gamma_2$  times before being interpolated to a finer grid. If the solution on the target grid level is not converged after the first V-cycle, the process is repeated until the desired level of accuracy is achieved. The W-cycle follows a similar but slightly more sophisticated procedure as illustrated in Figure 4.3, and the downward (restriction) and upward (interpolation) arrows carry the same meaning from the discussion above. Within a W-cycle, the error is restricted and interpolated between the three coarse grid levels more times to enhance error elimination before finally computing the solution on the target grid, and the process is repeated until convergence is reached. Returning to the earlier discussion of Figure 4.1, both the V-cycle and W-cycle multigrid methods were applied to compute the electrostatic potential on a large Cartesian grid for a single water molecule. The number of iterations on the target grid is reduced to  $\approx 75$  and  $\approx 35$  for the V-cycle and W-cycle methods, respectively, which is a remarkable improvement over the standard CG routine. Because of its superior performance, the fourth-order W-cycle multigrid method was employed for all VIPs reported in Chapter 3 using non-equilibrium PEQS.

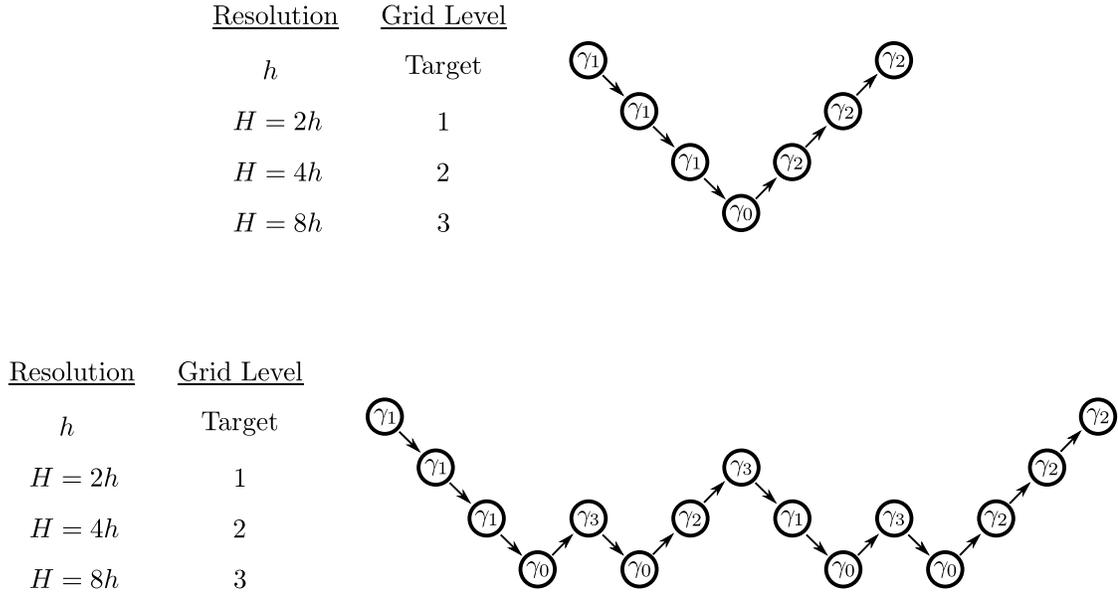


Figure 4.3: The flow diagrams for the fourth-order V-cycle (top) and W-cycle (bottom) multigrid methods are presented. Both methods employ a target grid with the finest resolution  $h$  and three coarse grids of resolutions  $H = 2h$ ,  $H = 4h$ , and  $H = 8h$ . The downward arrows represent restriction of the residual error vector from finer grids to coarser ones where the discretization error is relaxed  $\gamma_1$  times for the target and first two coarse grid levels and relaxed fully on the coarsest grid (denoted by  $\gamma_0$ ). The discretization error is then interpolated from coarser grids to back to finer ones, which is shown as an upward arrow, where it is further relaxed  $\gamma_2$  or  $\gamma_3$  times. After completing the restriction and interpolation operations, convergence of the solution on the target grid is tested, and the V- and W-cycles are repeated until the desired accuracy is reached. Figure adapted from Chapter 3 of Ref [12].

## Bibliography

- [1] B. Winter, R. Weber, I. V. Hertel, M. Faubel, P. Jungwirth, E. C. Brown, and S. E. Bradforth. Electron binding energies of aqueous alkali and halide ions: EUV photoelectron spectroscopy of liquid solutions and combined ab initio and molecular dynamics calculations. *J. Am. Chem. Soc.*, 127:7203–7214, 2005.
- [2] R. Seidel, B. Winter, and S. Bradforth. Valence electronic structure of aqueous solutions: Insights from photoelectron spectroscopy. *Annu. Rev. Phys. Chem.*, 67:283–305, 2016.
- [3] D. Luckhaus, Y. Yamamoto, T. Suzuki, and R. Signorell. Genuine binding energy of the hydrated electron. *Science Advances*, 3:1–5, 2017.
- [4] F. Uhlig, O. Marsalek, and P. Jungwirth. Electron at the surface of water: Dehydrated or not? *J. Phys. Chem. Lett.*, 4:338–343, 2013.
- [5] F. Uhlig, J. M. Herbert, M. P. Coons, and P. Jungwirth. Optical spectroscopy of the bulk and interfacial hydrated electron from ab initio calculations. *J. Phys. Chem. A*, 118:7507–7515, 2014.
- [6] K. Yagi, Y. Okano, T. Sato, Y. Kawashima, T. Tsuneda, and K. Hirao. Water cluster anions studied by long-range corrected density functional theory. *J. Phys. Chem. A*, 112:9845–9853, 2008.

- [7] L. D. Jacobson and J. M. Herbert. Polarization-bound quasi-continuum states are responsible for the ‘blue tail’ in the optical absorption spectrum of the aqueous electron. *J. Am. Chem. Soc.*, 132:10000–10002, 2010.
- [8] L. D. Jacobson and J. M. Herbert. Theoretical characterization of four distinct isomer types in hydrated-electron clusters, and proposed assignments for photoelectron spectra of water cluster anions. *J. Am. Chem. Soc.*, 133:19889–19899, 2011.
- [9] K. R. Siefertmann, Y. Liu, E. Lugovoy, O. Link, M. Faubel, U. Buck, B. Winter, and B. Abel. Binding energies, lifetimes and implications of bulk and interface solvated electrons in water. *Nat. Phys.*, 2:274–279, 2010.
- [10] G. Fisicaro, L. Genovese, O. Andreussi, N. Marzari, and S. Goedecker. A generalized Poisson and Poisson-Boltzmann solver for electrostatic environments. *J. Chem. Phys.*, 144:014103:1–12, 2016.
- [11] F. Uhlig, O. Marsalek, and P. Jungwirth. Unraveling the complex nature of the hydrated electron. *J. Phys. Chem. Lett.*, 3:3071–3075, 2012.
- [12] C. H. Venner and A. A. Lubrecht. *Multi-Level Methods in Lubrication*. Elsevier, 2000, 1st edition, 2000.
- [13] J. A. Pople. Quantum chemical models, December 1998. Nobel lecture.
- [14] A. Szabo and N. S. Ostlund. *Modern Quantum Chemistry*. Macmillan, New York, 1982.
- [15] P. Jungwirth and D. J. Tobias. Specific ion effects at the air/water interface. *Chem. Rev.*, 106:1259–1281, 2006.

- [16] M. Chaplin. Theory vs experiment: What is the surface charge of water? *Water*, 1:1–28, 2009.
- [17] B. Abel. Hydrated interfacial ions and electrons. *Annu. Rev. Phys. Chem.*, 64:533–552, 2013.
- [18] Y. Levin and A. P. dos Santos. Ions at hydrophobic interfaces. *J. Phys.: Condens. Matter*, 26:203101:1–11, 2014.
- [19] C. Bai and J. Herzfeld. Surface propensities of the self-ions of water. *ACS Cent. Sci.*, 2:225–231, 2016.
- [20] E. J. Hart and M. Anbar. *The Hydrated Electron*. Wiley-Interscience, 1970.
- [21] E. Alizadeh and L. Sanche. Precursors of solvated electrons in radiobiological physics and chemistry. *Chem. Rev.*, 112:5578–5602, 2012.
- [22] T. Ito, S. C. Baker, C. D. Stickley, J. G. Peak, and M. J. Peak. Dependence of the yield of strand breaks induced by  $\gamma$ -rays in DNA on the physical conditions of exposure: Water content and temperature. *Int. J. Radiat. Biol.*, 63:289–296, 1993.
- [23] J. M. Falcone, D. Becker, M. D. Sevilla, and S. G. Swarts. Products of the reactions of the dry and aqueous electron with hydrated DNA: Hydrogen and 5,6-dihydropyrimidines. *Radiat. Phys. Chem.*, 72:257–264, 2005.
- [24] E. Alizadeh, A. G. Sanz, G. García, and L. Sanche. Radiation damage to DNA: The indirect effect of low-energy electrons. *J. Phys. Chem. Lett.*, 4:820–825, 2013.

- [25] J. R. R. Verlet, A. E. Bragg, A. Kammrath, O. Cheshnovsky, and D. M. Neumark. Observation of large water-cluster anions with surface-bound excess electrons. *Science*, 307:93–96, 2005.
- [26] L. Turi, W.-S. Sheu, and P. J. Rossky. Characterization of excess electrons in water-cluster anions by quantum simulations. *Science*, 309:914–917, 2005.
- [27] J. R. R. Verlet, A. E. Bragg, A. Kammrath, O. Cheshnovsky, and D. M. Neumark. Comment on “Characterization of excess electrons in water-cluster anions by quantum simulations”. *Science*, 310:1769–1769, 2005.
- [28] L. Turi, W.-S. Sheu, and P. J. Rossky. Response to comment on “Characterization of excess electrons in water-cluster anions by quantum simulations”. *Science*, 310:1769–1769, 2005.
- [29] L. Ma, K. Majer, F. Chirot, and B. von Issendorff. Low temperature photoelectron spectra of water cluster anions. *J. Chem. Phys.*, 131:144303:1–6, 2009.
- [30] B. Winter and M. Faubel. Photoemission from liquid aqueous solutions. *Chem. Rev.*, 106:1176–1211, 2006.
- [31] R. Seidel, S. Thürmer, and B. Winter. Photoelectron spectroscopy meets aqueous solution: Studies from a vacuum liquid microjet. *J. Phys. Chem. Lett.*, 2:633–641, 2011.
- [32] Y. Tang, H. Shen, K. Sekiguchi, N. Kurahashi, T. Mizuno, Y. I. Suzuki, and T. Suzuki. Direct measurement of vertical binding energy of a hydrated electron. *Phys. Chem. Chem. Phys.*, 12:3653–3655, 2010.

- [33] A. T. Shreve, T. A. Yen, and D. M. Neumark. Photoelectron spectroscopy of hydrated electrons. *Chem. Phys. Lett.*, 493:216–219, 2010.
- [34] A. Lübcke, F. Buchner, N. Heine, I. V. Hertel, and T. Schultz. Time-resolved photoelectron spectroscopy of solvated electrons in aqueous NaI solution. *Phys. Chem. Chem. Phys.*, 12:14629–14634, 2010.
- [35] T. Horio, H. Shen, S. Adachi, and T. Suzuki. Photoelectron spectra of solvated electrons in bulk water, methanol, and ethanol. *Chem. Phys. Lett.*, 535:12–16, 2012.
- [36] F. Buchner, T. Schultz, and A. Lübcke. Solvated electrons at the water-air interface: Surface versus bulk signal in low kinetic energy photoelectron spectroscopy. *Phys. Chem. Chem. Phys.*, 14:5837–5842, 2012.
- [37] Y. Yamamoto, Y.-I. Suzuki, G. Tomasello, T. Horio, S. Karashima, R. Mitríc, and T. Suzuki. Time- and angle-resolved photoemission spectroscopy of hydrated electrons near a liquid water surface. *Phys. Rev. Lett.*, 112:187603:1–5, 2014.
- [38] Y. Yamamoto, S. Karashima, S. Adachi, and T. Suzuki. Wavelength dependence of UV photoemission from solvated electrons in bulk water, methanol, and ethanol. *J. Phys. Chem. A*, 120:1153–1159, 2016.
- [39] J. V. Coe, S. M. Williams, and K. H. Bowen. Photoelectron spectra of hydrated electron clusters vs. cluster size. *Int. Rev. Phys. Chem.*, 27:27–51, 2008.
- [40] L. D. Jacobson and J. M. Herbert. A one-electron model for the aqueous electron that includes many-body electron-water polarization: Bulk equilibrium

structure, vertical electron binding energy, and optical absorption spectrum. *J. Chem. Phys.*, 133:154106:1–19, 2010.

- [41] K. R. Siefermann and B. Abel. The hydrated electron: A seemingly familiar chemical and biological transient. *Angew. Chem. Int. Ed. Engl.*, 50:5264–5272, 2011.
- [42] B. Abel, U. Buck, A. L. Sobolewski, and W. Domcke. On the nature and signatures of the solvated electron in water. *Phys. Chem. Chem. Phys.*, 14:22–34, 2012.
- [43] M. Faubel, K. R. Siefermann, Y. Liu, and B. Abel. Ultrafast soft X-ray photoelectron spectroscopy at liquid water microjets. *Acc. Chem. Res.*, 45:120–130, 2012.
- [44] C.-R. Wang, J. Nguyen, and Q.-B. Lu. Bond breaks of nucleotides by dissociative electron transfer of nonequilibrium prehydrated electrons: A new molecular mechanism for reductive DNA damage. *J. Am. Chem. Soc.*, 131:11320–11322, 2009.
- [45] L. Turi and D. Borgis. Analytical investigations of an electron–water molecule pseudopotential. II. Development of a new pair potential and molecular dynamics simulations. *J. Chem. Phys.*, 117:6186–6195, 2002.
- [46] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, and J. Hermans. In B. Pullman, editor, *Intermolecular Forces*, page 331, Holland, 1981. Dordrecht.
- [47] P. Ren and J. W. Ponder. Polarizable atomic multipole water model for molecular mechanics simulation. *J. Phys. Chem. B*, 107:5933–5947, 2003.

- [48] L. D. Jacobson, C. F. Williams, and J. M. Herbert. The static-exchange electron-water pseudopotential, in conjunction with a polarizable water model: A new Hamiltonian for hydrated-electron simulations. *J. Chem. Phys.*, 130:124115:1–18, 2009.
- [49] L. Turi, Á. Madarász, and P. J. Rossky. Excess electron localization sites in neutral water clusters. *J. Chem. Phys.*, 125:014308:1–7, 2006.
- [50] D. Borgis, P. J. Rossky, and L. Turi. Quantized time correlation function approach to nonadiabatic decay rates in condensed phase: Application to solvated electrons in water and methanol. *J. Chem. Phys.*, 125:064501:1–13, 2006.
- [51] D. Borgis, P. J. Rossky, and L. Turi. Nuclear quantum effects on the nonadiabatic decay mechanism of an excited hydrated electron. *J. Chem. Phys.*, 127:174508:1–6, 2007.
- [52] Á. Madarász, P. J. Rossky, and L. Turi. Excess electron relaxation dynamics at water/air interfaces. *J. Chem. Phys.*, 126:234707:1–11, 2007.
- [53] Á. Madarász, P. J. Rossky, and L. Turi. Interior- and surface-bound excess electron states in large water cluster anions. *J. Chem. Phys.*, 130:124319:1–7, 2009.
- [54] K. A. Tay and A. Boutin. Hydrated electron diffusion: The importance of hydrogen-bond dynamics. *J. Phys. Chem. B*, 113:11943–11949, 2009.
- [55] L. D. Jacobson and J. M. Herbert. Comment on “Does the hydrated electron occupy a cavity?”. *Science*, 331:1387, 2011.

- [56] L. D. Jacobson and J. M. Herbert. A simple algorithm for determining orthogonal, self-consistent excited-state wave functions for a state-specific hamiltonian: Application to the optical spectrum of the aqueous electron. *J. Chem. Theory Comput.*, 7:2085–2093, 2011.
- [57] J. M. Herbert and L. D. Jacobson. Structure of the aqueous electron: Assessment of one-electron pseudopotential models in comparison to experimental data and time-dependent density functional theory. *J. Phys. Chem. A*, 115:14470–14483, 2011.
- [58] W. J. Glover, J. R. Casey, and B. J. Schwartz. Free energies of quantum particles: The coupled-perturbed quantum umbrella sampling method. *J. Chem. Theory Comput.*, 10:4661–4671, 2014.
- [59] J. R. Casey, B. J. Schwartz, and W. J. Glover. Free energies of cavity and non-cavity hydrated electrons near the instantaneous air/water interface. *J. Phys. Chem. Lett.*, 7:3192–3198, 2016.
- [60] E. R. Smith. Ewald summation for systems with slab geometry. *Proc. R. Soc. London A*, 375:475–505, 1981.
- [61] I.-C. Yeh and M. L. Berkowitz. Ewald summation for systems with slab geometry. *J. Chem. Phys.*, 111:3155–3162, 1999.
- [62] J. Tomasi, B. Mennucci, and R. Cammi. Quantum mechanical continuum solvation models. *Chem. Rev.*, 105:2999–3093, 2005.

- [63] J. M. Herbert and A. W. Lange. Polarizable continuum models for (bio)molecular electrostatics: Basic theory and recent developments for macromolecules and simulations. In Q. Cui, P. Ren, and M. Meuwly, editors, *Many-Body Effects and Electrostatics in Biomolecules*, chapter 11, pages 363–416. Pan Stanford, 2016.
- [64] O. Andreussi, I. Dabo, and N. Marzari. Revised self-consistent continuum solvation in electronic-structure calculations. *J. Chem. Phys.*, 136:064102:1–20, 2012.
- [65] M. P. Coons, Z.-Q. You, and J. M. Herbert. The hydrated electron at the surface of neat liquid water appears to be indistinguishable from the bulk species. *J. Am. Chem. Soc.*, 138:10879–10886, 2016.
- [66] J.-M. Mewes, Z.-Q. You, M. Wormit, T. Kriesche, J. M. Herbert, and A. Dreuw. Experimental benchmark data and systematic evaluation of two *a posteriori*, polarizable-continuum corrections for vertical excitation energies in solution. *J. Phys. Chem. A*, 119:5446–5464, 2015.
- [67] Z.-Q. You, J.-M. Mewes, A. Dreuw, and J. M. Herbert. Comparison of the Marcus and Pekar partitions in the context of non-equilibrium, polarizable-continuum reaction-field solvation models. *J. Chem. Phys.*, 143:234107:1–14, 2015.
- [68] A. W. Lange and J. M. Herbert. Symmetric versus asymmetric discretization of the integral equations in polarizable continuum solvation models. *Chem. Phys. Lett.*, 509:77–87, 2011.

- [69] D. M. Chipman. Simulation of volume polarization in reaction field theory. *J. Chem. Phys.*, 110:8012–8018, 1999.
- [70] D. M. Chipman. Reaction field treatment of charge penetration. *J. Chem. Phys.*, 112:5558–5565, 2000.
- [71] J. Tomasi and M. Persico. Molecular interactions in solution: An overview of methods based on continuous distributions of the solvent. *Chem. Rev.*, 94:2027–2094, 1994.
- [72] R. Cammi and J. Tomasi. Nonequilibrium solvation theory for the polarizable continuum model: A new formulation at the SCF level with application to the case of the frequency-dependent linear electric response function. *Int. J. Quantum Chem. Symp.*, 29:465–474, 1995.
- [73] M. Cossi and V. Barone. Separation between fast and slow polarizations in continuum solvation models. *J. Phys. Chem. A*, 104:10614–10622, 2000.
- [74] R. Improta, V. Barone, G. Scalmani, and M. J. Frisch. A state-specific polarizable continuum model time dependent density functional method for excited state calculations in solution. *J. Chem. Phys.*, 125:054103:1–9, 2006.
- [75] Y. Shao, Z. Gan, E. Epifanovsky, A. T. B. Gilbert, M. Wormit, J. Kussmann, A. W. Lange, A. Behn, J. Deng, X. Feng, D. Ghosh, M. Goldey, P. R. Horn, L. D. Jacobson, I. Kaliman, R. Z. Khaliullin, T. Kúš, A. Landau, J. Liu, E. I. Proynov, Y. M. Rhee, R. M. Richard, M. A. Rohrdanz, R. P. Steele, E. J. Sundstrom, H. L. Woodcock III, P. M. Zimmerman, D. Zuev, B. Albrecht, E. Alguire, B. Austin, G. J. O. Beran, Y. A. Bernard, E. Berquist,

K. Brandhorst, K. B. Bravaya, S. T. Brown, D. Casanova, C.-M. Chang, Y. Chen, S. H. Chien, K. D. Closser, D. L. Crittenden, M. Diedenhofen, R. A. DiStasio Jr., H. Do, A. D. Dutoi, R. G. Edgar, S. Fatehi, L. Fusti-Molnar, A. Ghysels, A. Golubeva-Zadorozhnaya, J. Gomes, M. W. D. Hanson-Heine, P. H. P. Harbach, A. W. Hauser, E. G. Hohenstein, Z. C. Holden, T.-C. Jagau, H. Ji, B. Kaduk, K. Khistyayev, J. Kim, J. Kim, R. A. King, P. Klunzinger, D. Kosenkov, T. Kowalczyk, C. M. Krauter, K. U. Lao, A. Laurent, K. V. Lawler, S. V. Levchenko, C. Y. Lin, F. Liu, E. Livshits, R. C. Lochan, A. Luenser, P. Manohar, S. F. Manzer, S.-P. Mao, N. Mardirossian, A. V. Marenich, S. A. Maurer, N. J. Mayhall, C. M. Oana, R. Olivares-Amaya, D. P. O'Neill, J. A. Parkhill, T. M. Perrine, R. Peverati, P. A. Pieniazek, A. Prociuk, D. R. Rehn, E. Rosta, N. J. Russ, N. Sergueev, S. M. Sharada, S. Sharma, D. W. Small, A. Sodt, T. Stein, D. Stück, Y.-C. Su, A. J. W. Thom, T. Tsuchimochi, L. Vogt, O. Vydrov, T. Wang, M. A. Watson, J. Wenzel, A. White, C. F. Williams, V. Vanovschi, S. Yeganeh, S. R. Yost, Z.-Q. You, I. Y. Zhang, X. Zhang, Y. Zhao, B. R. Brooks, G. K. L. Chan, D. M. Chipman, C. J. Cramer, W. A. Goddard III, M. S. Gordon, W. J. Hehre, A. Klamt, H. F. Schaefer III, M. W. Schmidt, C. D. Sherrill, D. G. Truhlar, A. Warshel, X. Xu, A. Aspuru-Guzik, R. Baer, A. T. Bell, N. A. Besley, J.-D. Chai, A. Dreuw, B. D. Dunietz, T. R. Furlani, S. R. Gwaltney, C.-P. Hsu, Y. Jung, J. Kong, D. S. Lambrecht, W. Liang, C. Ochsenfeld, V. A. Rassolov, L. V. Slipchenko, J. E. Subotnik, T. Van Voorhis, J. M. Herbert, A. I. Krylov, P. M. W. Gill, and M. Head-Gordon. Advances in molecular quantum chemistry contained in the q-chem 4 program package. *Mol. Phys.*, 113:184–215, 2015.

- [76] J. Wang and R. Luo. Assessment of linear finite-difference Poisson–Boltzmann solvers. *J. Comput. Chem.*, 31:1689–1698, 2010.
- [77] R. A. Marcus. On the theory of oxidation-reduction reactions involving electron transfer. I. *J. Chem. Phys.*, 24:966–978, 1956.
- [78] R. A. Marcus. Electrostatic free energy and other properties of states having nonequilibrium polarization. I. *J. Chem. Phys.*, 24:979–989, 1956.
- [79] J. E. Brady and P. W. Carr. An analysis of dielectric models of solvatochromism. *J. Phys. Chem.*, 89:5759–5766, 1985.
- [80] A. Klamt. Calculation of UV/Vis spectra in solution. *J. Phys. Chem.*, 100:3349–3353, 1996.
- [81] M. A. Aguilar, F. J. Olivares del Valle, and J. Tomasi. Nonequilibrium solvation: An *ab initio* quantum-mechanical method in the continuum cavity model approximation. *J. Chem. Phys.*, 98:7375–7384, 1993.
- [82] A. D. Becke. Density-functional exchange-energy approximation with correct asymptotic behavior. *Phys. Rev. A*, 38:3098–3100, 1988.
- [83] T. Tsuneda, T. Suzumura, and K. Hirao. A new one-parameter progressive Colle–Salvetti-type correlation functional. *J. Chem. Phys.*, 110:10664–10678, 2008.
- [84] C. Lee, W. Yang, and R. G. Parr. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B*, 37:785–789, 1988.

- [85] M. A. Rohrdanz and J. M. Herbert. Simultaneous benchmarking of ground- and excited-state properties with long-range-corrected density functional theory. *J. Chem. Phys.*, 129:034107:1–9, 2008.
- [86] M. A. Rohrdanz, K. M. Martins, and J. M. Herbert. A long-range-corrected density functional that performs well for both ground-state properties and time-dependent density functional theory excitation energies, including charge-transfer excited states. *J. Chem. Phys.*, 130:054112:1–8, 2009.
- [87] R. M. Richard and J. M. Herbert. Time-dependent density-functional description of the  $^1L_a$  state in polycyclic aromatic hydrocarbons: Charge-transfer character in disguise? *J. Chem. Theory Comput.*, 7:1296–1306, 2011.
- [88] R. Baer, E. Livshits, and U. Salzner. Tuned range-separated hybrids in density functional theory. *Annu. Rev. Phys. Chem.*, 61:85–109, 2010.
- [89] J. M. Herbert. The quantum chemistry of loosely bound electrons. In A. L. Parill and K. Lipkowitz, editors, *Reviews in Computational Chemistry*, volume 28, chapter 8, pages 391–517. Wiley-VCH, 2015.
- [90] B. M. Wong and T. H. Hsieh. Optoelectronic and excitonic properties of oligoacenes: Substantial improvements from range-separated time-dependent density functional theory. *J. Chem. Theory Comput.*, 6:3704–3712, 2010.
- [91] K. Garrett, X. A. S. Vazquez, S. B. Egri, J. Wilmer, L. E. Johnson, B. H. Robinson, and C. M. Isborn. Optimum exchange for calculation of excitation energies and hyperpolarizabilities of organic electro-optic chromophores. *J. Chem. Theory Comput.*, 10:3821–3831, 2014.

- [92] M. Feyereisen, G. Fitzgerald, and A. Komornicki. Use of approximate integrals in *ab initio* theory. An application in MP2 energy calculations. *Chem. Phys. Lett.*, 208:359–363, 1993.
- [93] Y. Jung, A. Sodt, P. M. W. Gill, and M. Head-Gordon. Auxiliary basis expansions for large-scale electronic structure calculations. *Proc. Natl. Acad. Sci. USA*, 102:6692–6697, 2005.
- [94] J. M. Herbert and M. Head-Gordon. Calculation of electron detachment energies for water cluster anions: An appraisal of electronic structure methods, with application to  $(\text{H}_2\text{O})_{20}^-$  and  $(\text{H}_2\text{O})_{24}^-$ . *J. Phys. Chem. A*, 109:5217–5229, 2005.
- [95] J. M. Herbert and M. Head-Gordon. Accuracy and limitations of second-order many-body perturbation theory for predicting vertical detachment energies of solvated-electron clusters. *Phys. Chem. Chem. Phys.*, 8:68–78, 2006.
- [96] C. F. Williams and J. M. Herbert. Influence of structure on electron correlation effects and electron–water dispersion interactions in anionic water clusters. *J. Phys. Chem. A*, 112:6171–6178, 2008.
- [97] E. Cancès and B. Mennucci. Comment on “Reaction field treatment of charge penetration” [J. Chem. Phys. 112, 5558 (2000)]. *J. Chem. Phys.*, 114:4744–4745, 2001.
- [98] E. Cancès and B. Mennucci. The escaped charge problem in solvation continuum models. *J. Chem. Phys.*, 115:6130–6135, 2001.
- [99] D. M. Chipman. Comparison of solvent reaction field representations. *Theor. Chem. Acc.*, 107:80–89, 2002.

- [100] C. J. Cramer and D. G. Truhlar. A universal approach to solvation modeling. *Acc. Chem. Res.*, 41:760–768, 2008.
- [101] A. Klamt, B. Mennucci, J. Tomasi, V. Barone, C. Curutchet, M. Orozco, and F. J. Luque. On the performance of continuum solvation methods. A comment on “Universal approaches to solvation modeling”. *Acc. Chem. Res.*, 42:489–492, 2009.
- [102] A. Pomogaeva and D. M. Chipman. Hydration energy from a composite method for implicit representation of the solvent. *J. Chem. Theory Comput.*, 10:211–219, 2014.
- [103] Y. Zhang and Z. Xu. Atomic radii of noble gas elements in condensed phases. *Amer. Miner.*, 80:670–675, 1995.
- [104] A. W. Lange and J. M. Herbert. A smooth, non-singular, and faithful discretization scheme for polarizable continuum models: The switching/Gaussian approach. *J. Chem. Phys.*, 133:244111:1–18, 2010.
- [105] M. E. Davis and J. A. McCammon. Dielectric boundary smoothing in finite difference solutions of the Poisson equation: An approach to improve accuracy and convergence. *J. Comput. Chem.*, 12:909–912, 1991.
- [106] J. A. Grant, B. T. Pickup, and A. Nicholls. A smooth permittivity function for Poisson–Boltzmann solvation methods. *J. Comput. Chem.*, 22:608–640, 2001.
- [107] T. Sommerfeld, A. DeFusco, and K. D. Jordan. Model potential approaches for describing the interaction of excess electrons with water clusters: Incorporation of long-range correlation effects. *J. Phys. Chem. A*, 112:11021–11035, 2008.

- [108] J. M. Herbert and L. D. Jacobson. Nature's most squishy ion: The important role of solvent polarization in the description of the hydrated electron. *Int. Rev. Phys. Chem.*, 30:1–48, 2011.
- [109] L. Turi and P. J. Rossky. Theoretical studies of spectroscopy and dynamics of hydrated electrons. *Chem. Rev.*, 112:5641–5674, 2012.
- [110] I. Carmichael. Spectral moments of solvated electrons. *Chem. Phys. Lett.*, 56:339–342, 1978.
- [111] J. V. Coe. Fundamental properties of bulk water from cluster ion data. *Int. Rev. Phys. Chem.*, 20:33–58, 2001.
- [112] U. Bovensiepen, C. Gahl, J. Stähler, M. Bockstedte, M. Meyer, F. Baletto, S. Scandolo, X.-Y. Zhu, A. Rubio, and M. Wolf. A dynamic landscape from femtoseconds to minutes for excess electrons at ice–metal interfaces. *J. Phys. Chem. C*, 113:979–988, 2009.
- [113] J. Stähler, J.-C. Deinert, D. Wegkamp, S. Hagen, and M. Wolf. Real-time measurement of the vertical binding energy during the birth of a solvated electron. *J. Am. Chem. Soc.*, 137:3520–3524, 2015.
- [114] C. Silva, P. K. Walhout, K. Yokoyama, and P. F. Barbara. Femtosecond solvation dynamics of the hydrated electron. *Phys. Rev. Lett.*, 80:1086–1089, 1998.
- [115] K. Yokoyama, C. Silva, D. H. Son, P. K. Walhout, and P. F. Barbara. Detailed investigation of the femtosecond pump–probe spectroscopy of the hydrated electron. *J. Phys. Chem. A*, 102:6957–6966, 1998.

- [116] J. Savolainen, F. Uhlig, S. Ahmed, P. Hamm, and P. Jungwirth. Direct observation of the collapse of the delocalized excess electron in water. *Nat. Phys.*, 6:697–701, 2014.
- [117] P. Stampfli. Theory for the electron affinity of clusters of rare gas atoms and polar molecules. *Phys. Rep.*, 255:1–77, 1995.
- [118] V. K. Voora, J. Ding, T. Sommerfeld, and K. D. Jordan. A self-consistent polarization potential model for describing excess electrons interacting with water clusters. *J. Phys. Chem. B*, 117:4365–4370, 2013.
- [119] L. Turi. On the applicability of one- and many-electron quantum chemistry models for hydrated electron clusters. *J. Chem. Phys.*, 144:154311:1–9, 2016.
- [120] A. H. C. West, B. L. Yoder, D. Luckhaus, C.-M. Saak, M. Doppelbauer, and R. Signorell. Angle-resolved photoemission of solvated electrons in sodium-doped clusters. *J. Phys. Chem. Lett.*, 6:1487–1492, 2015.
- [121] F.-Y. Jou and G. R. Freeman. Shapes of optical spectra of solvated electrons. Effect of pressure. *J. Phys. Chem.*, 81:909–915, 1977.
- [122] F.-Y. Jou and G. R. Freeman. Temperature and isotope effects on the shape of the optical absorption spectrum of solvated electrons in water. *J. Phys. Chem.*, 83:2383–2387, 1979.
- [123] T. R. Tuttle, Jr. and S. Golden. Resolution of optical spectra of solvated electrons in water. *J. Phys. Chem.*, 84:2457–2458, 1980.

- [124] T. R. Tuttle, Jr. and S. Golden. Shape stability of solvated-electron optical absorption bands. Part 1.—Experimental basis. *J. Chem. Soc. Faraday Trans. 2*, 77:873–888, 1981.
- [125] T. R. Tuttle, Jr., S. Golden, S. Lwenje, and C. M. Stupak. The shape as a characteristic property of solvated electron optical absorption bands. *J. Phys. Chem.*, 88:3811–3818, 1984.
- [126] D. M. Sagar, C. D. Bain, and J. R. R. Verlet. Hydrated electrons at the water/air interface. *J. Am. Chem. Soc.*, 132:6917–6919, 2010.
- [127] K. Matsuzaki, R. Kusaka, S. Nihonyanagi, S. Yamaguchi, T. Nagata, and T. Tahara. Partially hydrated electrons at the air/water interface observed by UV-excited time-resolved heterodyne-detected vibrational sum frequency generation spectroscopy. *J. Am. Chem. Soc.*, 138:7551–7557, 2016.
- [128] R. Signorell, M. Goldmann, B. L. Yoder, B. Andras, E. Chasovskikh, L. Lukas, and D. Luckhaus. Nanofocusing, shadowing, and electron mean free path in the photoemission from aerosol droplets. *Chem. Phys. Lett.*, 658:1–6, 2016.
- [129] A. P. Willard and D. C. Chandler. Instantaneous liquid interfaces. *J. Phys. Chem. B*, 114:1954–1958, 2010.
- [130] J. M. Herbert and M. Head-Gordon. First-principles, quantum-mechanical simulations of electron solvation by a water cluster. *Proc. Natl. Acad. Sci. USA*, 103:14282–14287, 2006.
- [131] J. Gu, Y. Xie, and H. F. Schaefer III. Electron attachment to nucleotides in aqueous solution. *ChemPhysChem*, 7:1885–1887, 2006.

- [132] J. Gu, Y. Xie, and H. F. Schaefer III. Near 0 eV electrons attach to nucleotides. *J. Am. Chem. Soc.*, 128:1250–1252, 2006.
- [133] A. Hübsch, R. G. Endres, D. L. Cox, and R. R. P. Singh. Optical conductivity of wet dna. *Phys. Rev. Lett.*, 94(17):178102–1:4, 2005.
- [134] M. Smyth and J. Kohanoff. Excess electron localization in solvated DNA bases. *Phys. Rev. Lett.*, 106:238108:1–4, 2011.
- [135] A. E. Bragg, J. R. R. Verlet, A. Kammrath, O. Cheshnovsky, and D. M. Neumark. Hydrated electron dynamics: From clusters to bulk. *Science*, 306:669–671, 2004.
- [136] A. E. Bragg, J. R. R. Verlet, A. Kammrath, O. Cheshnovsky, and D. M. Neumark. Electronic relaxation dynamics of water cluster anions. *J. Am. Chem. Soc.*, 127:15283–15295, 2005.
- [137] M. Born. Volumen und hydrationswärme der ionen. *Z. Phys.*, 1:45–48, 1920.
- [138] W. C. Still, A. Tempczyk, R. C. Hawley, and T. Hendrickson. Semianalytical treatment of solvation for molecular mechanics and dynamics. *J. Am. Chem. Soc.*, 112:6127–6129, 1990.
- [139] M. Schaefer and M. Karplus. A comprehensive analytical treatment of continuum electrostatics. *J. Phys. Chem.*, 100:1578–1599, 1996.
- [140] M. S. Lee, F. R. Salsbury, Jr., and C. L. Brooks III. Novel generalized Born methods. *J. Chem. Phys.*, 116:10606–10614, 2002.

- [141] D. Qiu, P. S. Shenkin, F. P. Hollinger, and W. C. Still. The GB/SA continuum model for solvation. A fast analytical method for the calculation of approximate Born radii. *J. Phys. Chem. A*, 101:3005–3014, 1997.
- [142] L. Onsager. Electric moments of molecules in liquids. *J. Am. Chem. Soc.*, 58:1486–1493, 1936.
- [143] J. G. Kirkwood. Theory of solutions of molecules containing widely separated charges with special application to zwitterions. *J. Chem. Phys.*, 2:351–361, 1934.
- [144] C. J. F. Böttcher and P. Bordewijk. *Theory of Electric Polarization*, volume II. Elsevier, New York, 2nd edition, 1973.
- [145] D. Rinaldi, A. Bouchy, J.-L. Rivail, and V. Dillet. A self-consistent reaction field model of solvation using distributed multipoles. i. energy and energy derivatives. *J. Chem. Phys.*, 120:2343–2350, 2004.
- [146] M. W. Wong, M. J. Frisch, and K. B. Wiberg. Solvent effects. 1. The mediation of electrostatic effects by solvents. *J. Am. Chem. Soc.*, 113:4776–4782, 1991.
- [147] M. W. Wong, K. B. Wiberg, and M. J. Frisch. Solvent effects. 2. Medium effect on the structure, energy, charge density, and vibrational frequencies of sulfamic acid. *J. Am. Chem. Soc.*, 114:523–529, 1992.
- [148] B. Mennucci. Polarizable continuum model. *WIREs Comput. Mol. Sci.*, 2:386–404, 2012.

- [149] E. Silla, F. Villar, O. Nilsson, J. L. Pascual-Ahuir, and O. Tapia. Molecular volumes and surfaces of biomacromolecules via GEOPOL: a fast and efficient algorithm. *J. Mol. Graph.*, 8:168–172, 1990.
- [150] E. Silla, I. Tuñòn, and J. L. Pascual-Ahuir. Geopol—an improved description of molecular-surfaces. 2. Computing the molecular area and volume. *J. Comput. Chem.*, 12:1077–1088, 1991.
- [151] J. L. Pascual-Ahuir, E. Silla, and I. Tuñòn. Geopol—an improved description of molecular-surfaces. 3. A new algorithm for the computation of a solvent-excluding surface. *J. Comput. Chem.*, 15:1127–1138, 1994.
- [152] E. Cancés, B. Mennucci, and J. Tomasi. A new integral equation formalism for the polarizable continuum model: Theoretical background and applications to isotropic and anisotropic dielectrics. *J. Chem. Phys.*, 107:3032–3041, 1997.
- [153] B. Mennucci, E. Cancés, and J. Tomasi. Evaluation of solvent effects in isotropic and anisotropic dielectrics and in ionic solutions with a unified integral equation method: Theoretical bases, computational implementation, and numerical applications. *J. Phys. Chem. B*, 101:10506–10517, 1997.
- [154] E. Cancès and B. Mennucci. New applications of integral equations methods for solvation continuum models: Ionic solutions and liquid crystals. *J. Math. Chem.*, 23:309–326, 1998.
- [155] S. Miertuš, E. Scrocco, and J. Tomasi. Electrostatic interaction of a solute with a continuum. a direct utilization of ab initio molecular potentials for the prevision of solvent effects. *Chem. Phys.*, 55:117–129, 1981.

- [156] M. Cossi, G. Scalmani, N. Rega, and V. Barone. New developments in the polarizable continuum model for quantum mechanical and classical calculations on molecules in solution. *J. Chem. Phys.*, 117:43–54, 2002.
- [157] S. I. Pekar. Research in electron theory of crystals. Technical Report AEC-tr-5575, U.S. Atomic Energy Commission, Division of Technical Information, 1963.
- [158] M. Cossi and V. Barone. Solvent effect on vertical electronic transitions by the polarizable continuum model. *J. Chem. Phys.*, 112:2427–2435, 2000.
- [159] M. A. Aguilar. Separation of the electric polarization into fast and slow components: A comparison of two partition schemes. *J. Phys. Chem. A*, 105:10393–10396, 2001.
- [160] D. M. Chipman. Vertical electronic excitation with a dielectric continuum model of solvation including volume polarization. I. Theory. *J. Chem. Phys.*, 131:014103:1–10, 2009.
- [161] J. W. Ponder. Tinker, v. 8.1.
- [162] D. J. Feller. The role of databases in support of computational chemistry calculations. *J. Chem. Phys.*, 17:1571–1586, 1996.
- [163] A. T. B. Gilbert, N. A. Besley, and P. M. W. Gill. Self-consistent field calculations of excited states using the maximum overlap method (MOM). *J. Phys. Chem. A*, 112:13164–13171, 2008.