MULTI-LAYER METHODS FOR QUANTUM CHEMISTRY IN THE CONDENSED PHASE: COMBINING DENSITY FUNCTIONAL THEORY, MOLECULAR MECHANICS, AND CONTINUUM SOLVATION MODELS

DISSERTATION

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

By

Adrian William George Lange, B.S.

Graduate Program In Chemistry

The Ohio State University

2012

Dissertation Committee:

John M. Herbert, Advisor
Terry A. Miller
Sherwin J. Singer
Abstract

We discuss the development and application of a number of theoretical physical models focused on improving our understanding of quantum chemical phenomena in condensed phase environments, especially aqueous solutions. The large number of atoms and molecules present in such systems precludes the application of the most advanced and accurate quantum chemistry theories available due to their exponential growth of required computational power with respect to the number of electrons in a system. As a feasible alternative, we opt to take a “multi-layer” approach, wherein the full chemical system is partitioned into different layers treated with varying levels of approximation, circumventing the exponential scaling computational cost. How this partitioning is performed and applied appropriately is the principal emphasis of this work.

Our main chemical system of interest is aqueous DNA and its excited electronic states. We examine applications of mixing quantum mechanics and classical molecular mechanics models, a multi-layer approach known as “QM/MM,” to simulate the electronic absorption spectrum of aqueous uracil as computed with Time-Dependent Density Functional Theory (TDDFT). We encounter a major issue of spurious charge-transfer (CT) states in TDDFT even at small uracil–water clusters. Applying Long-Range Corrected TDDFT (LRC-TDDFT), however, alleviates this issue and allows
us to investigate the absorption spectrum of aqueous DNA systems of up to as much as 8 nucleobases, providing some important clues to the initial dynamics of aqueous DNA excited by ultraviolet light and its possible ensuing damage. Then, to overcome certain computational limitations in modeling solvent by QM/MM alone, we turn to the methodology of polarizable continuum models (PCMs), which can be added on top of the QM/MM multi-layer approach as an “implicit” solvent model (in the sense that the average solvent charge density is approximated as a dielectric medium). We find that several extant PCM techniques are prone to numerical instabilities and discontinuous potential energy surfaces, and we propose ways to overcome such. Furthermore, we develop insights into the theory of PCM that yield an entirely new PCM for modeling the electrostatic effects of salty solutions. The culmination of our efforts is a cutting-edge QM/MM/PCM multi-layer approach for modeling quantum chemistry in the condensed phase.
To Samantha, my loving wife and raison d’être.
Acknowledgments

First and foremost, I gratefully acknowledge the marvelous love and gracious support of my parents, Werner and Roxanne, my siblings, Tishana, Alexander, Andreas, and Leanora, and my wife, Samantha, which have allowed me to pursue my education and achieve this, the highest academic degree in chemistry.

I sincerely thank Dr. John M. Herbert, my advisor, for accepting me—a naïve student who had just decided to depart from molecular biology but had no knowledge of physical chemistry—into his lab and from whom I have learned more about science than any other single source. I especially thank my former lab mate, Dr. Leif D. Jacobson (a.k.a. “Chet Glowers”), for being my amazing friend as well as role model for a graduate student and also for teaching me and/or clarifying many confusing details of physical chemistry and code writing. I also would like to recognize and thank the former Herbert group members, Dr. Christopher F. Williams and Dr. Mary A. Rohrdanz, for sharing their expertise and encouragement while our paths crossed.

In addition, several professors have also helped my education to grow and succeed. While I cannot hope to list them all, I would like to name four of the most outstanding, who have helped me in various special capacities: Dr. Irina Artsimovitch, Dr. Terry A. Miller, Dr. James V. Coe, and Dr. Sherwin Singer.
Finally, I express my deep gratitude to the Ohio State University for fostering both my undergraduate and graduate school education as well as for its financial support along the way, particularly for granting me the Ohio State Presidential Fellowship in my final year of graduate school.
Vita

1985 ........................................... Born, Warren, Ohio

2003 ........................................... Newton Falls High School

2007 ........................................... B.S. Chemistry, Minor in Microbiology, The Ohio State University

2007–2009 ................................. Graduate Teaching Associate, The Ohio State University

2009–2010 ................................. Graduate Research Associate, The Ohio State University

2011–present .............................. Presidential Fellow, The Ohio State University

Publications


(3) A. W. Lange and J. M. Herbert. Both Intra- and Interstrand Charge-Transfer Excited States in Aqueous B-DNA Are Present at Energies Comparable to or Just Above the $^{1}\pi\pi^{*}$ Excitonic Bright States. J. Am. Chem. Soc. 131, 3913-3922 (2009).


FIELDS OF STUDY

Major Field: Chemistry
Theoretical Physical Chemistry
# Table of Contents

ABSTRACT ........................................................................ ii
DEDICATION ..................................................................... iv
ACKNOWLEDGMENTS ....................................................... v
VITA ................................................................................ vii
LIST OF FIGURES ............................................................ xv
LIST OF TABLES .............................................................. xxvi

CHAPTER PAGE

1 Introduction ................................................................. 1

2 Simple methods to reduce charge-transfer contamination in time-dependent
density-functional calculations of clusters and liquids .................... 9

2.1 Introduction ............................................................ 9
2.2 Computational details ................................................. 11
2.3 Results and discussion ................................................ 14
  2.3.1 CT contamination in uracil–water clusters .................. 14
  2.3.2 QM/MM simulations of aqueous uracil ....................... 23
  2.3.3 Electronic absorption spectra .................................... 29
  2.3.4 Truncation of the TD-DFT excitation space ................ 35
2.4 Summary and conclusions ............................................ 42
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>Charge-transfer excited states in a $\pi$-stacked adenine dimer, as predicted using long-range-corrected time-dependent density functional theory</td>
<td>45</td>
</tr>
<tr>
<td>3.1</td>
<td>Introduction</td>
<td>45</td>
</tr>
<tr>
<td>3.2</td>
<td>Theory and methodology</td>
<td>48</td>
</tr>
<tr>
<td>3.3</td>
<td>Results</td>
<td>51</td>
</tr>
<tr>
<td>3.4</td>
<td>Discussion</td>
<td>57</td>
</tr>
<tr>
<td>4</td>
<td>Both intra- and interstrand charge-transfer excited states in aqueous B-DNA are present at energies comparable to, or just above, the $^1\pi\pi^*$ excitonic bright states</td>
<td>60</td>
</tr>
<tr>
<td>4.1</td>
<td>Introduction</td>
<td>60</td>
</tr>
<tr>
<td>4.2</td>
<td>Computational details</td>
<td>63</td>
</tr>
<tr>
<td>4.3</td>
<td>Assessment of TD-DFT</td>
<td>68</td>
</tr>
<tr>
<td>4.3.1</td>
<td>Failure of standard functionals</td>
<td>68</td>
</tr>
<tr>
<td>4.3.2</td>
<td>TD-LRC-DFT</td>
<td>76</td>
</tr>
<tr>
<td>4.4</td>
<td>Base stacking and base pairing effects</td>
<td>77</td>
</tr>
<tr>
<td>4.4.1</td>
<td>Absorption spectra</td>
<td>77</td>
</tr>
<tr>
<td>4.4.2</td>
<td>Charge-transfer states</td>
<td>81</td>
</tr>
<tr>
<td>4.4.3</td>
<td>Alternating base sequences</td>
<td>81</td>
</tr>
<tr>
<td>4.4.4</td>
<td>CIS(D) results for $A_2:T_2$</td>
<td>85</td>
</tr>
<tr>
<td>4.5</td>
<td>Multimers in aqueous solution</td>
<td>87</td>
</tr>
<tr>
<td>4.5.1</td>
<td>TD-LRC-DFT results for $A:T$ and $A_2$</td>
<td>87</td>
</tr>
<tr>
<td>4.5.2</td>
<td>CIS(D) results for $A:T$ and $A_2$</td>
<td>91</td>
</tr>
<tr>
<td>4.5.3</td>
<td>TD-LRC-DFT results for $(ApA)^-$</td>
<td>91</td>
</tr>
<tr>
<td>4.6</td>
<td>Conclusions</td>
<td>94</td>
</tr>
<tr>
<td>5</td>
<td>Polarizable continuum reaction-field solvation models affording smooth potential energy surfaces</td>
<td>97</td>
</tr>
<tr>
<td>6</td>
<td>A smooth, non-singular, and faithful discretization scheme for polarizable continuum models: The switching/Gaussian approach</td>
<td>114</td>
</tr>
<tr>
<td>6.1</td>
<td>Introduction</td>
<td>114</td>
</tr>
<tr>
<td>6.2</td>
<td>Reaction-field theory</td>
<td>118</td>
</tr>
<tr>
<td>6.2.1</td>
<td>Electrostatic interactions</td>
<td>119</td>
</tr>
<tr>
<td>6.2.2</td>
<td>Variational conditions</td>
<td>122</td>
</tr>
</tbody>
</table>
7.5.3 Dependence of the surface charge on the electrostatic potential 210
7.5.4 Definition of $D_{ii}$ ........................................... 213
7.5.5 Conductor limit .............................................. 216
7.6 Conclusions ..................................................... 219

8 A simple polarizable continuum solvation model for electrolyte solutions 221

8.1 Introduction .................................................... 221
8.2 Theory .......................................................... 226
  8.2.1 The Debye-Hückel model problem .......................... 226
  8.2.2 Boundary conditions and integral operators ............... 229
  8.2.3 Conductor-like screening model ............................ 232
  8.2.4 Debye-Hückel-like screening model ......................... 235
  8.2.5 Screened SS(V)PE and screened IEF-PCM ................. 236
8.3 Discretization .................................................. 238
  8.3.1 Overview of the SWIG approach ........................... 239
  8.3.2 Electrostatic solvation energy and gradient ............... 240
  8.3.3 Matrix elements for SWIG discretization ................. 243
  8.3.4 Electrostatic potentials and electric fields ............... 247
  8.3.5 DESMO analytic gradient .................................. 251
8.4 Numerical Tests ............................................... 253
  8.4.1 Comparison to the Kirkwood model ......................... 255
  8.4.2 Comparison to the model of Lotan and Head-Gordon .... 260
  8.4.3 Comparison to results from the adaptive
        Poisson-Boltzmann solver .................................. 265
8.5 Summary ....................................................... 269

9 Conclusion ....................................................... 271

APPENDICES

A Supporting Information for: “Both intra- and interstrand charge-transfer
excited states in aqueous B-DNA are present at energies comparable to,
or just above, the $1\pi\pi^*$ excitonic bright states” ..................... 273

  A.1 TD-DFT basis set dependence ............................... 273
  A.2 Benchmarking the LRC functionals .......................... 277
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>A.3 Supplementary gas-phase data</td>
<td>281</td>
</tr>
<tr>
<td>A.4 MD simulations and QM/MM models</td>
<td>286</td>
</tr>
<tr>
<td>A.5 Supplementary solution-phase data</td>
<td>289</td>
</tr>
<tr>
<td>B Supporting information for “Polarizable continuum reaction-field solvation models affording smooth potential energy surfaces”</td>
<td>295</td>
</tr>
<tr>
<td>B.1 Additional information on the SWIG formalism</td>
<td>295</td>
</tr>
<tr>
<td>B.1.1 Derivation</td>
<td>295</td>
</tr>
<tr>
<td>B.1.2 Switching function</td>
<td>300</td>
</tr>
<tr>
<td>B.2 Computational details</td>
<td>301</td>
</tr>
<tr>
<td>B.2.1 (Adenine)(H$<em>2$O)$</em>{52}$ optimization and frequencies</td>
<td>301</td>
</tr>
<tr>
<td>B.2.2 NaCl dissociation</td>
<td>302</td>
</tr>
<tr>
<td>C Supplementary data for “Symmetric versus asymmetric discretization of the integral equations in polarizable continuum solvation models”</td>
<td>304</td>
</tr>
<tr>
<td>C.1 Additional data</td>
<td>304</td>
</tr>
<tr>
<td>C.2 Comparison of FIXPVA Implementations</td>
<td>309</td>
</tr>
<tr>
<td>C.3 FIXPVA discontinuity</td>
<td>311</td>
</tr>
<tr>
<td>Bibliography</td>
<td>311</td>
</tr>
</tbody>
</table>
## List of Figures

<table>
<thead>
<tr>
<th>FIGURE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>18</td>
</tr>
<tr>
<td>2.2</td>
<td>18</td>
</tr>
<tr>
<td>2.3</td>
<td>20</td>
</tr>
<tr>
<td>2.4</td>
<td>22</td>
</tr>
<tr>
<td>2.5</td>
<td>25</td>
</tr>
<tr>
<td>2.6</td>
<td>31</td>
</tr>
</tbody>
</table>

2.1 Typical examples of spurious CT excitations in small uracil–water clusters: (a) water-to-uracil CT, and (b) uracil-to-water CT. Each excitation may be conceptualized as a rearrangement of the electron detachment density on the left into an attachment density on the right.

2.2 Typical examples of spurious CT excitations in a uracil–(H$_2$O)$_{25}$ cluster: (a) water-to-uracil CT, and (b) water-to-water CT.

2.3 Excitation energies, oscillator strengths, and detachment densities for the lowest six TD-PBE0/6-31+G* excited states of a uracil–(H$_2$O)$_7$ cluster.

2.4 TD-PBE0/6-31+G* excitations for the $d = 4.0$ Å cluster, illustrating intensity borrowing by spurious CT states.

2.5 (a) Number of excited states within 6 eV of the ground state, and (b) excitation energy of the 40th excited state, each as a function of the radius $d$ of the QM region.

2.6 Electronic absorption spectra (thick lines, scale on the left) and densities of states (thin lines, scale on the right) from QM/MM calculations at the TD-B3LYP/6-31+G* and TD-PBE0/6-31+G* levels. As described in the text, the calculations in (a) and (b) utilize a uracil-only QM region, (c) and (d) employ a microhydrated QM region, while (e) and (f) use a full solvation shell for the QM region. Dotted vertical lines show the positions of the first two band maxima in the experimental absorption spectrum of aqueous uracil (from Ref. 1).
2.7 TD-B3LYP/6-31+G* calculations on uracil–water clusters using (a) a microhydrated QM region, and (b) a full solvation shell in the QM region. Electronic absorption spectra (scale on the left) are plotted for both the QM region only (thick, solid line) and for the QM/MM calculation (dotted line). The thin, solid line is the density of states (scale on the right) for the QM-only cluster calculation. 34

2.8 (a) Absorption spectra and (b) densities of states, from TD-PBE0/6-31+G* QM/MM calculations with a full QM solvation shell, using \( \kappa_{\text{occ}} = 0.2 \) (solid lines) and \( \kappa_{\text{occ}} = 0.0 \) (broken lines). 39

2.9 (a) Absorption spectra and (b) densities of states, from TD-B3LYP/6-31+G* calculations of the full QM solvation shell in the absence of MM point charges, using \( \kappa_{\text{occ}} = 0.2 \) (solid lines) and \( \kappa_{\text{occ}} = 0.0 \) (broken lines). 41

3.1 Vertical excitation energies for \( A_2 \), calculated at the TD-LRC-PBE0/6-31+G* level as a function of the range-separation parameter \( \mu \). The black lines are adiabatic excitation energies (alternating solid and broken, for clarity). The CT and \( \pi\pi^* \) diabatic states are also indicated [cf. Fig. 3.2(a)]. 52

3.2 Diabatized vertical excitation energies as a function of the range-separation parameter \( \mu \), for (a) stacked adenine dimer, \( A_2 \), and (b) \( A_2(H_2O)_{27} \). The two monomers are labeled 3’ and 5’, and the \( \pi\pi^*(-) \) and \( \pi\pi^*(+) \) labels indicate linear combinations of localized \( \pi \rightarrow \pi^* \) excitations (“Frenkel excitons”). Two \( n\pi^* \) states are plotted in (a), but they are indistinguishable on this scale. 53

3.3 TD-LRC-PBE0 detachment and attachment densities for \( A_2 \) using (a) \( \mu = 0 \), (b) \( \mu = 0.10 \ a_0^{-1} \), and (c) \( \mu = 0.6 \ a_0^{-1} \). Each plot represents an isocontour that encompasses 70% of the total density. Also shown are the excitation energies and the amount of intermolecular CT (\( \Delta q \)) upon excitation, as quantified by natural population analysis. 55

3.4 TD-LRC-PBE0 diabatic potential energy curves for \( A_2 \) as a function of the intermolecular distance, using (a) \( \mu = 0 \) and (b) \( \mu = 0.35 \ a_0^{-1} \). The two \( n\pi^* \) potential curves overlap one another in both plots, and the CT states in (b) lie above 6.2 eV and are not shown. 56
4.1 Plot of the number of CT states below the brightest $^1\pi\pi^*$ exciton state, computed at the TD-PBE0/6-31G* level for a sequence of single-stranded adenine multimers in their canonical B-DNA geometries, with backbone atoms removed. ................................................. 72

4.2 Natural transition orbitals of a spurious end-to-end CT state appearing at 5.6 eV above the ground state, 0.1 eV below brightest $\pi\pi^*$ state. The calculation is performed at the TD-PBE0/6-31G* level, for $\pi$-stacked A$_7$ multimer. This particular particle/hole pair represents 99.9% of the transition density for the state in question. ................................................. 75

4.3 Absorption spectra for A$_2$ (broken curves) and for adenine monomer (solid curves), computed by applying a 0.3 eV gaussian broadening to the gas-phase vertical excitation energies, weighted by their respective oscillator strengths. The monomer spectra are labeled as “2*A” to indicate that these oscillator strengths are weighted by an additional factor of two. Both calculations employ the 6-31G* basis set and use canonical B-DNA geometries. ................................................. 78

4.4 Natural transition orbitals (NTOs) corresponding to the excited state with largest oscillator strength in (a) A$_3$:T$_3$ and (b) ATA:TAT. In (a), the exciton is localized almost entirely on the adenine strand and consists of two significant NTO particle/hole pairs, whereas the bright state in ATA:TAT is mostly a localized monomer-like excitation. ................................................. 80

4.5 Attachment densities (in blue) and detachment densities (in red) for delocalized, interstrand CT states in A$_3$:T$_3$ and A$_4$:T$_4$, in which there is significant CT between nucleobases that are not hydrogen bonded. Each calculation was performed at the TD-LRC-$\omega$PBE/6-31G* level, and the surfaces shown encapsulate 90% of density. Excited-state Mulliken charges on each monomer are also provided. (The monomers are essentially neutral in the ground state.) ................................................. 82
4.6 Natural transition orbitals (NTOs) corresponding to the state with the largest oscillator strength in (a) A$_5$ and (b) ATATA. The two particle/hole NTO pairs with largest amplitude are shown in either case. In A$_5$, the exciton state couples $\pi\pi^*$ excitations that are four bases removed in sequence, and furthermore displays a characteristic nodal structure resulting from a linear combination of localized $\pi\pi^*$ excitations. As such, there are several significant excitation amplitudes, even in the NTO basis. In contrast, the bright state in ATATA is predominantly localized on a single adenine monomer, and is well-described by a single NTO particle/hole pair.

4.7 Absorption spectra for aqueous A:T and A$_2$ obtained from a TD-DFT/6-31G* QM/MM calculation. The LRC-$\omega$PBE functional is used in (a) and (b), whereas the LRC-$\omega$PBEh functional is used in (c) and (d). To avoid congestion, the optically-weak $^1n\pi^*$ states are omitted. Gaussian distributions are obtained from averages over solvent configuration; for the CT states, the stick spectra are shown as well. The optically-bright CT states around 6.9 eV borrow intensity from the second $\pi\pi^*$ absorption band, which is not shown.

4.8 Absorption spectra for (a) hydrated A:T and (b) hydrated A$_2$, obtained from a SCS-CIS(D)/6-311+G* QM/MM calculation, using the same configurational snapshots used to generate Fig. A.5. Gaussian distributions were obtained from averages over solvent configuration (using CIS oscillator strengths), though for the CT states the stick spectra are shown as well. In A:T there is considerable mixing between the second $\pi\pi^*$ band and the CT states, lending significant oscillator strength to the latter.

4.9 Absorption spectrum of aqueous (ApA)$^-$ calculated at the TD-LRC-$\omega$PBE/6-31G* level, including corrective shifts as discussed in the text. Weakly-absorbing $n\pi^*$ states are omitted, for clarity.
5.1 Molecular mechanics geometry optimization of (adenine)(H$_2$O)$_{52}$ in bulk water, using three different implementations of COSMO. The vertical scale represents the cluster binding energy, including the electrostatic free energy of solvation. Also shown are the solute cavity surfaces for the optimized structures. Each grid point $r_i$ is depicted as a sphere whose radius is proportional to $a_i$, and colored according to the charge $q_i$. .......................................................... 105

5.2 Vibrational spectra of the FIXPVA- and SWIG-COSMO optimized (adenine)(H$_2$O)$_{52}$ structures. (The inset is an enlarged view of the region up to 4,000 cm$^{-1}$.) Harmonic frequencies were calculated by finite difference of analytic energy gradients and convolved with 10 cm$^{-1}$ gaussians, weighted by the computed intensities. Arrows indicate FIXPVA peaks with no obvious SWIG analogues. ................................. 107

5.3 Total energy (solid curves, scale at left) and Na-atom gradient (dashed curves, scale at right) for NaCl dissociation, computed at the UHF/6-31+G$^*$/SS(V)PE level including non-electrostatic contributions to the PCM energy. A horizontal dotted line indicates where the gradient is zero. Panel (c) shows two sets of results for the SWIG model, corresponding to a switching functions $F_i^p$ with two different values of $p$. ................................................................. 108

5.4 Non-electrostatic contributions to the PCM energy, for the UHF/SS(V)PE NaCl potential curves from 5.3. ................................. 110

6.1 NaCl dissociation in water ($\varepsilon = 78.39$), computed at the AMBER99/SS(V)PE level using the “subSWIG” discretization scheme in which $D_{ii}$ is defined using the sum rule in Eq. (6.51). Panel (a) plots the solvation energy, $E_{pol}$, and its gradient with respect to Cl displacement, while panel (b) plots the largest eigenvalue of $Q$, along with the number of surface grid points for which $F_i > 10^{-8}$. Data points were calculated at 0.01 Å intervals. For clarity, the vertical scale has been truncated in both panels, i.e., some of the sharp spikes are off of the scale that is used. ................................................................. 145
6.2 NaCl dissociation in water ($\varepsilon = 78.39$), computed at the AMBER99/SS(V)PE level using the SWIG discretization scheme. Panel (a) plots the solvation energy, $E_{pol}$, and its gradient with respect to Cl displacement, while panel (b) plots the largest eigenvalue of $Q$, along with the number of surface grid points for which $F_i > 10^{-8}$. Data points were calculated at 0.01 Å intervals.

6.3 Solvation energy, $E_{pol}$, and eigenvalues of $Q$, near a singular point of the subSWIG potential energy surface for NaCl dissociation. (Data points are calculated every $10^{-4}$ Å.) The inset shows a larger range of internuclear distances, and suggests that singularities may go unnoticed unless the spacing between data points is extremely small.

6.4 Rotational variance, $\Delta_{rot}$, and Gauss’ Law error, $\Delta_{GL}$, as a function of the number of grid points per atom, for a set of 20 amino acids in water ($\varepsilon = 78.39$). Error bars represent one standard deviation about the mean. For clarity, the error bars are omitted for $\Delta_{GL}$.

6.5 Plots of the C-PCM[SWIG] discretization error, $W_N - W_{1202}$, (solid lines) and the RMSE of the gradient (broken lines), as a function of the number of Lebedev grid points per atomic sphere. Data points represent averages over the set of 20 amino acids in water ($\varepsilon = 78.39$), with error bars representing one standard deviation on either side of the mean.

6.6 Harmonic vibrational spectrum of Arg-Asp in water ($\varepsilon = 78.39$), computed by finite difference of analytic energy gradients at the B3LYP/6-31G*/C-PCM level, with either (a) FIXPVA or (b) SWIG discretization. Stick spectra were convolved with 20 cm$^{-1}$ Gaussians. Arrows indicate peaks in the FIXPVA spectrum that have no obvious analogues in the SWIG spectrum at nearby frequencies.

6.7 Harmonic vibrational spectra of glycerol computed at the HF/6-31G* level in (a) the gas phase; (b) liquid glycerol ($\varepsilon = 42.7$), described at the SS(V)PE[SWIG] level; and (c) liquid glycerol, described at the SS(V)PE[subSWIG] level. Stick spectra were convolved with 20 cm$^{-1}$ Gaussians. Arrows indicate spurious peaks in the subSWIG spectrum.
6.8 Fluctuations in the energy during an MD simulation of single-stranded d(GACT) in water, described at the AMBER99/C-PCM level. The inset shows a close-up view of the energy fluctuations obtained in the gas phase and with C-PCM[SWIG]. The time step is 1.0 fs.

6.9 Fluctuations in the energy during an MD simulation of glycine in water, described at the PBE0/6-31+G*/SS(V)PE level. The inset shows a close-up view of the energy fluctuations obtained in the gas phase and with SS(V)PE[ISWIG] over the entire simulation. The time step is 0.97 fs.

6.10 Solvation energy, $E_{\text{pol}}$, of glycine in water, during an MD simulation performed at the PBE0/6-31+G*/SS(V)PE[ISWIG] level. The regions around $-15$ kcal/mol represent the carboxylic acid tautomer, whereas lower-energy regions (around $-50$ kcal/mol) represent the zwitterion. The inset shows a longer amount of simulation time, and indicates that proton transfer occurs multiple times. The time step in this simulation is 0.97 fs.

7.1 Relative energies of the amino acids in water, obtained at (a) the AMBER99 level and (b) the HF/6-31+G* level. The cavity surface is discretized using 590 Lebedev points per atomic sphere, and solution-phase energies obtained using $X = \text{DAS}$ or $X = \text{SAD}^\dagger$ in Eq. (7.11) are reported relative to the energy obtained using a symmetrized form of $X$. The GBO discretization uses Gaussian blurring only, whereas SWIG and ISWIG use a switching function in conjunction with Gaussian blurring.

7.2 Energies of the amino acids in water, relative to results obtained using GBO discretization, for (a) the $X = \text{DAS}$ and (b) the $X = \text{SAD}^\dagger$ form of $K$. Cavity surfaces were discretized using 590 Lebedev points per atomic sphere.

7.3 Convergence of the polarization energy (axis at left), as a function of the number of Lebedev grid points per atomic sphere, for histidine described at the AMBER99 level with SWIG discretization. Results for three alternative $K$ matrices are shown, along with the norm of the matrix $M = \text{DAS} - \text{SAD}^\dagger$ (axis at right).
7.4 Relative energies of the amino acids in water, obtained at the HF/6-31+G* level. The cavity surface is discretized using 60 GE POL points per atomic sphere, and solution-phase energies obtained using \( X = \text{DAS} \) or \( X = \text{SAD}^\dagger \) in Eq. (7.11) are reported relative to the energy obtained using a symmetrized form of \( X \). For comparison, the dashed line at -10 kcal/mol indicates the lower bound in Fig. 7.1. The PC-GE POL \( X = \text{SAD}^\dagger \) data point for amino acid H lies at -2544 kcal/mol, far outside the range of this figure.

7.5 Geometrical definitions used to define the FIXPVA switching functions.

7.6 Deviations from Gauss’ Law for the nuclear contribution to the induced surface charge, and for the total (nuclear + electronic) induced surface charge, for two different versions of IEF-PCM. (The inset shows an enlarged view of the nuclear charge error for \( X = \text{DAS} \).) The solutes are a series of homonuclear diatomic molecules described at the HF/6-31+G* level. The bond lengths, solute cavities, and discretization points are the same for each molecule. SWIG, ISWIG, and GBO discretization produce essentially identical results, so only SWIG results are shown here.

7.7 Histogram of \( D_{ii}a_i \) values (unitless) collected from the Lebedev grids (590 points per atom) across all twenty of the amino acids in the HF calculations. These values are plotted (a) using Eq. (7.17) and (b) using Eq. (7.16). The inset in panel (a) shows a magnified view of the narrow range in which all the values fall.

7.8 Approach to the conductor limit, for histidine described at the AMBER99 level using SWIG discretization. The same cavity is used in each case and is discretized using 110 Lebedev points per atomic sphere.

8.1 Schematic representation of the spherical Debye-Hückel model system. The point charge \( q \) is contained within a sphere of radius \( b \), and the shaded region represents the ion exclusion layer.
8.2 Difference between the PCM solvation energy and that predicted by the analytic LHG formulas, for two point charges centered in spherical cavities with $\varepsilon = 4$ and $\kappa = 0$. The energy difference is plotted as a function of the center-to-center distance between the two spheres, the sum of whose radii is 2.99 Å. (Note that both the horizontal and vertical scales are logarithmic.)

A.1 TD-LRC-\(\omega\)PBE excitation energies (as a function of \(\omega\), with \(C_{\text{HF}} = 0\)) for the benchmark set of molecules in Table A.2. All calculations were performed with the cc-pVDZ basis set and SG-1 quadrature grid.

A.2 Vertical excitation energies for the \(1\pi\pi^*\) and \(1\text{CT}\) states of \(A_2\), as computed at the LRC-\(\omega\)PBE/6-311+G* level (\(C_{\text{HF}} = 0\)) as a function of the LRC range parameter, \(\omega\). CC2/TZVPP results are shown as horizontal dotted lines.

A.3 Vertical excitation energies for the \(1\pi\pi^*\) and \(1\text{CT}\) states of \(A_2\), as computed at the LRC-\(\omega\)PBEh/6-311+G* level (with \(C_{\text{HF}} = 0.2\)) as a function of the LRC range parameter, \(\omega\). CC2/TZVPP results are shown as horizontal dotted lines.

A.4 Stick spectra of various gas-phase \(A_n, T_n,\) and \(A_n:T_n\) multimers composed of nucleobases arranged in the canonical B-DNA geometries. All calculations are performed at the LRC-\(\omega\)PBE/6-31G* level. To avoid congestion, optically-weak \(n\pi^*\) excitations are omitted from these spectra, as are Rydberg excitations (which are largely absent in this energy range, due to the omission of diffuse basis functions). The \(1\pi\pi^*\) states in the multimers (\(n \geq 2\)) are delocalized excitons. The lowest adenine \(\rightarrow\) thymine CT states of A:T appear at 6.5 eV, out of the range depicted here.

A.5 Absorption spectra for aqueous A:T and \(A_2\) at the LRC-\(\omega\)PBE/6-31G* level using three different models of aqueous solvation. (To avoid congestion, the optically-weak \(n\pi^*\) states are omitted.) Gaussian distributions are obtained from averages over solvent configuration; for the CT states, the stick spectra are shown as well.
A.6 Absorption spectra for hydrated A:T (left panels) and hydrated A₂ (right panels), obtained at the CIS(D)/6-311+G* level using QM/MM Model 1. Gaussian distributions were obtained from averages over solvent configuration (using CIS oscillator strengths), though for the CT states the stick spectra are shown as well. In A:T there is considerable mixing between the second ππ∗ band and the CT states, lending significant oscillator strength to the latter. ............................. 290

A.7 TD-LRC-DFT absorption spectra for aqueous A:T and A₂, in which individual excitation energies have been shifted according to estimated errors as described in the manuscript. The LRC-ωPBE functional is used in (a) and (b), whereas the LRC-ωPBEh functional is used in (c) and (d). Oscillator strengths for the CT states are not reliable, as a result of intensity borrowing at the original, calculated excitation energies. ................................................................. 291

A.8 Absorption spectrum of A₂:T₂ obtained from a TD-LRC-ωPBE/6-31G* QM/MM calculation. The QM region is depicted in shown in the figure. ................................. 292

B.1 Number of non-negligible (Fᵢ > 10⁻⁸) Lebedev grid points for NaCl, as a function of the Na–Cl distance. ................................. 299

C.1 Relative energies of the amino acids in water (ε = 78.39), obtained at (a) the amber99 level and (b) the HF/6-31+G* level. The cavity surface was discretized using 110 Lebedev points per atomic sphere, and solution-phase energies are reported relative to the energy obtained using the symmetric form of K. ................................. 306

C.2 Comparison of total energies obtained for amber99 solutes, computed by numerical solution of the Poisson-Boltzmann equation (using the APBS software) to those obtained from two different forms of IEF-PCM with GBO discretization. The APBS and IEF-PCM solute cavities are identical. APBS calculations used a 193×193×193 grid with a grid resolution of 0.1 Å, whereas IEF-PCM calculations used N = 590 Lebedev points per atomic sphere. ................................. 308

xxiv
C.3 Difference in total energy between the \( X = \text{SAD}^\dagger \) and \( X = \text{DAS} \) forms of \( \text{K} \), for successive ionization of histidine described at the HF/6-31+G* level. (The dashed curve is a quadratic fit to the four data points.) The cavity surface is discretized with SWIG using \( N = 590 \) Lebedev points per atomic sphere, and \( \varepsilon = 78.39 \).  

C.4 Illustration of FIXPVA discontinuity along NaCl dissociation at HF/6-31+G* run with GAMESS. Electrostatic solvation energy \( (E_{pol}) \) in kcal/mol is plotted with the solid line, and the surface area in Å\(^2\) is plotted with the dashed line. A GEPOL grid of 60 points per sphere is used. The atomic radii are Na = 1.2 Å and Cl = 1.4 Å, which are marked with a light vertical line and coincide with the observed discontinuities. Data points are plotted at 0.01 Å intervals.
# List of Tables

<table>
<thead>
<tr>
<th>TABLE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Lowest valence TD-DFT excitation energies $\omega$ for a gas-phase isomer of uracil–(H$_2$O)$_4$, at its PBE0/6-31+G* geometry.</td>
</tr>
<tr>
<td>2.2</td>
<td>Summary of TD-PBE0/6-31+G* calculations on uracil–water clusters extracted from a single snapshot of an aqueous-phase MD simulation.</td>
</tr>
<tr>
<td>2.3</td>
<td>Summary of TD-PBE0/6-31+G* QM/MM calculations on aqueous uracil, as a function of the size of the QM region.</td>
</tr>
<tr>
<td>2.4</td>
<td>Summary of TD-BLYP/6-31+G* calculations on uracil–water clusters.</td>
</tr>
<tr>
<td>2.5</td>
<td>TD-PBE0/6-31+G* excitation energies obtained using truncated excitation spaces.</td>
</tr>
<tr>
<td>3.1</td>
<td>Measures of CT contamination in TD-DFT calculations of (uracil)(H$_2$O)$_N$. The TD-PBE0 data are from Ref. 2.</td>
</tr>
<tr>
<td>4.1</td>
<td>Vertical excitation energies (in eV) for the low-lying singlet excited states of adenine (A), thymine (T), and dimers thereof, in their canonical B-DNA geometries.</td>
</tr>
<tr>
<td>4.2</td>
<td>Vertical excitation energies (using the 6-311G* basis) for the lowest intermolecular CT state between two adenine monomers (MP2/6-311++G** geometries) separated by 20 Å and given a twist angle of 36°, as in B-DNA.</td>
</tr>
<tr>
<td>4.3</td>
<td>Vertical excitation energies (in eV) for adenine (Ade) dinucleotide computed at the TD-DFT/6-31G* level.</td>
</tr>
</tbody>
</table>
4.4 CIS(D)/6-311+G* vertical excitation energies (in eV) for the low-energy excited states of \(A_2\cdot T_2\). .................................................. 85

5.1 Definitions of the matrices appearing in Eq. (1) for the two PCMs considered here, using the notation defined in Ref. 3. .................... 98

6.1 Definitions of the matrices in Eq. (8.43), for the PCMs considered here. The matrix \(A\) is diagonal and consists of the surface element areas, \(a_i\), while the matrices \(S\) and \(D\) are defined in the text. The quantity \(\varepsilon\) represents the dielectric constant of the medium. ....................... 126

6.2 Gradients of the matrices in Table 6.1, for the two PCMs considered here. For brevity, we have defined \(M^x = D^xAS + DA^xS + DAS^x\) and \(f_{\epsilon,\pi} = \left(\frac{\varepsilon-1}{\varepsilon+1}\right) \left(\frac{1}{2\pi}\right)\). ................................................................. 129

6.3 Matrix elements of \(S\) and \(D\) for the FIXPVA and CSC methods. In Ref. 4, the CSC matrix elements are defined for spheres of unit radius, but we have generalized them here to arbitrary areas, \(a_i\). We have also generalized the FIXPVA approach of Ref. 5 for use with IEF-PCM/SS(V)PE, as described in the text. In the expressions for \(S_{ii}\), \(R_i\) is the radius of the \(i\)th sphere \((i \in I)\), and the constant \(C_S\) is a self-energy factor that depends upon the choice of surface grid: \(C_S = 1.0694\) for GEPOL grids\(^6\) and \(C_S = 1.104\) for Lebedev grids.\(^7\) .... 139

6.4 Comparison of solvation energies (in kcal/mol), using a solute cavity consisting of a single sphere whose radius is adjusted in order to reproduce the solvation energies reported in Ref. 8, where an isodensity contour was used to define the cavity surface. ................................. 150

6.5 Comparison of solvation energies (in kcal/mol) using a vDW cavity composed of atomic spheres. Note that SWIG and subSWIG discretization procedures are equivalent for C-PCM/GCOSMO. ....... 152

6.6 Definitions of \(w_{\text{displ}}\) for various continuum models. The quantity \(\varepsilon\) represents the dielectric constant of the medium. .......................... 170
7.1 Errors in the total cavity surface area, for the amino acid data set using 
\( N = 590 \) Lebedev points per atom (SWIG, ISWIG, and FIXPVA) as 
well as using \( N = 60 \) GEPOL points per atom. The mean signed error 
(MSE), root mean square error (RMSE), and maximum signed error 
(Max) are listed, in percent, taking PC results with Lebedev grids as 
the benchmark. ..................................................... 199

7.2 Gauss' Law error statistics for IEF-PCM. Listed are the mean absolute 
error (MAE), the root mean square error (RMSE), and the maximum 
error (Max), evaluated over the amino acid data set, with \( \varepsilon = 78.39 \) 
and \( N = 590 \). All values are given in atomic charge units. .......... 201

7.3 Statistics for relative energies of the amino acids in water, obtained 
at the HF/6-31+G* level, when using the sum rule of Eq. (7.16). The 
mean absolute difference (MAD), root mean square difference (RMSD), 
and the maximum absolute difference (Max) in the solvation energies 
(in kcal/mol) are tabulated for the two asymmetric \( X \) matrices, relative 
to the symmetric version. ........................................ 205

8.1 Definitions of \( K \) and \( R \) for screened PCM. ....................... 242

8.2 Definitions of \( K^x \) and \( R^x \) for screened PCMs. ............. 244

8.3 Ionic strengths (in mol/L) for each pair of parameters \( \varepsilon \) and \( \lambda \) explored 
here, assuming \( T = 298 \) K. Note that physiological ionic strengths are 
\( \sim 0.1–0.2 \) mol/L. ................................. 254

8.4 Reaction field energies, \( E_{pol} \), for a +e point charge centered in a cavity 
of radius 2 Å, along with errors in various PCMs. ..................... 256

8.5 Salt shifts, \( E_{pol}(\kappa) - E_{pol}(\kappa = 0) \), for a +e point charge centered in a 
cavity of radius 2 Å. ........................................... 257

8.6 Reaction field energies, \( E_{pol} \), for a Kirkwood model of a dipolar solute 
(\( \mu = 9.6 \) debye inside a cavity of radius 3 Å), along with errors in 
various PCMs. .................................................. 258

xxviii
8.7 Salt shifts, $E_{\text{pol}}(\kappa) - E_{\text{pol}}(\kappa = 0)$, for a Kirkwood model of a zwitterion ($\mu = 9.6$ debye inside a cavity of radius 3 Å). ........................................... 259

8.8 Reaction field energies, $E_{\text{pol}}$, obtained from the LHG model\textsuperscript{10} for a set of twenty randomly-positioned charges in spherical cavities (see the text for details). PCM results are given in terms of the error relative to the LHG result. ........................................... 261

8.9 Salt shifts, $E_{\text{pol}}(\kappa) - E_{\text{pol}}(\kappa = 0)$, obtained from the LHG model\textsuperscript{10} and from various PCMs, for a set of twenty randomly-positioned charges in spherical cavities (see the text for details). ........................................... 262

8.10 Reaction field energies, $E_{\text{pol}}$, for alanine dipeptide, obtained using the APBS software. PCM results are given in terms of the error relative to the APBS result. ........................................... 266

8.11 Salt shifts, $E_{\text{pol}}(\kappa) - E_{\text{pol}}(\kappa = 0)$, for alanine dipeptide, obtained using the APBS software and various PCMs. ........................................... 267

8.12 Reaction field energies, $E_{\text{pol}}$, for thymine dinucleotide, obtained using the APBS software. PCM results are given in terms of the error relative to the APBS result. ........................................... 268

8.13 Salt shifts, $E_{\text{pol}}(\kappa) - E_{\text{pol}}(\kappa = 0)$, for thymine dinucleotide, obtained using the APBS software and various PCMs. ........................................... 269

A.1 Basis set dependence of vertical excitation energies (in eV) of simple nucleobase systems, calculated at the LRC-ωPBE level. Oscillator strengths are shown in parentheses. ........................................... 274

A.2 Vertical excitation energies (in eV) for a set of localized (L) and charge-transfer (CT) excitations, including root-mean-square errors (RMSEs) and mean signed errors (MSEs). ........................................... 276
A.3 Vertical excitation energies (in eV) and oscillator strengths for low-lying singlet excited states of simple nucleobase systems. All methods used the 6-311+G* basis set except CC2, where the TZVP basis was used. Geometries correspond to canonical B-DNA. Some of the CC2 relaxed oscillator strengths failed to converge and are therefore not reported here. For the CIS(D) and SCS-CIS(D) excitation energies, the oscillator strengths are CIS values.

A.4 LRC-ωPBE/6-31G* vertical excitation energies (VEEs) and oscillator strengths of ATATA. The monomers are labeled by number in 5′ to 3′ order.

A.5 LRC-ωPBE/6-31G* vertical excitation energies (VEEs) and oscillator strengths of ATA:TAT. The monomers are labeled by number in 5′ to 3′ order and additionally according to their strand (α = ATA, β = TAT).

C.1 Possible definitions for the matrices K and R in Eq. 7.

C.2 Maximum (signed) difference, in kcal/mol, between the IEF-PCM energy computed using $X = \text{DAS}$ (or $X = \text{SAD}^\dagger$) and that computed using the symmetric form of $X$. The data set is the amino acids in water. Three different values of $N$, the number of Lebedev points per atomic sphere, and three different discretization procedures are compared.

C.3 FIXPVA on GEPOL grids with 240 points per atom versus Lebedev grids with 194 points per atom. Total energies ($W$), surface area (SA), and Gauss’ law error for nuclear charges in a.u. ($\Delta_{\text{Gauss}}$) are compared. Calculations are run at HF/6-31+G*//C-PCM with Bondi radii scaled by 1.2 and $\varepsilon = 78.39$. GEPOL calculations are run with GAMESS, and Lebedev calculations are run with our locally modified version of Q-Chem.
The field of quantum chemistry involves modeling atoms and molecules with the central equation of quantum mechanics, the Schrödinger equation (SE),

\[ i\hbar \frac{\partial}{\partial t} \Psi = \hat{H} \Psi , \]  

(1.1)

where \( \Psi \) is the wavefunction and \( \hat{H} \) is the Hamiltonian operator. The SE is deceptively simple, as it becomes very difficult to solve in practice for the many particles (i.e., electrons and nuclei) present in molecules. Even “small” molecules composed of just a few atoms are a challenge. If one desires to model molecular systems dissolved in a solvent environment, which is the main focus of this work, the difficulty of solving the SE is compounded much further as there are many more particles for which to account. In order to make quantum chemistry at all feasible, one must make a series of smart approximations. It is not our intention to delve into the details of all of these here. Rather, we briefly review what some of the most important approximations are and why they are necessary to make.

---

1.1 We assume a non-relativistic regime throughout this work, as is typically the case for the sort of chemistry we will be concerned with here.
First, it is typical to use separation of variables to transform Eq. (1.1) into a
time-dependent equation and a time-independent equation. The time-independent
SE (TISE),
\[ \hat{H} \Psi = E \Psi , \]  
(1.2)
is an eigenvalue equation that can be solved to provide the total energy, \( E \), of a
quantum system for each quantum state. One can then make the \textit{ansatz} that the total
wavefunction of a molecule, \( \Psi \), is separable into a product of a nuclear wavefunction,
\( \psi_{nuc} \), and an electronic wavefunction, \( \psi_{elec} \).
\[ \Psi = \psi_{elec} \psi_{nuc} . \]  
(1.3)
This separation is not an approximation insofar as a complete basis is employed for
both \( \psi_{elec} \) and \( \psi_{nuc} \). In practice, though, this is usually not attainable, and finite
basis sets are employed. Solving the TISE simultaneously for both \( \psi_{elec} \) and \( \psi_{nuc} \) is
rarely performed in practice because of the various complexities involved in computing
couplings between the two. Thus, one appeals to the Born-Oppenheimer Approxima-
tion (BOA), which, primarily by virtue of the much larger masses of the nuclei in
molecules, allows one to assume that the nuclei behave like “classical” (\textit{i.e.}, obeying
Newton’s equations of motion) point charges, thereby forgoing the complication of
solving the TISE for \( \psi_{nuc} \). The BOA allows us to write the energy as a function of
nuclear coordinates, \( E(\vec{R}) \), and immediately affords the extremely important concept
of a “potential energy surface” (PES). That is, one can envision a multi-dimensional
surface for \( E(\vec{R}) \) that describes the nature of a given molecular system as a function
of how its nuclear geometry is arranged. A PES provides a wealth of information
about chemical bonding, vibrations, rotations, reactivity, etc. A PES could also be used to propagate nuclear dynamics classically via Newton’s equations of motion, or it could even be used to solve for a $\psi_{\text{nuc}}$ and propagate dynamics quantum mechanically according to Eq. (1.1). However, obtaining a complete PES is highly non-trivial. It involves solving the TISE for the electronic wavefunction at every possible $3N$ nuclear degrees of freedom, with $N$ being the number of nuclei. \(^1\)\(^2\) As our molecular system of interest gains more nuclei, computing a complete PES rapidly becomes intractable.

Even still, solving for $\psi_{\text{elec}}$ at just a single point of nuclear coordinates is very challenging due to the many interactions of a molecule’s electrons, and further approximations must be made. The particular field of quantum chemistry associated with solving the TISE for $\psi_{\text{elec}}$ is known as electronic structure theory (EST). EST provides several methodologies for obtaining $\psi_{\text{elec}}$ at varying levels of approximation, including Hartree-Fock (HF) theory, Density Functional Theory (DFT), Coupled Cluster theory, and Møller-Plesset perturbation theory, just to name a few popular ones. EST methods are fairly complex and require considerable computational power to carry out, scaling exponentially with respect to the number of basis functions (or, number of electrons) employed in solving $\psi_{\text{elec}}$. Computational expense inevitably grows with increasingly accurate EST methodology, yet even the “cheapest” EST methods, like HF or DFT, have an exponential scaling (in terms of CPU time and/or required core memory) somewhere between quadratic and cubic. Such scaling is a profound barrier.

\(^1\)\(^2\)One usually assumes in quantum chemistry that translations and rotations can be treated classically such that it is only necessary to sample the $3N - 6$ (or, $3N - 5$ for a linear molecule) internal nuclear degrees of freedom.
to the application of EST for modeling the chemistry of “large” molecular systems. For instance, suppose a single point HF calculation \((i.e., \text{just calculating } E(\vec{R}) \text{ for the ground electronic state at a single nuclear geometry})\) on one water molecule takes 10 seconds to finish on a computer. If we then want to perform the same calculation on a cluster of 25 waters—a far cry from the \(\sim 6 \times 10^{23}\) (Avogadro’s number) water molecules present in real bulk water—the calculation will take \((25 \times 10^2 s)^2 \sim 17\) hours, assuming only quadratic scaling. Cubic scaling amounts to \(\sim 181\) days! Theories more accurate than HF have even higher scaling and will take much longer. Clearly, the exponential scaling of EST methods prohibits studying the condensed phase \((i.e. \text{not gas phase})\), wherein many molecules and atoms are present, with a full, all-electron, quantum chemistry-only approach.

It is therefore imperative to develop further approximations to enable the application of EST to the condensed phase, albeit sacrificing some accuracy in the process, in order to understand the role of quantum chemistry in these environments. Introducing a concept of mixing a variety of levels of approximation together becomes extremely useful here. The primary idea is that instead of treating the whole molecular system under the same, single level of theory \((e.g., \text{a whole water cluster is treated at the HF level})\), we assume that a certain part of the system can be treated with one certain level of theory while another part can be treated with a different level of theory. We refer to this as a “multi-layer” approach in the sense that each “layer” of
the full chemical system assumes different treatments and approximations.\textsuperscript{1,3} Perhaps the most well-know example of such a multi-layer method is mixed quantum mechanics/molecular mechanics (QM/MM).\textsuperscript{11} In QM/MM, a small region of the total system is assumed to behave quantum mechanically and is treated with EST, while the rest of the surrounding environment is assumed to behave classically and is treated at the level of molecular mechanics (MM). MM is based on easily computed, analytical formulas that make a classical approximation to the quantum mechanical nature of the electrostatics, bonds, and vibrational motions present in a molecular system. QM/MM provides an avenue to avoid the exponential scaling of EST by only focusing on the small region in which one expects quantum behavior to be important, such as the reactive center of an enzyme where chemical bonds can break or form, while the remainder of the system is treated more approximately via simple MM. The cost of a QM/MM calculation is therefore only about as computationally expensive as the EST calculation on the small QM region.

The surrounding environment, of course, plays a major role in chemistry, and QM/MM incorporates such effects predominantly through embedding the QM region in the electrostatic field of the classical MM region. For example, one can model the effects of a solvent on a given solute by surrounding it with MM molecules, which will polarize the solute’s electronic wavefunction and give rise to an electrostatic stabilization (or perhaps a destabilization) relative to the gas phase. This electrostatic solvation effect is one of the largest contributors to chemical phenomena in solution.

\textsuperscript{1,3}One could argue that we have already made a multi-layer approximation in EST by treating the nuclei as classical particles while maintaining the electrons as quantum particles.
describing even such familiar chemistry as table salt (sodium chloride) dissolving in water. Other than solvent, though, the MM region could be used, for example, to describe a protein environment or the double helix scaffolding of deoxyribonucleic acid (DNA).

A problem that arises in the modeling of solvent, though, is that the solvent molecules are dynamical, and a single point calculation represents only one snapshot of the realistic dynamical motions of the solvent. In practice, one can sample many configurations of solvent geometry (e.g., using molecular dynamics or Monte Carlo techniques) and gather statistics across them to obtain an average description of the solvation effects. However, even with QM/MM, configurational sampling is a computationally expensive endeavor, requiring several EST calculations to obtain a single average. Thus, we are once again faced with the need to make an approximation in order to create a tractable model of chemistry in solvent.

To avoid such unfeasible sampling involved in the “explicit” inclusion of surrounding solvent molecules, one can take a so-called “implicit” solvent approach, wherein one approximates the configurationally sampled solvent electrostatic charge density as a classical continuum of dielectric media. Thereby, individual solvent molecules are not actually present, but the mean-field electrostatics of the bulk solvent are nonetheless retained via the dielectric electrostatics. The concept and success of implicit solvent modeling can be traced back to the seminal work of Max Born in 1920 (roughly the same time as when the SE was discovered) who famously described the
solvation energy of ionic chemical species with a very simple formula,\(^\text{12}\)

\[
E_{\text{solv}}^{\text{Born}} = -\left(\frac{\varepsilon - 1}{\varepsilon}\right) \frac{q^2}{2R},
\]

(1.4)

where \(\varepsilon\) is the dielectric constant (i.e., relative permittivity) of the bulk solvent, \(q\) is the charge of the ionic species, and \(R\) is a radius characteristic of the molecule/atom size. The Born ion model later underwent significant elaborations by the likes of several well-known chemists, including Kirkwood,\(^\text{13}\) Debye & Hückel,\(^\text{14}\) and Onsager,\(^\text{15}\) who coined the term “reaction-field” to describe the electrostatic response of the dielectric solvent in implicit solvation. Implicit solvation can be used as another layer on top of QM/MM to sidestep the above issue of solvent configurational averaging while still incorporating the very important reaction-field electrostatics of solvent.

Once all of the above approximations have been made, ranging from the fundamental SE to the dielectric continuum treatment of solvent, we may finally be able to model quantum chemistry in the condensed phase. The path taken through these approximations, though, is paramount to the accuracy of the resulting multi-layer model. As we describe thoroughly in the remainder of this work, one must pay special attention to the details of each layer in the construction of a multi-layer model in order to avoid catastrophic contamination with artifacts. With care taken, the resulting model is a very powerful tool for investigating and understanding the chemistry of molecules in a condensed phase environment.

Our primary molecular system of interest is electronically excited DNA in aqueous solution. Our investigation begins with perhaps the simplest possible related
system, the low-lying electronic states of aqueous uracil. Despite its apparent simplicity, though, this system already poses significant challenges to theoretical modeling, namely the presence of spurious charge-transfer electronic states as predicted by Time-Dependent Density Functional Theory (TDDFT). We find that QM/MM tends to alleviate some of these difficulties, but, in order to grow our system size toward biologically relevant DNA strands, we must address the intrinsic error of TDDFT that gives rise to these spurious charge-transfer states. We then examine so-called Long Range Corrected TD-DFT (LRC-TDDFT) methodology and find that it mitigates the issue of spurious CT, allowing us to investigate the absorption spectrum for electronic excitations of aqueous DNA containing up to as much as 8 nucleobases. Realizing the difficulty of configurational sampling of solvent molecules, though, we appeal to the incorporation of the implicit solvation methodology known as Polarizable Continuum Models (PCMs). However, we find an alarming number of numerical instabilities and inaccuracies in several popular approaches to PCMs, which motivates us to develop a new method to ameliorate the sources of such. In addition, our investigation of PCM facilitates the derivation of an entirely new PCM for modeling the effects of salt, which are naturally present in biological environments like that surrounding DNA. Ultimately, we arrive at a robust multi-layer model capable of describing quantum chemistry in solvent, including our DNA system of interest.
CHAPTER 2

Simple methods to reduce charge-transfer contamination in time-dependent density-functional calculations of clusters and liquids

2.1 Introduction

Time-dependent density functional theory (TD-DFT) is currently the most popular method for calculating excited electronic states of gas-phase molecules with ~10–200 atoms, owing to its favorable computational scaling (cubic or better with respect to system size\textsuperscript{16,17}) and reasonable accuracy (0.2–0.3 eV for the lowest few valence excitations\textsuperscript{17–19}). Condensed-phase TD-DFT calculations, on the other hand, are beset by serious contamination from spurious, low-energy charge-transfer (CT) excited states,\textsuperscript{20–26} the proximate cause of which is TD-DFT’s tendency to underestimate long-range CT excitation energies.\textsuperscript{17,18,27–30} Although this problem is present already in the gas phase (and will manifest itself in TD-DFT calculations of well-separated molecules,\textsuperscript{29} or even sufficiently large single molecules\textsuperscript{30–32}), it is much more pervasive in liquids and clusters.

\textsuperscript{2,1}This chapter appeared as a full article in the Journal of Chemical Theory and Computation, in 2007, volume 3, pages 1680–1690.
Underestimation of long-range CT energetics is a consequence of incorrect asymptotic behavior on the part of the exchange-correlation potential,\textsuperscript{29} and several long-range correction schemes have been developed recently in an attempt to alleviate this problem.\textsuperscript{33–38} These corrections appear to mitigate CT problems for well-separated molecules in the gas phase, though only one of them has been tested in a cluster environment.\textsuperscript{26} Furthermore, these corrections do not rectify all of the problems associated with the long-range behavior of existing density functionals,\textsuperscript{39} and moreover the improved asymptotic behavior sometimes comes at the expense of diminished accuracy for ground-state properties.\textsuperscript{40} In the present work, we explore some alternative methods for reducing CT contamination that are different from (though fully compatible with) these long-range correction schemes.

Several previous assessments of the performance of TD-DFT in liquids and clusters have focused exclusively on weakly-allowed $n \rightarrow \pi^*$ excitations in systems such as aqueous acetone\textsuperscript{20,21,25,26} and aqueous formamide.\textsuperscript{24} In the case of acetone in liquid water\textsuperscript{20,21} or water clusters,\textsuperscript{25} spurious CT bands overlap the lowest $n \rightarrow \pi^*$ band at 4.5 eV when non-hybrid (but gradient-corrected) density functionals are employed. Hartree–Fock exchange \textit{does} have the correct long-range behavior for CT states,\textsuperscript{29} and hybrid functionals with 20–25\% Hartree–Fock exchange are found to remove CT contamination from the lowest valence band, by pushing the offending CT states to $\sim 1$ eV higher in energy.\textsuperscript{21}

In the present work, we use uracil as a typical example of a molecule possessing both bright states ($^1\pi\pi^*$) and dark states ($^1n\pi^*$). Our results for uracil–water
clusters demonstrate that hybrid functionals alone do not guarantee that the lowest valence band will be free of CT contamination; clusters as small as uracil–(H$_2$O)$_4$ exhibit spurious CT states at energies comparable to or below the lowest $n \rightarrow \pi^*$ and $\pi \rightarrow \pi^*$ excitation energies. These extra states significantly increase the cost of the calculations, in both time and memory, and for a large cluster like uracil–(H$_2$O)$_{37}$, the memory bottleneck precludes us from calculating any states at all above 6 eV.

Two simple procedures to reduce CT contamination are examined here. First, we demonstrate that a mixed quantum mechanics/molecular mechanics (QM/MM) formalism significantly reduces the number of spurious CT states, as compared to calculations performed on the gas-phase QM region. This is true even for large QM regions, and allows us to calculate a full electronic absorption spectrum for a QM region consisting of uracil–(H$_2$O)$_{37}$. In conjunction with a liquid-phase QM/MM calculations, or on its own in the gas phase, spurious CT states can also be removed by omitting TD-DFT excitation amplitudes that correspond to long-range CT. For the present systems, this typically increases the valence excitation energies by $\lesssim 0.1$ eV.

2.2 Computational details

As the only long-range component of contemporary density functionals, Hartree–Fock exchange is known to reduce contamination from long-range CT excited states by pushing these states to higher excitation energies.$^{21,28,29,32}$ As such, our study will focus primarily the hybrid functionals B3LYP$^{41,42}$ and PBE0$^{43-45}$ though for comparison we present a few results obtained with the non-hybrid functional BLYP$^{42,46}$
The PBE0 functional (also known as PBE1PBE\textsuperscript{44}) consists of PBE correlation in conjunction with 25\% Hartree–Fock exchange and 75\% PBE exchange, and has been specifically recommended for excited-state calculations.\textsuperscript{47,48} While a larger fraction of Hartree–Fock exchange—for example, Becke’s “half and half” mixture of Hartree–Fock and Slater exchange,\textsuperscript{49} in conjunction with LYP\textsuperscript{42} correlation—can reduce the overall number of CT states even further, this functional is less accurate for valence excitation energies,\textsuperscript{32} as well as for ground-state thermochemistry.\textsuperscript{50} Newer, highly-parametrized functionals that include full Hartree–Fock exchange may be superior in these respects,\textsuperscript{51} but such functionals are not yet widely available, nor have they been widely tested. We shall restrict our attention to the popular hybrids B3LYP and PBE0.

All TD-DFT calculations reported here employ the Tamm–Dancoff approximation\textsuperscript{52} and were performed using Q-Chem.\textsuperscript{53} Only singlet excitations are considered. Density plots were rendered with the Visual Molecular Dynamics program\textsuperscript{54} using a contour value of 0.001 a.u. in all cases.

The basis-set dependence of the lowest \(n \rightarrow \pi^*\) and \(\pi \rightarrow \pi^*\) excitation energies in uracil–water clusters appears to be very mild, as demonstrated by benchmark calculations for uracil–(H\textsubscript{2}O)\textsubscript{4} that are listed in Table 2.1. For both B3LYP and PBE0, excitation energies obtained with the 6-31+G* basis set differ by no more than 0.1 eV from those obtained with much larger basis sets. As such, all TD-DFT calculations will employ 6-31+G*, along with the SG-0 quadrature grid.\textsuperscript{55}
Table 2.1: Lowest valence TD-DFT excitation energies $\omega$ for a gas-phase isomer of uracil–(H$_2$O)$_4$, at its PBE0/6-31+G* geometry.

<table>
<thead>
<tr>
<th>Functional</th>
<th>Basis set</th>
<th>$\omega$/eV</th>
<th>$n\rightarrow\pi^*$</th>
<th>$\pi\rightarrow\pi^*$</th>
</tr>
</thead>
<tbody>
<tr>
<td>PBE0</td>
<td>6-31+G*</td>
<td>5.06</td>
<td>5.54</td>
<td></td>
</tr>
<tr>
<td>PBE0</td>
<td>6-311+(2d,2p)</td>
<td>5.02</td>
<td>5.46</td>
<td></td>
</tr>
<tr>
<td>PBE0</td>
<td>aug-cc-pVDZ</td>
<td>5.00</td>
<td>5.44</td>
<td></td>
</tr>
<tr>
<td>PBE0</td>
<td>aug-cc-pVTZ</td>
<td>5.00</td>
<td>5.45</td>
<td></td>
</tr>
<tr>
<td>B3LYP</td>
<td>6-31+G*</td>
<td>4.93</td>
<td>5.43</td>
<td></td>
</tr>
<tr>
<td>B3LYP</td>
<td>6-311+(2d,2p)</td>
<td>4.89</td>
<td>5.34</td>
<td></td>
</tr>
<tr>
<td>B3LYP</td>
<td>aug-cc-pVDZ</td>
<td>4.87</td>
<td>5.33</td>
<td></td>
</tr>
<tr>
<td>B3LYP</td>
<td>aug-cc-pVTZ</td>
<td>4.87</td>
<td>5.34</td>
<td></td>
</tr>
</tbody>
</table>

Our interest lies in liquid-phase environments, and thus we wish to employ uracil–water geometries representative of aqueous uracil rather than a gas-phase cluster. We obtain such geometries from a molecular dynamics (MD) simulations of aqueous uracil at constant temperature (298 K) and density (0.9989 g/cm$^3$). Uracil was added to a pre-equilibrated, 25 Å × 25 Å × 25 Å periodic box of flexible water molecules, which was then re-equilibrated using 300 ps of MD. The AMBER99$^{56}$ and TIP3P$^{57}$ force fields (as implemented in the TINKER$^{58}$ software package) were used for uracil and for water, respectively. Following equilibration, uracil–water clusters were extracted from the simulation based on distance criteria that are described in Section 7.4. Water molecules near the uracil (according to these criteria) are included explicitly in the
TD-DFT calculations, while additional water molecules up to 20.0 Å away (about 2300 molecules) are incorporated, in some cases, as TIP3P point charges.

2.3 Results and discussion

2.3.1 CT contamination in uracil–water clusters

In an effort to understand just how “long range” the long-range CT problem in TD-DFT really is, we performed TD-DFT calculations on a sequence of increasingly large uracil–water clusters extracted from the MD simulation described in Section 4.2, by selecting all water molecules having at least one atom within a specified distance $d$ of any uracil atom. All other water molecules were discarded. All clusters were generated from the same MD snapshot, so that each successively larger cluster contains the smaller clusters as its core, and these clusters range in size from bare uracil (when $d = 1.5$ Å) to uracil–$(\text{H}_2\text{O})_{37}$ (when $d = 4.5$ Å).

For each cluster in this sequence, we calculated the first 40 TD-PBE0/6-31+G* excited states. Table 2.2 summarizes the results, including two simple measures of the extent of CT contamination: the excitation energy $\omega_{40}$ of the 40th state above the ground state, and the number of excited states within 6 eV of the ground state. (In these clusters, the second electronic absorption band typically consists of a few states in the 6.0–6.5 eV range, so 6 eV provides a lower bound to the number of TD-DFT excited states that must be calculated in order to reach this second band.)

At the TD-PBE0/6-31+G* level, bare uracil possesses five excited states below 6 eV, the lowest two of which are an $n\pi^*$ dark state (at 4.56 eV) and a $\pi\pi^*$ bright state
Table 2.2: Summary of TD-PBE0/6-31+G* calculations on uracil–water clusters extracted from a single snapshot of an aqueous-phase MD simulation.

<table>
<thead>
<tr>
<th>$d/\text{Å}$</th>
<th>No. water molecules</th>
<th>No. states below 6 eV</th>
<th>$\omega_{40}/\text{eV}$</th>
<th>First $^1\pi\pi^*$ state</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a$</td>
<td>1.5</td>
<td>0</td>
<td>9.05</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>2.0</td>
<td>4</td>
<td>8.06</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>2.5</td>
<td>7</td>
<td>7.46</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>3.0</td>
<td>15</td>
<td>6.60</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>3.5</td>
<td>18</td>
<td>6.50</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>4.0</td>
<td>25</td>
<td>6.22</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>4.5</td>
<td>37</td>
<td>5.65</td>
<td>18</td>
</tr>
</tbody>
</table>

*a*Distance threshold for selecting water molecules  
*b*Excitation energy of the 40th state above the ground state  
*c*Indicates where the state appears in the TD-DFT excitation manifold

(at 5.33 eV). There are also two more dark states of mixed $n\pi^*$/Rydberg character, plus one optically-allowed $n\pi^*$ state whose oscillator strength is 30% of that associated with the $\pi\pi^*$ state. (Uracil is slightly non-planar in the geometries extracted from the MD simulation, so we use “bright” and “dark” as qualitative descriptions of transition intensities. The “optically allowed” $n\pi^*$ state, for example, correlates in a planar chromophore to an excitation out of an $a''$ lone pair orbital.)

The 40 excitations calculated for bare uracil reach 9 eV above the ground state, but due to the appearance of spurious CT states, $\omega_{40}$ drops as cluster size increases, while at the same time the number of states below 6 eV increases. By the time the cluster size reaches $d = 4.5$ Å [uracil–(H$_2$O)$_{37}$], the first 40 excited states reach only
5.65 eV, well below the energy of the second absorption band. At these energies, the density of excited states is $\sim 60$ states/eV, and using Q-Chem on a machine with 4 Gb of memory, we are unable to calculate enough states to reach 6 eV. Excluding core orbitals from the TD-DFT excitation space (which changes the excitation energies by $< 10^{-4}$ eV) reduces the required memory for the Davidson iterations\textsuperscript{59} by a factor of $N_{\text{core}}/N_{\text{occupied}} \approx 0.21$ and (just barely) allows us to calculate the 59 states that are required to reach 6 eV, by which point the density of states has reached $\sim 80$ states/eV. (For comparison, multireference calculations of gas-phase uracil find a total of eight $n\pi^*$ and $\pi\pi^*$ states in the 5.0–7.0 eV range.\textsuperscript{60})

The results in Table 2.2 are for PBE0, but B3LYP paints a similar picture (with even a slightly larger number of spurious CT, consistent with its slightly smaller fraction of Hartree–Fock exchange). We conclude that, despite their success for acetone in liquid water,\textsuperscript{21} in certain systems the popular hybrid functionals B3LYP and PBE0 may still suffer from considerable CT contamination at or below the lowest valence excitation energies. Whereas a $\pi\pi^*$ bright state ought to be either the first or second excited state (depending on the order of the $n\pi^*$ and $\pi\pi^*$ states, which changes as a function of cluster size and geometry), we see from Table 2.2 that clusters as small as uracil–(H$_2$O)$_4$ exhibit spurious states below the first bright state. Apparently, the “long range” CT problem in TD-DFT can manifest even at hydrogen-bond distances, and even when using hybrid functionals with up to 25% Hartree–Fock exchange. That said, it should be emphasized that the problem is dramatically worse for non-hybrids—a BLYP calculation on the $d = 2.5$ Å cluster, for example, yields
more than 40 states below 6 eV, even though there are only seven water molecules, while at \( d = 3.0 \text{ Å} \), CT states appear starting at 2.85 eV and the first \( \pi\pi^* \) state is not even among the first 40 excited states!

In further contrast to the case of acetone in water, where no significant hybridization is observed between the water molecules and the acetone lone pairs,\textsuperscript{20} we do observe hybridization between water and the carbonyl lone pairs of uracil. Consequently, the real \( n\pi^* \) states (and sometimes even \( \pi\pi^* \) states with some \( n\pi^* \) character) are sometimes difficult to discern from the spurious CT states simply on the basis of the TD-DFT excitation amplitudes and Kohn–Sham molecular orbitals (MOs). Such ambiguity is avoided by instead examining electron attachment and detachment densities obtained from the eigenvectors of the difference density matrix between the ground and excited states.\textsuperscript{61} The detachment density represents the part of the density that is removed from the ground state and rearranged in the excited state to form the attachment density.\textsuperscript{17} We make exclusive use of these densities in identifying the qualitative character of the excited states.

Typical examples of low-energy CT states appearing in small uracil–water clusters are illustrated in Fig. 2.1, while Fig. 2.2 depicts some typical CT states in a larger cluster. In small clusters, the CT states below about 5.5 eV are almost exclusively water-to-uracil CT states of the type depicted in Fig. 2.1(a), where the detachment density is dominated by the out-of-plane lone pair on a single water molecule. Such states appear in larger clusters as well [Fig. 2.2(a)], where the water molecule in question tends to be located at the surface of the cluster.
Figure 2.1: Typical examples of spurious CT excitations in small uracil–water clusters: (a) water-to-uracil CT, and (b) uracil-to-water CT. Each excitation may be conceptualized as a rearrangement of the electron detachment density on the left into an attachment density on the right.

Figure 2.2: Typical examples of spurious CT excitations in a uracil–(H₂O)₂₅ cluster: (a) water-to-uracil CT, and (b) water-to-water CT.
The appearance of these states is easy to understand. First note that the out-of-plane lone pairs on the water molecules (except possibly those at the center of a large cluster) are the highest occupied MOs (HOMOs) in the system, while the lowest unoccupied MO (LUMO) is always a uracil $\pi^*$ orbital. In the limit of large separation between an occupied and a virtual MO, and absent any component of Hartree–Fock exchange, TD-DFT will predict CT between these orbitals at an excitation energy equal to the difference in their Kohn–Sham eigenvalues. Thus, if any frontier occupied MOs are spatially separated from low-lying virtual MOs, then one will obtain spurious, low-energy CT excitations, unless a large component of Hartree–Fock exchange (greater than 25%, evidently) is employed. Such states should be anticipated in most condensed-phase systems.

In addition to the uracil $\pi^*$ LUMO, larger uracil–water clusters also possess low-lying virtual MOs localized on the solvent that are not present in small clusters. This opens up another avenue to spurious CT in large clusters and affords water-to-water CT excitations such as that depicted in Fig. 2.2(b). Occasionally these states have some uracil-to-water CT character as well, but mostly the uracil-to-water CT states appear at energies above 5.5 eV, though they proliferate rapidly at higher excitation energies.

The small-cluster CT states are intriguing, because a cluster like uracil–(H$_2$O)$_7$ might not immediately come to mind upon mention of “long-range” CT in TD-DFT. To emphasize that small clusters are indeed susceptible to CT contamination, we examine in detail the excited states of the $d = 2.5$ Å cluster, of which there are
Figure 2.3: Excitation energies, oscillator strengths, and detachment densities for the lowest six TD-PBE0/6-31+G* excited states of a uracil–(H\textsubscript{2}O)\textsubscript{7} cluster.

six within 5.5 eV of the ground state. Detachment densities for these six states are depicted in Fig. 2.3, along with excitation energies ($\omega_i$) and oscillator strengths ($f_i$). Attachment densities are not shown, as each one is dominated by the LUMO and resembles the attachment density shown in Fig. 2.1(a). The detachment densities identify states 3 and 5 are the first $n\pi^*$ and $\pi\pi^*$ states, respectively, whereas the remaining states below 5.5 eV involve water-to-uracil CT of the type discussed above.

With regard to the oscillator strengths for excitations, we note that the $n\pi^*$ state borrows sufficient intensity to achieve an oscillator strength 25% as large as that of the nominal bright state, whereas the CT excitations are mostly dark, consistent with nearly non-overlapping attachment and detachment densities. In system configurations where $n\pi^*/\pi\pi^*$ intensity borrowing is less significant, however, oscillator
strengths for the low-energy CT states sometimes exceeds that of the \( n\pi^* \) state. Thus the real dark states cannot be identified simply from a list of excitation energies and oscillator strengths, but only by careful analysis of the MOs or (better yet) attachment/detachment densities.

In larger clusters, however, CT states can undergo a type of ersatz intensity borrowing that greatly complicates interpretation of the vertical excitation spectrum. A hint as to this behavior is the overall decrease in the oscillator strength of the first \( \pi\pi^* \) as a function of cluster size (see Table 2.2), though the trend is not monotonic—the \( \pi\pi^* \) intensity recovers at \( d = 4.5 \) Å case, at least for this one particular cluster geometry. The reason for this diminished intensity is that, as the density of spurious CT states increases, there appear CT states with energies comparable to that of the \( \pi\pi^* \) state, and these spurious excitations borrow intensity from the real bright state. Since oscillator strengths out of the ground state are positive and sum to a constant (the Thomas–Reiche–Kuhn sum rule\(^62\)), this decreases the oscillator strength of the real bright state. (This explanation is only qualitative, since the sum rule is not exactly fulfilled within the Tamm–Dancoff approximation that we employ here.\(^17\))

In larger clusters, this form of intensity borrowing actually makes it difficult to determine which excitation is the real bright state. The \( d = 4.0 \) Å cluster, for example, exhibits five excited states between 5.05 and 5.20 eV that have significant intensity (states 7–11 in the excitation manifold), which are depicted in Fig. 2.4. With the exception of state 11 (which has the smallest oscillator strength of the five), each of the detachment densities has a significant uracil \( \pi \) component, but in all cases there
Figure 2.4: TD-PBE0/6-31+G* excitations for the $d = 4.0$ Å cluster, illustrating intensity borrowing by spurious CT states.

is a significant contribution from a water lone pair as well. All five of the attachment densities are dominated by the uracil $\pi^*$ LUMO. State 9 is selected as the $\pi\pi^*$ state in Table 2.2 because its TD-DFT eigenvector contains a larger component of the uracil $\pi \rightarrow \pi^*$ excitation that any of the other four states, but note that this is not the strongest transition of the five, as in this case the spurious CT states have borrowed the majority of the oscillator strength of the $\pi\pi^*$ bright state. In reporting a vertical excitation spectrum, then, it is not appropriate simply to report the transition with the largest oscillator strength as “the” bright state.
2.3.2 QM/MM simulations of aqueous uracil

In a recent TD-PBE0 study of hydrated uracil, it was found that a uracil–(H₂O)₄ complex embedded in a polarizable continuum induces a redshift of only 0.1 eV in the first uracil $\pi \rightarrow \pi^*$ excitation, whereas the experimentally-measured solvatochromatic shift is about 0.5 eV. In fact, the polarizable continuum accounts for the entirety of the calculated shift; the four explicit water molecules do not modify the gas-phase excitation energy at all. (A recent TD-BLYP study of s-tetrazine in aqueous solution also found that those water molecules that are directly hydrogen-bonded to the chromophore do not suffice to explain the observed solvatochromatic shift.) As a next step, it seems natural to consider QM/MM simulations of aqueous uracil, using a QM region substantially larger than uracil–(H₂O)₄. Such calculations are discussed in the present section.

To make comparison with results in the previous section, we first consider a sequence of calculations whose QM regions are precisely the same series of increasingly large uracil–water clusters described in Table 2.2 of Section 2.3.1. The MM region in these new calculations consists of all additional water molecules extracted from our MD simulation, out to a distance of 20.0 Å away from uracil. These MM water molecules (about 2300 in all) are incorporated as TIP3P point charges. Table 2.3, which is analogous to Table 2.2 in the previous section, summarizes the results of TD-PBE0/6-31+G* calculations on these QM/MM systems.

Addition of the MM solvent region has a very small effect on the excitation energy for the first $\pi\pi^*$ state, inducing a shift of no more than 0.07 eV, even in cases where
Table 2.3: Summary of TD-PBE0/6-31+G* QM/MM calculations on aqueous uracil, as a function of the size of the QM region.

<table>
<thead>
<tr>
<th>$d/\text{Å}^a$</th>
<th>No. QM water molecules</th>
<th>No. states below 6 eV</th>
<th>$\omega_{40}/\text{eV}$</th>
<th>First $^{1}\pi\pi^*$ state</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.5</td>
<td>0</td>
<td>3</td>
<td>9.73</td>
<td>2 5.31 0.1396</td>
</tr>
<tr>
<td>2.0</td>
<td>4</td>
<td>5</td>
<td>8.66</td>
<td>2 5.25 0.1168</td>
</tr>
<tr>
<td>2.5</td>
<td>7</td>
<td>5</td>
<td>8.16</td>
<td>2 5.22 0.1707</td>
</tr>
<tr>
<td>3.0</td>
<td>15</td>
<td>8</td>
<td>7.59</td>
<td>2 5.15 0.1616</td>
</tr>
<tr>
<td>3.5</td>
<td>18</td>
<td>7</td>
<td>7.44</td>
<td>1 5.13 0.1041</td>
</tr>
<tr>
<td>4.0</td>
<td>25</td>
<td>8</td>
<td>7.15</td>
<td>1 5.09 0.1422</td>
</tr>
<tr>
<td>4.5</td>
<td>37</td>
<td>10</td>
<td>6.91</td>
<td>1 5.10 0.1624</td>
</tr>
</tbody>
</table>

$^a$Distance criterion for selecting the QM region

the QM region consists only of uracil, or of uracil plus only a few water molecules. There is also no clear trend in the direction of this shift.

The MM solvent region does have one tremendously important effect, however: it dramatically reduces the number of spurious CT states at all values of $d$, the distance threshold for selecting QM water molecules. With the addition of point charges, even a large QM region like $d = 4.5$ Å (which is 13–14 Å across, and contains 123 atoms) affords only 10 excited states within 6 eV of the ground state. Absent the TIP3P charges, the same QM region affords an estimated 60 states below 6 eV. Figure 2.5(a) plots the growth in the number of low-energy excited states as a function of $d$, for TD-PBE0 calculations with and without MM point charges. For gas-phase clusters the number of states rises rapidly with cluster size, but this growth is very sluggish in the presence of MM point charges.
Figure 2.5: (a) Number of excited states within 6 eV of the ground state, and (b) excitation energy of the 40th excited state, each as a function of the radius $d$ of the QM region.
Although MM point charges eliminate many low-energy CT states (for reasons explained below), Fig. 2.5(b) reveals a steady decrease in $\omega_{40}$ as a function of $d$, even for the QM/MM calculations, though the falloff is sharper in the absence of point charges. The decrease in $\omega_{40}$ indicates that the MM charges do not remove all spurious CT states, especially at higher excitation energies. While the QM/MM calculation at $d = 4.5$ Å yields only 10 states below 6 eV, there are another 30 states (mostly spurious) between 6.00 eV and 6.91 eV. On the other hand, a far greater number of spurious states appear in this energy régime when the MM charges are removed, and this is an important practical consideration, given that the number of excited states requested in a TD-DFT calculation determines the memory required for the Davidson iterations. In fact, we are unable to locate the second electronic absorption band in gas-phase uracil–(H$_2$O)$_3$ ($d = 4.5$ Å) due to the large number of states required. The total memory requirement for such a calculation exceeds 4 Gb, even when core orbitals are excluded from the TD-DFT excitation space.

Examination of the low-energy excited states for the $d = 2.5$ Å cluster—which may be compared to the $d = 2.5$ Å results of the previous section that are depicted in Fig. 2.3—provides a clue to the origin of this reduction in CT states. For the $n \to \pi^*$ and $\pi \to \pi^*$ excitations (states 3 and 5 in Fig. 2.3), we find that the attachment densities, detachment densities, and excitation energies are nearly unchanged by addition of the point charges. Each of the CT excitations (states 1, 2, 4, and 6 in Fig. 2.3) is also unchanged in its qualitative character, but is shifted to $\sim 1$ eV higher in energy. What was $\omega_1 = 4.48$ eV in the absence of point charges becomes $\omega_3 = 5.48$ eV in the
QM/MM calculation.

In fact, for the QM/MM calculations we find that the $n\pi^*$ and $\pi\pi^*$ states are always the first and second excited states, irrespective of the size of the QM region. (Interestingly, the order of these two states changes as a function of solvent configuration, something that could not have been deduced from polarizable continuum models.\textsuperscript{63}) Because there are few low-energy CT states, there is also no intensity-borrowing problem of the sort discussed in Section 2.3.1. (The dip in the $\pi\pi^*$ oscillator strength that is observed at $d = 3.5$ Å results from substantial intensity borrowing on the part of the $n\pi^*$ state.)

To understand why the point charges wield such an influence on CT excitation energies, recall that in the larger clusters of Section 2.3.1, only the water molecules on the surface of the cluster contribute to the lowest-energy CT states (see Fig. 2.2). Addition of the MM point charges has the effect of stabilizing the lone pair orbitals on these water molecules, lowering their Kohn–Sham eigenvalues and thereby increasing the excitation energy associated with water-to-uracil CT. Importantly, this stabilization is sufficient to remove low-energy CT states only in conjunction with a hybrid functional; TD-BLYP calculations are still beset by numerous CT states at low energies, even within a QM/MM framework.

To emphasize this point, Table 2.4 summarizes TD-BLYP calculations on our sequence of uracil–water clusters, both with and without point charges. (These data are plotted alongside TD-PBE0 results in Fig. 2.5). Although the MM solvent does reduce the number of states below 6 eV, the number of such states remains large, even
in the QM/MM calculations. Using BLYP, attempts to locate the second absorption band quickly become intractable as cluster size increases.

These observations clarify the results of Bernasconi, Sprik, and Hutter,\textsuperscript{20,21} who simulated electronic absorption spectra of aqueous acetone using plane-wave Car–Parrinello MD. Using BLYP, these authors find that the lowest valence ($n\pi^*$) band is buried beneath a much broader and more intense CT band, comprised of several spurious CT states. In contrast, the B3LYP and PBE0 functionals shift the CT band upwards by $\sim 1$ eV, well above the valence $n\pi^*$ band.\textsuperscript{21} Our results show that this is only partially attributable to the use of hybrid functionals. Equally important is

<table>
<thead>
<tr>
<th>$d/\text{Å}$</th>
<th>Includes TIP3P charges?</th>
<th>No. states below 6 eV</th>
<th>$\omega_{40}/$ eV</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.5</td>
<td>No</td>
<td>12</td>
<td>7.89</td>
</tr>
<tr>
<td>2.0</td>
<td>No</td>
<td>28</td>
<td>6.48</td>
</tr>
<tr>
<td>2.5</td>
<td>No</td>
<td>$&gt;40$</td>
<td>5.68</td>
</tr>
<tr>
<td>3.0</td>
<td>No</td>
<td>$&gt;40$</td>
<td>4.67</td>
</tr>
<tr>
<td>1.5</td>
<td>Yes</td>
<td>7</td>
<td>8.41</td>
</tr>
<tr>
<td>2.0</td>
<td>Yes</td>
<td>17</td>
<td>7.22</td>
</tr>
<tr>
<td>2.5</td>
<td>Yes</td>
<td>26</td>
<td>6.50</td>
</tr>
<tr>
<td>3.0</td>
<td>Yes</td>
<td>$&gt;40$</td>
<td>5.72</td>
</tr>
</tbody>
</table>

\textsuperscript{a}Denotes the size of the cluster or the size of the QM region

\textsuperscript{b}Point charges were used in some calculations to represent additional water molecules out to $d = 20.0$ Å

Table 2.4: Summary of TD-BLYP/6-31+G* calculations on uracil–water clusters.
the fact that Bernasconi et al. use plane-wave DFT (hence periodic boundary conditions), which means that there were no “surface” water molecules present that might contribute low-energy CT excitations. On the other hand, Hartree–Fock exchange is incredibly expensive to evaluate in a plane-wave basis. Gaussian-orbital-based electronic structure theory, in conjunction with MM point charges to model an extended solvent network, thus represents a useful, affordable alternative.

2.3.3 Electronic absorption spectra

To this point, all calculations have used geometries taken from the same MD snapshot, which allows us to discuss trends with respect to cluster size. Solvent and chromophore geometry, however, play important roles in modulating the excitation energies, modifying the order and relative intensities of the valence excitations at least, and possibly the CT excitations as well. In order to take these effects into account, we next discuss electronic absorption spectra simulated as averages over a total of 100 configurations extracted from an MD simulation at intervals of 1.0 ps. (A two-fold reduction in the number of configurations has a modest effect on the spectra, but the rough spectral envelopes appear to be converged with respect to configuration sampling.) As with the QM/MM calculations discussed above, water molecules near the uracil (according to criteria described below) are included in the QM region, while additional water molecules out to 20.0 Å are incorporated as TIP3P point charges.

Absorption spectra are obtained by constructing histograms of the TD-DFT excitation energies (using a bin width of 0.1 eV), wherein the excitations are summed
according to their oscillator strengths; these spectra are plotted, at both the TD-
B3LYP and TD-PBE0 levels, in Fig. 2.6. Electronic densities of states (also plotted
in the figure) are obtained in similar fashion, by assigning equal weight to each excited
state. Once again PBE0 predicts slightly fewer low-energy CT states but its overall
behavior is very similar to that of B3LYP.

Far more important is the size of the QM region, and we compare three crite-
ria for affecting the QM/MM separation. In Figs. 2.6(a) and 2.6(b)—representing
B3LYP and PBE0, respectively—the QM region consists only of uracil, and all water
molecules are modeled as point charges. In Figs. 2.6(c) and 2.6(d), the QM region
includes all water molecules having at least one atom within 2.5 Å of one of uracil’s
hydrogen bonding sites (i.e., the N–H hydrogen atoms and the C=O oxygen atoms).
This amounts to an average of 5.5 water molecules in the QM region and so we refer
to this case as the “microhydrated” QM region. Finally, in Figs. 2.6(e) and 2.6(f), the
QM region contains all water molecules having at least one atom within 5.0 Å of the
uracil center of mass, for an average of 18.7 water molecules in the QM region. This is
sufficient to form a full solvation shell around uracil, so we refer to these calculations
as the “full solvation shell” QM region.

In order to obtain sensible averages, it is important that each individual TD-DFT
calculation determine enough excited states to reach a given energy threshold, at
which the histograms will terminate. For the largest QM region, the first 40 excited
states consistently reach 6.8 eV, which we thus choose as our energy cutoff. For the
uracil-only and microhydrated QM regions, 10 and 20 excited states, respectively, are
Figure 2.6: Electronic absorption spectra (thick lines, scale on the left) and densities of states (thin lines, scale on the right) from QM/MM calculations at the TD-B3LYP/6-31+G* and TD-PBE0/6-31+G* levels. As described in the text, the calculations in (a) and (b) utilize a uracil-only QM region, (c) and (d) employ a microhydrated QM region, while (e) and (f) use a full solvation shell for the QM region. Dotted vertical lines show the positions of the first two band maxima in the experimental absorption spectrum of aqueous uracil (from Ref. 1).
required to reach 6.8 eV.

In the case of a uracil-only QM region [Figs. 2.6(a) and 2.6(b)], both the first and second absorption bands appear to be mostly free of CT contamination. It is not obvious \textit{a priori} that this should be the case, despite the fact that there are no explicit solvent molecules present to support long-range CT states, as there remains the possibility of anomalously-low Rydberg excitations. Using BLYP, one does in fact observe Rydberg states below 6 eV for a uracil-only QM region (see the $d = 1.5$ Å results in Table 2.4). For B3LYP and PBE0, such states are completely absent within the first electronic absorption band, as evident from the density of states, which has a value of approximately 4 states/eV over the span of the first absorption band, which is about 0.5 eV wide. On average, then, this band must consist of two states, namely, the first $n\pi^*$ and $\pi\pi^*$ states. This is consistent with multireference calculations for gas-phase uracil that find one $n\pi^*$ state and one $\pi\pi^*$ state below 5.5 eV.\textsuperscript{60} Within the second absorption band, the density of states ranges from 4–7 states/eV over a band that is about 1 eV wide, indicating that on average there are 5 or 6 states within this band. This is also consistent with the aforementioned multireference calculations, which find a total of six $n\pi^*$ and $\pi\pi^*$ states in the 5.5–7.0 eV range.\textsuperscript{60}

Examining next the results for the microhydrated QM region, Figs. 2.6(c) and 2.6(d), we see that the density of states within the first absorption band is largely unchanged and, importantly, decays nearly to zero around 5.5 eV, in between the first and second absorption bands. (The experimental spectrum also decays nearly to zero around 5.5 eV.\textsuperscript{1}) Unlike the case of a uracil-only QM region, however, the
density of states shows no sign of dropping in the tail of the second absorption band and, in the B3LYP case at least, appears to be increasing above 6.5 eV, even as the spectral intensity decays. Embedding the QM region in an MM solvent pushes the CT threshold up to about 6.0 eV, with many more spurious states above 6.5 eV.

Finally there is the QM region consisting of a full solvation shell, Figs. 2.6(e) and 2.6(f). Here, the threshold for observing a substantial number of CT states creeps down somewhat from the 6.0 eV observed above, and consequently the density of states no longer decays to zero at 5.5 eV. In addition, a small number of system configurations exhibit CT states around 4.5 eV, below the first absorption band.

Regarding the solvatochromatic shifts, we note that gas-phase TD-B3LYP/6-31+G* and TD-PBE0/6-31+G* calculations put the $\pi \rightarrow \pi^*$ excitation energy at 5.4 eV and 5.5 eV, respectively. Based on the absorption band maxima in Figs. 2.6(a) and 2.6(b), addition of a point-charge-only solvent environment induces, on average, a 0.2 eV redshift in this excitation energy. (For comparison, a polarizable continuum model induces a 0.1 eV redshift.\textsuperscript{63}) A further redshift of about 0.1 eV is obtained using a microhydrated QM region, but incorporating a full solvation shell into the QM region does not change the position of the band maximum, though it does slightly increase the relative intensity of the low-energy side of the distribution. Given that TD-DFT calculations of acetone–(H$_2$O)$_N$ clusters\textsuperscript{25} find that the lowest $n \rightarrow \pi^*$ excitation energy changes by $<$ 0.05 eV between $N = 10$ and $N = 250$, the remaining 0.2 eV solvatochromatic shift not recovered in our largest QM/MM calculations probably represents the intrinsic accuracy limit of TD-DFT.
Figure 2.7: TD-B3LYP/6-31+G* calculations on uracil–water clusters using (a) a microhydrated QM region, and (b) a full solvation shell in the QM region. Electronic absorption spectra (scale on the left) are plotted for both the QM region only (thick, solid line) and for the QM/MM calculation (dotted line). The thin, solid line is the density of states (scale on the right) for the QM-only cluster calculation.
Finally, we revisit the role of the MM point charges, this time in the context of configurationally-averaged absorption spectra and densities of states. Figure 2.7 compares the QM/MM absorption spectra to those obtained upon removal of the MM solvent. In the latter case, of course, the configurational averages contain a far larger number of spurious CT states, as is clear from the densities of states plotted in Fig. 2.7, which extend down to 4.0 eV for the microhydrated QM region and down to 3.0 eV for the larger QM region, whereas the QM/MM calculations have virtually no states below 4.8 eV (cf. Fig. 2.6). Nevertheless, absorption profiles obtained with and without the MM charges are quite similar. Partly this reflects the fact that most of the CT excitations are spectroscopically dark, but it furthermore reflects the fact that summing the oscillator strengths (as opposed, say, to finding the configurationally-averaged excitation energy of the most intense transition) gathers up the intensity of any CT states that may borrow intensity from the $\pi\pi^*$ state.

### 2.3.4 Truncation of the TD-DFT excitation space

Although an MM embedding provides a simple and affordable means to reduce CT contamination (and for large systems is a method of choice in its own right), in some cases one might be interested in a gas-phase cluster rather than a proper liquid. In this section, we examine a separate (though compatible) procedure, whereby CT states are eliminated by ansatz, by removing from the linear-response eigenvalue equation those occupied-to-virtual ($|i\rangle\rightarrow|a\rangle$) excitation amplitudes that correspond to long-range CT. Automated criteria for performing this truncation of the excitation space have been developed by Besley,\textsuperscript{24} whose procedure we adopt here. Truncation of the
excitation space has also been explored, within the context of plane-wave DFT, by Odelius et al.\textsuperscript{23}

According to Besley’s procedure,\textsuperscript{24} one first identifies a subset of the atoms as belonging to the chromophore, for which we choose all of the uracil atoms. Excitation amplitudes $x_{ia}$ are then removed unless the occupied Kohn–Sham orbital $|i\rangle$ contains a significant contribution from basis functions centered on chromophore atoms, as measured by the contribution that these basis functions make to the Mulliken population of $|i\rangle$. We denote the threshold contribution as $\kappa_{\text{occ}}$; if the chromophore-centered basis functions do not contribute at least $\kappa_{\text{occ}}$ electrons to $|i\rangle$, then the $x_{ia}$ are omitted, for all $a$. (Values given for $\kappa_{\text{occ}}$ in this work are total populations, including both spins.)

Besley\textsuperscript{24} suggests additional truncation based on a second threshold $\kappa_{\text{virt}}$, according to which the sum of squares of the MO coefficients, $\sum_\mu |c_{\mu a}|^2$, is used to measure the contribution that the set of chromophore-centered basis functions $\{|\mu\rangle\}$ makes to the virtual orbital $|a\rangle$. As a result of the diffuse functions present in our basis set, however, we find that this sum is quite similar for each of the low-lying virtual MOs, whether or not they are localized around the uracil molecule. One way to circumvent this problem is to employ a mixed basis set, eliminating diffuse functions on the chromophore so that uracil-centered basis functions no longer contribute significantly to virtual MOs localized on the solvent. In practice, we find that useful results can be obtained without any truncation of the virtual space, so we retain 6-31+G* for all atoms and truncate the excitation space based solely on the occupied orbital criterion.

Table 2.5 lists TD-PBE0 excitation energies for the first and second $\pi\pi^*$ states of
<table>
<thead>
<tr>
<th>$d/\text{Å}$</th>
<th>Includes TIP3P charges?</th>
<th>$\kappa_{\text{occ}}$</th>
<th>First $\pi\pi^*$</th>
<th>Second $\pi\pi^*$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>$\omega$/eV</td>
<td>State no.$^a$</td>
<td>$\omega$/eV</td>
</tr>
<tr>
<td>2.5</td>
<td>No</td>
<td>0.0</td>
<td>5.28</td>
<td>5</td>
</tr>
<tr>
<td>2.5</td>
<td>No</td>
<td>0.2</td>
<td>5.28</td>
<td>2</td>
</tr>
<tr>
<td>2.5</td>
<td>No</td>
<td>0.4</td>
<td>5.32</td>
<td>2</td>
</tr>
<tr>
<td>2.5</td>
<td>No</td>
<td>1.0</td>
<td>5.39</td>
<td>2</td>
</tr>
<tr>
<td>2.5</td>
<td>Yes</td>
<td>0.0</td>
<td>5.22</td>
<td>2</td>
</tr>
<tr>
<td>2.5</td>
<td>Yes</td>
<td>0.2</td>
<td>5.24</td>
<td>1</td>
</tr>
<tr>
<td>2.5</td>
<td>Yes</td>
<td>0.4</td>
<td>5.27</td>
<td>1</td>
</tr>
<tr>
<td>2.5</td>
<td>Yes</td>
<td>1.0</td>
<td>5.34</td>
<td>1</td>
</tr>
<tr>
<td>4.5</td>
<td>No</td>
<td>0.0</td>
<td>5.06</td>
<td>18</td>
</tr>
<tr>
<td>4.5</td>
<td>No</td>
<td>0.2</td>
<td>5.18</td>
<td>5</td>
</tr>
<tr>
<td>4.5</td>
<td>No</td>
<td>0.4</td>
<td>5.21</td>
<td>3</td>
</tr>
<tr>
<td>4.5</td>
<td>No</td>
<td>1.0</td>
<td>5.21</td>
<td>3</td>
</tr>
<tr>
<td>4.5</td>
<td>Yes</td>
<td>0.0</td>
<td>5.10</td>
<td>1</td>
</tr>
<tr>
<td>4.5</td>
<td>Yes</td>
<td>0.2</td>
<td>5.21</td>
<td>1</td>
</tr>
<tr>
<td>4.5</td>
<td>Yes</td>
<td>0.4</td>
<td>5.33</td>
<td>1</td>
</tr>
<tr>
<td>4.5</td>
<td>Yes</td>
<td>1.0</td>
<td>5.52</td>
<td>1</td>
</tr>
</tbody>
</table>

$^a$Indicates where the state appears in the TD-DFT excitation manifold

Table 2.5: TD-PBE0/6-31+G* excitation energies obtained using truncated excitation spaces.
two different uracil–water clusters, using several different values of $\kappa_{\text{occ}}$ ranging from $\kappa_{\text{occ}} = 0$ (a full excitation space) to $\kappa_{\text{occ}} = 1$ (the value used in Besley’s benchmark calculations$^{24}$). Results are presented both with and without MM point charges. In the smaller of these two clusters, uracil–$(\text{H}_2\text{O})_7$ ($d = 2.5$ Å), we find that the accuracy of the excitation energies degrades rather slowly as a function of $\kappa_{\text{occ}}$. Already at $\kappa_{\text{occ}} = 0.2$, all CT states below the second $\pi\pi^*$ state are eliminated, meanwhile no significant error is incurred in the $\pi\rightarrow\pi^*$ excitation energies. Errors of $\lesssim 0.04$ eV were also reported by Besley$^{24}$ using $\kappa_{\text{occ}} = 1.0$ and $\kappa_{\text{virt}} = 0.8$ for a formamide–$(\text{H}_2\text{O})_4$ cluster; larger clusters were not considered in that study.

The larger of the two clusters is uracil–$(\text{H}_2\text{O})_{37}$ ($d = 4.5$ Å), and in this case the valence excitation energies are more sensitive to the value of $\kappa_{\text{occ}}$. Even $\kappa_{\text{occ}} = 0.2$ engenders errors of 0.1 eV and 0.2 eV, respectively, in the first and second $\pi\rightarrow\pi^*$ excitation energies. As before, $\kappa_{\text{occ}} = 0.2$ is sufficient to remove the water-to-uracil CT states, therefore further increase of $\kappa_{\text{occ}}$ is of no benefit. The spurious states that remain involve water-to-water and uracil-to-water CT, and elimination of these states would require truncation based on the virtual orbitals.

To assess the accuracy of truncation over a range of geometries, we recalculate the optical spectrum for the large (full solvation shell) QM/MM calculations, using a truncation threshold of $\kappa_{\text{occ}} = 0.2$, and in Fig. 2.8(a) we compare this spectrum to that obtained using a full excitation space. Truncation produces virtually no change in the overall absorption envelope, except that it shifts the entire spectrum (both the first and second absorption bands) to slightly higher energy. The magnitude of
Figure 2.8: (a) Absorption spectra and (b) densities of states, from TD-PBE0/6-31+G* QM/MM calculations with a full QM solvation shell, using $\kappa_{\text{occ}} = 0.2$ (solid lines) and $\kappa_{\text{occ}} = 0.0$ (broken lines).
this overall shift is something less than the bin width of the histogram, 0.1 eV. (Test
calculations on smaller systems indicate that additional configurations are required in
order to achieve better than 0.1 eV resolution.) We conclude that truncation affords
a consistent level of accuracy across many system configurations.

The density of states for this calculation, Fig. 2.8(b), shows that truncation does
not remove any CT states within the first absorption band—these were removed
already by the introduction of MM point charges. As the excitation energy increases,
however, calculations in the full excitation space predict an increasingly large number
of spurious states, relative to results obtained with \( \kappa_{\text{occ}} = 0.2 \). The practical upshot
is that the latter calculations consistently require only 15 excited states to reach
6.8 eV (the energy cutoff in Fig. 2.8), whereas 30–40 states are required when a full
excitation space is employed. This represents an approximately two-fold reduction
in the memory required for the Davidson iterations, which is roughly proportional to
the number of excited states requested.

Truncation of the excitation space is similarly accurate for gas-phase clusters, as
demonstrated when we remove the MM point charges from the QM/MM calculations
discussed above. Figure 2.9(a) compares spectra obtained with \( \kappa_{\text{occ}} = 0.2 \) to those
calculated with a full excitation space, while Fig. 2.9(b) compares the corresponding
densities of states. In the absence of truncation, we obtain excited states all the way
down to 3 eV, and 40 excited states are required just to reach \( \omega = 5.4 \) eV. Thus the
spectra in Fig. 2.9(a) include only the first absorption band. When \( \kappa_{\text{occ}} = 0.2 \), this
energy cutoff is reach consistently with only the first 25 excited states. As before, the
Figure 2.9: (a) Absorption spectra and (b) densities of states, from TD-B3LYP/6-31+G* calculations of the full QM solvation shell in the absence of MM point charges, using $\kappa_{occ} = 0.2$ (solid lines) and $\kappa_{occ} = 0.0$ (broken lines).
spectrum calculated with the truncated excitation space is shifted to higher energy by $\lesssim 0.1$ eV, with little change in the overall absorption envelope.

Finally, we note that the success of orbital truncation as a means to reduce CT contamination is contingent upon use of a hybrid density functional. Even for a fairly small, $d = 3.0$ Å cluster, with MM point charges included and using a truncation threshold of $\kappa_{\text{occ}} = 0.2$, a TD-BLYP calculation yields 17 excited states below the first $\pi\pi^*$ state. Keeping $\kappa_{\text{occ}} = 0.2$ but omitting the point charges, CT becomes so prevalent that is impossible to discern the identity of the $\pi\pi^*$ state simply on the basis of oscillator strengths, as so much of this intensity has bled into the spurious CT states.

2.4 Summary and conclusions

TD-DFT calculations in both small and large molecular clusters are beset by a legion of spurious CT excitations at or below the lowest valence excitation energies, even when hybrid functionals such as B3LYP or PBE0 are employed. These spurious excitations proliferate rapidly as cluster size increases.

If only the optically-bright states are of interest, TD-DFT still affords useful information, with the proviso that one should sum oscillator strengths over states with comparable excitation energies, since near-degeneracies between a bright state and one or more spurious CT states may lead to anomalous intensity borrowing, robbing the bright state of intensity. If one is interested in optically-forbidden transitions, however, then the spurious CT states present a formidable problem, as they are often
difficult to distinguish from real excited states whose oscillator strengths are small. In any case, the spurious CT states introduce a severe memory bottleneck.

Our results demonstrate that CT contamination is substantially reduced (though not eliminated) using either of two simple measures, neither of which requires modification of existing density functionals. One option is to embed the system of interest within a larger electrostatic medium, via a QM/MM calculation. This procedure substantially reduces CT contamination by stabilizing occupied Kohn–Sham orbitals on the edge of QM region and, to a lesser extent, destabilizing virtual MOs at the QM/MM interface. (A dielectric or polarizable continuum model for the surroundings would probably have a similar effect.) The QM/MM embedding removes most low-energy CT states, even when the QM region is rather large (>120 atoms), but works only in conjunction with hybrid functionals. Although the same effect can be achieved without MM point charges by using plane-wave DFT with periodic boundary conditions, the plane-wave calculations are prohibitively expensive because hybrid functionals are still required.\textsuperscript{21}

As an alternative to, or in conjunction with, a QM/MM embedding, CT states can also be eliminated by removing certain excitation amplitudes from the TD-DFT linear response equations, according to an automated procedure.\textsuperscript{24} This procedure must be used with extreme caution, as it eliminates CT states (real or spurious) by ansatz. In cases where no real CT is expected, however, this technique significantly reduces the number of spurious states while introducing errors in the valence excitation energies that are typically smaller than 0.1 eV, at least for the examples considered here. Once
again, the success of this technique is contingent upon use of a hybrid functional. Using BLYP, serious CT contamination persists, despite either of the aforementioned measures.

Finally, we note that the aforementioned procedures are intended only to remove CT “contamination”, that is, the appearance of spurious CT states at low energies. Where real CT states are present (whose energies will of course be grossly underestimated by standard TD-DFT), the prescribed truncation of the excitation manifold will eliminate these as well. Electrostatic embedding, on the other hand, will modulate the energetics of real CT states, but it cannot be expected to compensate for the fundamentally incorrect way in which these states are described by contemporary TD-DFT. The simple procedures described here are therefore most applicable to studies of optically-bright, valence excitations in large molecular systems.
CHAPTER 3

Charge-transfer excited states in a π-stacked adenine dimer, as predicted using long-range-corrected time-dependent density functional theory

3.1 Introduction

Half a century ago, Eisinger concluded that the lowest excited states of oligonucleotides and DNA are excimer-like, with relatively strong interactions between chromophores in the excited state, including some degree of charge transfer (CT) between nucleobases. Fifty years later we are still debating the role of charge-transfer states in these systems. Much theoretical work on DNA employs so-called Frenkel exciton models, in which CT is neglected and excited states consist exclusively of linear combination of localized π→π* excitations. However, femtosecond transient absorption experiments have revealed the existence of long-lived, optically-dark “trap” states that are suggested to involve CT between nucleobases on the same DNA strand.

Detailed quantum-chemical calculations may ultimately shed light on this issue, but some problems must first be surmounted. One problem is system size: calculating excited states of a single purine base is already a taxing problem for correlated wavefunction methods. Although two- and three-base \( \pi \) stacks have been attacked with CASPT2 and CASSCF calculations, respectively\(^7\) serious compromises must be made regarding the choice of active space and basis set, and larger systems remain intractable. In contrast, the configuration interaction singles (CIS) method and time-dependent density functional theory (TD-DFT) can be applied to much larger systems, but already for \( \pi \)-stacked dimers, the predictions made by these methods are controversial. Low-lying CT states (\(< 4 \text{ eV above the ground state}\)) have been reported in two- and three-base \( \pi \) stacks on the basis of TD-DFT calculations,\(^7\) but subsequent CIS and semi-empirical ZINDO calculations failed to locate such states,\(^7,9\) and the TD-DFT results were thus ascribed to that method’s well-known proclivity toward severe underestimation of CT excitation energies.\(^2,18,27,29,30\)

The appearance of intra- and intermolecular CT states, whose excitation energies may be underestimated by 10 eV or more, is a serious problem in small-molecule TD-DFT calculations, but a \textit{catastrophic} problem in larger systems,\(^2,20,21,25,26\) where they ultimately form a near-continuum of spurious, optically-dark excited states.\(^2\) Thus, there are reasons to be suspicious of TD-DFT calculations in \( \pi \)-stacked nucleobase multimers (or hydrogen-bonded base pairs, for the same reason). Indeed, in one of the aforementioned TD-DFT studies of \( \pi \)-stacked dimers,\(^7\) the authors mention the appearance of low-lying Rydberg states when diffuse basis functions are employed. Like
CT states, Rydberg excitations are exquisitely sensitive to the asymptotic behavior of the exchange–correlation potential, and while Rydberg states may be unimportant in solution their appearance at low excitation energies in the gas phase is a symptom of a deeper problem. Omission of diffuse basis functions pushes the Rydberg states to higher energies, thus masking TD-DFT’s inherent inability to describe such states.

Over the past few years, several groups have pursued long-range-correction (LRC) techniques in order to ameliorate this problem.\textsuperscript{38,80–83} Recognizing that Hartree–Fock (HF) exchange affords the correct distance dependence for long-range CT excitations,\textsuperscript{29} these methods attempt to graft HF exchange onto local (and therefore short-range) DFT, without mucking up the delicate parameterization that yields high-quality results for properties that do not depend on the tail of the exchange–correlation potential. This is accomplished by means of a switching function (typically the error function) applied to the Coulomb operator. The short-range part of the exchange interaction is then treated with local DFT and the long-range part with full HF exchange. This modification rectifies intermolecular CT energies, at least for well-separated monomers in the gas phase.\textsuperscript{34}

In the present work, we apply this scheme to TD-DFT calculations of hydrated uracil clusters, a π-stacked adenine dimer in the gas phase, and the same π-stacked dimer in a water cluster.
3.2 Theory and methodology

We have implemented the LRC procedure described by Hirao and co-workers,\textsuperscript{80,81} for both pure and hybrid density functionals, within a developers’ version of Q-Chem.\textsuperscript{53} The LRC consists of using the error function \( \text{erf}(\mu r) \), where \( \mu \) is an adjustable parameter, to split the Coulomb operator into short- and long-range components:

\[
\frac{1}{r} = \frac{1 - \text{erf}(\mu r)}{r} + \frac{\text{erf}(\mu r)}{r},
\]

(3.1)

The first term on the right decays to zero on a length scale of \( \sim 1/\mu \), and the second term is a long-range background. Consider a density functional whose exchange–correlation energy can be written as

\[
E_{xc} = E_c + (1 - C_{\text{HF}})E_{x,\text{local}} + C_{\text{HF}} E_{x,\text{HF}},
\]

(3.2)

where \( C_{\text{HF}} \) represents the coefficient of HF exchange (\( e.g. \), \( C_{\text{HF}} = 0.20 \) for B3LYP and \( C_{\text{HF}} = 0.25 \) for PBE0). Then the LRC exchange–correlation energy for this functional is

\[
E_{xc}^{\text{LRC}} = E_c + (1 - C_{\text{HF}})E_{x,\text{local}}^{\text{SR}} + C_{\text{HF}} E_{x,\text{HF}}^{\text{SR}} + E_{x,\text{HF}}^{\text{LR}},
\]

(3.3)

where \( E_{x,\text{HF}}^{\text{SR}} \) and \( E_{x,\text{HF}}^{\text{LR}} \) denote the HF exchange energy evaluated using the short- and long-range parts of the Coulomb operator, respectively, and \( E_{x,\text{local}}^{\text{SR}} \) denotes the local exchange energy evaluated using the short-range part of the Coulomb operator. The correlation contribution \( E_c \) is not modified.
Exchange integrals over modified Coulomb operators, necessary to evaluate $E_{x,\text{HF}}^{\text{SR}}$ and $E_{x,\text{HF}}^{\text{LR}}$, were implemented in Q-Chem by Gill and co-workers.\textsuperscript{84} Short-range exchange is given by

$$E_{x,\text{local}}^{\text{SR}} = -\frac{1}{2} \sum_{\sigma} \int d\mathbf{r} \rho_{\sigma}^{4/3}(\mathbf{r}) K_{\sigma}^{\text{SR}}(\mathbf{r}) ,$$

where $K_{\sigma}^{\text{SR}}$ is a short-range version of the exchange kernel $K_{\sigma}$ for $\sigma$ spin, which has a particular form for each generalized gradient approximation (GGA). Hirao and co-workers\textsuperscript{81} provide a general prescription for obtaining $K_{\sigma}^{\text{SR}}$ from $K_{\sigma}$ for any GGA (see also Ref. 85).

We have implemented LRC versions of Becke88 exchange\textsuperscript{46} and Perdew-Burke-Ernzerhof (PBE) exchange,\textsuperscript{43} from which LRC variants of many common functionals can be constructed. We focus primarily on LRC-PBE0 in this work, since the PBE0 functional has been recommended for its accuracy in TD-DFT calculations\textsuperscript{47} and has been used to study $\pi$-stacked dimers of 9-methyladenine, where low-energy CT states were reported.\textsuperscript{86} All calculations discussed here utilize the 6-31+G* basis set and the SG-1 quadrature grid.\textsuperscript{87}

The TD-LRC-PBE0 method employed here has one adjustable parameter with two important limits: as $\mu \to 0$, the method reduces to TD-PBE0, and as $\mu \to \infty$ it reduces to an exchange functional of the form

$$E_{xc} = E_{c} + E_{x,\text{HF}} ,$$

which amounts to time-dependent HF (TD-HF) theory, also known as the random-phase approximation plus some additional DFT correlation energy, $E_{c}$. (The Tamm–Dancoff approximation is not used here.) In short, we have a method that can
smoothly interpolate between two previous calculations, which reached opposite conclusions regarding the presence of low-energy CT states in π-stacked nucleobase dimers.

We consider three test systems: hydrated-uracil clusters, taken from our previous work;\textsuperscript{2} adenine dimer (A\textsubscript{2}) in its B-DNA geometry (3.4 Å separation and 36° twist angle), without phosphate or sugar moieties; and an A\textsubscript{2}(H\textsubscript{2}O)\textsubscript{27} cluster. In the latter, the water molecules were equilibrated by an aqueous-phase molecular dynamics simulation at $T = 300$ K, constraining A\textsubscript{2} to retain its B-DNA geometry. All water molecules within 3 Å of A\textsubscript{2} were retained in the cluster. All geometries are available in Supporting Information.\textsuperscript{88}

<table>
<thead>
<tr>
<th>N</th>
<th>No. states below 6 eV</th>
<th>40th excitation energy (eV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>H\textsubscript{2}O</td>
<td>PBE0</td>
<td>LRC-PBE</td>
</tr>
<tr>
<td>0</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>6</td>
<td>2</td>
</tr>
<tr>
<td>7</td>
<td>11</td>
<td>2</td>
</tr>
<tr>
<td>15</td>
<td>19</td>
<td>2</td>
</tr>
<tr>
<td>18</td>
<td>20</td>
<td>2</td>
</tr>
<tr>
<td>25</td>
<td>29</td>
<td>2</td>
</tr>
<tr>
<td>37</td>
<td>59</td>
<td>2</td>
</tr>
</tbody>
</table>

Table 3.1: Measures of CT contamination in TD-DFT calculations of (uracil)(H\textsubscript{2}O)\textsubscript{N}. The TD-PBE0 data are from Ref. 2.
3.3 Results

Standard TD-DFT calculations in large molecules and clusters are plagued by a high density of spurious CT excitations, not just at long range but even at van der Waals contact distances. We have investigated this problem previously using (uracil)(H₂O)ₙ clusters, and in Table 3.1 we compare TD-PBE0 results from that study to TD-LRC-PBE results. The latter use μ = 0.47 a₀⁻¹ (where a₀ is the Bohr radius), a value recommended for the LRC-BOP functional. The LRC method predicts just two excited states below 6 eV (consistent with multireference calculations), regardless of cluster size. Notably, when using a non-hybrid functional such as PBE, and absent any LRC, we were previously unable to perform TD-DFT calculations beyond N = 15, due to the large number of states required to reach 6 eV.

Turning now to the A₂ and A₂(H₂O)₂₇ calculations that are the main topics of this work, Fig. 3.1 shows that TD-DFT vertical excitation energies for A₂ increase monotonically as a function of μ, consistent with the observation that the CIS and TD-HF methods (μ = ∞) tend to overestimate all excitation energies. For A₂, the TD-HF method predicts only a few states within ~7 eV of the ground state, but the density of states increases significantly as μ→0.

By inspecting excited-state Mulliken charges, TD-DFT excitation amplitudes, natural transition orbitals, and electron attachment/detachment densities, we are able to construct diabatic states along which the excited states change smoothly as
Figure 3.1: Vertical excitation energies for A₂, calculated at the TD-LRC-PBE0/6-31+G* level as a function of the range-separation parameter μ. The black lines are adiabatic excitation energies (alternating solid and broken, for clarity). The CT and ππ* diabatic states are also indicated [cf. Fig. 3.2(a)].
Figure 3.2: Diabatized vertical excitation energies as a function of the range-separation parameter $\mu$, for (a) stacked adenine dimer, $A_2$, and (b) $A_2(H_2O)_{27}$. The two monomers are labeled 3$'$ and 5$'$, and the $\pi\pi^*$($-$) and $\pi\pi^*$($+$) labels indicate linear combinations of localized $\pi\rightarrow\pi^*$ excitations (“Frenkel excitons”). Two $n\pi^*$ states are plotted in (a), but they are indistinguishable on this scale.
a function of $\mu$. Several such states are shown in Fig. 3.2, including a pair of intermolecular CT states (one for each possible direction of intermolecular CT), a pair of Frenkel exciton states (plus/minus linear combinations of localized $\pi\rightarrow\pi^*$ excitations), and a pair of localized $n\rightarrow\pi^*$ states. The $n\pi^*$ states, which are significantly blue-shifted by solvation, are not included for $A_2(H_2O)_{27}$ due to the high cost of calculating enough excited states to reach them, at small values of $\mu$. This difficulty is a direct consequence of the proliferation of low-lying excited states as $\mu \rightarrow 0$.

As expected, the CT states are far more sensitive to the value of $\mu$ than are the valence states. [Rydberg states are also very sensitive to $\mu$, but these appear above the valences states of $A_2$ at all values of $\mu$, even in the gas phase, and are thus omitted from Fig. 3.2(a) and from our discussion.] In the TD-PBE0 ($\mu \rightarrow 0$) limit, the CT excitation energies appear at or below 5 eV and represent the lowest-energy excitations in both $A_2$ and $A_2(H_2O)_{27}$. The water cluster stabilizes the CT states by $\sim 0.5$ eV (roughly independent of $\mu$) as compared to the gas phase, nevertheless a TD-HF calculation ($\mu = \infty$) places these states above 7 eV—to too high to be biologically relevant at the Franck–Condon geometry. This is consistent with a previous CIS study of $\pi$-stacked dimers of nucleobases and base analogues, which reported no evidence of CT in the lowest few excited states.\(^{78}\)

Figure 3.3 depicts electron attachment and detachment densities for the lower of the two CT states, at several values of $\mu$. For $\mu = 0$, this state corresponds to a net transfer of $\Delta q = 0.94$ electrons between adenine molecules, as quantified by natural population analysis.\(^{91}\) (Similar values of $\Delta q$ are obtained from Mulliken population
Figure 3.3: TD-LRC-PBE0 detachment and attachment densities for $A_2$ using (a) $\mu = 0$, (b) $\mu = 0.10 \ a_0^{-1}$, and (c) $\mu = 0.6 \ a_0^{-1}$. Each plot represents an isocontour that encompasses 70% of the total density. Also shown are the excitation energies and the amount of intermolecular CT ($\Delta q$) upon excitation, as quantified by natural population analysis.
Figure 3.4: TD-LRC-PBE0 diabatic potential energy curves for $A_2$ as a function of the intermolecular distance, using (a) $\mu = 0$ and (b) $\mu = 0.35 \ a_0^{-1}$. The two $n\pi^*$ potential curves overlap one another in both plots, and the CT states in (b) lie above 6.2 eV and are not shown.

analysis.) The CT nature of this state is clearly evident at $\mu = 0$, but the character of this state is more mixed at larger values of $\mu$.

In Fig. 3.4, we plot diabatic potential energy surfaces for gas-phase $A_2$, as a function of the monomer separation $R$, using $\mu = 0$ and also $\mu = 0.35 \ a_0^{-1}$. (The diabats in Fig. 3.2 correspond to vertical excitation at $R = 3.4 \ \text{Å}$. For $\mu = 0$ [TD-PBE0, Fig. 3.4(a)], the $S_1$ state is unambiguously a CT state at the Franck–Condon point, while the $S_1$ minimum appears at a smaller value of $R$—and is much deeper than—the shallow $S_0$ minimum. These two features are the classic signatures of an
excimer state, though we should note that both features are present also in the $S_1$ state at $\mu = 0.35 \, \text{\AA}^{-1}$ [Fig. 3.4(b)], for which $S_1$ is not a CT state but rather a $\pi\pi^*$ state. This state appears at higher energy than does the lowest $\pi\pi^*$ state for $\mu = 0$, because the LRC has the effect of pushing all of the excited states to higher energy (albeit the CT states more so than the valence states). For reasons that we do not yet understand, the LRC also considerably deepens the ground-state minimum (which is too shallow at the PBE0 level, due to lack of dispersion interactions), and also reduces the minimum-energy separation to 3.4 Å in the ground state.

### 3.4 Discussion

Variation of $C_{\text{HF}}$ in a standard (non-LRC) density functional [Eq. (3.2), with the PBE functional$^{43}$ used for both $E_c$ and $E_{x,\text{local}}$, for example] tends to shift all excitation energies upward as $C_{\text{HF}}$ increases, at roughly the same rate (see the Supporting Information)$^{88}$. This is to be contrasted with variation of $\mu$, which affects the CT (and Rydberg) states to a much greater extent than the valence states, since short-range exchange–correlation effects are largely preserved by the LRC but long-range exchange is fundamentally different. Furthermore, the character of the CT states changes somewhat as a function of $\mu$, as is evident from the attachment/detachment densities in Fig. 3.3. For $\mu \approx 0$, the lowest CT state involves transfer of $\gtrsim 0.9$ electrons from the HOMO of the supersystem (which is localized on one adenine monomer) to the LUMO of the supersystem (localized on the other monomer), and this one excitation amplitude accounts for $\sim 85\%$ of the norm of the TD-DFT eigenvector.
However, starting around $\mu \approx 0.1$ bohr$^{-1}$ (where many of the adiabats in Fig. 3.1 begin to shift rapidly as a function of $\mu$), the character of this nominal CT state becomes more mixed, even while it continues to involve significant CT between the two adenine molecules. In the $\mu \to \infty$ limit, no single excitation amplitude accounts for more than 12% of the norm of the TD-HF eigenvector.

Santoro et al.$^{86}$ have reported intermolecular CT states appearing just below 5 eV in a $\pi$-stacked dimer of 9-methyladenine, using TD-PBE0 in conjunction with a polarizable continuum model of aqueous solvation. This calculation is comparable to our $A_2(H_2O)_{27}$ calculation with $\mu = 0$. Our results demonstrate that the solvent stabilization is only partially responsible for the low energy of the CT states. Certainly, the solvent does preferentially stabilize these states, relative to the valence states, but the more important factor in their appearance at low energies is the tuning of the range parameter $\mu$.

Thus the crucial question is which $\mu$ value to use. For the LRC-BOP functional$^{89}$ (a re-parameterization of LRC-BLYP) and the LRC-\(\omega\)PBE functional$^{40}$ (a different formulation of LRC-PBE), optimization of $\mu$ with respect to ground-state benchmarks affords $\mu = 0.47$ \(a_0^{-1}\) and $\mu = 0.40$ \(a_0^{-1}\), respectively.$^{40,81}$ However, the statistically-optimal value of $\mu$ for ground-state properties may not be the best value for excitation energies, so we have optimized $\mu$ in LRC-PBE0 so as to reproduce the SAC-CI potential energy curve$^{34}$ for the lowest CT excitation in the $C_2H_4 \cdots C_2F_4$ heterodimer. (Energies and geometries from this calculation are available in Supporting Information.$^{88}$) Even if the SAC-CI benchmarks are all shifted by $\pm 0.2$ eV, this procedure
affords optimal $\mu$ values that lie between the two values suggested above. For $\mu$ in this range, the CT states in $A_2(H_2O)_{27}$ lie more than 1 eV above the $\pi\pi^*$ states, at least at the Franck-Condon point.

The present work makes clear that one must be extremely wary of TD-DFT predictions of intermolecular CT states, even at base-stacking distances. A more comprehensive study of this issue, including a simultaneous optimization of $\mu$ against both ground- and excited-state benchmarks, is currently underway in our group.
CHAPTER 4

Both intra- and interstrand charge-transfer excited states in aqueous B-DNA are present at energies comparable to, or just above, the \(^1\pi\pi^*\) excitonic bright states\(^{4.1}\)

4.1 Introduction

A detailed understanding of the photochemistry and photophysics precipitated by UV excitation of nucleobases, nucleotides, and oligonucleotides is crucial to elucidating not only the mechanisms that underlie DNA’s intrinsically high photostability,\(^{92}\) but also those mechanisms that do sometimes lead to DNA damage, including strand cleavage and photo-isomerization products.\(^{92,93}\)

UV excitation of the lowest \(^1\pi\pi^*\) state of an isolated nucleobase in aqueous solution results in radiationless decay back to \(S_0\) on a variety of time scales,\(^{94-98}\) the shortest of which is sub-picosecond,\(^{94,95}\) suggesting the existence of one or more conical intersections near the point of initial excitation. In deoxyribo-oligonucleotides, the initially-excited \(^1\pi\pi^*\) Frenkel exciton state decays on a time scale \(\tau_1 \sim 0.2-0.4\) ps (as

\(^{4.1}\)This chapter appeared as a full article in the *Journal of the American Chemical Society*, in 2009, volume 131, page 3913.
determined by time-resolved fluorescence experiments\textsuperscript{99–103}, but femtosecond time-resolved absorption spectroscopy reveal that ground-state recovery occurs much more slowly, on a time scale $\tau_2 \gtrsim 100$ ps.\textsuperscript{72,73,104–108} Kohler and co-workers\textsuperscript{73,104–107} attribute this discrepancy between absorption and fluorescence measurements to the existence of optically-dark “trap” states that are populated from an excitonic bright state on a time scale $\tau_1$, but survive for a time $\tau_2$ before decaying back to $S_0$. Recently, long-lived decay components have been observed also in time-resolved fluorescence experiments.\textsuperscript{109,110}

Based on the similarity between decay transients obtained for single- versus double-stranded oligonucleotides,\textsuperscript{73} as well as the correlation between ground-state recovery time constants $\tau_2$ and charge-transfer (CT) propensities (as estimated from the ionization energies and electron affinities of the constituent nucleobases),\textsuperscript{105} Kohler and co-workers hypothesize that the aforementioned trap states are excimers characterized by intrastrand CT between adjacent nucleobases.\textsuperscript{73,104–107} This experimental identification of \textit{intra}-strand (rather than \textit{inter}-strand) CT states is especially interesting in view of substantial evidence that the lowest $1\pi\pi^*$ state of the guanine–cytosine base pair (G:C) decays via interstrand CT coupled to proton transfer, in both the gas phase\textsuperscript{111–113} and in aqueous solution.\textsuperscript{114–117} (In particular, gas-phase calculations explain the experimental observation that the Watson–Crick tautomer of G:C exhibits a much shorter $S_1$ lifetime than other tautomers.\textsuperscript{118}) Together, these observations suggest that the excited-state dynamics of the nucleobases changes qualitatively in the presence of $\pi$-stacking interactions.
Ab initio quantum chemistry stands to play an important role in unraveling the complicated electronic structure of oligonucleotides, although for systems containing multiple nucleobases, ab initio calculations are primarily limited to the level of time-dependent density functional theory (TD-DFT). While computationally feasible in rather large molecules or clusters, TD-DFT calculations are often plagued by spurious, low-energy CT states. Consequently, standard TD-DFT is problematic (at best) for calculations involving large molecules, especially those containing multiple chromophores, and is essentially useless if one’s goal is to identify real CT states.

Fortunately, recently-developed long-range correction (LRC) procedures, which asymptotically incorporate nonlocal Hartree–Fock (HF) exchange into existing density functionals, have been shown to push spurious CT states out of the spectral region occupied by valence $n\pi^*$ and $\pi\pi^*$ states. The LRC introduces one adjustable parameter (namely, the length scale on which the theory switches from density-functional exchange to HF exchange), and CT excitation energies are exquisitely sensitive to the value of this parameter. Careful parameterization, however, can afford a functional that is just as accurate for CT excitations energies as it is for localized excitation energies.

In the present work, we examine the low-lying excited states of a variety of nucleic acid systems, primarily using TD-DFT, after first verifying that TD-DFT excitation energies compare reasonably well to results obtained at the CIS(D), CC2, and
CASPT2 levels, for nucleobase dimers. In larger systems, we find that TD-DFT calculations using standard (non-LRC) density functionals are saddled with pervasive CT contamination, especially if one attempts to include the sugar/phosphate backbone, solvent molecules, or counter-ions in the TD-DFT calculation. The LRC procedure, however, removes those CT states that are obviously spurious (e.g., $\text{PO}_4^-$ → Na$^+$ CT states at $\sim$4 eV above the ground state), yet both intra- and interstrand CT states remain, at energies comparable to—or only slightly above—the $^1\pi\pi^*$ excitation energies. Previous reports of CT states well below the $^1\pi\pi^*$ and $^1n\pi^*$ states$^{76,77,86,122}$ appear to be artifacts. On the other hand, uncorrelated configuration-interaction singles (CIS) calculations$^{78}$ place the CT states much too high in energy, well above the $^1\pi\pi^*$ states.

In aqueous B-DNA, our calculations invariably predict both inter- and intrastrand CT states at comparable excitation energies, regardless of how the LRC is parameterized. This suggests that a realistic model of the excited states of DNA must consider both base-stacking and base-pairing interactions. A single Watson–Crick base pair fails to capture the complexity of DNA’s electronic structure.

### 4.2 Computational details

In this work, we confine our attention to nucleic acid systems composed of adenine and thymine bases. Unless otherwise noted, the geometries of all DNA oligomers correspond to the canonical B-DNA conformation.$^{4,2}$ The sugar/phosphate backbone

$^{4,2}$Canonical B-DNA geometries were generated using the NUCLEIC program, which is part of the TINKER software package (TINKER, v. 4.2, @http://dasher.wustl.edu/tinker@).
is removed in most cases, and the glycosidic terminus capped with hydrogen, as this modification alters the excitation energies by only about 0.1 eV.

We model solvation through a mixed quantum mechanics/molecular mechanics (QM/MM) approach, using geometries obtained from molecular dynamics simulations at 298 K. In most cases, the nucleobases in these simulations were held rigid in their canonical B-DNA geometries, in which case the simulation amounts to an average over solvent geometries. (The sole exception is the simulation of adenine dinucleotide discussed in Section 4.5.3, in which the adenine monomers are allowed to relax.) Ten configurational snapshots, each separated by at least 5 ps, were extracted from the simulation for use in the QM/MM calculations. All water molecules within 2.5 Å of any nucleobase atom were included explicitly in the QM region, while water molecules up to 10 Å away were included as TIP3P point charges. For calculation of the absorption spectrum, this procedure appears to be converged with respect to the size of the QM region. Additional simulation details, including tests of various QM/MM models, can be found in the Supporting Information (Appendix A).

TD-DFT calculations are performed using two LRC variants of the generalized gradient approximation (GGA) of Perdew, Burke, and Ernzerhof (PBE),\textsuperscript{43} which we denote as LRC-ωPBE (as implemented in Ref. 123) and LRC-ωPBEh (as implemented in Ref. 121). These two functionals employ somewhat different range-separation procedures, and while this distinction is important for the simultaneous description of both ground-state thermochemistry and TD-DFT excitation energies, it makes little difference excitation energies alone.\textsuperscript{121} The important difference between these two
functionals is that the former contains no short-range HF exchange, and employs a Coulomb attenuation parameter $\omega = 0.3 \, a_0^{-1}$ that is optimal for excitation energies, in the absence of short-range HF exchange. The LRC-$\omega$PBEh functional contains 20% short-range HF exchange and uses $\omega = 0.2 \, a_0^{-1}$. Here, the “h” indicates that the functional is a hybrid, even at short range, and both $\omega$ and the fraction of short-range HF exchange have been optimized for TD-DFT excitation energies (and also for ground-state thermochemistry).

Benchmark studies for a variety of molecules indicate that both TD-LRC-$\omega$PBEh and TD-LRC-$\omega$PBE calculations, with the parameters specified above, exhibit mean errors of 0.2–0.3 eV for both valence and CT excitation energies (see the Supporting Information [Appendix A] as well as Ref. 121). In the Supporting Information, we also study the $\omega$-dependence of TD-LRC-$\omega$PBE and TD-LRC-$\omega$PBEh excitation energies for the $\pi$-stacked adenine dimer, which reveals that these functionals slightly overestimate localized excitation energies, while LRC-$\omega$PBEh slightly underestimates CT excitation energies. These errors will be taken into account when we present solution-phase absorption spectra in Section 4.5.

For comparison, we will also present results obtained from the widely-used PBE0 hybrid functional (also known as PBE1PBE). TD-PBE0 calculations afford mean errors of $\sim 0.3$ eV for localized valence excitation energies, but severely underestimates CT excitation energies.
Although basis set superposition error (BSSE) is extremely important in calculations of dimerization energies for \( \pi \)-stacked systems, simple estimates of the excited-state BSSE in the \( \pi \)-stacked cytosine dimer suggest that the \( S_0 \) and \( S_1 \) BSSEs differ by less than 0.03 eV,\textsuperscript{127} and therefore vertical excitation energies are hardly affected. Moreover, the aug-cc-pVDZ and aug-cc-pVTZ basis sets afford virtually identical results for excitation energies, suggesting that complete-basis extrapolation (whereupon BSSE should vanish) would not greatly affect the excitation energies. As we are interested in excitation energies rather than binding energies, no attempt is made here to correct for BSSE.

For nucleobase dimers, we compare TD-DFT excitation energies to those obtained using various correlated wave function models, including CC2,\textsuperscript{128} CIS(D),\textsuperscript{129} and a spin-component-scaled (SCS) version of CIS(D).\textsuperscript{130} In all three cases, we employ the resolution-of-identity (RI) version of the method.\textsuperscript{131,132} In principle, the highest level method employed here is CC2, and for single-reference problems with closed-shell ground states, benchmark studies indicate that CC2 excitation energies are typically within 0.3 eV of full configuration interaction,\textsuperscript{133,134} CASPT2,\textsuperscript{135} or experiment.\textsuperscript{136} In particular, CC2 results for each of the nucleobase monomers lie within 0.3 eV of CASPT2 results, for all \( n\pi^* \) and \( \pi\pi^* \) states within 7 eV of the ground state, and these CC2 excitation energies also compare favorably to “best estimates” obtained either from experiment or from higher-level \textit{ab initio} calculations.\textsuperscript{135} A recent study of of 4-(dimethylamino)benzonitrile found CC2 results for CT states to be in excellent agreement with higher-level methods that include triple excitations.\textsuperscript{137} In comparison,
CIS(D) excitation energies are of slightly lower quality, and typically fall within 0.2–0.5 eV of experiment,\textsuperscript{130,138} with slightly better performance for SCS-CIS(D).\textsuperscript{130}

All TD-DFT calculations employ the Tamm–Dancoff approximation,\textsuperscript{139} and only singlet excitations are considered. The LRC-DFT functionals have been implemented in a locally-modified version of Q-Chem,\textsuperscript{53} which we also use for the CIS(D) calculations. CC2 calculations were performed using Turbomole v. 5.9.\textsuperscript{140} To characterize the excited states, we examine attachment/detachment densities,\textsuperscript{61} natural transition orbitals (NTOs),\textsuperscript{90} and (unrelaxed) excited-state Mulliken charges.

For the CIS(D) and CC2 calculations, we employ the largest basis sets that are feasible given our available hardware, namely, 6-311+G* for CIS(D) and TZVP in the case of CC2. For the TD-DFT calculations, a careful study of the basis-set dependence (as reported in the Supporting Information [(Appendix A)] reveals that enlarging the basis set from 6-31G* to 6-31(2+,2+)G** tends to reduce the excitation energies by 0.1–0.4 eV but maintains both the ordering of the states and their relative oscillator strengths. As our interest in this work lies in the relative positions of the $\pi\pi^*$ and CT states, in large systems with multiple nucleobases and explicit solvent, most of our TD-DFT calculations employ the 6-31G* basis set.
4.3 Assessment of TD-DFT

4.3.1 Failure of standard functionals

Contemporary GGAs, and also hybrid functionals with less than 100% HF exchange, do not reproduce the correct asymptotic distance dependence for CT excitation energies. In TD-DFT calculations on large molecules, clusters, or liquids, this manifests as a near-continuum of spurious, low-energy CT states. Anomalously low CT excitation energies can be found between donor and acceptor orbitals that are localized on two molecules connected by a hydrogen bond, i.e., on length scales much shorter than the 3.4 Å base-stacking distance.

LRC density functionals attempt to rectify this problem by smoothly introducing HF exchange (and attenuating GGA exchange) in the asymptotic limit. This switching is accomplished by an Ewald-type partition of the Coulomb operator and governed by a range-separation parameter, \( \omega \), where \( 1/\omega \) represents the length scale for attenuation of GGA exchange. Based on thermochemical benchmarks, values of \( \omega \sim 0.3\text{–}0.4 \, a_0^{-1} \) have been proposed for this parameter. For TD-LRC-\( \omega \)PBE, \( \omega = 0.3 \, a_0^{-1} \) affords the best performance for excitation energies, and is therefore the value used in most of our calculations.

Table 4.1 compares vertical excitation energies, at several levels of theory, for the gas-phase adenine and thymine monomers as well as the Watson–Crick A:T base pair and the \( \pi \)-stacked adenine dimer, \( A_2 \), in its B-DNA geometry. We label the two low-energy \( ^1\pi\pi^* \) states of adenine as either “bright” (B) or “weak” (W), according to their relative oscillator strengths. In \( A_2 \), these states are “Frenkel excitons”,

68
<table>
<thead>
<tr>
<th>Excited State</th>
<th>Method</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PBE0&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>Adenine Monomer</td>
</tr>
<tr>
<td>nπ*</td>
<td>5.19</td>
</tr>
<tr>
<td>ππ* (W, L&lt;sub&gt;b&lt;/sub&gt;)</td>
<td>5.46</td>
</tr>
<tr>
<td>ππ* (B, L&lt;sub&gt;a&lt;/sub&gt;)</td>
<td>5.56</td>
</tr>
<tr>
<td></td>
<td>Thymine Monomer</td>
</tr>
<tr>
<td>nπ*</td>
<td>5.26</td>
</tr>
<tr>
<td>ππ*</td>
<td>5.59</td>
</tr>
<tr>
<td></td>
<td>A:T Base Pair</td>
</tr>
<tr>
<td>Thy nπ*</td>
<td>4.87</td>
</tr>
<tr>
<td>Ade nπ*</td>
<td>5.35</td>
</tr>
<tr>
<td>Thy ππ*</td>
<td>5.14</td>
</tr>
<tr>
<td>Ade ππ* (W)</td>
<td>5.41</td>
</tr>
<tr>
<td>Ade ππ* (B)</td>
<td>5.52</td>
</tr>
<tr>
<td>Ade → Thy CT</td>
<td>4.45</td>
</tr>
<tr>
<td></td>
<td>A&lt;sub&gt;2&lt;/sub&gt; π-Stacked Dimer</td>
</tr>
<tr>
<td>5′ nπ*</td>
<td>5.15</td>
</tr>
<tr>
<td>3′ nπ*</td>
<td>5.16</td>
</tr>
<tr>
<td>ππ* (W−)</td>
<td>5.30</td>
</tr>
<tr>
<td>ππ* (W+)</td>
<td>5.43</td>
</tr>
<tr>
<td>ππ* (B−)</td>
<td>5.51</td>
</tr>
<tr>
<td>ππ* (B+)</td>
<td>5.56</td>
</tr>
<tr>
<td>3′-Ade → 5′-Ade CT</td>
<td>4.95</td>
</tr>
<tr>
<td>5′-Ade → 3′-Ade CT</td>
<td>5.09</td>
</tr>
<tr>
<td>MAD (nπ* and ππ*)&lt;sup&gt;c&lt;/sup&gt;</td>
<td>0.11</td>
</tr>
<tr>
<td>MAD (CT)&lt;sup&gt;d&lt;/sup&gt;</td>
<td>1.35</td>
</tr>
</tbody>
</table>

<sup>a</sup>6-311+G<sup>*</sup> basis  
<sup>b</sup>TZVPP basis  
<sup>c</sup>Mean absolute deviation with respect to CC2, for the 16 valence excitations  
<sup>d</sup>Mean absolute deviation with respect to CC2, for the 3 CT excitations

Table 4.1: Vertical excitation energies (in eV) for the low-lying singlet excited states of adenine (A), thymine (T), and dimers thereof, in their canonical B-DNA geometries.
i.e., symmetric (“+” in Table 4.1) or antisymmetric (“−”) linear combinations of the localized $^{1}\pi\pi^{*}$ excitations on either monomer. For the monomers, the excitation energies at all levels of theory are 0.3–0.5 eV higher than those obtained using multireference calculations,\textsuperscript{141–145} most of which are in reasonable agreement with experimental band maxima in the gas phase.\textsuperscript{146,147} Differences between these multireference calculations and the excitation energies reported in Table 4.1 reflect both differences in the monomer geometries (we have not optimized the geometries for the gas phase), as well as the comparatively lower quality of our methods. The important comparison for our purposes is how TD-DFT compares to CIS(D) and CC2 at the geometries used here.

For the $\pi\pi^{*}$ and $n\pi^{*}$ excitations, PBE0 is in good agreement with the wave function methods, but this functional places CT states in $A_{2}$ more than 1 eV below what is predicted by wave function methods. In A:T, PBE0 places CT states more than 2 eV below what the wave function methods predict. This seems suspicious, and leads us to consider the lowest intermolecular CT excitation energy of two gas-phase adenine monomers separated by $R = 20 \text{ Å}$. The lowest intermolecular CT excitation energy, $\Delta E_{\text{CT}}$, is bounded below according to\textsuperscript{29}

$$\Delta E_{\text{CT}} > \text{IP} + \text{EA} - 1/R \quad (4.1)$$

(in atomic units), where IP and EA denote the ionization potential and electron affinity of adenine, respectively. We can thus estimate $\Delta E_{\text{CT}}$ using either experimental data (IP = 8.45 eV,\textsuperscript{148} EA = −0.54 eV\textsuperscript{149}) or high-level \textit{ab initio} data (IP = 8.37 eV,\textsuperscript{150} EA = −0.91 eV\textsuperscript{151}) for gas-phase adenine. These estimates for
\[ \Delta E_{\text{CT}} \] are presented in Table 4.2, alongside the results of \textit{ab initio} calculations on \( A_2 \) at \( R = 20 \, \text{Å} \).

<table>
<thead>
<tr>
<th>Method</th>
<th>( \Delta E_{\text{CT}} / \text{eV} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>TD-PBE0</td>
<td>5.55</td>
</tr>
<tr>
<td>TD-LRC-( \omega )PBE</td>
<td>8.81</td>
</tr>
<tr>
<td>TD-LRC-( \omega )PBEh</td>
<td>8.35</td>
</tr>
<tr>
<td>CIS(D)</td>
<td>8.36</td>
</tr>
<tr>
<td>SCS-CIS(D)</td>
<td>8.45</td>
</tr>
<tr>
<td>Eq. (4.1), expt. data</td>
<td>7.19</td>
</tr>
<tr>
<td>Eq. (4.1), \textit{ab initio} data</td>
<td>6.74</td>
</tr>
</tbody>
</table>

Table 4.2: Vertical excitation energies (using the 6-311G* basis) for the lowest intermolecular CT state between two adenine monomers (MP2/6-311++G** geometries) separated by 20 Å and given a twist angle of 36°, as in B-DNA.

Comparison to CIS(D) calculations reveals that Eq. (4.1), which overestimates the Coulomb stabilization of the ionized monomers, is not a particularly close lower bound, nevertheless the TD-PBE0 result for adenine \( \rightarrow \) adenine CT lies substantially below the minimum value established by Eq. (4.1). The two LRC functionals, in contrast, provide fairly good agreement with CIS(D) results for \( \Delta E_{\text{CT}} \).

In larger DNA oligomers, where TD-DFT is the only feasible \textit{ab initio} method, CT contamination proliferates rapidly as the length of the oligomer increases, when standard functionals such as PBE0 are employed. This is illustrated in Fig. 4.1, which plots the number of CT states appearing below the brightest \( \pi\pi^* \) state in the first absorption band, for a sequence of single-stranded \( A_n \) homologues of increasing
Figure 4.1: Plot of the number of CT states below the brightest $^1\pi\pi^*$ exciton state, computed at the TD-PBE0/6-31G* level for a sequence of single-stranded adenine multimers in their canonical B-DNA geometries, with backbone atoms removed.
length. At the TD-PBE0 level, the number of low-energy CT states increases as $\sim n^2$, whereas the LRC-ωPBEh functional, for example, predicts no CT states below the bright exciton state, in any of these $A_n$ systems. The growth in CT states predicted by PBE0 can be understood by noting that most of the predicted CT states—regardless of the functional that is employed—involves donor and acceptor orbitals that are localized on individual nucleobases. If there are $n$ nucleobases then there are $n(n - 1)$ ways to transfer an electron from one base to another, and a close inspection of Fig. 4.1 reveals that the TD-PBE0 method puts essentially all monomer-to-monomer CT states below the bright $\pi\pi^*$ state. This includes states such as the one depicted in Fig. 4.2, in which an electron in $A_7$ is transferred from one end of the strand to the other.

Although these calculations correspond to gas-phase $A_n$, Improta and co-workers have examined $A_n$ multimers at the TD-PBE0 level, using a polarizable continuum model of aqueous solvation. They, too, report low-lying CT states between non-adjacent nucleobases, and they raise the possibility that these states might be artifacts of the method. The absence of such states in TD-LRC-DFT calculations confirms this suspicion.

When backbone atoms are introduced, TD-PBE0 calculations lose all semblance of plausibility. As an example, we consider the dinucleotide (ApA)$^-$, where the “p” denotes the phosphate/sugar backbone, again in its canonical B-DNA geometry. TD-PBE0 predicts 18 CT states below the first bright state in this system (see Table 4.3), 14 of which involve a significant amount of $\text{PO}_4^- \rightarrow$ adenine CT. The situation is
<table>
<thead>
<tr>
<th>Excited State</th>
<th>PBE0</th>
<th>LRC-ωPBE</th>
<th>LRC-ωPBEh</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gas Phase (ApA)^−</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Brightest $\pi\pi^*$</td>
<td>5.72</td>
<td>5.93</td>
<td>5.86</td>
</tr>
<tr>
<td>Lowest $\text{PO}_4^−$ → Ade CT</td>
<td>3.93</td>
<td>6.11</td>
<td>5.77</td>
</tr>
<tr>
<td>Lowest Ade → Ade CT</td>
<td>5.17</td>
<td>6.41</td>
<td>6.12</td>
</tr>
<tr>
<td>No. CT states below bright state</td>
<td>18</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

| Na^+(ApA)^−(H_2O)_{17}        |      |          |           |
| Brightest $\pi\pi^*$          | >5.26| 5.83     | 5.75      |
| Lowest Ade → Ade CT           | 4.78 | 6.03     | 5.70      |
| No. CT states below bright state | >20 | 1        | 2         |

Table 4.3: Vertical excitation energies (in eV) for adenine (Ade) dinucleotide computed at the TD-DFT/6-31G* level.
Figure 4.2: Natural transition orbitals of a spurious end-to-end CT state appearing at 5.6 eV above the ground state, 0.1 eV below brightest $\pi\pi^*$ state. The calculation is performed at the TD-PBE0/6-31G* level, for $\pi$-stacked A$_7$ multimer. This particular particle/hole pair represents 99.9% of the transition density for the state in question.

even worse when nearby water molecules and a Na$^+$ counter-ion are included in the calculation [which we accomplish using a Na$^+$(ApA)$^-(H_2O)_{47}$ cluster], in which case each of the first 20 excited states is a CT state in which a water molecule on the surface of the cluster either donates or accepts an electron. (In this example, the solvent stabilizes the phosphate orbitals to such an extent that PO$_4^-$ $\rightarrow$ adenine CT states are not observed among the first 20 excited states.)

As we have shown previously,$^2$ spurious CT states cluster calculations can be reduced in number—though not totally eliminated—by addition of MM point charges beyond the QM region, and presumably also by polarizable continuum models, for the same reason. However, such techniques do not rectify the underlying problem, namely,
a qualitatively incorrect description of long-range CT by standard density functionals.
Standard functionals cannot be used to explore the complex DNA systems of interest here.

4.3.2 TD-LRC-DFT

Examining Table 4.1, we see that both LRC functionals afford TD-DFT excitation energies in good agreement with CC2 theory (in principle, the highest level of calculation reported here), with an average deviation from CC2 of just more than 0.1 eV. For the $n\pi^*$ and $\pi\pi^*$ states, the TD-PBE0 method is in similar agreement with the CC2 results, but this method underestimates each of the CT excitation energies by more than 1 eV. For the LRC-ωPBEh functional, we note that the largest deviations occur for the two CT states in $A_2$, which are each underestimated (compared to CC2) by about 0.3 eV, although the TD-LRC-ωPBE results are in excellent agreement with CC2. In a recent study of the LRC-ωPBEh functional,\textsuperscript{121} the statistical error in vertical excitation energies was found to be about 0.3 eV (with respect to CASPT2 and other high-level benchmarks), for both valence excitations and CT excitations. The present results are in line with this, if we accept the accuracy of the CC2 benchmarks, and thus it seems safe to proceed to larger systems using these functionals. We note that in complex systems such as Na$^+$(ApA)$^-$($H_2O$)$_{47}$, the LRC functionals move essentially all of the CT states to energies comparable to, or higher than, the $\pi\pi^*$ states comprising the first absorption band (see Table 4.3). The same is true in $\pi$-stacked $A_n$. 

76
4.4 Base stacking and base pairing effects

4.4.1 Absorption spectra

Consistent with other calculations, we find that the first absorption band of adenine monomer is comprised of two closely-spaced $^1\pi\pi^*$ states, with the higher-energy state ("$^1L_a$") possessing the larger oscillator strength. In oligomers, the Coulomb coupling between localized $^1\pi\pi^*$ excitations on individual nucleobases leads to delocalized "excitonic" states, and Platt’s $L_a/L_b$ notation loses its meaning. When we need a nomenclature for the $^1\pi\pi^*$ exciton states, we will label them as simply "bright" (B) or "weak" (W), in reference to their relative oscillator strengths.

In the Supporting Information (Appendix A), we show how the gas-phase excitation spectra evolve in $A_n$, $T_n$, and $A_n:T_n$, from $n = 1$ to $n = 4$, when the monomers are assembled in the B-DNA configuration. As the number of stacked bases increases, the higher-energy bright state shifts to slightly higher energy, while the weaker, lower-lying $\pi\pi^*$ states shift slightly to the red. The net result is a blue-shifted absorption spectrum (relative to the monomer’s spectrum) with a red tail, consistent with experimental steady-state absorption measurements. These effects arise due to coupling introduced by base stacking, which causes higher-energy, symmetric combinations of localized $\pi\pi^*$ excitations to exhibit larger oscillator strengths than lower-energy, antisymmetric combinations, leading to the observed red tail.

In fact, all of the qualitative changes in the absorption spectrum of $A_n$, relative to the monomer spectrum, can be rationalized from these crude gas-phase calculations, by applying a gaussian broadening to the gas-phase stick spectra and weighting the
Figure 4.3: Absorption spectra for A$_2$ (broken curves) and for adenine monomer (solid curves), computed by applying a 0.3 eV gaussian broadening to the gas-phase vertical excitation energies, weighted by their respective oscillator strengths. The monomer spectra are labeled as “2*A” to indicate that these oscillator strengths are weighted by an additional factor of two. Both calculations employ the 6-31G* basis set and use canonical B-DNA geometries.
excitation energies according to their oscillator strengths. The resulting spectra for adenine monomer and dimer are shown in Fig. 4.3. The aforementioned blue shift in the peak absorption intensity is clearly evident, as is the (very slight) red tail, and the hypochromic effect, i.e., the decrease in absorption intensity engendered by π-stacking.

Spectra are computed in Fig. 4.3 at both the TD-PBE0 and TD-LRC-ωPBE levels. Apart from an overall solvatochromatic red shift of \( \sim 0.2 \) eV, the TD-PBE0 spectra in are quite similar to those computed previously (at the same level of theory) for 9-methyladenine and its π-stacked dimer, using a polarizable continuum model of aqueous solvation in conjunction with the same broadening procedure employed here. In these earlier studies, the low-energy tail of the \( A_2 \) spectrum was attributed to weakly-absorbing CT states, but our own TD-PBE0 calculations put the oscillator strengths of the CT excitations at \(< 0.005\), and as such these states do not contribute to the spectra shown in Fig. 4.3. (At the TD-LRC-ωPBE level, the CT oscillator strengths are only 0.01, as compared to π\( \pi^* \) oscillator strengths as large as 0.41.) This demonstrates that it is not necessary to invoke low-energy CT states in order to explain stacking-induced changes in the absorption spectrum; excitonic coupling suffices.

Lastly, we note that excitons in \( A_n:T_n \) oligomers tend to be localized on a single strand, as shown for example in Fig. 4.4(a), owing to the energy-gap dependence of the excitonic coupling and the mismatch between adenine and thymine monomer
Figure 4.4: Natural transition orbitals (NTOs) corresponding to the excited state with largest oscillator strength in (a) A₃:T₃ and (b) ATA:TAT. In (a), the exciton is localized almost entirely on the adenine strand and consists of two significant NTO particle/hole pairs, whereas the bright state in ATA:TAT is mostly a localized monomer-like excitation.
excitations. This result, which arises naturally from TD-DFT calculations, is consistent with results obtained from exciton model Hamiltonians.\textsuperscript{71,154}

4.4.2 Charge-transfer states

For the two LRC functionals discussed in this work, the lowest adenine → thymine CT state in gas-phase A:T appears at 6.50 eV above the ground state, while the lowest thymine → adenine interstrand CT state is not among the first 20 excited states, placing it at least 7.8 eV above the ground state. These observations can be rationalized based upon the IPs and EAs of adenine and thymine,\textsuperscript{148–151} which strongly favor adenine as the electron donor and thymine as the acceptor. We do not observe thymine → adenine CT in the energy range of interest here.

In duplex A\textsubscript{n}:T\textsubscript{n} systems, we observe interstrand CT states in which the particle and/or the hole is delocalized over several nucleobases, as depicted in Fig. 4.5. In certain cases, including those shown in the figure, such states may involve adenine → thymine CT between two bases that are not hydrogen-bonded to one another \textit{i.e., from the \textit{k}th adenine on one strand to the (\textit{k} ± 1)st thymine on the opposite strand}. In larger multimers, delocalization appears to be the norm; for \textit{n} ≥ 3, we rarely observe localized, interstrand CT states.

4.4.3 Alternating base sequences

The discussion so far has been limited to multimers constructed from homopolymers of adenine and thymine. Here, we briefly investigate the excited states found in alternating sequences of adenine and thymine. Owing to energy mismatch between
Figure 4.5: Attachment densities (in blue) and detachment densities (in red) for delocalized, interstrand CT states in A$_3$:T$_3$ and A$_4$:T$_4$, in which there is significant CT between nucleobases that are not hydrogen bonded. Each calculation was performed at the TD-LRC-$\omega$PBE/6-31G* level, and the surfaces shown encapsulate 90% of density. Excited-state Mulliken charges on each monomer are also provided. (The monomers are essentially neutral in the ground state.)
Figure 4.6: Natural transition orbitals (NTOs) corresponding to the state with the largest oscillator strength in (a) $A_5$ and (b) ATATA. The two particle/hole NTO pairs with largest amplitude are shown in either case. In $A_5$, the exciton state couples $\pi\pi^*$ excitations that are four bases removed in sequence, and furthermore displays a characteristic nodal structure resulting from a linear combination of localized $\pi\pi^*$ excitations. As such, there are several significant excitation amplitudes, even in the NTO basis. In contrast, the bright state in ATATA is predominantly localized on a single adenine monomer, and is well-described by a single NTO particle/hole pair.
the monomer $\pi\pi^*$ excitations of the two nucleobases, the strong excitonic coupling observed in $A_n$ and $T_n$ is absent in $(AT)_n$, and consequently the bright states in the latter system are localized excitations. This is illustrated for ATATA in Fig. 4.6. Intrastrand CT states in this single-stranded heteropolymer are much closer in energy to the brightest adenine $\pi\pi^*$ states than was seen in $A_n$. (Excitation energies for ATATA are available in the Supporting Information [Appendix A].)

Hybridization of ATA with TAT, to form the duplex ATA:TAT, leads to some coupling between nucleobases on opposite strands, as evident from the NTOs depicted in Fig. 4.4(b), but the coupling is weak enough that it does not induce formation of delocalized interstrand CT states, as were observed in $A_n:T_n$. Interstrand adenine $\rightarrow$ thymine CT on the central base pair of ATA:TAT appears at 6.5 eV, about 0.3 eV higher than in $A_3:T_3$, and nearly equivalent to the CT excitation energy in A:T. This observation underscores the stability of delocalized, interstrand CT in $A_n:T_n$, which results primarily from the delocalized character of the virtual orbitals along the thymine strand. In ATA:TAT, energy mismatch of the monomer orbitals precludes such delocalization, and in ATA:TAT we see no evidence of delocalized, interstrand CT states within 6.7 eV of the ground state. *Intra*-strand CT states, on the other hand, appear at energies just below the adenine absorption peak (see the Supporting Information [Appendix A]), as observed also in ATATA.

The experimental results of Crespo-Hernández *et al.*\textsuperscript{73} suggest that CT states are formed in high yield in $(AT)_{n/2}:(TA)_{n/2}$, just as they are in single-stranded $A_n$ and double-stranded $A_n:T_n$, however the lifetime of these states is a factor of two
shorter in the case of alternating sequences.\textsuperscript{73} In addition, the experimental results of Buchvarov \textit{et al.}\textsuperscript{108} suggest that \((\text{AT})_{n/2}:(\text{TA})_{n/2}\) oligomers do not form delocalized excitons to the same extent as in the homopolymers. Our data directly support the latter suggestion, and are at least consistent with the observations of Crespo-Hernández \textit{et al.} Weak excitonic coupling between adenine and thymine prevents the formation of delocalized CT states in ATA:TAT.

### 4.4.4 CIS(D) results for A\textsubscript{2}:T\textsubscript{2}

<table>
<thead>
<tr>
<th>Excited State</th>
<th>CIS(D)</th>
<th>SCS-CIS(D)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thy $\pi\pi^*$ (−)</td>
<td>5.19</td>
<td>5.05</td>
</tr>
<tr>
<td>Thy $\pi\pi^*$ (+)</td>
<td>5.34</td>
<td>5.21</td>
</tr>
<tr>
<td>Ade $\pi\pi^*$ (W−)</td>
<td>5.28</td>
<td>5.10</td>
</tr>
<tr>
<td>Ade $\pi\pi^*$ (W+)</td>
<td>5.38</td>
<td>5.18</td>
</tr>
<tr>
<td>Ade $\pi\pi^*$ (B−)</td>
<td>5.55</td>
<td>5.27</td>
</tr>
<tr>
<td>Ade $\pi\pi^*$ (B+)</td>
<td>5.58</td>
<td>5.45</td>
</tr>
<tr>
<td>Ade $\rightarrow$ Ade CT</td>
<td>6.00</td>
<td>6.00</td>
</tr>
<tr>
<td>Thy $\rightarrow$ Thy CT</td>
<td>6.34</td>
<td>6.40</td>
</tr>
<tr>
<td>Ade $\rightarrow$ Thy CT (localized)</td>
<td>6.41</td>
<td>6.36</td>
</tr>
<tr>
<td>Ade $\rightarrow$ Thy CT (delocalized)</td>
<td>6.10</td>
<td>6.20</td>
</tr>
</tbody>
</table>

Table 4.4: CIS(D)/6-311+G* vertical excitation energies (in eV) for the low-energy excited states of A\textsubscript{2}:T\textsubscript{2}.

Although we are unable to extend the CC2 calculations beyond two nucleobases, the CIS(D) calculations can be extended to slightly larger systems, and in Table 4.4
we list the low-energy excited states of $A_2\cdot T_2$ computed at the CIS(D) level. Consistent with the results for $A_2$ and $A:T$ (Table 4.1), we find intrastrand adenine $\rightarrow$ adenine CT states 0.4–0.5 eV above the brightest excitonic states, with interstrand CT states a bit higher. [Note that the comparison to CC2 in Table 4.1 indicates that CIS(D) may overestimate the adenine $\rightarrow$ thymine CT excitation energies.] Perhaps the most interesting feature of these $A_2\cdot T_2$ calculations is the existence of an adenine $\rightarrow$ thymine CT state that is delocalized over all four nucleobases, which appears only 0.1–0.2 eV above the adenine $\rightarrow$ adenine CT state and well below another adenine $\rightarrow$ thymine CT state that is localized on a single base pair. This example demonstrates how the electronic structure may change significantly as the model system is extended beyond two nucleobases.

The overall picture that emerges from these CIS(D) calculations is consistent with that obtained from TD-LRC-DFT. For $A_2\cdot T_2$, both methods predict small red shifts, relative to $A_2$, in the adenine-localized $\pi\pi^*$ exciton energies as well as the lowest adenine $\rightarrow$ adenine CT excitation. At the CIS(D) level, base stacking stabilizes the interstrand adenine $\rightarrow$ thymine CT state by 0.6 eV relative to the unstacked $A:T$ base pair, and stabilizes the delocalized adenine $\rightarrow$ thymine CT state to an even greater extent. In the gas phase and using the canonical B-DNA geometry, it appears that the CT states are are significantly higher in energy than the $\pi\pi^*$ exciton states comprising the first absorption band. It is also apparent, however, that whenever interstrand CT states are found, intrastrand CT states are always present at comparable energies.
4.5 Multimers in aqueous solution

Up to this point, we have considered only gas-phase multimers, in order to identify the electronic effects of base stacking and pairing, as distinct from solvent interactions. We next consider solution-phase multimers, using a QM/MM model of aqueous solvation that treats the first solvation shell with DFT and bulk water with MM point charges, as detailed in the Supporting Information (Appendix A).

4.5.1 TD-LRC-DFT results for A:T and A$_2$

Figure A.5 displays TD-LRC-DFT absorption spectra for aqueous A:T and A$_2$, obtained by averaging over solvent configurations. The duplex A$_2$:T$_2$ is examined in the Supporting Information (Appendix A). (Apart from a $\sim$0.2 eV reduction in the adenine $\rightarrow$ thymine CT excitation energies, the A$_2$:T$_2$ spectrum is largely a superposition of the A$_2$ and T$_2$ spectra examined here.)

Each excitonic $\pi\pi^*$ state in Fig. A.5 is represented by a gaussian centered at the mean excitation energy and weighted by the mean oscillator strength, with a width (standard deviation) obtained from configurational averaging. The two LRC functionals examined here afford nearly identical results for the $\pi\pi^*$ states, which are red-shifted by about 0.2 eV relative to gas-phase calculations, in good agreement with experimentally-measured solvatochromatic shifts. The narrow widths of the gaussian distributions indicate that these states are largely unaffected by solvent fluctuations.

The CT states are far more sensitive to solvent configuration, and span a range of
Figure 4.7: Absorption spectra for aqueous A:T and A\textsubscript{2} obtained from a TD-DFT/6-31G* QM/MM calculation. The LRC-ωPBE functional is used in (a) and (b), whereas the LRC-ωPBEh functional is used in (c) and (d). To avoid congestion, the optically-weak \textit{1}nπ\textsuperscript{*} states are omitted. Gaussian distributions are obtained from averages over solvent configuration; for the CT states, the stick spectra are shown as well. The optically-bright CT states around 6.9 eV borrow intensity from the second ππ\textsuperscript{*} absorption band, which is not shown.
more than 1 eV in both A:T and A₂. To emphasize this fact, stick spectra for the CT states are included in Fig. A.5, along with gaussian fits as described above. For A₂, in particular, the CT states sometimes overlap with the \( \pi \pi^* \) states (or, less frequently, fall slightly below the \( \pi \pi^* \) states), but there is also a broad tail out to significantly higher energies, 0.5–1.0 eV above the \( \pi \pi^* \) absorption band.

On average, the CT states lie about 0.1 eV below the corresponding gas-phase values, though this number belies the breadth of the CT distribution in solution. As a check of our averaging procedure, however, we also calculated the solvatochromatic shift for the lowest CT state of A₂, using the SS(V)PE solvation model\(^\text{156}\) at the LRC-\( \omega \)PBE/6-31G* level. [Because SS(V)PE is not yet available for TD-DFT, the CT excitation energy was instead calculated by using the maximum overlap method\(^\text{157}\) to find an excited-state self-consistent field solution corresponding to charge transfer.] This procedure also predicts a 0.1 eV solvatochromatic red shift, in perfect agreement with the average QM/MM result. A more conventional TD-DFT/polarizable continuum calculation of the solvent shift predicts a very small red shift of < 0.1 eV.\(^\text{122}\) We take these results as an affirmation the validity of the QM/MM averaging procedure.

In addition to being sensitive to solvent configuration, the CT excitation energies are also far more sensitive to the choice of LRC functional than are the \( \pi \pi^* \) states. Consistent with gas-phase results, the LRC-\( \omega \)PBEh functional predicts a systematic 0.3–0.4 eV reduction in the CT excitation energies, as compared to those obtained using LRC-\( \omega \)PBE. For aqueous A₂, this is enough to move the CT band from being just above the \( \pi \pi^* \) band [in the case of LRC-\( \omega \)PBE, Fig. A.5(b)] into a spectral
region that largely overlaps the $\pi \pi^*$ band [for LRC-$\omega$PBEh, Fig. A.5(d)], leading to substantial configuration mixing and intensity borrowing in the latter case. In A:T, both functionals predict that the CT states mostly lie above the $\pi \pi^*$ states.

These differences between functionals prompted us to re-examine the benchmark TD-LRC-DFT data of Rohrdanz et al.$^{121}$ When compared to a set of high-level ab initio benchmarks,$^{126}$ both functionals afford the same root-mean-square error, 0.3 eV, for both localized ($n\pi^*$ and $\pi\pi^*$) excitations and CT excitations, as summarized in the Supporting Information (Appendix A). The LRC-$\omega$PBEh functional, however, almost always overestimates the CT excitation energies, while LRC-$\omega$PBE exhibits positive and negative errors with approximately equal frequency. For CT excitation energies, the mean signed errors for these two functionals are $-0.2$ eV (for LRC-$\omega$PBEh) and $+0.1$ eV (for LRC-$\omega$PBE). For $n\pi^*$ and $\pi\pi^*$ excitations, both functionals exhibit a mean signed error of $+0.2$ eV, indicating that localized excitation energies are over-estimated, on average.

Assuming that the benchmarks in Ref. 126 are representative, this suggests shifting the CT states in Fig. A.5 by $-0.1$ eV and $+0.2$ eV for LRC-$\omega$PBE and LRC-$\omega$PBEh, respectively, while shifting the $\pi\pi^*$ states by $-0.2$ eV in both cases. This modification removes the discrepancy between the two LRC functionals, leaving a spectrum in which the CT band is centered just above the $\pi\pi^*$ band, with only a few low-energy outliers that overlap the $\pi\pi^*$ band. The shifted spectra may be found in the Supporting Information (Appendix A). (Note that oscillator strengths in the shifted spectra are not reliable, especially for the CT states, owing to intensity borrowing.
when TD-DFT predicts CT/\pi\pi^* quasi-degeneracies.)

4.5.2 CIS(D) results for A:T and A_2

Absorption spectra for A_2 and A:T, computed at the SCS-CIS(D) level, are shown in Fig. A.6. [Spectra obtained at the CIS(D) can be found in the Supporting Information. (Appendix A)] Like the TD-LRC-DFT calculations, this method predicts a 0.2 eV solvatochromatic red shift in the \pi\pi^* excitation energies, although CT states at the CIS(D) and SCS-CIS(D) levels appear at slightly higher energies, as compared to TD-LRC-DFT predictions. Based on comparison to CC2 results in gas-phase A_2 and A:T, it appears that the former methods slightly overestimate the intrastrand CT excitation energies, and significantly overestimate interstrand CT excitation energies, whereas TD-LRC-DFT results are much closer to CC2.

4.5.3 TD-LRC-DFT results for (ApA)^-

Lastly, we consider adenine dinucleotide, (ApA)^-, in aqueous solution, using the same QM/MM approach as above, except that we allow the adenine monomer geometries to relax during the molecular dynamics simulation. (The backbone atoms are still held rigid, to maintain \pi-stacking.) The TD-LRC-\omega PBE absorption spectrum is shown in Fig. 4.9, which includes the corrective shifts discussed above. Consistent with other calculations of backbone-induced shifts,\textsuperscript{158} the bright exciton peaks (B±) are red-shifted by \sim0.1 eV relative to those in A_2. Relaxation of the adenine monomer geometries broadens the \pi\pi^* absorption peaks, and also lowers the CT excitation
Figure 4.8: Absorption spectra for (a) hydrated A:T and (b) hydrated A₂, obtained from a SCS-CIS(D)/6-311+G* QM/MM calculation, using the same configurational snapshots used to generate Fig. A.5. Gaussian distributions were obtained from averages over solvent configuration (using CIS oscillator strengths), though for the CT states the stick spectra are shown as well. In A:T there is considerable mixing between the second ππ* band and the CT states, lending significant oscillator strength to the latter.
Figure 4.9: Absorption spectrum of aqueous (ApA)^− calculated at the TD-LRC-ωPBE/6-31G* level, including corrective shifts as discussed in the text. Weakly-absorbing $n\pi^*$ states are omitted, for clarity.
energies somewhat, resulting in more significant overlap between the two bands. As in Na\(^+\)(ApA\(^-\))(H\(_2\)O)\(_4\), the presence of water molecules pushes the PO\(_4\)\(^-\) → adenine CT states to at least 6.5 eV in solution, and it therefore appears unlikely that the phosphate group plays a significant role in the low-energy photophysics of B-DNA.

The picture established by these solution-phase absorption spectra contrasts sharply with the results from previous TD-DFT studies of A\(_2\) and related systems,\(^{76,77,86,122}\) where CT states were found well below the bright states. In light of the foregoing discussion, we feel confident in ascribing these previous TD-DFT results to artifacts of standard TD-DFT’s underestimation of CT excitation energies. In this context, we note that it is nearly impossible to repeat our aqueous (ApA)\(^-\) calculations using PBE0, owing to an inordinately large number of low-energy CT states, many of them involving the phosphate group and/or the solvent.

### 4.6 Conclusions

Carefully-calibrated TD-LRC-DFT calculations correct the severe underestimation (\(\gtrsim 1\) eV) of CT excitation energies exhibited by most standard density functionals, and provide excitation energies within \(\sim 0.3\) eV of high-level benchmarks such as CC2 theory, for both valence \((n\pi^*\) and \(\pi\pi^*)\) and CT excitation energies.

In this work, we used TD-LRC-DFT to investigate the effects of base stacking and base pairing on DNA multimers composed of adenine and thymine, using a QM/MM model of aqueous solvation that includes a full solvation shell of QM water molecules.
The location of CT excited states is exquisitely sensitive to the instantaneous configuration of the water molecules, and both the inter- and intrastrand CT states span a range of more than 1 eV, as a function of solvent configuration. (Notably, the average CT excitation energy, as predicted by continuum solvation models, differs from the gas-phase value by no more than 0.1 eV.) Intrastrand adenine → adenine CT states are, on average, slightly lower in energy than interstrand adenine → thymine CT states, though there is significant overlap between the two bands. The low-energy tail of the intrastrand CT band overlaps the high-energy part of the brightest exciton state in the first $\pi\pi^*$ absorption band.

The errors in our calculated excitation energies could easily be $\sim 0.3$ eV, which (depending on the sign) could have the effect of shifting the CT states into greater overlap with the $\pi\pi^*$ states, or shifting the entirety of the CT band to a position just above the $\pi\pi^*$ band. We feel confident, however, that the CT states are neither well below, nor significantly above, the $\pi\pi^*$ band. As such, these results stand in marked contrast to previous TD-DFT studies of $\pi$-stacked DNA multimers,\textsuperscript{76,77,86,122} where CT states more than 1 eV below the $\pi\pi^*$ states were reported, using functionals such as B3LYP and PBE0 that are known to underestimate CT excitation energies. We regard these purported, low-energy CT states as artifacts. On the other hand, uncorrelated CIS calculations place the CT states 1–2 eV above the $\pi\pi^*$ states,\textsuperscript{78} but this energy gap narrows to 0.5 eV or less when electron correlation is introduced.
The fact that both intra- and interstrand CT states appear at comparable energies suggests that both base-stacking and base-pairing interactions must be considered simultaneously in a reasonable simulation of the excited-state dynamics of DNA, whereas previous simulations have tended to focus on hydrogen-bonded base pairs. Additional work, including more realistic base sequences and a more thorough consideration of the effect of DNA dynamics, is needed. TD-LRC-DFT methods—with careful calibration of CT excitation energies—appears to be a promising way to build upon the simulations reported here.
CHAPTER 5

Polarizable continuum reaction-field solvation models affording smooth potential energy surfaces\textsuperscript{5.1}

Reaction-field models are an efficient means to incorporate certain bulk solvent effects into electronic structure and molecular mechanics (MM) calculations, without the need for configurational averaging over explicit solvent molecules.\textsuperscript{159} Among the most popular such methods are apparent surface charge (ASC) approaches, commonly known as polarizable continuum models (PCMs), wherein electrostatic interactions with the continuum are modeled by a charge density, $\sigma(s)$, at the surface of a solute cavity. Given the solute charge density, $\rho(r)$, an integral equation to determine $\sigma(s)$ can be formulated based upon approximate solution of Poisson’s equation, subject to cavity boundary conditions.\textsuperscript{159,160}

Definition of the cavity boundary is a crucial aspect of ASC PCMs. Most often, the solute cavity is constructed from a union of atomic spheres,\textsuperscript{161,162} possibly augmented by some additional spheres to smooth out any crevasses.\textsuperscript{163} (Isocontours of $\rho$ have also been used,\textsuperscript{3,164} but this approach significantly complicates the formulation of

\footnote{\textsuperscript{5.1}This chapter appeared as a full article in the \textit{Journal of Physical Chemistry Letters}, in 2010, volume 1, page 556.}
analytic energy gradients, and is not considered here.) In any case, it is necessary to discretize the cavity surface into a set of surface elements centered at points \( r_i \) and having areas \( a_i \). Upon discretization, the integral equation for \( \sigma \) is replaced by a set of coupled linear equations,

\[
Kq = Rv,
\]

that determine a vector \( q \) of surface charges located at the points \( r_i \), whose interaction with \( \rho(\mathbf{r}) \) represents the electrostatic part of the continuum solvent effect. The vector \( v \) in Eq. (8.43) consists of the solute’s electrostatic potential at the surface discretization points, while the matrices \( K \) and \( R \) characterize a given PCM. Chipman\(^3,8,160\) has shown how a variety of PCMs may be cast into the form of Eq. (8.43); in particular, the conductor-like screening model (known as C-PCM or COSMO),\(^{165,166}\) the “integral equation formalism” (IEF-PCM),\(^{167}\) and Chipman’s “surface and simulation of volume polarization for electrostatics” [SS(V)PE] model\(^{160}\) all have basically this form. The precise forms of \( K \) and \( R \) are detailed in Ref. 3 and summarized in 6.1.

The issue addressed in the present work is that a straightforward implementation
of Eq. (8.43) inevitably leads to discontinuities in the potential energy surface of the solute, because certain surface points $r_i$ will disappear within—or emerge from—the interior of the solute cavity, as the solute atoms are displaced. These discontinuities hinder geometry optimizations, or prevent them from converging at all; lead to energy drift and other instabilities in molecular dynamics simulations; and may introduce serious artifacts in vibrational frequencies, especially when the latter are calculated by finite difference, as is often the case for high-level electronic structure methods.

There have been a few previous attempts to eliminate these discontinuities, mostly within the context of COSMO, by introducing switching functions to attenuate the surface elements as they pass through certain buffer regions surrounding each solute atom.\textsuperscript{5,168–170} We have observed, however, that certain artifacts persist in some of these ostensibly smooth PCMs. The key result of the present work is a reformulation and generalization of one of these methods, to yield an implementation of Eq. (8.43) that affords smooth potential energy surfaces, and whose gradients are well behaved even as two atoms are pulled apart.

In our approach, the $i$th surface discretization point is attenuated using a switching function

$$F_i = \prod_{J, i \notin J} f(\hat{r}_{iJ})$$

(5.2)

that consists of a product of elementary switching functions, $0 \leq f(\hat{r}_{iJ}) \leq 1$. The dimensionless quantity $\hat{r}_{iJ}$ describes the location of the $i$th grid point within a buffer region around the $J$th atom. (Additional details are provided in the Supporting Information.) The function $F_i$ is identical to the switching function used in the
In addition to discontinuities, ASC PCMs may suffer from singularities in the Coulomb interactions between surface charges, if the separation $r_{ij}$ between two charges is small. Switching functions actually exacerbate this problem, by allowing closer approach of $r_i$ and $r_j$. Again following York and Karplus, we avoid such singularities by representing the charge around the point $r_i$ as

$$q_i(\zeta_i^2/\pi)^{3/2} \exp(-\zeta_i^2 |r - r_i|^2),$$

where $\zeta_i$ is a fixed parameter. The matrix $S$ that represents the Coulomb interactions among the surface elements is taken to be

$$S_{ij} = \begin{cases} (2/\pi)^{1/2} \zeta_i F_i^{-1} & i = j \\ \text{erf}(\zeta_{ij} r_{ij})/r_{ij} & i \neq j \end{cases}$$

(5.3)

where $\zeta_{ij} = \zeta_i \zeta_j / (\zeta_i^2 + \zeta_j^2)^{1/2}$. Note that $S_{ij}$ is finite, even as $r_{ij} \to 0$. As discussed in the Supporting Information, the factor of $F_i^{-1}$ is introduced into $S_{ii}$ in order to ensure that $K^{-1}$ has a null space corresponding to any “switched off” grid points ($F_i = 0$), so that $F_i$ functions to attenuate the surface charges, albeit indirectly, via the $S$ matrix. As such, the dimension of the matrices $S$, $A$, $D$, and $K$ can be reduced to include only those grid points for which $F_i > \delta$, where $\delta$ is some finite drop tolerance.

We refer to the attenuation scheme based upon Eqs. (B.11) and (5.3) as the “Switching/Gaussian” (SWIG) approach. The SWIG-COSMO method is identical to the S-COSMO method of York and Karplus, except that we correct an error in the nuclear gradients appearing in Refs. 168 and 171. The correct derivative of $F_i$ in the nuclear gradients appearing in Refs. 168 and 171. The correct derivative of $F_i$
with respect to a perturbation of the $M$th nucleus is

$$\nabla_M F_i = \sum^K_{K} \frac{\partial f(\hat{r}_{iK})}{\partial \hat{r}_{iK}} (\nabla_M \hat{r}_{iK}) \prod_{J \neq K} f(\hat{r}_{ij}) ,$$

(5.4)

whence

$$\nabla_M S_{ii} = -S_{ii} \left( \sum^K_{K} \frac{1}{f(\hat{r}_{iK})} \frac{\partial f(\hat{r}_{iK})}{\partial \hat{r}_{iK}} \nabla_M \hat{r}_{iK} \right) .$$

(5.5)

York and co-workers\textsuperscript{171} later suggested replacing the switching function $F_i$ with $F_i^p$, where $p$ is an adjustable parameter for which the value $p = 0.25$ was suggested. In our experience, however, values of $p \neq 1$ introduce unwanted oscillations into the energy gradient, as demonstrated below.

Extension of the SWIG procedure to SS(V)PE and IEF-PCM, which is reported here for the first time, requires construction of the matrix $D$ in 6.1. The $D$ and $S$ matrices are related according to\textsuperscript{159}

$$D_{ij} = -\hat{n}_j \cdot \frac{\partial S_{ij}}{\partial r_j} ,$$

(5.6)

where $\hat{n}_j$ is a unit vector normal to the cavity surface, at the point $r_j$. Using $S_{ij}$ from Eq. (5.3), one obtains

$$D_{ij} = \left( \text{erf}(\zeta_{ij} r_{ij}) - \frac{2\zeta_{ij}}{\sqrt{\pi}} e^{-\zeta_{ij}^2 r_{ij}^2} \right) \hat{n}_j \cdot \frac{\hat{r}_{ij}}{r_{ij}^3} ,$$

(5.7)

where $r_{ij} = r_i - r_j$. The gradient with respect to nuclear displacements is

$$\nabla_M D_{ij} = - \left[ D_{ij} - \left( \frac{4\zeta_{ij}^3}{\sqrt{\pi}} e^{-\zeta_{ij}^2 r_{ij}^2} \right) \hat{n}_j \cdot r_{ij} \right]$$

$$\times \left( \frac{r_{ij}}{r_{ij}^2} \right) \cdot \nabla_M r_{ij} - \left( S_{ij} - \frac{2\zeta_{ij}}{\sqrt{\pi}} e^{-\zeta_{ij}^2 r_{ij}^2} \right)$$

$$\times \left( \frac{2\hat{n}_j \cdot r_{ij}}{r_{ij}^4} r_{ij} - \hat{n}_j \hat{r}_{ij} \right) \cdot \nabla_M r_{ij} .$$

(5.8)
The diagonal elements $D_{ii}$ are less straightforward to define for the SWIG procedure. When the solute density $\rho(\mathbf{r})$ consists of point charges, as in MM calculations, these elements have been defined so as to preserve an exact geometric sum rule, but this sum rule is not applicable here because so-called surface points may actually lie inside of the cavity, within a narrow switching region of the cavity surface. In practice, we find that attempts to enforce the aforementioned sum rule sometimes compromise the positive-definiteness of $K$, resulting in singularities that prohibit convergence of the method. As an alternative, Tomasi et al. have shown how to define $D_{ii}$ in terms of $S_{ii}$, which avoids this issue. The simplest approach, however, is to note that

$$\lim_{r_{ij} \to 0} D_{ij} = 0 ,$$

which follows from Eq. (5.7) and implies that $D_{ii} \to 0$ as $a_i \to 0$. For a sufficiently dense discretization grid, one is therefore justified in taking $D_{ii} \equiv 0$, and we have made this choice for all of the calculations reported here. Numerical tests demonstrate that the results are virtually identical to those obtained when $D_{ii}$ is defined in terms of $S_{ii}$.

Application of the SWIG procedure to ASC PCMs results in potential energy surfaces that are rigorously smooth, in the mathematical sense of possessing continuous gradients. (See the Supporting Information for a proof.) Physically unreasonable fluctuations could still exist, however, as smoothness in the colloquial sense (“chemical smoothness”) is a much more demanding criterion. This will be examined in the numerical calculations that follow.
We have implemented the SWIG-COSMO and SWIG-SS(V)PE models within a locally-modified version of the Q-Chem electronic structure program,\textsuperscript{53} which we have further modified to perform stand-alone MM and QM/MM calculations. We have developed an efficient biconjugate gradient solver for Eq. (8.43) and an efficient implementation of the analytic energy gradient, which is especially important for MM and QM/MM applications, where the number of surface grid points may be quite large. The details of our implementation will be reported in a future publication, but all analytic gradients have been validated against finite-difference results. We discretize the solute cavity surface using atom-centered Lebedev grids, with gaussian parameters $\zeta_i$ taken from Ref. 168 and standard values for the atomic radii. (Details are provided in the Supporting Information.) We set $\varepsilon = 78.39$ (water) for all calculations.

For comparison to the SWIG model introduced here, we have also implemented the “fixed-point, variable area” (FIXPVA) algorithm, another smooth version of COSMO introduced recently by Su and Li.\textsuperscript{5} [We have generalized FIXPVA, in a straightforward way, for use with any PCM having the form of Eq. (8.43).] The FIXPVA approach uses a point-charge discretization of $\sigma(s)$ and an alternative switching function to attenuate the areas, $a_i$, of the surface elements. As a baseline, we compare both SWIG and FIXPVA to the “variable tesserae number” (VTN) method of Li and Jensen,\textsuperscript{175} a very simple implementation of ASC PCMs. The VTN approach ameliorates some of the problems associated with discontinuities in the potential surface,\textsuperscript{175} although it does not completely eliminate these discontinuities.
For an accurate description of solvation using PCMs, it is often necessary to retain some explicit solvent molecules, e.g., in the first solvation shell. With this in mind, our first illustrative application is an MM geometry optimization of (adenine)(H$_2$O)$_{52}$, using the AMBER99 force field$^{176,177}$ for adenine and the 52 explicit water molecules, and COSMO for bulk water. 5.1 plots the energy as a function of optimization step for the VTN, FIXPVA, and SWIG implementations of COSMO. Not surprisingly, the VTN method suffers from numerous Coulomb singularities and discontinuities, but the optimization does eventually converge. However, the optimized structure itself exhibits a Coulomb singularity owing to the presence of two nearby surface grid points. This results in a pair of surface charges whose magnitude dwarfs that of all the others, as evident from the surface charge density depicted in 5.1. The two point charges in question are more than twice as large as any of the charges in the FIXPVA or SWIG calculations, and likely distort the VTN solvation energy and gradient.

The FIXPVA-COSMO optimization also converges to a minimum, albeit in a somewhat larger number of steps, but there are two problems. First, the cavity surface at the optimized geometry actually exhibits holes (clearly evident in 5.1), wherein all of the surface element areas have been scaled to zero. The second problem with the FIXPVA optimization is that the energy curve, like that obtained for VTN-COSMO, exhibits numerous sharp spikes. Unlike VTN, these spikes cannot result from discontinuities, since the FIXPVA potential surface is rigorously smooth. Rather, they indicate that the energy changes incredibly rapidly in certain regions of the FIXPVA potential surface, and consequently, an optimization step selected using
Figure 5.1: Molecular mechanics geometry optimization of (adenine)(H₂O)₅₂ in bulk water, using three different implementations of COSMO. The vertical scale represents the cluster binding energy, including the electrostatic free energy of solvation. Also shown are the solute cavity surfaces for the optimized structures. Each grid point \( r_i \) is depicted as a sphere whose radius is proportional to \( a_i \), and colored according to the charge \( q_i \).
local gradient information occasionally moves the system to a much higher-energy geometry. Corroborating this explanation is the fact that the energy spikes disappear if we decrease the maximum allowed step size by a factor of ten, although in this case the optimization fails to converge within 5,000 steps. This is our first example of physically unrealistic fluctuations in a mathematically smooth potential surface; subsequent examples will suggest that the FIXPVA gradient exhibits rapid fluctuations as a result of instabilities attributable to the use of surface point charges. The SWIG-COSMO optimization, in contrast, exhibits monotonic convergence.

5.2 shows harmonic vibrational spectra, computed via finite difference of analytic energy gradients, for the FIXPVA- and SWIG-COSMO geometries of (adenine)(H₂O)₅₂ that were optimized above. (The corresponding VTN calculation resulted in several imaginary frequencies and is not shown. To obtain strictly real frequencies with FIXPVA, it was necessary to reduce Q-Chem’s default finite-difference step size by a factor of ten.) The FIXPVA and SWIG approaches are in good agreement for the majority of the peaks, however, FIXPVA predicts several peaks with impossibly large frequencies, ranging from \( \sim 5,000 \text{ to } 16,000 \text{ cm}^{-1} \). Each of these spurious peaks is associated with vibration of a water molecule near the cavity surface, where close approach of surface charges leads to rapid variation in the gradient.

As an electronic structure example, we next consider the dissociation reaction \( \text{NaCl} \rightarrow \text{Na}^+ + \text{Cl}^- \) (a standard test case\(^5,171\)) at the unrestricted Hartree–Fock (UHF)/6-31+G*/SS(V)PE level. Potential energy curves and gradients are shown in 5.3 for the VTN, FIXPVA, and SWIG discretization procedures. Discontinuities
Figure 5.2: Vibrational spectra of the FIXPVA- and SWIG-COSMO optimized (adenine)(H$_2$O)$_{52}$ structures. (The inset is an enlarged view of the region up to 4,000 cm$^{-1}$.) Harmonic frequencies were calculated by finite difference of analytic energy gradients and convolved with 10 cm$^{-1}$ gaussians, weighted by the computed intensities. Arrows indicate FIXPVA peaks with no obvious SWIG analogues.
Figure 5.3: Total energy (solid curves, scale at left) and Na-atom gradient (dashed curves, scale at right) for NaCl dissociation, computed at the UHF/6-31+G*/SS(V)PE level including non-electrostatic contributions to the PCM energy. A horizontal dotted line indicates where the gradient is zero. Panel (c) shows two sets of results for the SWIG model, corresponding to a switching functions $F_i^p$ with two different values of $p$. 
in the VTN energy and gradient are clearly evident, even near the equilibrium geometry. The FIXPVA energy and gradient are continuous, but the latter exhibits sizable oscillations that manifest as several very shallow minima between at highly-stretched bond lengths. In this particular example, there are no sign changes in the second derivative near the equilibrium geometry, and consequently FIXPVA affords a reasonable vibrational frequency. However, the large number of inflection points observed at larger bond lengths, as the atomic spheres begin to separate, suggest that in larger systems where numerous atomic spheres overlap, spurious inflection points are likely the origin of anomalous vibrational frequencies observed, e.g., for (adenine)(H₂O)₅₂.

The SWIG-SS(V)PE gradient exhibits only mild oscillations, and the corresponding potential surface exhibits only a single minimum, at least for the switching function $F_i$ in Eq. (B.11). However, substitution of $F_i^p$ in place of $F_i$, and using the value $p = 0.25$ suggested by York and co-workers,\textsuperscript{171} introduces oscillations in the gradient that are even more rapid than those observed for FIXPVA [see 5.3(c)]. Interestingly, if one uses the incorrect gradient expressions provided in Ref. 171, one obtains a much less oscillatory gradient for the $p = 0.25$ case, although the $p = 1$ case changes little. In our experience, values of $p \neq 1$ only serve to introduce unwanted oscillations in the gradient.

In addition to solute/continuum electrostatic interactions, the NaCl potential curves in 5.3 also include some standard non-electrostatic terms representing cavitation\textsuperscript{161} and dispersion/repulsion\textsuperscript{178} interactions. Within the PCM formalism, each
of these interactions is a function of the surface area of the cavity. For the SWIG model, the total cavity surface area is

\[
\sum_{i}^{\text{Lebedev points}} a_i = \sum_{J}^{\text{atoms}} R_J^2 \sum_{i \in J} w_i F_i
\]

where the \(w_i\) are the Lebedev weights and the \(R_J\) are the radii of the atomic spheres.

Figure 5.4: Non-electrostatic contributions to the PCM energy, for the UHF/SS(V)PE NaCl potential curves from 5.3.

5.4 depicts the non-electrostatic energy as a function of Na–Cl distance, for the UHF/SS(V)PE calculation discussed above. The step-like behavior of the VTN energy is a consequence of abrupt changes in the number of surface elements, which introduce step-like discontinuities in the cavity surface area. (The VTN scheme employs surface tesserae having fixed areas, which appear abruptly as their centers...
emerge from the cavity interior.) That the discontinuities persist even at very large separations is a consequence of using a larger, “solvent-accessible” cavity surface to compute the dispersion and repulsion energies.\textsuperscript{178}

The FIXPVA method scales all areas $a_i$ by a switching function, and consequently the FIXPVA surface area is bounded from above by the VTN surface area. As such, atomic radii and switching parameters developed for the electrostatic interactions will probably underestimate non-electrostatic contributions, necessitating separate sets of switching parameters for the electrostatic, cavitation, and dispersion/repulsion energies.\textsuperscript{179} Re-parameterization, however, will not eliminate the rather large oscillations in the FIXPVA surface area as a function of Na–Cl distance. The SWIG approach, in contrast, affords a nearly monotonic increase in surface area as a function of distance, and essentially interpolates between the discontinuous steps present in the VTN calculation. Given that the number of significant grid points, and hence the dimension of $K$, changes numerous times as the atoms are pulled apart (see the Supporting Information for a plot), this is an impressive demonstration that the SWIG procedure provides “chemically smooth” potential surfaces.\textsuperscript{174}

The poor behavior of the VTN approach in each of these applications illustrates the importance of a switching function in ASC PCM calculations, while problems encountered with the FIXPVA method indicate that the details of the attenuation procedure are important. Both the FIXPVA and SWIG methods guarantee continuous potential surfaces and gradients, but the latter offers distinct advantages. Because FIXPVA employs point charges to discretize the surface charge density $\sigma(s)$,
this method must rapidly scale \( a_i \rightarrow 0 \) within the buffer region, in order to avoid Coulomb singularities (which, in the end, are not always avoided). This rapid scaling leads to unwanted oscillations in the solute potential energy surface, and a poor representation of the cavity surface. Tremendous errors in vibrational frequencies may result when calculated by finite difference methods, despite the fact that the gradients are smooth.

The SWIG method uses spherical gaussians centered at Lebedev grid points to represent \( \sigma(s) \), thus the surface Coulomb interactions are free of singularities, even as the surface grid points pass through the switching region in close proximity to one another. As such, the switching function can act more slowly, eliminating unphysical fluctuations in the energy gradient. In principle, the alternative switching function employed in the FIXPVA method could be combined with gaussian surface charges, which might alleviate some of the problems with FIXPVA, though we have not pursued such an approach.

In summary, we have introduced a Switching/Gaussian (“SWIG”) discretization procedure for ASC PCMs, based upon a reformulation and generalization of the S-COSMO method of York and Karplus,\textsuperscript{168} which we have extended to sophisticated PCMs including Chipman’s SS(V)PE model.\textsuperscript{160} Discretization of the cavity surface is accomplished using Lebedev grids, rather than more elaborate surface tessellation schemes,\textsuperscript{163} which avoids the need to implement complicated geometrical derivatives of the tesserae areas.\textsuperscript{180} As such, the method is easy to implement within existing codes. SWIG-PCM potential surfaces and gradients are rigorously smooth,
in the mathematical sense, and moreover appear to be free of unphysical fluctuations. As such, vibrational frequencies can safely be calculated by finite difference of analytic gradients. Cavity surface areas, and therefore surface-area-dependent non-electrostatic interactions, vary smoothly as a function of solute geometry, without spurious oscillations. In future, work we will report efficient implementations of the SWIG-COSMO and SWIG-SS(V)PE analytic gradients, along with further tests of these methods.

Supporting information for the computational details regarding cavity construction and switching functions is provided in Appendix B.
CHAPTER 6

A smooth, non-singular, and faithful discretization scheme for polarizable continuum models: The switching/Gaussian approach

6.1 Introduction

This work focuses on a particular class of reaction-field models that are intended to describe a solute molecule immersed in a structureless dielectric medium. The solute has a charge distribution that might be calculated using quantum chemistry, or taken from a molecular mechanics (MM) force field, whereas the effects of the dielectric continuum, which is polarized by the solute’s charge distribution, are represented by a charge density, $\sigma_{pol}(\vec{s})$, at the surface of a cavity that represents the solute/continuum boundary. Given the solute’s charge density, $\rho_0(\vec{r})$, equations to define $\sigma_{pol}(\vec{s})$ are obtained by approximate solution of Poisson’s equation, subject to cavity boundary conditions. Various levels of approximation have been reviewed recently, and Chipman has shown how all of these “apparent surface charge” (ASC) models can be cast within a common conceptual and computational framework. We adopt his notation as much as possible.

---

6.1 This chapter appeared as a full article in the Journal of Chemical Physics, in 2010, volume 133, pages 244111:1–18.
Nowadays, the ASC approach is virtually synonymous with the term “polarizable continuum model” (PCM), and encompasses such variants as the conductor-like model (C-PCM, also known as as GCOSMO\textsuperscript{181,182} as well as the “integral equation formalism” (IEF-PCM),\textsuperscript{183} which is also known by the acronym SS(V)PE (“surface and simulation of volume polarization for electrostatics”).\textsuperscript{160} The IEF-PCM and SS(V)PE methods, which are formally equivalent,\textsuperscript{184} provide an exact solution for the ASC arising from the charge density within the solute cavity, and an approximate solution for the volume polarization due to the “escaped charge”, \textit{i.e.}, that part of $\rho_0$ that penetrates beyond the solute cavity.

For ASC models, one must specify a surface that defines the continuum boundary. When the solute is described using quantum chemistry, a conceptually appealing choice is to use an isodensity contour, $\rho_0(\vec{s}) = \text{constant},$\textsuperscript{3,164} but this choice significantly complicates the formulation of analytic energy gradients. More often, the cavity is constructed from atom-centered (or functional-group-centered) spheres, whose radii might be related to atomic van der Waals radii, or might simply be taken as empirical parameters in the model. With carefully-parameterized united-atom radii, solvation free energies calculated using this simple prescription may actually be more accurate than those obtained from an isodensity construction.\textsuperscript{162,185}

The difficulty with with all of these approaches, and the focus of the present work, lies in the discretization of the cavity surface into a finite set of surface elements, within which $\sigma_{\text{pol}}(\vec{s})$ is assumed to be constant. Thus, $\sigma_{\text{pol}}(\vec{s})$ is approximated using a finite set of surface grid points, $\vec{s}_i$, each with a surface area $a_i$, and $\sigma_{\text{pol}}(\vec{s})$ is
replaced by a set of charges, $q_i$, located at these grid points. Given such a formulation, the potential energy surface of the solute molecule is inherently discontinuous, because the number of surface grid points changes discontinuously as a function of molecular geometry, with certain points disappearing into the solute cavity and others emerging from the cavity, onto the surface. These discontinuities hinder geometry optimizations, or prevent them from converging at all, introduce artifacts into vibrational frequency calculations, and pose problems for energy conservation in molecular dynamics simulations.\textsuperscript{186}

Several schemes have been introduced to alleviate this problem,\textsuperscript{4,5,168,170,175,186} although none appears to be in widespread use at the present time. These methods differ in their details, but the unifying idea is to introduce a switching function that smoothly attenuates the contribution of a given surface element as it passes into the interior of the solute cavity. We have recently shown, however, that this is insufficient to produce a potential energy surface that is “smooth enough” for many chemical applications, even if it is rigorously smooth in the mathematical sense of possessing continuous derivatives.\textsuperscript{186} We find that the details of the potential surface are extremely sensitive to the balance between the switching function and the singular Coulomb potentials of the point charges that represent $\sigma_{pol}$; naive adoption of a switching function may lead to spurious oscillations in the energy or gradient.\textsuperscript{186}

A more sophisticated smoothing approach is that pioneered by York and Karplus (YK),\textsuperscript{168} who represent $\sigma_{pol}$ using spherical Gaussian charges, thereby eliminating the singularity in the Coulomb potential. The present work represents a generalization
of the YK scheme that we call the “switching/Gaussian” (SWIG) method. Whereas YK consider only the conductor-like screening model (COSMO),\textsuperscript{165} our formulation is applicable to all ASC PCMs. A preliminary version of this work was reported in Ref. 186, and here we present a slightly modified version that is more faithful to the underlying integral equation theory. In addition, the present work provides additional details of the implementation, and a more thorough battery of tests and exemplary applications. Subsequent to our initial report,\textsuperscript{186} a similar approach was reported by Scalmani and Frisch,\textsuperscript{4} who refer to their implementation as the “continuous surface charge” (CSC) approach, a terminology that we shall adopt to refer to their implementation. The CSC method also attempts to generalize the YK approach, but differs in important details from the methodology presented here.

We will show that the SWIG approach possesses three critically important features: (i) it yields potential energy surfaces that are rigorously smooth, and also free of unwanted oscillations in the energy, the energy gradient, and the cavity surface area; (ii) the matrix formulation of SWIG is free of the singularities that pervade other smooth PCM formulations; and (iii) the SWIG discretization is faithful to the accuracy of the underlying integral equation theory, \textit{i.e.}, it does not greatly perturb the energetics. We believe that this is the first PCM discretization scheme to possess all three of these important properties.

The outline of this paper is as follows. In Section 6.2, we review the fundamental
reaction-field equations that lead to the general discretized PCM equations as formulated by Chipman.\textsuperscript{160} The SWIG formalism is presented in Section 6.3, and in Section 6.4 we analyze this approach in comparison with other discretization methods, pointing out several hitherto unnoticed problems with existing methods, including non-variational solvation energies. In Section 6.5, we present numerical results using both the C-PCM/GCOSMO and IEF-PCM/SS(V)PE solvation models. In some cases the solute is described using an MM force field, while in other examples the solute is described using quantum mechanics, at the level of self-consistent field (SCF) theory. In Section 6.6 we demonstrate the versatility and robustness of the SWIG discretization procedure with applications involving geometry optimization, calculation of vibrational spectra, and molecular dynamics simulations.

6.2 Reaction-field theory

In this section, we present an overview of the PCM reaction-field formalism for the solute–continuum electrostatic interactions, then show how the resulting integral equations are discretized for numerical solution.

Non-electrostatic interactions such as dispersion, repulsion, or cavitation are sometimes grafted onto ASC PCMs,\textsuperscript{159} but will not be discussed here, except to note that the commonly-used expressions for these interactions are functions of the total surface area of the cavity.\textsuperscript{161,178} As such, the continuity and smoothness of these non-electrostatic interaction terms depends upon ensuring that the cavity surface area is a smooth function of the solute geometry. We have considered non-electrostatic...
interactions in previous work, and found that the SWIG discretization suppresses spurious oscillations in the cavity surface area, which are sometimes observed using other smooth discretization schemes.

### 6.2.1 Electrostatic interactions

We begin by considering the electrostatic interactions of an arbitrary charge density embedded in a linear isotropic dielectric medium. We separate the total charge density into two components,

\[ \rho_{\text{tot}}(\vec{r}) = \rho_0(\vec{r}) + \rho_{\text{pol}}(\vec{r}), \]  

(6.1)

where \( \rho_0 \) is the solute’s charge density in the absence of the dielectric and \( \rho_{\text{pol}} \) is the density of bound charges that arises from polarization of the dielectric by the electric field due to \( \rho_0 \). Similarly, the total electrostatic potential is decomposed according to

\[ \phi_{\text{tot}}(\vec{r}) = \phi_0(\vec{r}) + \phi_{\text{pol}}(\vec{r}). \]  

(6.2)

The quantity \( \phi_{\text{pol}} \) is called the reaction-field potential, and originates with the charge distribution \( \rho_{\text{pol}} \) that is induced by the “reaction” of the dielectric in response to \( \phi_0 \), the field due to \( \rho_0 \). Such a response corresponds to displacement of bound charges within the dielectric, implying that a certain amount of work, \( w_{\text{displ}} \), is required to induce \( \rho_{\text{pol}} \).

The total electrostatic energy of the solute + continuum supersystem, \( W \), is equal to the electrostatic interaction energy of the solute’s charge density with the polarized dielectric, plus the work required to displace the bound charges. We denote the latter
quantity by \( w_{\text{displ}} \), and therefore

\[
W = w_{\text{displ}} + \frac{1}{2} \int \rho_{\text{tot}}(\vec{r})\phi_{\text{tot}}(\vec{r}) d^3\vec{r}.
\]  

(6.3)

This energy can be separated into two components,

\[
W = E_0 + E_{\text{pol}},
\]  

(6.4)

where \( E_0 \) is the electrostatic self-energy associated with \( \rho_0 \),

\[
E_0 = \frac{1}{2} \int \rho_0(\vec{r})\phi_0(\vec{r}) d^3\vec{r}.
\]  

(6.5)

The other component of \( W \) is the electrostatic solvation energy,

\[
E_{\text{pol}} = \frac{1}{2} \int \rho_0(\vec{r})\phi_{\text{pol}}(\vec{r}) d^3\vec{r}
\]

\[
+ \frac{1}{2} \int \rho_{\text{pol}}(\vec{r})\phi_0(\vec{r}) d^3\vec{r}
\]

\[
+ \frac{1}{2} \int \rho_{\text{pol}}(\vec{r})\phi_{\text{pol}}(\vec{r}) d^3\vec{r} + w_{\text{displ}}.
\]  

(6.6)

The last two terms in this equation constitute the total work of inducing \( \rho_{\text{pol}} \).

The appropriate microscopic expression for \( w_{\text{displ}} \) is not immediately obvious. However, by virtue of conservation of energy, the total work to induce \( \rho_{\text{pol}} \) must be equal and opposite to the energy gained by the dielectric in the interaction of \( \rho_{\text{pol}} \) with \( \phi_0 \):

\[
-\frac{1}{2} \int \rho_{\text{pol}}(\vec{r})\phi_0(\vec{r}) d^3\vec{r}
\]

\[
= \frac{1}{2} \int \rho_{\text{pol}}(\vec{r})\phi_{\text{pol}}(\vec{r}) d^3\vec{r} + w_{\text{displ}}.
\]  

(6.7)
This relation implies that $w_{\text{displ}}$ can be cast in terms of the electrostatic interactions between the solute and polarization charges. In addition, it leads to the usual condensed expression for the electrostatic solvation energy,\textsuperscript{187}

\begin{equation}
E_{\text{pol}} = \frac{1}{2} \int \rho_0(\vec{r}) \phi_{\text{pol}}(\vec{r}) \, d^3\vec{r} .
\end{equation}

This equation suggests that $E_{\text{pol}}$ can be determined without explicit knowledge of $w_{\text{displ}}$, although we show in Appendix 6.8 that explicit expressions for $w_{\text{displ}}$ can be derived for various PCMs.

In ASC PCM methods, it is assumed that $\rho_0$ occupies a cavity within the dielectric, and that $\rho_{\text{pol}}$ can be recast into an \textit{apparent} charge density, $\sigma_{\text{pol}}$, that resides entirely on the two-dimensional cavity surface. This surface charge density approximates the electrostatic potential due to $\rho_{\text{pol}}$:

\begin{equation}
\int \frac{\sigma_{\text{pol}}(\vec{s})}{|\vec{r} - \vec{s}|} \, d^2\vec{s} \approx \phi_{\text{pol}}(\vec{r}) .
\end{equation}

Given this approximation, and replacing $\rho_{\text{pol}}(\vec{r})$ with $\sigma_{\text{pol}}(\vec{s}) \delta(\vec{r} - \vec{s})$ in Eq. (6.6), the electrostatic solvation energy can be rewritten entirely in terms of surface integrals,

\begin{equation}
E_{\text{pol}} = w_{\text{displ}} + \int \sigma_{\text{pol}}(\vec{s}) \phi_0(\vec{s}) \, d^2\vec{s}
+ \frac{1}{2} \int \sigma_{\text{pol}}(\vec{s}) \hat{S} \sigma_{\text{pol}}(\vec{s}) \, d^2\vec{s} .
\end{equation}

The integral operator $\hat{S}$ in this equation is defined such that $\hat{S} \sigma_{\text{pol}}(\vec{s}) = \phi_{\text{pol}}(\vec{s})$; see Appendix 6.8 for details. In deriving Eq. (6.10), we have used the identity

\begin{equation}
\int \rho_0(\vec{r}) \phi_{\text{pol}}(\vec{r}) \, d^3\vec{r} = \int \rho_{\text{pol}}(\vec{r}) \phi_0(\vec{r}) \, d^3\vec{r} ,
\end{equation}

121
which follows from the symmetry of the Coulomb interaction.

As an alternative to Eq. (6.10), \( E_{\text{pol}} \) can be written in the more traditional form

\[
E_{\text{pol}} = \frac{1}{2} \int \sigma_{\text{pol}}(\vec{s}) \phi_0(\vec{s}) \, d^2\vec{s}, \tag{6.12}
\]

using Eqs. (6.8) and (6.11). According to Eq. (6.12), calculation of \( E_{\text{pol}} \) is reduced to computation of \( \sigma_{\text{pol}} \), which is the hallmark of ASC PCMs.

If \( \rho_0 \) is a polarizable charge density, as it is in quantum-mechanical (QM) calculations, then \( \sigma_{\text{pol}} \) may polarize and therefore modify \( \rho_0 \), which in turn modifies \( \sigma_{\text{pol}} \). In such cases, one must solve for \( E_0 \) and \( E_{\text{pol}} \) self-consistently, by incorporating the polarization charge \( \sigma_{\text{pol}} \) into the SCF iterations.

### 6.2.2 Variational conditions

While Eq. (6.12) is convenient for computing the solvation energy, it is not very useful in deriving variational conditions for PCMs, because it masks the full functional dependence of \( E_{\text{pol}} \) \([\text{cf. Eq. (6.10)}]\). Here, we consider variation of Eq. (6.10).

In order for the total energy, \( W \), to be a stationary point that is minimized with respect to variation of the surface charge density, the following relationship must hold:

\[
0 = \frac{\delta W}{\delta \sigma_{\text{pol}}(\vec{s})} = \phi_0(\vec{s}) + \hat{S}\sigma_{\text{pol}}(\vec{s}) + \frac{\delta w_{\text{displ}}}{\delta \sigma_{\text{pol}}(\vec{s})}. \tag{6.13}
\]

To ensure that \( W \) is in fact minimized with respect to variation of \( \sigma_{\text{pol}} \), we require that

\[
\frac{\delta^2 W}{\delta \sigma_{\text{pol}}(\vec{s}) \delta \sigma_{\text{pol}}(\vec{s}')} = \hat{S} + \frac{\delta^2 w_{\text{displ}}}{\delta \sigma_{\text{pol}}(\vec{s}) \delta \sigma_{\text{pol}}(\vec{s}')} \tag{6.14}
\]
is a positive-definite operator.

From Eq. (6.13), one obtains

$$-\phi_0(\vec{s}) = \hat{S}\sigma_{pol}(\vec{s}) + \frac{\delta w_{displ}}{\delta \sigma_{pol}(\vec{s})}\,,$$

(6.15)

which provides an equation relating $\phi_0$ at the cavity surface to the apparent surface charge. As shown in Appendix 6.8, Eq. (6.15) is a general equation that can be used to solve for $\sigma_{pol}$ in PCM methods.

### 6.2.3 Discretization

The analytical formulation of reaction-field theory presented above can be solved exactly only in special cases, such as the Born ion model. In order to apply reaction field theory to arbitrary charge densities and cavity shapes, the analytical formulation must be discretized and solved numerically. In this section we review how solute cavities are constructed and how the integral equations are transformed into finite-dimensional matrix equations.

**Cavity shape**

The first step in discretization is to define the shape of the cavity that defines the solute/continuum interface. Cavity shape is not unambiguous and certain shapes may be well-suited for some applications but not for others. Various approaches have been devised to construct cavities, ranging from a single sphere that encompasses the entire solute, to a more realistic shape consisting of a union of spheres centered at each solute nucleus.
Typically, a set of atomic van der Waals radii, such as those deduced from crystallographic data by Bondi,\textsuperscript{188} are used to define the spheres for each solute atom, although united-atom approaches have also been proposed.\textsuperscript{162} A surface composed of such spheres is called a van der Waals (vdW) surface. The vdW surface may exhibit crevices where finite-size solvent molecules should not be able to penetrate, hence the radii of the spheres is often augmented by adding a solvent probe radius to the vdW surface, resulting in a so-called “solvent accessible surface” (SAS).\textsuperscript{159} Alternatively, the vdW radii might simply be scaled by a factor, typically 1.2,\textsuperscript{165,189} to mimic the same effect. Another common choice is the “solvent excluding surface” (SES), also known as a Connolly surface,\textsuperscript{190} which smooths over the cusps of intersecting spheres in the vdW or SAS surface, by adding additional surface points. It has been noted,\textsuperscript{4} however, that certain algorithms for constructing the SES are inherently discontinuous, owing to the introduction of extra surface points. For this reason, we consider only SAS and vdW surfaces.

**Cavity surface grid**

With a cavity shape in hand, the next step is to divide the continuous surface into a grid of surface elements centered at points $\mathbf{s}_i$, with corresponding surface areas $a_i$. The most commonly used prescriptions are the GEPOL algorithm\textsuperscript{163} and Lebedev quadrature.\textsuperscript{191} GEPOL uses a regular 60-sided polyhedron to approximate a sphere, with surface elements that are triangles with areas $\tilde{a}_i = 4\pi R^2/60$, for a sphere of radius $R$. (More advanced extensions of the GEPOL grid have also been used.)\textsuperscript{180} The
geometric gradients of $\tilde{a}_i$ with respect to the nuclear coordinates are quite complicated, however, and we therefore use Lebedev grids exclusively, in order to simplify the formulation of analytic energy gradients. Lebedev grids, which are used almost universally to perform the numerical integration steps in density-functional theory calculations, are spherical grids having octahedral symmetry, with quadrature weights $w_i$ that are formulated to provide exact integrals for spherical harmonics, $Y_{lm}$, up to a given value of $l$. When a Lebedev grid is used to discretize a sphere of radius $R$, the individual surface elements have areas

$$\tilde{a}_i = w_i R^2 .$$

(6.16)

The trouble with these discretization procedures is that surface elements will emerge from and/or vanish into the interior of the cavity as the solute nuclei move. As a result, both $E_{pol}$ and the cavity surface area are discontinuous functions of the coordinates of the nuclei. Historically, this issue has been largely ignored, although several groups have recently worked on methods designed to ameliorate the discontinuities. Common to all of these methods is the introduction of a geometry-dependent switching function, $F_i$, for each surface element, such that the area of the $i$th surface element is

$$a_i = \tilde{a}_i F_i ,$$

(6.17)

where $0 \leq F_i \leq 1$. 125
Table 6.1: Definitions of the matrices in Eq. (8.43), for the PCMs considered here. The matrix $A$ is diagonal and consists of the surface element areas, $a_i$, while the matrices $S$ and $D$ are defined in the text. The quantity $\varepsilon$ represents the dielectric constant of the medium.

<table>
<thead>
<tr>
<th>Method</th>
<th>Matrix $K$</th>
<th>Matrix $R$</th>
</tr>
</thead>
<tbody>
<tr>
<td>C-PCM/GCOSMO</td>
<td>$S - (\frac{\varepsilon - 1}{\varepsilon}) (DAS + SAD^t)$</td>
<td>$- (\frac{\varepsilon - 1}{\varepsilon}) I$</td>
</tr>
<tr>
<td>IEF-PCM/SS(V)PE</td>
<td>$S - (\frac{\varepsilon - 1}{\varepsilon}) (DAS + SAD^t)$</td>
<td>$- (\frac{\varepsilon - 1}{\varepsilon}) (I - \frac{1}{2\pi} DA)$</td>
</tr>
</tbody>
</table>

Electrostatic energy

The fundamental assumption of the discretization procedure is that the electrostatic environment ($\phi_0$, $\sigma_{pol}$, etc.) varies slowly within each surface element, so that $\sigma_{pol}(\vec{s})$ may be assumed to be constant over a given area, $a_i$. Naturally, this approximation improves as the density of the surface grid increases and the individual areas become small. Discretization converts the surface charge density, $\sigma_{pol}(\vec{s})$, into a vector $q$, whose elements $q_i$ are the surface charges for the various surface area elements. In most PCM implementations, $q_i$ is treated as a point charge located at $\vec{s}_i$. Likewise, $\phi_0(\vec{s})$ is given by a vector $v$, where $v_i = \phi_0(\vec{s}_i)$.

Chipman and Dupuis have shown that discretization of the cavity surface results in a set of linear equations,

$$Kq = Rv,$$

that determine the vector $q$. The matrices $K$ and $R$ depend upon the particular PCM, and are written in terms of two other matrices, $S$ and $D$, that are discussed
in Section 6.3.1 and in Appendix 6.8. The forms of $K$ and $R$ for C-PCM/GCOSMO and IEF-PCM/SS(V)PE are listed in Table 6.1.

An equivalent form of Eq. (8.43) is

$$q = Qv$$  \hspace{1cm} (6.19)

where $Q = K^{-1}R$ is often called the response matrix, since the action of $Q$ on $v$ affords the surface charges $q$ that are induced in response to the electrostatic potential, $v$. Formally, computation of $q$ requires inversion of $K$, although iterative algorithms that avoid matrix diagonalization have been developed.\textsuperscript{192} Upon discretizing the integral that defines $E_{pol}$ in Eq. (6.12), one can express the electrostatic solvation energy in a variety of equivalent forms:

$$E_{pol} = \frac{1}{2}q^\dagger v = \frac{1}{2}q^\dagger Q^{-1}q = \frac{1}{2}v^\dagger Qv \right.$$  \hspace{1cm} (6.20)

The total energy is then

$$W = E_0 + \frac{1}{2}q^\dagger v \right.$$  \hspace{1cm} (6.21)

The variational principle analogous to Eq. (6.13) is

$$0 = \frac{\partial W}{\partial q^\dagger} = v - Q^{-1}q \right.$$  \hspace{1cm} (6.22)

which is equivalent to Eq. (8.44). However, the derivation of Eq. (6.22) assumes that $Q^\dagger = Q$, which is true within C-PCM/GCOSMO but not for IEF-PCM/SS(V)PE. Fortunately, $E_{pol}$ is invariant to the following symmetrization:

$$E_{pol} = \frac{1}{4}v^\dagger \left[ K^{-1}R + R^\dagger(K^\dagger)^{-1} \right] v \right.$$  \hspace{1cm} (6.23)
As such, we may replace $Q$ in the energy expressions with a symmetric matrix

$$
\tilde{Q} = \frac{1}{2} (Q + Q^\dagger) .
$$

(6.24)

This symmetrization is crucial for efficient evaluation of energy gradients.

To ensure that $W$ is minimized by the induced surface charges $q$, let us compute the second variation of $W$ with respect to $q$. Using Eq. (6.22), one obtains

$$
\frac{\partial^2 W}{\partial q^\dagger q} = -Q^{-1} .
$$

(6.25)

It follows that in order for $W$ to be a local minimum, the matrix $-Q^{-1}$ must be positive-definite, hence $Q$ must be negative-definite. If so, then it follows from Eq. (6.20) that $E_{pol} < 0$ for any $v$ and any $q$. The physical implication of this result is that the reaction field always lowers the total energy. We will revisit this important point in Section 6.4, in regard to the choice of matrix elements for the smooth version of IEF-PCM/SS(V)PE.

**Electrostatic gradients**

PCM analytic energy gradients have been reported previously, but we briefly summarize the formalism here, in order to make contact with the reaction-field formalism presented above. Differentiating the total energy with respect to a perturbation, $x$, and using the notation $W^x = \partial W/\partial x$, we have

$$
W^x = E_0^x + E_{pol}^x .
$$

(6.26)

The quantity $E_0^x$ is simply the gas-phase gradient of the solute, computed using a density matrix that is converged in the presence of the reaction field. The polarization
energy gradient is

\[ E_{pol}^{x} = \frac{1}{2} \left[ (v^\dagger)^x Qv + v^\dagger Q^x v + v^\dagger Qv^x \right]. \tag{6.27} \]

The first and third terms on the right are equivalent if \( Q \) is symmetric, but as indicated above, we may replace \( Q \) in this expression with the symmetric matrix \( \tilde{Q} \) from Eq. (6.24). This is required in order to avoid the appearance of the density matrix derivative, the calculation of which would require solution of coupled-perturbed equations.\(^{195}\) To avoid this costly endeavor, it is necessary to perform the symmetrization of \( Q \) in Eq. (6.24) or, alternatively, to calculate two sets of charges, \( q' = Qv \) and \( q'' = Q^\dagger v \). Combining the two sets as \( q = \frac{1}{2} (q' + q'') \) is equivalent to performing the symmetrization of Eq. (6.24).

Using Eq. (6.23) we can rewrite \( E_{pol}^{x} \) as

\[ E_{pol}^{x} = \frac{1}{2} v^\dagger \left( K^{-1} R \right)^x v + qv^x. \tag{6.28} \]

The physical interpretation of this equation is that the surface charges experience the full electric field of the solute but only half that of the the surface charge density, a consequence of the work required to induce the cavity surface charge. The second term
on the right in Eq. (6.28) is straightforward to compute, since \( \mathbf{q} \) must be computed anyway in order to obtain \( E_{\text{pol}} \), and \( \mathbf{v}^x \) is simply the electric field at a given surface grid point, a standard quantity at both MM and QM levels of theory. The remaining term in \( E_{\text{pol}}^x \) can be rearranged and simplified into a general form for all PCMs:

\[
\mathbf{v}^\dagger (K^{-1}R)^x \mathbf{v} = \mathbf{v}^\dagger K^{-1} (R^x - K^x K^{-1} R) \mathbf{v}
\]  

(6.29)

Further manipulation of this term requires us to choose a particular PCM. Expressions for \( K^x \) and \( R^x \) for both C-PCM/GCOSMO and IEF-PCM/SS(V)PE are provided in Table 6.2.

### 6.3 The switching/Gaussian method for cavity discretization

We now come to the main topic of this work: the switching/Gaussian (SWIG) cavity discretization procedure.

#### 6.3.1 Surface charge representation and matrix elements

An important distinction between various cavity discretization methods is the manner in which the surface charge density, \( \sigma_{\text{pol}}(\mathbf{s}) \), is represented, leading to different definitions of the matrices \( \mathbf{S} \) and \( \mathbf{D} \) upon which \( \mathbf{K} \) and \( \mathbf{R} \) depend (see Table 6.1). The matrices \( \mathbf{S} \) and \( \mathbf{D} \) are the discrete forms of certain integral operators, \( \hat{S} \) and \( \hat{D} \), that act on \( \sigma_{\text{pol}}(\mathbf{s}) \), and which are briefly reviewed in Appendix 6.8. In discrete form, the action of \( \mathbf{S} \) on \( \mathbf{q} \) produces the electrostatic potential due to \( \mathbf{q} \), evaluated at the set of points \( \{\mathbf{s}_i\} \). The action of \( \mathbf{D} \) on \( \mathbf{q} \) produces the surface-dipole (double-layer) potential, again evaluated at the points \( \{\mathbf{s}_i\} \). The action of \( \mathbf{D}^\dagger \) on \( \mathbf{q} \) affords
the surface normal component of the electric field. Multiplying by the area $a_i$ for each surface point provides an approximate normal electric field over the $i$th surface element, and a compact notation is obtained by introducing a diagonal matrix, $A$, with elements $A_{ij} = a_i \delta_{ij}$ (see Table 6.1).

The matrix elements of $S$ and $D$ are derived from the Coulomb operator between surface charges, and thus depend upon how this surface charge is represented. The most common choice is to use point charges $q_i$ located at the discretization points $\vec{s}_i$, but following YK,\textsuperscript{168} we will represent $\sigma_{pol}(\vec{s})$ using spherical Gaussians centered at the points $\vec{s}_i$:

$$g_i(\vec{r}) = q_i (\zeta_i^2 / \pi)^{3/2} \exp(-\zeta_i^2 |\vec{r} - \vec{s}_i|^2) .$$ \hspace{1cm} (6.30)

The amplitude $q_i$ in this equation is precisely the charge that appears in the vector $\mathbf{q}$, and the exponent $\zeta_i$ is allowed to depend upon the surface area of the $i$th surface element. Specifically, if we denote the radius of the $I$th atomic sphere by $R_I$, then for a Lebedev quadrature point $i \in I$ located on this sphere, the Gaussian exponent, $\zeta_i$, is chosen to be\textsuperscript{168}

$$\zeta_i = \frac{\zeta}{R_I \sqrt{w_i}} .$$ \hspace{1cm} (6.31)

Here, $\zeta$ is a width parameter that is optimized, for each particular Lebedev grid, in order to reproduce the Born solvation energy of a conductor and a uniform surface charge distribution.\textsuperscript{168,197} (Numerical values of $\zeta$ are taken from Ref. 168.)

The Gaussian charge basis eliminates the Coulomb singularity present in point charge interactions, so that the surface charge interactions remain finite at all geometries. This is important, as we have found that the use of a switching function can
exacerbate problems with singularities in the PCM equations, by allowing surface grid points to approach more closely than they would in standard (discontinuous) discretization schemes.\textsuperscript{186}

Having selected a representation for the surface charge, the off-diagonal matrix elements of $S$ and $D$ follow quite readily. The off-diagonal element $S_{ij}$ is simply the Coulomb interaction between $g_i$ and $g_j$, which can be evaluated analytically. The result is

$$S_{ij} = \frac{\text{erf}(\zeta_{ij}r_{ij})}{r_{ij}},$$  \hspace{1cm} (6.32)

where $r_{ij} = |\vec{r}_i - \vec{r}_j|$ and $\zeta_{ij} = \zeta_i \zeta_j / (\zeta_i^2 + \zeta_j^2)^{1/2}$. The off-diagonal element $D_{ij}$ is related to $S_{ij}$ according to the relation\textsuperscript{159}

$$D_{ij} = \vec{n}_j \cdot \frac{\partial S_{ij}}{\partial \vec{r}_j},$$  \hspace{1cm} (6.33)

where $\vec{n}_j$ is the outward pointing unit vector normal to the cavity surface, at the point $\vec{s}_j$. Using Eqs. (B.4) and (8.50), we obtain

$$D_{ij} = \left( \text{erf}(\zeta_{ij}r_{ij}) - \frac{2\zeta_{ij}e^{-\zeta_{ij}^2}}{\sqrt{\pi}} e^{-\zeta_{ij}^2 r_{ij}^2} \right) \frac{\vec{n}_j \cdot \vec{r}_{ij}}{r_{ij}^3},$$  \hspace{1cm} (6.34)

where $\vec{r}_{ij} = \vec{r}_i - \vec{r}_j$.

The diagonal matrix elements $S_{ii}$ and $D_{ii}$, which are termed the \textit{self-potential} and the \textit{self-field} interactions, respectively,\textsuperscript{4} are not straightforward to define because they involve Coulomb interactions within a discretized surface element. At the same time, these self-interactions turn out to be critically important to obtaining a smooth potential energy surface as well as solvation energies that are faithful to the integral equation PCM theory. Because Gaussian charges eliminate the singularity in the
Coulomb potential, an obvious choice for the self-potential interaction is based upon
the observation that
\[
\lim_{r_{ij} \to 0} \frac{\text{erf}(\zeta_i r_{ij})}{r_{ij}} = \zeta_i \sqrt{2/\pi} .
\] (6.35)
We use this limit to define \( S_{ii} \). To ensure smoothness, however, \( S_{ii} \) must be scaled
by the inverse of the switching function \( F_i \), consistent with the original YK prescrip-
tion.\(^{168}\) Thus, we define
\[
S_{ii} = \frac{\zeta_i \sqrt{2/\pi}}{F_i} .
\] (6.36)
The precise nature of the switching function is discussed in Section 6.3.2. For now, it
suffices to note that the presence of \( F_i \) in the denominator of \( S_{ii} \) ensures continuity
of the potential surface, as shown in Appendix 6.9.

The quantity \( D_{ii} \) is related to the interaction of \( \sigma_{pol}(\vec{s}_i) \) with its own normal
electric field component, and the product \( D_{ii} a_i \) gives the total interaction over the
entire area \( a_i \). In our previous work,\(^{186}\) we noted that \( \lim_{r_{ij} \to 0} D_{ij} = 0 \), and used this
as justification to set \( D_{ii} = 0 \). While this choice is certainly valid in the limit that
\( a_i \to 0 \), in practice it tends to degrade the accuracy of solvation energies,\(^6\) and our
subsequent experience has shown that smooth PCMs with \( D_{ii} = 0 \) cannot reproduce
Born ion solvation energies across a range of dielectric constants. In the present work,
we define
\[
D_{ii} = -\frac{\zeta_i \sqrt{2/\pi}}{2R_I} ,
\] (6.37)
where \( R_I \) is the radius of the atomic sphere on which the point \( \vec{s}_i \) resides. This choice
is based upon the definition \( D_{ii} = S_{ii}/2R_I \) from Ref. 198. We find that this approach
does reproduce Born-ion solvation energies
Having defined $S$ and $D$, the analytic gradients for SWIG can be derived by taking the gradient of the matrix elements with respect to the $M$th solute nucleus. For $S_{ij}$ with $i \neq j$, we obtain

$$\hat{\nabla}_M S_{ij} = - \left( \text{erf}(\zeta_{ij} r_{ij}) - \frac{2 \zeta_{ij} r_{ij} e^{-\zeta_{ij}^2 r_{ij}^2}}{\sqrt{\pi}} \right) \hat{\nabla}_M r_{ij} r_{ij}^2,$$  

(6.38)

whereas

$$\hat{\nabla}_M S_{ii} = - \frac{\zeta_i \sqrt{2/\pi}}{F_i^2} \hat{\nabla}_M F_i.$$  

(6.39)

The switching function gradient, $\hat{\nabla}_M F_i$, is given in Appendix 6.10. For the diagonal elements of $D$, we have

$$\hat{\nabla}_M D_{ii} = 0,$$  

(6.40)

whereas for $i \neq j$ we obtain

$$\hat{\nabla}_M D_{ij} = \left( \frac{4 r_{ij}^2 \zeta_i^3 e^{-r_{ij}^2 \zeta_i^2}}{\sqrt{\pi}} \right) \nabla_{ij} - \left( \text{erf}(\zeta_{ij} r_{ij}) - \frac{2 \zeta_{ij} r_{ij} e^{-\zeta_{ij}^2 r_{ij}^2}}{\sqrt{\pi}} \right) \hat{\nabla}_M r_{ij} r_{ij}^2,$$  

(6.41)

An equivalent form of Eq. (6.41) was given in Ref. 186.

The quantity $\hat{\nabla}_M r_{ij}$, which appears in both Eq. (6.38) and Eq. (6.41), vanishes unless either $i \in M$ or $j \in M$. We can express this derivative as

$$\hat{\nabla}_M r_{ij} = \frac{\vec{r}_i - \vec{r}_j}{r_{ij}} (\delta_{iM} - \delta_{jM}),$$  

(6.42)

where $\delta_{iM} = 1$ if $i \in M$, else $\delta_{iM} = 0$. 

134
6.3.2 Switching functions

We have already introduced a dimensionless switching function, \( F_i \), for the \( i \)th discretization point, although its precise form has not yet been specified. The function \( F_i \) should smoothly attenuate the \( i \)th surface point’s contribution to the solvation energy, as \( \vec{s}_i \) passes into or out of the cavity. The cavity consists of a union of spheres centered at points \( \{ \vec{r}_J \} \), and a reasonable ansatz for \( F_i \) is a product of elementary switching functions,\(^{168,186}\)

\[
F_i = \prod_{J, i \notin J}^{\text{atoms}} f(\vec{s}_i, \vec{r}_J) .
\]

(6.43)

The elementary switching functions, \( f(\vec{s}_i, \vec{r}_J) \), should vary smoothly and monotonically between zero and one. Equation (6.43) ensures that \( F_i \) depends upon the nuclear coordinates of the entire solute molecule, and will be zero if even a single \( f(\vec{s}_i, \vec{r}_J) = 0 \). In practice, we set \( f(\vec{s}_i, \vec{r}_J) \) to zero whenever this function drops below \( 10^{-8} \). In our experience, this does not produce any numerically detectable discontinuities.

There are probably many different choices for the function \( f \) that would work in this context; we will describe two choices that we have found to work well in practice. The first of these, originally proposed by YK,\(^{168}\) is based upon definition of a switching region around each atomic sphere. Using \( R_{\text{sw},J} \) to denote the width of the switching region around the \( J \)th atom, the inner and outer boundaries of the switching region for atom \( J \) are defined by radii

\[
R_{\text{in},J} = R_J - \alpha_J R_{\text{sw},J}
\]

(6.44)
and
\[ R_{\text{out},J} = R_J + (1 - \alpha_J)R_{\text{sw},J}, \quad (6.45) \]
respectively, where \( \alpha_J \) is an adjustable parameter \((0 < \alpha_J < 1)\). The definitions of the parameters \( \alpha_J \) and \( R_{\text{sw},J} \) depend upon \( R_J \) and also the number of Lebedev grid points used to discretize the \( J \)th sphere; these definitions, which we take from Ref. 168, are provided in Appendix B.

The extent to which the \( i \)th grid point penetrates into the switching region surrounding the \( J \)th sphere is measured using the dimensionless quantity
\[ d_{iJ} = \frac{|\vec{s}_i - \vec{r}_J| - R_{\text{in},J}}{R_{\text{sw},J}}. \quad (6.46) \]
The elementary switching function used by YK is then given by
\[ f(\vec{s}_i, \vec{r}_J) = h(d_{iJ}), \quad (6.47) \]
where
\[ h(x) = \begin{cases} 
0 & x < 0 \\
 x^3(10 - 15x + x^2) & 0 \leq x \leq 1 \\
1 & x > 1 
\end{cases}. \quad (6.48) \]

In later work, York and co-workers\(^ {171} \) replaced \( F_i \) in Eq. (6.43) with \( F_i^p \), where the exponent \( p \) was taken to be an additional adjustable parameter. However, we have demonstrated that values of \( p \neq 1 \) can lead to unwanted oscillations in the energy gradient,\(^ \text{186} \) so we take \( p = 1 \) throughout this work.

The switching function defined above is certainly not unique, and even within this \textit{ansatz}, parameters such as \( R_{\text{sw},J} \) and \( \alpha_J \) are not unique either. York and co-workers\(^ {168,171} \) have presented some arguments in favor of simple formulas that define
these parameters, and we have adopted these definitions because they produce good results in a variety of tests. It is likely that a somewhat different switching function could be found that, in conjunction with Gaussian surface charges, also affords good results.

To this end, we have explored some alternative switching functions, to make comparison and perhaps to diminish some of the arbitrariness in this aspect of the SWIG method. One such alternative is to exploit the spatial extent of the spherical Gaussian charges to determine the penetration of a surface element into the cavity. The idea is to compute the fraction of the Gaussian charge distribution $g_i$ that exists inside of the $J$th sphere. Restricting ourselves to one dimension for simplicity, we obtain the resulting elementary switching function

$$f(\vec{s}_i, \vec{r}_J) = 1 - \frac{1}{2} \left\{ \operatorname{erf}[\zeta_i (R_J - r_{iJ})] + \operatorname{erf}[\zeta_i (R_J + r_{iJ})] \right\},$$

where $R_J$ is the radius of the $J$th sphere and $r_{iJ} = |\vec{s}_i - \vec{r}_J|$. The term in curly braces measures how much normalized charge density exists inside the $J$th sphere. Thus, $f(\vec{s}_i, \vec{r}_J) = 0$ when the $i$th Gaussian is contained entirely within the $J$th sphere. To distinguish between the elementary switching function in Eq. (6.47), versus that defined in Eq. (6.49), we henceforth refer to the latter approach as the “improved SWIG” (ISWIG) method, whereas “SWIG” will imply that the YK switching function in Eq. (6.47) is used. In all other respects, the two methods are identical.

A potential advantage of ISWIG is that it accounts for differences in the relative
extents of the Gaussian charges, whereas the SWIG approach uses a switching function that is independent of the Gaussian exponents, $\zeta_i$. Using ISWIG, a Gaussian charge with a small width will be rapidly attenuated as it penetrates into an atomic sphere, whereas SWIG will attenuate this charge more slowly, if the switching region is large relative to the Gaussian charge width. ISWIG should also avoid situations in which a broad Gaussian charge is centered just outside of a switching region, with a substantial amount of its charge distribution inside the cavity, yet is fully switched on. On the other hand, the SWIG switching function is somewhat more efficient computationally, since its switching regions extends over a fairly narrow region around each atomic sphere. In numerical tests performed so far, both SWIG and ISWIG afford similar results.

6.4 Comparison to other discretization methods

We next compare and contrast our discretization methods to two others that have recently been proposed as solutions to the PCM discontinuity problem, namely the “fixed points with variable areas” (FIXPVA) method of Su and Li, and the “continuous surface charge” (CSC) method of Scalmani and Frisch. All three of these methods claim to yield smooth potential energy surfaces through the use of a switching function, yet there exist subtle but important differences between them.

6.4.1 Continuity

Matrix elements for the FIXPVA and CSC methods are given in Table 6.3. Of these two approaches, CSC is more similar to SWIG, and in fact uses the same off-diagonal
<table>
<thead>
<tr>
<th>Matrix element</th>
<th>FIXPVA</th>
<th>CSC</th>
</tr>
</thead>
<tbody>
<tr>
<td>$S_{ij}$</td>
<td>$1/r_{ij}$</td>
<td>$\text{erf}(\zeta_{ij}r_{ij})/r_{ij}$</td>
</tr>
<tr>
<td>$S_{ii}$</td>
<td>$C_S \sqrt{\frac{4\pi}{a_i}}$</td>
<td>$\left(4\pi R_I - \sum_{j \neq i} S_{ij}a_j\right)/a_i$</td>
</tr>
<tr>
<td>$D_{ij}$</td>
<td>$\tilde{n}<em>j \cdot \tilde{r}</em>{ij}/r_{ij}^3$</td>
<td>$\left(\text{erf}(\zeta_{ij}r_{ij}) - \frac{2\zeta_{ij}r_{ij}}{\sqrt{\pi}} e^{-\zeta_{ij}^2 r_{ij}^2}\right) \tilde{n}<em>j \cdot \tilde{r}</em>{ij}/r_{ij}^3$</td>
</tr>
<tr>
<td>$D_{ii}$</td>
<td>$-\left(2\pi + \sum_{j \neq i} D_{ij}a_j\right)/a_i$</td>
<td>$-\left(2\pi + \sum_{j \neq i} D_{ij}a_j\right)/a_i$</td>
</tr>
</tbody>
</table>

Table 6.3: Matrix elements of $S$ and $D$ for the FIXPVA and CSC methods. In Ref. 4, the CSC matrix elements are defined for spheres of unit radius, but we have generalized them here to arbitrary areas, $a_i$. We have also generalized the FIXPVA approach of Ref. 5 for use with IEF-PCM/SS(V)PE, as described in the text. In the expressions for $S_{ii}$, $R_I$ is the radius of the $I$th sphere ($i \in I$), and the constant $C_S$ is a self-energy factor that depends upon the choice of surface grid: $C_S = 1.0694$ for GEPOL grids\textsuperscript{6} and $C_S = 1.104$ for Lebedev grids.\textsuperscript{7}

Elements $S_{ij}$ and $D_{ij}$, in conjunction with Gaussian charges, and the same switching function as that defined in Eq. (6.47). The FIXPVA approach uses point charges, and this choice necessitates the use of an alternative switching function,\textsuperscript{5} as close approach of these point charges must be avoided.

The switching function does not appear explicitly in the matrix elements $S_{ii}$ for the FIXPVA and CSC methods, but is instead folded into the areas $a_i$. However, we can cast the SWIG expression for $S_{ii}$ [Eq. (6.36)] into something that more closely resembles the expressions in Table 6.3 by using Eq. (6.31) to relate $\zeta_i$ to $a_i = w_i R_I^2 F_i$:

$$S_{ii} = \zeta_i \sqrt{\frac{2}{\pi a_i F_i}}.$$  \hfill (6.50)

Thus, SWIG, CSC, and FIXPVA all attenuate surface points via the diagonal of $S$, 139
and in Appendix 6.9 we show that this is sufficient to obtain a smooth potential energy surface. As such, all three methods afford potential energy surfaces that are smooth, in the mathematical sense.

6.4.2 Sum rules

As compared to the off-diagonal elements of $S$ and $D$, formulation of the self-energy and self-field within a given surface element, and therefore the diagonal matrix elements $S_{ii}$ and $D_{ii}$, is less straightforward. Traditionally, $D_{ii}$ has been defined using a sum rule,

$$D_{ii} = -\frac{1}{a_i} \left(2\pi + \sum_{j \neq i} D_{ij} a_j \right) ,$$

which avoids the need to determine $D_{ii}$ by ill-conditioned numerical integration over the $i$th surface element, and furthermore provides more accurate energetics than simply setting $D_{ii} = 0$.\(^6,172\) This sum rule was originally derived by Purisima and Nilar,\(^173\) starting from the discretized expression for $\hat{D}$,

$$\sum_j \int \frac{(\vec{r}_j - \vec{r}_i) \cdot \vec{n}_j}{r_{ij}^3} d^2 s_j = -\sum_j D_{ij} a_j .$$

These authors recognized that the integrals appearing in Eq. (6.52) could be rewritten as a solid angle integral over each surface element, from which it follows that

$$-\sum_j D_{ij} a_j = \sum_j \int d\Omega_j = 2\pi .$$

Here, $d\Omega_j$ is the solid angle subtended by the $j$th surface element, from the vantage point $\vec{s}_i$ on the the cavity surface.

Equation (6.51) is used to define $D_{ii}$ in the CSC method.\(^4\) In addition, we have used this sum rule to extend the FIXPVA discretization method to the
IEF-PCM/SS(V)PE approach (see Table 6.3), whereas FIXPVA discretization was previously introduced only in the context of C-PCM/GCOSMO.\footnote{5}

For $S_{ii}$, the FIXPVA method employs the widely-used formula $S_{ii} = C_S (4\pi/a_i)^{1/2}$, where $C_S$ is a numerical constant representing the average value of the Coulomb self-repulsion integral over surface elements of various sizes.\footnote{6,7,165} The CSC method, in contrast, uses a sum rule

$$S_{ii} = \frac{1}{a_i} \left( 4\pi R_I - \sum_{j \neq i} S_{ij} a_j \right) \quad (6.54)$$

that is derived\footnote{4} by minimizing an approximate functional for the Born solvation energy of a spherical ion in a conductor, with respect to the Coulomb interaction between surface charges. Only the CSC approach uses the sum rule in Eq. (6.54).

Because the surface charge is assumed to be constant over the area $a_i$, the self-interaction over $a_i$ is strictly repulsive. Therefore, the quantities $q_i^2 S_{ii}$ and $-q_i^2 D_{ii}$ must both be positive. (The matrix $D$ is traditionally defined such that $-D_{ii}$ is the actual contribution of the self-field to the PCM equations.) In the SWIG approach developed here, these matrix elements are defined such that $S_{ii} > 0$ and $D_{ii} < 0$ by construction, and $S_{ii} > 0$ is also guaranteed in the FIXPVA approach. When sum rules are used to define these matrix elements, however, there is no guarantee that the correct sign is obtained. (In earlier work,\footnote{186} this problem led us to set $D_{ii} = 0$.) For the CSC discretization scheme, which uses sum rules to define both $S_{ii}$ and $D_{ii}$, this is potentially a problem for all variants of PCM theory, including the C-PCM/GCOSMO method where the $D$ matrix is absent.
From Eq. (6.51) we see that \( D_{ii} > 0 \) if \( \sum_{j \neq i} D_{ij} a_j < -2\pi \), for some particular cavity geometry. This condition occurs readily when point charges are used in conjunction with a switching function, due to close approach of point charges leading to some \( D_{ij} \ll 0 \). Even when Gaussians are employed, however, we have found that the requirement \( D_{ii} < 0 \) is regularly violated for any solute cavity consisting of more than one sphere.

Clearly, Eq. (6.51) is no longer exact within these smooth PCM approaches. One reason is that the relationship between the integral in Eq. (6.52) and the solid angle integral in Eq. (6.53) is derived based on a pointwise discretization, whereas our surface elements are spherical Gaussians. More importantly, Eq. (6.53) is valid only if the surface elements and vantage points all reside on a closed surface, but when switching functions are used, some of the “surface” points \( \vec{s}_i \) actually reside inside the cavity. For vantage points inside of the cavity, the total solid angle subtended by the surface is \( \sum_j \int d\Omega_j = 4\pi \).\(^{173}\) Thus, if \( \vec{s}_i \) lies inside of the cavity surface, we should expect that

\[
-\sum_{j \neq i} D_{ij} a_j \approx 4\pi ,
\]

which we have verified in numerical calculations. It follows from Eqs. (6.51) and (6.55) that \( D_{ii} a_i \approx 2\pi \) in this case, in violation of the condition \( D_{ii} < 0 \). We have not inspected the sum rule for \( S_{ii} \), Eq. (6.54), but it may be subject to similar issues.

Breakdown of the sum rules has serious consequences. In our initial report of SWIG,\(^{186}\) we mentioned the sum rule issue for \( D_{ii} \) and pointed out that it can compromise the positive-definiteness of the matrix \( \mathbf{K} \), causing numerical instabilities. We
have since determined that loss of positive-definiteness of $K$ is not the fundamental origin of these instabilities. Instead, they arise due to violation of the negative-definiteness of $Q$, resulting in non-variational surface charge densities and singularities in $E_{pol}$. A numerical exploration of this issue is presented in Section 6.5.1.

6.5 Numerical tests

In this section, we present numerical comparisons of various discretization methods, along with numerical tests of convergence with respect to the Lebedev grid density. We have implemented the SWIG, ISWIG, and FIXPVA discretization schemes with a locally-modified version of Q-Chem. To make contact with the CSC approach, we have also implemented a variant of SWIG that uses the CSC (sum rule) definition of $D_{ii}$ in Eq. (6.51), as opposed the SWIG definition in Eq. (6.37). We refer to this modified approach as “subSWIG”.

Except for one set of calculations in Section 6.5.2, where spherical solute cavities are used, all calculations use cavities constructed from a union of atomic spheres. For QM solutes, we use Bondi’s values for the vdW radii, except for hydrogen where a vdW radius of 1.1 Å is used. The vdW radii are then scaled by a factor of 1.2 for use in cavity construction. For MM solutes, unscaled Lennard-Jones radii from the AMBER99 force field are used to construct the cavity.

6.5.1 Bond breaking

Dissociation of NaCl in water provides an illustrative example of the consequences that accompany breakdown of the sum rule that is used to define $D_{ii}$ in the
CSC/subSWIG approach. Here, we compute various quantities along the Na–Cl dissociation coordinate, using the AMBER99 force field for the NaCl solute, and the IEF-PCM/SS(V)PE solvation model for the aqueous solvent.

Figure 6.1 shows some results using subSWIG discretization. A single, sharp spike in $E_{pol}$ (for which $E_{pol} > 0$) can be seen in the solvation energy curve, while the gradient of $E_{pol}$ is rapidly oscillatory. Figure 6.1(b) shows that rapid oscillations in the gradient are correlated with geometries for which the $Q$ matrix exhibits one or more positive eigenvalues. In contrast, SWIG discretization (Fig. 6.2) produces comparatively minor oscillations in the gradient, and $Q$ remains rigorously negative-definite at all solute geometries.

The failure of subSWIG to preserve negative-definiteness of $Q$ opens the possibility of non-variational solvation energies ($E_{pol} > 0$), although in Fig. 6.1(a) we see only a single instance where $E_{pol} > 0$, and it occurs at an unrealistically short Na–Cl distance, where one sphere completely envelops the other. However, the benign appearance of Fig. 6.1(a) turns out to be an artifact of the relatively large spacing (0.01 Å) between the data points. If we take much smaller steps, as in Fig. 6.3, we discover sharp spikes in the solvation energy at other internuclear distances, which are correlated in each case with positive eigenvalues of $Q$. These singularities exist despite the fact that subSWIG employs both a switching function and Gaussian surface charges, and is equivalent to SWIG except that it uses a sum rule to define $D_{ii}$.

One could reasonably argue that the use of AMBER99 point charges is inappropriate for bond-breaking, and that ionic radii for Na$^+$ and Cl$^-$ should not be used when the Na–Cl distance is small. However, our intention here is not to develop an accurate model of bond-breaking in solution, but rather to explore the singularities that can occur as atomic spheres interpenetrate, and this simple example suffices for that purpose.
Figure 6.1: NaCl dissociation in water ($\varepsilon = 78.39$), computed at the AMBER99-SS(V)PE level using the “subSWIG” discretization scheme in which $D_{ii}$ is defined using the sum rule in Eq. (6.51). Panel (a) plots the solvation energy, $E_{\text{pol}}$, and its gradient with respect to Cl displacement, while panel (b) plots the largest eigenvalue of $Q$, along with the number of surface grid points for which $F_i > 10^{-8}$. Data points were calculated at 0.01 Å intervals. For clarity, the vertical scale has been truncated in both panels, i.e., some of the sharp spikes are off of the scale that is used.
Figure 6.2: NaCl dissociation in water ($\varepsilon = 78.39$), computed at the AMBER99-SS(V)PE level using the SWIG discretization scheme. Panel (a) plots the solvation energy, $E_{pol}$, and its gradient with respect to Cl displacement, while panel (b) plots the largest eigenvalue of $Q$, along with the number of surface grid points for which $F_i > 10^{-8}$. Data points were calculated at 0.01 Å intervals.
The regions where $E_{pol} > 0$ are highly localized, so it is not surprising that the energy curve in Fig. 6.1(a), where the data points are 0.01 Å apart, fortuitously avoids these anomalies. Nevertheless, we will see that these singularities are encountered in realistic calculations on polyatomic solutes, and that they pose real problems.

Examining the largest eigenvalue of $Q$ in the case of SWIG discretization [Fig. 6.2(b)], one might worry about what appear to be sharp jumps in this eigenvalue, but in fact these oscillations are perfectly natural. As shown in Appendix 6.9, the switching function causes $Q$ to exhibit a null space corresponding to those surface elements for which $F_i = 0$, so that the largest (i.e., least negative) eigenvalue of $Q$ must approach zero as any $F_i \to 0$. Figure 6.2(b) shows that rapid oscillations in this eigenvalue are indeed correlated with changes in the number of grid points for which $F_i$ exceeds our drop tolerance of $10^{-8}$.

### 6.5.2 Solvation energies

Certain PCM theories and cavity shapes may be more or less accurate than others for predicting solvation free energies, but ideally we would like to separate this aspect of the model from the smoothing procedure, so that the smoothed model is faithful to the underlying PCM and does not significantly perturb solvation energies, relative to those obtained using established discretization schemes. As a representative example of the latter, we choose the “variable tesserae number” (VTN) discretization method,\textsuperscript{175} which uses point charges and does not employ a switching function. The VTN scheme was developed as a “less discontinuous” discretization method, and
Figure 6.3: Solvation energy, $E_{pol}$, and eigenvalues of $Q$, near a singular point of the subSWIG potential energy surface for NaCl dissociation. (Data points are calculated every $10^{-4}$ Å.) The inset shows a larger range of internuclear distances, and suggests that singularities may go unnoticed unless the spacing between data points is extremely small.
serves as a baseline against which to compare the SWIG, CSC, and FIXPVA methods. The VTN method uses the same matrix elements as FIXPVA (see Table 6.3), but does not scale the areas with a switching function. Instead, $F_i$ in Eq. (6.17) is replaced with a delta function, according to whether the $i$th grid point is inside of the cavity or not. VTN calculations are therefore subject to Coulomb singularities as well as discontinuities.

As a first set of tests, we examine the molecules H$_2$O, CH$_3$CONH$_2$, NO$^+$, and CN$^-$, which were previously used by Chipman$^8$ to test his isodensity implementation of the SS(V)PE model. We do not employ an isodensity cavity construction, but instead place each molecule inside of a spherical cavity centered at the molecule’s center of mass, and discretized using 1202 Lebedev grid points. The radius of this cavity is selected in order to replicate the solvation energies reported by Chipman,$^8$ computed in our case using SS(V)PE[VTN]. Using a spherical cavity allows us to sidestep issues of continuity and smoothness in this set of tests, and thereby examine the extent to which the Gaussian surface charges perturb solvation energies, as compared to the baseline VTN discretization. (Note that the switching function is irrelevant here, since the cavity consists of a single sphere.)

Solvation energies for the aforementioned molecules are reported in Table 6.4, where they were computed using the same procedure as in Ref. 8. Specifically, each molecule was optimized in the gas phase at the Hartree-Fock (HF) level to obtain the gas-phase energy, $E_{\text{gas}}$. The 6-31G** basis set was used for all molecules except CN$^-$, for which we used the 6-31+G* basis. Solvation energies, $E_{\text{slvn}} = W - E_{\text{gas}}$, were
<table>
<thead>
<tr>
<th>Theory</th>
<th>Method</th>
<th>H$_2$O</th>
<th>CH$_3$CONH$_2$</th>
<th>NO$^+$</th>
<th>CN$^-$</th>
</tr>
</thead>
<tbody>
<tr>
<td>C-PCM</td>
<td>VTN</td>
<td>−8.69</td>
<td>−10.91</td>
<td>−88.47</td>
<td>−67.32</td>
</tr>
<tr>
<td></td>
<td>SWIG</td>
<td>−8.69</td>
<td>−10.93</td>
<td>−88.49</td>
<td>−67.30</td>
</tr>
<tr>
<td>IEF-PCM</td>
<td>VTN</td>
<td>−8.58</td>
<td>−10.81</td>
<td>−88.47</td>
<td>−67.31</td>
</tr>
<tr>
<td></td>
<td>SWIG</td>
<td>−8.62</td>
<td>−10.84</td>
<td>−88.48</td>
<td>−67.29</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Theory</th>
<th>Method</th>
<th>H$_2$O</th>
<th>CH$_3$CONH$_2$</th>
<th>NO$^+$</th>
<th>CN$^-$</th>
</tr>
</thead>
<tbody>
<tr>
<td>C-PCM</td>
<td>VTN</td>
<td>−4.90</td>
<td>−6.43</td>
<td>−52.30</td>
<td>−39.56</td>
</tr>
<tr>
<td></td>
<td>SWIG</td>
<td>−4.90</td>
<td>−6.45</td>
<td>−52.31</td>
<td>−39.55</td>
</tr>
<tr>
<td>IEF-PCM</td>
<td>VTN</td>
<td>−3.92</td>
<td>−5.01</td>
<td>−52.21</td>
<td>−39.35</td>
</tr>
<tr>
<td></td>
<td>SWIG</td>
<td>−3.92</td>
<td>−5.02</td>
<td>−52.22</td>
<td>−39.34</td>
</tr>
</tbody>
</table>

Table 6.4: Comparison of solvation energies (in kcal/mol), using a solute cavity consisting of a single sphere whose radius is adjusted in order to reproduce the solvation energies reported in Ref. 8, where an isodensity contour was used to define the cavity surface.
then computed at the gas-phase geometry, where \( W \) [Eq. (6.3)] is the solution-phase energy for some particular choice of PCM and discretization method. We performed these calculations in both a high-dielectric solvent (water) and a low-dielectric solvent (toluene). The results in Table 6.4 demonstrate that SWIG discretization reproduces VTN energies to within 0.04 kcal/mol, in either solvent, using both the IEF-PCM/SS(V)PE solvation model and the C-PCM/GCOSMO model. We take these results to indicate that Gaussian smoothing of the surface charges has a negligible effect on the energetics of the underlying PCM.

To ascertain the effects of the switching function on \( E_{\text{solv}} \), we need to use nonspherical cavities, so we repeated the calculations on \( \text{H}_2\text{O}, \text{CH}_3\text{CONH}_2, \text{NO}^+, \) and \( \text{CN}^- \) using a vdW cavity constructed from a union of atomic spheres, as described above. Each sphere was discretized using a 1202-point Lebedev grid, discarding any points for which \( F_i < 10^{-8} \). Solvation energies are reported in Table 6.5 for the VTN, SWIG, ISWIG, and subSWIG discretization methods.

These various discretization schemes afford very similar solvation energies at the C-PCM/GCOSMO level, but results computed using IEF-PCM/SS(V)PE vary over a somewhat wider range. In particular, the SWIG and subSWIG discretization schemes differ by as much as 1.3 kcal/mol for \( \text{NO}^+ \) and \( \text{CN}^- \) in water, the cases where the solvation energy is largest. VTN solvation energies are quite close to those obtained using other discretization methods, except in the case of \( \text{CH}_3\text{CONH}_2 \) in water, where discrepancies as large as 1.6 kcal/mol are observed.

While it is not completely clear what is the most accurate value of \( E_{\text{solv}} \), we
Table 6.5: Comparison of solvation energies (in kcal/mol) using a vdW cavity composed of atomic spheres. Note that SWIG and subSWIG discretization procedures are equivalent for C-PCM/GCOSMO.
are inclined to prefer the SWIG and ISWIG results, which agree with the other
discretization schemes at the C-PCM/GCOSMO level, where the $D$ matrix is absent,
yet are free of singularities at the IEF-PCM/SS(V)PE level. The SWIG and ISWIG
solvation energies in Table 6.5 differ from one another by no more than 0.6 kcal/mol,
which is comparable to (or smaller than) the intrinsic error in solvation energies
computed using models of this type.\textsuperscript{200}

\subsection{6.5.3 Discretization errors}

Like most cavity discretization schemes, the Lebedev grids used here do not preserve
rotational invariance of the solvation energy. In our implementation, this is because
each octahedral Lebedev grid is constructed in the laboratory frame, then translated
to the atomic center in question. Although the energy will be rotationally invariant
in the limit of an infinitely-dense discretization grid, it is desirable to know what is
the smallest number of Lebedev grid points that affords tolerable errors.

We assess rotation invariance for the SWIG discretization scheme using the twenty
standard amino acids as a data set. We choose these molecules because ($i$) MM
parameters are readily available for them; ($ii$) their side chains exhibit a variety
of chemical properties (polar and non-polar, charged and neutral, etc.); and ($iii$)
proteins are often investigated with implicit solvent models. The geometries of the
amino acids were generated, in their zwitterionic forms, using the TINKER program,\textsuperscript{58}
and were not further optimized. For each amino acid, we calculated the energy at
each of ten randomly-generated rotations of the TINKER structure, and quantified the

153
Figure 6.4: Rotational variance, $\Delta_{\text{rot}}$, and Gauss’ Law error, $\Delta_{\text{GL}}$, as a function of the number of grid points per atom, for a set of 20 amino acids in water ($\varepsilon = 78.39$). Error bars represent one standard deviation about the mean. For clarity, the error bars are omitted for $\Delta_{\text{GL}}$.

Rotational invariance according to the quantity

$$
\Delta_{\text{rot}} = \frac{1}{20} \sum_{A} \left[ \frac{1}{10} \sum_{i=1}^{10} (W_{A,i} - \overline{W}_{A})^2 \right]^{1/2},
$$

(6.56)

where $W_{A,i}$ represents the energy of the $i$th orientation of amino acid $A$, and $\overline{W}_{A}$ is the average over the ten values of $W_{A,i}$.

Figure 6.4 plots $\Delta_{\text{rot}}$ versus the number of Lebedev grid points per atomic sphere, for both QM (HF/6-31+G*) and MM (AMBER99) solutes, at the C-PCM[SWIG] level. Even for the sparsest grid that we tested (26 grid points per atomic sphere), $\Delta_{\text{rot}} < 0.1 \text{ kcal/mol}$. The error is solute-dependent, of course, but combined with
small rotational errors observed by YK for a different set of solutes,\textsuperscript{168} these results indicate that errors in rotational invariance are essentially negligible when Lebedev grids are used.

We also calculate the error in Gauss’ Law for this same data set. Restricting our attention to the amber case, for which $\rho_0$ is contained entirely within the cavity, we can write Gauss’ Law as

$$\int \sigma_{pol}(\vec{s}) d^2\vec{s} = -\left(\frac{\varepsilon - 1}{\varepsilon}\right) \int \rho_0(\vec{r}) d^3\vec{r}. \quad (6.57)$$

The deviation from this exact equality can be quantified as

$$\Delta_{GL} = \frac{1}{20} \sum_{\text{amino acids}} \left\{ \frac{1}{10} \sum_{i=1}^{10} \left[ Q_{A,i}^{\text{surf}} + \left(\frac{\varepsilon - 1}{\varepsilon}\right) Q_A \right]^2 \right\}^{1/2}, \quad (6.58)$$

where $Q_{A,i}^{\text{surf}}$ is the total surface charge on amino acid $A$ in its $i$th orientation, and $Q_A$ net charge on this amino acid. The dependence of $\Delta_{GL}$ on the number of grid points is shown in Figure 6.4, and as with $\Delta_{\text{rot}}$, we find that $\Delta_{GL}$ is reasonably small, even for sparse grids.

To investigate the convergence of the total energy, $W$, with respect to the Lebedev grid density, we used this same set of amino acids (but only one orientation for each), and examined the energy differences $W_N - W_{1202}$, where the subscript indicates the number of grid points per atomic sphere. These differences are plotted, as a function of $N$, in Figure 6.5, for both amber99 and HF/6-31+G* solutes. (In the latter case, the discretization errors that we observe using are comparable, as a function of $N$, to those obtained using an isodensity cavity discretized with a single-center Lebedev grid.\textsuperscript{8}) The energy discretization error is significant for the smallest grids
Figure 6.5: Plots of the C-PCM[SWIG] discretization error, $W_N - W_{1202}$, (solid lines) and the RMSE of the gradient (broken lines), as a function of the number of Lebedev grid points per atomic sphere. Data points represent averages over the set of 20 amino acids in water ($\varepsilon = 78.39$), with error bars representing one standard deviation on either side of the mean.
but is reduced to less than 0.5 kcal/mol, on average, for \( N = 110 \) (AMBER99) or \( N = 194 \) (HF/6-31+G*). That the HF calculations are more sensitive to \( N \) likely reflects the more complicated topography of the electrostatic potential, relative to that generated by a collection of point charges.

Figure 6.5 also plots the root mean square error (RMSE) in the energy gradient, as compared to the result obtained using 1202 Lebedev points per atomic sphere. When \( N \) is small, the RMSE is comparable in magnitude to the gradient itself (\( \sim 10^{-2} \) a.u. for the unoptimized amino acid structures). However, for \( N = 110 \) (AMBER99) and \( N = 194 \) (HF/6-31+G*), the RMSE in the gradient drops below \( 10^{-4} \) a.u.

### 6.6 Sample applications

In this section we use SWIG in several applications that demand smooth potential energy surfaces, and compare the results to those obtained using FIXPVA and sub-SWIG discretization. All calculations reported in the section assume that the solute and solvent remain in equilibrium at all times; see Ref. 159 for a discussion of nonequilibrium solvation. Solute cavities in these applications are constructed as described in Section 6.5.

#### 6.6.1 Geometry optimization and vibrational frequency analysis

In our initial report of the SWIG method,\(^{186}\) we demonstrated that certain discretization schemes give rise to spurious harmonic frequencies, when computed by finite difference of analytic energy gradients, as is often the required when the solute is
described using a correlated wavefunction. On the other hand, the “solute” in the example of Ref. 186 consisted of adenine plus 52 explicit MM water molecules, and is arguably not representative of most applications of PCMs. Here, we consider some more pedestrian examples.

**Comparison to FIXPVA**

As a first application, we compute harmonic vibrational frequencies for the dipeptide arginine–aspartate (Arg-Asp), in aqueous solution at the B3LYP/6-31G* level, using C-PCM in conjunction with either SWIG or FIXPVA discretization. The initial structure of the dipeptide, in its zwitterionic form, was generated using the tinker program, and subsequently optimized in the gas phase. The gas-phase structure was then optimized with C-PCM[SWIG] and, separately, with C-PCM[FIXPVA]. (Atomic spheres were discretized using 110 Lebedev grid points.) Unfortunately, these two solution-phase optimizations converged to quite different structures, making for an unfair comparison. Thus, the FIXPVA optimization was restarted from the SWIG optimized structure, and ultimately similar optimized geometries were obtained from both methods. Harmonic frequencies were then computed by finite difference of analytic energy gradients, with a step size of $10^{-3}$ Å. The resulting spectra are shown in Fig. 6.6.

Because the SWIG and FIXPVA discretizations afford somewhat different potential energy surfaces, one should not expect these methods to yield identical vibrational spectra. Nevertheless, it is worrisome that there are a few peaks present in the FIXPVA spectrum, largely involving N–H and C–H stretching modes, that have
Figure 6.6: Harmonic vibrational spectrum of Arg-Asp in water ($\varepsilon = 78.39$), computed by finite difference of analytic energy gradients at the B3LYP/6-31G*/C-PCM level, with either (a) FIXPVA or (b) SWIG discretization. Stick spectra were convolved with 20 cm$^{-1}$ Gaussians. Arrows indicate peaks in the FIXPVA spectrum that have no obvious analogues in the SWIG spectrum at nearby frequencies.
no obvious analogues in the SWIG spectrum. Although these peaks do appear in an appropriate spectral region for such modes (unlike some of the spurious peaks in the adenine–water example from Ref. 186), upon closer examination we find that the vibrations in question are associated with finite-difference steps that rapidly turn on or off certain surface elements, causing a sharp change in the FIXPVA energy and gradient and a large perturbation in the harmonic frequency. On the other hand, the SWIG spectrum appears to be free of peaks associated with dramatic changes in the cavity surface.

The FIXPVA potential energy surface is rigorously smooth, mathematically speaking, so discontinuities cannot explain spurious peaks in the vibrational spectrum. Instead, the rapid fluctuations in the gradient that give rise to these peaks result from the fact that the FIXPVA switching function attenuates the surface elements much more rapidly than the switching functions used in the SWIG and ISWIG methods. Rapid switching is necessary within the FIXPVA approach, in order to avoid singularities arising from close approach of surface point charges. However, this rapid scaling of the surface element areas can lead to “holes” in the cavity surface, in regions of high surface grid density. 186 These holes lead to a poor representation of $\sigma_{pol}(\vec{s})$ and an underestimate of the cavity surface area. 5,186 Although the FIXPVA switching function prevents $r_{ij}$ from every being exactly zero (by attenuating the surface elements in question as $r_{ij} \to 0$), the delicate balance between the switching function and the singular Coulomb potential is sufficient to cause unwanted oscillations in energy and gradient, even for the relatively small perturbations used in finite-difference
calculations. (Such oscillations can also lead to the appearance of spurious maxima and/or minima in the potential energy surface.\textsuperscript{186})

Although stable geometry optimizations in large molecules have been reported using FIXPVA,\textsuperscript{5} these calculations were carried out using a GEPOL cavity surface, and are not directly comparable to what is reported here. It is possible that Lebedev grids exacerbate the difficulties with FIXPVA, because the Lebedev grid points are distributed less uniformly than are the GEPOL tesserae, and may therefore be closer together in some cases. In the present work, we took parameters for the FIXPVA switching function from Ref. 5, where they were determined for use with GEPOL. Re-parameterization for Lebedev grids might mitigate some of the problems that we observe using FIXPVA, but will not eliminate the delicate balance between the switching function and the singular Coulomb potential that requires the use of a rapidly-varying switching function. We avoid this requirement here, by means of Gaussian surface charges.

**Comparison to CSC/subSWIG**

Strong Coulomb interactions are not the only possible cause of spurious lines in vibrational spectra. To illustrate, we compute the harmonic vibrational spectrum of a glycerol molecule in liquid glycerol solution ($\varepsilon = 42.7$) at the HF/6-31G*/SS(V)PE level using both the SWIG and subSWIG discretizations. (Recall that these methods are the same, except that subSWIG uses the CSC sum rule to define the $D_{ii}$ matrix elements.) Details of the calculations were the same as for the Arg-Asp calculations described above, except that in this case, both discretization schemes afford essentially
identical solution-phase geometries, and this geometry is very close to that obtained in the gas phase. Vibrational spectra are depicted in Fig. 6.7.

Although the gas-phase spectrum and SS(V)PE[SWIG] spectrum are similar (as one might expect, given that the PCM alters the geometry very little), the SS(V)PE-[subSWIG] spectrum exhibits peaks that are clearly absent in these other spectra. Two of these features, which are associated with C–H stretching modes, have frequencies above 4700 cm$^{-1}$, and are clearly artifacts. These artifacts arise as a consequence of violations in the variational principle that result when $D_{ii}$ is defined using a sum rule. It is worth noting that not all of the peaks are affected by this problem; the O–H stretching peak around 4000 cm$^{-1}$, for example, is quite similar in all three spectra shown in Fig. 6.7.

6.6.2 Molecular dynamics

A molecular dynamics (MD) calculation is an especially stringent test of PCM cavity discretization, because such a calculation may explore a broad swath of the solute’s potential energy surface, and because conservation of energy places demands on how rapidly the energy may change as a function of solute geometry. Here, we report MD simulations in a PCM, using both QM and MM descriptions of the solute.

Molecular mechanics

We propagated molecular dynamics for a single-stranded DNA oligomer, d(GACT), using the AMBER99 force field to describe the DNA and the C-PCM model to describe
Figure 6.7: Harmonic vibrational spectra of glycerol computed at the HF/6-31G* level in (a) the gas phase; (b) liquid glycerol ($\varepsilon = 42.7$), described at the SS(V)PE[SWIG] level; and (c) liquid glycerol, described at the SS(V)PE[subSWIG] level. Stick spectra were convolved with 20 cm$^{-1}$ Gaussians. Arrows indicate spurious peaks in the subSWIG spectrum.
Figure 6.8: Fluctuations in the energy during an MD simulation of single-stranded d(GACT) in water, described at the AMBER99/C-PCM level. The inset shows a close-up view of the energy fluctuations obtained in the gas phase and with C-PCM[SWIG]. The time step is 1.0 fs.

The d(GACT) molecule is fairly flexible, and undergoes considerable geometric rearrangement over 10 ps, substantially altering the cavity shape. The FIXPVA approach is unable to cope with these changes, and exhibits a catastrophic failure to conserve the aqueous solvent ($\epsilon = 78.39$). A time step of 1.0 fs was used, with initial velocities sampled from a Boltzmann distribution at $T = 300$ K.

Figure 6.8 compares the energy fluctuations observed using SWIG and FIXPVA discretization, as well as those from a gas-phase MD simulation of the same solute.
energy, with two jumps of \( \gtrsim 10 \text{ kcal/mol} \) within the first 3 ps of simulation. Eventually, one of these jumps simply crashes the simulation, due to an overflow error in the atomic velocities. As in the case of FIXPVA geometry optimizations, these abrupt changes in energy originate in the too-close approach of surface charges as grid points vanish into, or emerge from, the interior of the cavity. The SWIG discretization, on the other hand, exhibits energy fluctuations of only \( \sim 0.2 \text{ kcal/mol} \) over the entire 10 ps of the simulation, which is comparable to what is observed in the gas phase.
Ab initio MD

As a second MD example, we performed an ab initio MD simulation on glycine in water, with the solute described at the PBE0/6-31+G* level, using the SS(V)PE model. We compare the ISWIG discretization approach to “subISWIG”, which substitutes the sum rule in Eq. (6.51) as a definition of $D_{ii}$. The simulations begin at the gas-phase optimized geometry and use initial velocities corresponding to the zero-point energy of the gas phase vibrational modes, with a time step of 0.97 fs. Energy fluctuations are shown in Fig. 6.9. Within the first 15 fs, the subSWIG simulation encounters a singularity, leading to a jump in the energy of nearly 100 kcal/mol. After several more large jumps, the simulation eventually reaches a geometry for which the SCF procedure fails to converge. Three separate attempts with subSWIG were made, using slightly different initial conditions, but each suffered the same fate. In contrast, the ISWIG discretization conserves energy just as well as the gas phase simulation, and is stable through at least 5 ps of simulation.

In aqueous solution, the lowest-energy tautomer of glycine is the zwitterion $\text{+NH}_3\text{CH}_2\text{CO}_2^-$, whereas the neutral form $[\text{NH}_2\text{CH}_2\text{C(O)OH}]$ is $\sim 11$ kcal/mol higher in energy, but in the gas phase the neutral form is more stable by $\sim 18$ kcal/mol. Our SS(V)PE[ISWIG] simulations bear this out; starting from the gas-phase geometry, in which the carboxylic acid moiety is protonated, the molecule eventually undergoes an intramolecular proton transfer to form the zwitterion, a process that is not observed in the gas-phase MD simulation. Figure 6.10 shows that the solvation energy changes smoothly as this proton transfer occurs.
Figure 6.10: Solvation energy, $E_{pol}$, of glycine in water, during an MD simulation performed at the PBE0/6-31+G*/SS(V)PE[ISWIG] level. The regions around $-15$ kcal/mol represent the carboxylic acid tautomer, whereas lower-energy regions (around $-50$ kcal/mol) represent the zwitterion. The inset shows a longer amount of simulation time, and indicates that proton transfer occurs multiple times. The time step in this simulation is 0.97 fs.
6.7 Summary

This work introduces the Switching/Gaussian (SWIG) discretization method for apparent surface charge PCMs, which generalizes a method originally introduced by York and Karplus\textsuperscript{168} in the context of the conductor-like screening model. The approach developed here is applicable not only to C-PCM/GCOSMO calculations, but also to more sophisticated PCMs such as IEF-PCM/SS(V)PE. Both QM and MM solutes are possible.

Comparison to other “smooth discretization” approaches reveals that while these alternative methods may afford potential energy surfaces that are rigorously smooth, in a mathematical sense, they suffer from Coulomb singularities when point charges are used to represent the surface charge. This problem can be eliminated by using spherical Gaussian functions to represent the surface charge, but in this case, the approximate nature of certain sum rules that are traditionally used in these models can lead to non-variational singularities in the solvation energy.

The SWIG approach eliminates both of these problems, by using Gaussian surface charges but avoiding the use of sum rules to define the matrix elements of the PCM equations. The result is a method that affords smooth potential energy surfaces and is largely free of unwanted oscillations in the energy gradient, so that molecular dynamics simulations are stable and robust, and harmonic vibrational frequencies can safely be calculated by finite difference of analytic energy gradients. SWIG discretization also appears to be faithful to the energetics of the underlying integral-equation PCM. As such, there seems to be no reason not to use this discretization method for
all PCM calculations.

6.8 Integral equation formulation

Here, we show how different PCMs are represented in the context of the reaction-field equations presented in Section 6.2. The relevant integral operators are described in detail elsewhere.\textsuperscript{8,167} Briefly, the operator $\hat{S}$ acts on an arbitrary surface charge density, $\sigma(\vec{s})$, to generate the corresponding electrostatic potential on the surface:

$$\hat{S}\sigma(\vec{s}) = \int \frac{\sigma(\vec{s}')}{|\vec{s} - \vec{s}'|} d^2\vec{s}' = \phi(\vec{s}) . \quad (6.59)$$

The operator $\hat{D}^\dagger$ acts on $\sigma(\vec{s})$ to produce the negative of the normal component of the electric field,

$$\hat{D}^\dagger \sigma(\vec{s}) = \int \sigma(\vec{s}') \frac{\vec{n}_s \cdot (\vec{s} - \vec{s}')}{|\vec{s} - \vec{s}'|^3} d^2\vec{s}' = -\vec{F}(\vec{s}) \cdot \vec{n}_s . \quad (6.60)$$

(Here, $\vec{n}_s$ is the outward-pointing unit vector normal to the cavity surface, at the point $\vec{s}$.) We also define $\hat{D}$, the adjoint of $\hat{D}^\dagger$, and note that

$$\hat{D}\hat{S} = \hat{S}\hat{D}^\dagger . \quad (6.61)$$

(This equality, however, is generally not preserved when these operators are discretized to yield finite-dimensional matrices.\textsuperscript{8,203})

In Section 6.2.1 we introduced $w_{\text{displ}}$, the work required to displace the bound charges within the dielectric. Our goal here is to obtain an expression for this quantity, in terms of the electrostatic interactions at the cavity surface. Various PCMs afford
Continuum Model $w_{\text{displ}}$

<table>
<thead>
<tr>
<th>Model</th>
<th>$w_{\text{displ}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conductor</td>
<td>0</td>
</tr>
<tr>
<td>C-PCM/GCOSMO</td>
<td>$\frac{1}{2} \left( \frac{1}{\varepsilon - 1} \right) \int \sigma_{\text{pol}}(\vec{s}) \hat{S} \sigma_{\text{pol}}(\vec{s}) d^2 \vec{s}$</td>
</tr>
<tr>
<td>IEF-PCM/SS(V)PE</td>
<td>$\left( \frac{1}{\varepsilon - 1} \right) \int \sigma_{\text{pol}}(\vec{s}) \left[ \hat{I} - \frac{1}{2\pi} \hat{D} \right]^{-1} \hat{S} \sigma_{\text{pol}}(\vec{s}) d^2 \vec{s}$</td>
</tr>
</tbody>
</table>

Table 6.6: Definitions of $w_{\text{displ}}$ for various continuum models. The quantity $\varepsilon$ represents the dielectric constant of the medium.

Different expressions for $w_{\text{displ}}$, and although the original derivations of these models did not explicitly invoke the charge-displacement work, we will show that all of them can be recast into the framework of Eq. (6.10), whence an expression for $w_{\text{displ}}$ is obtained.

The simplest possible continuum model is obtained by assuming that the medium is a conductor. In this case, $w_{\text{displ}} = 0$ because charges in the continuum are unbound. Expressions for $w_{\text{displ}}$ in finite-dielectric models are not as obvious but can be obtained after some algebra. The resulting expressions for $w_{\text{displ}}$ in C-PCM/GCOSMO and IEF-PCM/SS(V)PE are given in Table 6.6. Note that in each case, $w_{\text{displ}} \to 0$ as $\varepsilon \to \infty$.

As a check of these expressions for $w_{\text{displ}}$, we insert them into the variational condition, Eq. (6.13). Upon rearranging the result, one obtains the condition

$$\hat{S} \sigma_{\text{pol}}(\vec{s}) = - \left( \frac{\varepsilon - 1}{\varepsilon} \right) \phi_0(\vec{s})$$

(6.62)
in the case of C-PCM/GCOSMO, whereas for IEF-PCM/SS(V)PE, the variational condition can be written

\[
\hat{I} - \left( \frac{\varepsilon - 1}{\varepsilon + 1} \right) \frac{1}{2\pi} \hat{D} \hat{\sigma}_{\text{pol}}(\vec{s}) = - \left( \frac{\varepsilon - 1}{\varepsilon + 1} \right) \left[ \hat{I} - \frac{1}{2\pi} \hat{D} \right] \phi_0(\vec{s}) .
\] (6.63)

These two equations are precisely the standard PCM working equations for these two models, which can be expressed in the general form \[3,8,160\] [cf. Eq. (8.43)]

\[
\hat{K}_{\text{pol}}(\vec{s}) = \hat{R}_{\phi_0}(\vec{s}) .
\] (6.64)

Discretization of Eqs. (6.62) and (6.63) affords the matrix forms of \(K\) and \(R\) that appear in Table 6.1. [Note the symmetrization that is performed in the IEF-PCM/SS(V)PE case, since \(D_{\text{AS}} \neq S_{\text{AD}}^\dagger\).]

Defining \(\hat{Q} = \hat{K}^{-1} \hat{R}\), it follows from Eqs. (6.64) and (6.15) that

\[
-\hat{Q}^{-1} \sigma_{\text{pol}}(\vec{s}) = \hat{S}_{\sigma_{\text{pol}}}(\vec{s}) + \frac{\delta W_{\text{displ}}}{\delta \sigma_{\text{pol}}(\vec{s})} .
\] (6.65)

Invoking the condition that \(W\) should be minimized with respect to variation of \(\sigma_{\text{pol}}\) [Eq. (6.14)], it follows that the operator \(\hat{Q}\) must be negative-definite. (This same result was demonstrated, in a somewhat different fashion, in Ref. 203.)

### 6.9 Continuity

Here we show how the switching function in SWIG and similar methods ensures that the solute’s potential energy surface is continuous. Our proof generalizes the one given in Ref. 168 for the YK version of COSMO.
We begin by decomposing the $K$ matrix from Eq. (8.43) into a sum of its diagonal and off-diagonal contributions, $K = K_{\text{diag}} + K_{\text{off}}$. We can then express $K$ as

$$K = K_{\text{diag}}^{1/2} L K_{\text{diag}}^{1/2}$$

where

$$L = I + K_{\text{diag}}^{-1/2} K_{\text{off}} K_{\text{diag}}^{-1/2}$$

and $I$ is a unit matrix. Inserting Eq. (6.66) into Eq. (6.20), one may express the solvation energy as

$$E_{\text{pol}} = \frac{1}{2} y^\dagger K_{\text{diag}}^{-1/2} L^{-1} K_{\text{diag}}^{-1/2} y,$$

where $y = Rv$.

The key point is that every term in the diagonal matrix element $K_{ii}$ contains a factor of $S_{ii}$, for each of the solvation models considered here (see Table 6.1.) Thus, by placing the switching function $F_i$ in the denominator of $S_{ii}$ [Eq. (6.36)], we ensure that the $i$th diagonal element of $K^{-1/2}$ goes to zero as $F_i \to 0$. For the FIXPVA and CSC methods, $a_i \to 0$ as $\vec{s}_i$ enters the cavity, which has the same effect, according to Eq. (6.50).

It then follows from Eqs. (6.66) and (6.67) that $K^{-1}$ has a null space corresponding to those surface elements for which $F_i = 0$, and that $L$ acts as a unit matrix within this null space. As such, the dimension of $q = K^{-1} y$ can be reduced without approximation to include only those surface grid points for which $F_i > 0$. For the same reason, the $i$th grid point’s contribution to $E_{\text{pol}}$ vanishes as $F_i \to 0$. Since $F_i$ is a smooth function of the nuclear coordinates, so is $E_{\text{pol}}$. 

172
6.10 Switching function gradient

The gradient of the switching function $F_i$ with respect to the $M$th solute nucleus is

$$\hat{\nabla}_M F_i = F_i \sum_j^{\text{atoms}} \frac{\hat{\nabla}_M f(\vec{s}_i, \vec{r}_J)}{f(\vec{s}_i, \vec{r}_J)}.$$  \hspace{1cm} (6.69)

For the elementary switching function $f(\vec{s}_i, \vec{r}_J)$ that is defined in Eq. (6.47), the derivative with respect to a perturbation of the $M$th nucleus is

$$\hat{\nabla}_M f(\vec{s}_i, \vec{r}_J) = \frac{\partial h(d_{iJ})}{\partial d_{iJ}} \hat{\nabla}_M d_{iJ}.$$  \hspace{1cm} (6.70)

The derivative $\partial h(d_{iJ})/\partial d_{iJ}$ is easily derived from Eq. (6.48), while the other term in Eq. (6.70) is given by

$$\hat{\nabla}_M d_{iJ} = \left( \frac{\vec{r}_i - \vec{r}_J}{r_{iJ}} \right) \frac{\delta_{iM} - \delta_{JM}}{R_{sw,J}}.$$ \hspace{1cm} (6.71)

For the alternative elementary switching function defined in Eq. (6.49), the gradient is

$$\hat{\nabla}_M f(\vec{s}_i, \vec{r}_J) = \frac{\partial f(\vec{r}_i, \vec{r}_J)}{\partial r_{iJ}} \hat{\nabla}_M r_{iJ},$$ \hspace{1cm} (6.72)

where

$$\frac{\partial f(\vec{s}_i, \vec{r}_J)}{\partial r_{iJ}} = \frac{\zeta_i}{\sqrt{\pi}} \left\{ \exp[-\zeta_i^2(R_J - r_{iJ})^2]) + \exp[-\zeta_i^2(R_J + r_{iJ})^2]) \right\} \hspace{1cm} (6.73)$$

and

$$\hat{\nabla}_M r_{iJ} = \frac{\vec{r}_J - \vec{s}_i}{r_{iJ}}(\delta_{JM} - \delta_{iM}).$$ \hspace{1cm} (6.74)

For Lebedev discretization, the $i$th quadrature point contributes $a_i = w_i R_i^2 F_i$ to the total surface area. From Eq. (6.69), it follows that

$$\hat{\nabla}_M a_i = a_i \sum_j^{\text{atoms}} \frac{\hat{\nabla}_M f(\vec{s}_i, \vec{r}_J)}{f(\vec{s}_i, \vec{r}_J)}.$$ \hspace{1cm} (6.75)
In view of this result, the expression for the gradient of the total cavity surface area is considerably simpler than it is in either the GEPOL surface tessellation approach\textsuperscript{180} or the FIXPVA method.\textsuperscript{5} Although we do not consider non-electrostatic solute–continuum interactions in the present work, Eq. (6.75) shows that the SWIG approach leads to simple expressions for the gradients of typical non-electrostatic interaction terms,\textsuperscript{161,178} which are explicit functions of the cavity surface area.

6.11 Addendum

Subsequent to this work, a comment on the accuracy of claims pertaining to the CSC method made herein was published in Ref. 204. In summary, the authors note that the CSC method is not equivalent to the subSWIG method because the summation used in the CSC method for the elements $D_{ii}$ is restricted to only that atom to which $i$ belongs, not the full molecular surface as implied in the traditional sum rule. The summation is performed prior to switching off grid points in CSC and therefore should not be prone to the issue of breaking negative definiteness of the $Q$ matrix.

In our response,\textsuperscript{205} we recognize the restriction on the summation issue, however we did not explicitly state anywhere that subSWIG and CSC are indeed equivalent. The subSWIG method still serves as an important example of how the sum rule can cause an otherwise smooth discretization of PCM to fail. Ultimately, we are legally unable to make any direct comparisons of SWIG and the implementation of CSC in the \textsc{Gaussian09}\textsuperscript{206} software, despite a proposition of sharing information published in Ref. 204, due to clauses in the license agreement for \textsc{Gaussian09} at the Ohio
State University, which, out of commercial competition with the Q-Chem\textsuperscript{53} software, specifically state that neither John M. Herbert nor his students may use or obtain any data from the GAUSSIAN09 software.
CHAPTER 7

Symmetric versus asymmetric discretization of the integral equations in polarizable continuum solvation models

7.1 Introduction

Within the quantum chemistry community, the most widely-used class of dielectric continuum solvation models are “apparent surface charge” (ASC) methods, better known as polarizable continuum models (PCMs). Given a definition for what constitutes the boundary between the solute and the dielectric continuum (i.e., the “solute cavity”), these models afford a prescription for calculating an “apparent” charge density on the cavity surface, whose electrostatic interaction with the solute approximates the solute/continuum electrostatic interaction. Because this interaction is represented using a two-dimensional surface charge density, ASC methods are far more computationally efficient than numerical solution of Poisson’s equation, which requires a three-dimensional volume integration.

The development of PCM theory originated thirty years ago with the work of Tomasi and co-workers, and has since been revised and elaborated manyfold into

---

7.1 This chapter appeared as a full article in Chemical Physics Letters, in 2011, volume 509, pages 77–87.
its modern form; see Refs. 207 and 159 for a history of this development. Nowadays, numerous levels of approximation are available within the PCM formalism, including the conductor-like screening model 165, which is typically known within the PCM framework as C-PCM 166 or GCOSMO 181, 182, as well as the more sophisticated “integral equation formulation” (IEF-PCM) 167, 183, 198, 210. The latter is formally equivalent 184, at the level of integral equations, to the “surface and simulation of volume polarization for electrostatics” [SS(V)PE] method 160. The IEF-PCM/SS(V)PE approach affords an exact solution for the solute/continuum electrostatic interaction that arises from whatever part of the solute’s charge density exists within the cavity, and furthermore provides an approximate solution for the “volume polarization” that arises from any “escaped charge”, i.e., that part of the solute’s charge density that penetrates beyond the cavity, into the dielectric medium.

The IEF-PCM/SS(V)PE approach, which is the focus of the present work, is derived in terms of certain integral operators that act on the surface charge and surface electric potential 160, 167, 210–213. Numerical solution of the integral equations that define this model requires discretization of the cavity surface, resulting in finite-dimensional matrix equations. The matrix forms of the integral operators, however, fail to preserve certain properties of the corresponding integral operators. Consequently, the proper definition of the matrix equations is ambiguous, because the integral equations can be written in various equivalent forms that yield inequivalent matrix equations.

This issue has been noted in the past, but only very modest numerical analysis and discussion have been devoted to it. Two previous studies 3, 6 report that solvation
energies differ by no more than 0.01 kcal/mol amongst the variants of IEF-PCM/SS(V)PE that are considered here, although the data sets in these studies consisted of no more than four small molecules. (Lipparini et al.\textsuperscript{203} also assert that different forms of the matrix equations yield solvation energies that differ by only $\sim$1 kcal/mol, but their numerical tests are unpublished.) Other properties, including solute dipole moments, molecular orbital eigenvalues, and total surface charge are also reported to differ negligibly amongst these variants\textsuperscript{6,214}. In light of these results, the importance of the particular form of the matrix equations has been largely ignored in the literature. As a result, no consensus actually exists in regard to which version of the matrix equations is best.

In contrast to previous studies, we present a thorough examination of various matrix formulations of IEF-PCM/SS(V)PE, from both formal and numerical points of view. We find cases in which these variants afford substantial differences in electrostatic solvation energies and total surface charge. This work also constitutes the first assessment of these differences for recently-developed “smooth PCMs”\textsuperscript{5,213}, in which switching functions are used to guarantee that the solute’s potential energy surface is smooth.
7.2 Theory

7.2.1 ASC PCM equations

The working equations for a variety of different ASC PCMs can be written in the general form\textsuperscript{3,8,160,213}

\[
\hat{K}\sigma(\vec{s}) = \hat{R}\phi_0(\vec{s}),
\]  

(7.1)

where \(\sigma(\vec{s})\) is the ASC density and \(\phi_0(\vec{s})\) is the solute’s electrostatic potential at the cavity surface. Different PCMs correspond to different choices for the integral operators \(\hat{K}\) and \(\hat{R}\). For the IEF-PCM/SS(V)PE model, which is the focus of the present work,

\[
\hat{K} = \left(\hat{I} - \frac{f_\varepsilon}{2\pi}\hat{D}\right)\hat{S},
\]  

(7.2)

and

\[
\hat{R} = -f_\varepsilon\hat{\mathcal{Y}},
\]  

(7.3)

where

\[
f_\varepsilon = \frac{\varepsilon - 1}{\varepsilon + 1},
\]  

(7.4)

and

\[
\hat{\mathcal{Y}} = \hat{I} - \frac{1}{2\pi}\hat{D}.
\]  

(7.5)

In these equations, \(\varepsilon\) denotes the dielectric constant that characterizes the continuum and \(\hat{I}\) is the identity operator. Definitions of the surface-Coulomb operator, \(\hat{S}\), and surface-dipole operator, \(\hat{D}\), can be found in the literature\textsuperscript{160,167,183,198,210,213}. At present, it suffices to note that

\[
\hat{D}\hat{\mathcal{S}} = \hat{\mathcal{S}}\hat{D}^\dagger.
\]  

(7.6)
Violation of this exact condition, within standard discretization schemes, leads to the ambiguity that is addressed here. This issue does not arise for the conductor-like models (COSMO, GCOSMO, and C-PCM), for which the operator $\hat{D}$ is absent in the definitions of $\hat{K}$ and $\hat{R}^{160,213}$.

Upon discretization, the cavity surface is replaced by a finite set of surface grid points, $\{\vec{s}_i\}$, with surface areas $\{a_i\}$. Equation (7.1) is thereby transformed into a finite-dimensional matrix equation,

$$Kq = Rv \ .$$ (7.7)

The electrostatic potential vector, $v$, has elements $v_i = \phi_0(\vec{s}_i)$, and the vector $q$ consists of point charges $q_i$. If we define a diagonal matrix $A$, with $A_{ij} = a_i \delta_{ij}$, then the discretized forms of $\hat{K}$ and $\hat{R}$ are

$$K = \left( I - \frac{f_\varepsilon}{2\pi DA} \right) S$$ (7.8)

and

$$R = -f_\varepsilon Y$$ (7.9)

$$= -f_\varepsilon \left( I - \frac{1}{2\pi DA} \right) .$$

Appropriate definitions for the matrix elements $D_{ij}$ and $S_{ij}$ has been discussed extensively in the literature$^{3,4,159,183,186,198,213}$. (In particular, several choices have been suggested for the diagonal elements, the definition of which is not entirely straightforward owing to singularities in certain integrands, when evaluated over a single surface element.) We use the definitions suggested in our previous work$^{213}$, which
are appropriate for use with the smooth discretization schemes introduced below. We will revisit the topic of diagonal matrix elements in Section 7.2.2.

In general, discretization preserves the symmetry of \( \hat{S} \), i.e., \( S^\dagger = S \), and if the solute cavity consists of a single sphere then \( D^\dagger = D \) as well\(^{160,214} \). However, realistic solute cavities are non-spherical, and for non-spherical cavities discretization fails to preserve Eq. (7.6). In other words,

\[
DAS \neq SAD^\dagger
\]

(7.10)

for realistic cavity shapes. As a consequence of this inequality, the step in going from Eq. (7.2) to Eq. (7.8) is not entirely justified, because we could just as well have written \( \hat{K} \) in a form that involves \( \hat{S} \hat{D}^\dagger \) rather than \( \hat{D} \hat{S} \). Had we done so, then the factor of \( DAS \) in Eq. (7.8) would be replaced by \( SAD^\dagger \) upon discretization. In fact, one could use Eq. (7.6) to justify replacing \( DAS \) in Eq. (7.8) with any linear combination \( c_1 DAS + c_2 SAD^\dagger \) such that \( c_1 + c_2 = 1 \).

In view of this ambiguity, let us replace the definition of \( K \) in Eq. (7.8) with the more general form

\[
K = S - \frac{f_c}{2\pi} X.
\]

(7.11)

Equation (7.8) is recovered if \( X = DAS \). Taking \( X = SAD^\dagger \) is also justified, as is the symmetric choice \( X = (DAS + SAD^\dagger)/2 \). Other linear combinations are possible but are not considered here. Chipman uses the symmetric form consistently\(^8,160,215 \), in a method that he calls SS(V)PE; this form facilitates somewhat more efficient solution of Eq. (8.43). On the other hand, IEF-PCM calculations using all three of
the aforementioned variants have been reported\textsuperscript{6,203,214}, along with other asymmetric variants that are not considered here\textsuperscript{167,198}.

Whatever the definition of $K$, ASC PCM calculations consist of solving Eq. (8.43) for the surface charge vector, $q$, given the solute’s electrostatic potential. The surface charges can then be used to evaluate the polarization energy\textsuperscript{213},

$$E_{pol} = \frac{1}{2} q^\dagger v = \frac{1}{2} v^\dagger Qv , \quad (7.12)$$

which is the electrostatic contribution to the solvation energy. In Eq. (7.12), we have introduced the so-called solvent response matrix,

$$Q = K^{-1} R . \quad (7.13)$$

The total energy is

$$W = E_0 + E_{pol} , \quad (7.14)$$

where $E_0$ represents the internal energy of the solute, computed in the field of the ASC. When the solute is described using quantum mechanics, the solute’s wavefunction must be converged in the presence of the surface charges, so the system of linear equations in Eq. (8.43) must be solved at each self-consistent field cycle.

### 7.2.2 Smooth discretization

We have recently developed a general discretization procedure (applicable to C-PCM, IEF-PCM, and related PCMs) that we call the “switching/Gaussian” (SWIG) procedure\textsuperscript{186,213}. This approach rectifies a long-standing problem with ASC PCMs, namely, that straightforward pointwise discretization or tessellation (as in the widely-used
GEPOL algorithm\textsuperscript{216–218} often leads to discontinuities in the solute’s potential energy surface. These may arise due to tessellation algorithms that fail to treat nuclear perturbations in a symmetric fashion\textsuperscript{170,219}, or else due to the appearance or disappearance of grid points as the solute geometry changes\textsuperscript{5,168,175,186,213,219}, which is equivalent to a discontinuous change in cavity surface area and in the dimension of the system of equations in Eq. (8.43). Discontinuities of the latter type can hinder the convergence of geometry optimizations (or prevent them from converging at all), and can lead to other artifacts including spurious vibrational frequencies, non-variational solvation energies, and catastrophic failure to conserve energy in molecular dynamics simulations\textsuperscript{186,213}.

A variety of “smooth” discretization methods have been proposed in recent years, in attempt to alleviate these problems\textsuperscript{4,5,168–170,186,213,219}. We have documented numerical problems with several of these smooth discretization procedures, and shown that these problems are absent in the SWIG approach and a related “ISWIG” procedure that is described below\textsuperscript{186,213}. A detailed description of these PCM implementations, including formulas for the matrix elements $D_{ij}$ and $S_{ij}$, can be found in Ref.\textsuperscript{213}. The essential aspects are summarized here.

The SWIG discretization algorithm has two key features: a switching function to attenuate the point $\vec{s}_i$’s contribution to Eq. (8.43), as $\vec{s}_i$ passes through a buffer region surrounding the cavity surface; and Gaussian blurring of the surface charges $q_i$, in order to eliminate numerical problems associated with the singular Coulomb potential. Such problems are exacerbated when a switching function is introduced to
ensure continuity, because this function allows surface grid points to approach one another more closely than would be possible if all interior grid points were simply discarded. The close approach of point charges can lead to unwanted fluctuations in the solvation energy or its gradient, as a function of the nuclear coordinates\textsuperscript{186,213}.

This sort of numerical instability is more often associated with the $D$ matrix rather than with the $S$ matrix, because $D_{ij} \propto r_{ij}^{-3}$ whereas $S_{ij} \propto r_{ij}^{-1}$, where $r_{ij} = |\vec{s}_i - \vec{s}_j|\textsuperscript{213}$. Thus, $D_{ij}$ fluctuates more rapidly than $S_{ij}$ when $r_{ij}$ is small. As such, IEF-PCM and related methods that utilize $D$ may be more susceptible to numerical instabilities than is C-PCM, where $D$ is absent. This fact was previously noted in Ref.\textsuperscript{167}, where it was suggested that a modification of the GEPOL tessellation algorithm is required in the context of IEF-PCM. Alternatively, we will show that instabilities can be avoided using Gaussian blurring of the point charges, as this technique modifies the $D$ matrix elements such that $D_{ij} \rightarrow 0$ as $r_{ij} \rightarrow 0\textsuperscript{213}$.

To avoid numerical instabilities and unwanted energy fluctuations, the SWIG algorithm uses the charges $q_i$ as amplitudes for spherical Gaussian functions\textsuperscript{168,213},

$$g_i(\vec{r}) = q_i(\zeta_i^2/\pi)^{3/2} \exp\left(-\zeta_i^2|\vec{r} - \vec{s}_i|^2\right). \quad (7.15)$$

The functions $g_i$ are centered at the surface grid points, $\vec{s}_i$, and we use atom-centered Lebedev grids\textsuperscript{191} to generate these points. Lebedev grids are already widely used in quantum chemistry codes, and this form of discretization does not increase the complexity of analytic gradient expressions, in the way that surface tessellation schemes sometimes do\textsuperscript{180}. The Gaussian exponents, $\zeta_i$, are taken from Ref.\textsuperscript{197}, where they were optimized to reproduce solvation energies for the Born ion model. The use of
surface Gaussians instead of point charges ensures that the representation of $\sigma(s')$ is continuous even though a discrete set of grid points is used.

The switching function is the second key feature of SWIG. In Ref. $^{213}$, we presented two different switching functions, the first of which was originally proposed by York and Karplus $^{168}$, who developed a smooth version of COSMO. We use the term “SWIG” to refer to this particular switching function, used in conjunction with Gaussian blurring. An alternative switching function was introduced in Ref. $^{213}$, which uses both the radii of the atomic spheres and the Gaussian exponents to determine how rapidly a surface point is attenuated. Consistent with the terminology introduced in Ref. $^{213}$, we call this the “improved SWIG” (ISWIG) method. Both SWIG and ISWIG discretization are examined here.

We also examine several other discretization schemes that will aid in understanding the origin of certain discrepancies amongst the $K$-matrix variants. A simple point-charge (PC) discretization is obtained by placing point charges $q_i$ at the Lebedev grid points and discarding any grid points that reside within the cavity. Alternatively, we might modify these point charges using Eq. (7.15), and we refer to this discretization procedure as “Gaussian blurring only” (GBO). Neither the PC nor the GBO procedure employs a switching function, hence both approaches are apt to produce discontinuities in the potential surface. As such, these procedures are not recommended for chemical applications, but they do shed light on the role of the switching function and the Gaussian blurring procedure.

In addition, we carry out PC and GBO computations with GEPOL grids $^{216-218}$ to
investigate if any unforeseen anomalies arise from the use of Lebedev grids. We will refer to these calculations as “PC-GEPOL” and “GBO-GEPOL”, respectively. While the GEPOL algorithm has the ability to add extra spheres to approximate solvent-excluded surfaces, we only consider GEPOL grids without these extra spheres, so that these calculations should be directly comparable to the PC and GBO calculations using Lebedev grids.

Another discretization scheme that we shall examine is the fixed-point, variable area (FXPVA) algorithm\textsuperscript{5}, which uses a switching function in conjunction with a point-charge representation of $\sigma(\vec{s})$. Unlike the SWIG and ISWIG approach to attenuating grid points, the FXPVA procedure does not allow grid points to penetrate into the interior of the cavity. Instead, FXPVA uses the switching function to scale the individual surface areas $a_i$. An unfortunate side effect of this approach is that it sometimes completely scales away important surface elements, thus significantly underestimating the cavity surface area\textsuperscript{5,179,186} and—even worse—leaving holes in the cavity surface\textsuperscript{186}. This matter is analyzed in more detail herein.

Finally, we investigate the effect of the diagonal elements of the $D$ matrix. We note that two definitions for $D_{ii}$ have been commonly reported and used in the past\textsuperscript{159,219}. One of the common definitions was originally derived by Purisima and Nilar\textsuperscript{173}, and amounts to a sum rule for each row of $D$:

$$D_{ii} = -\frac{1}{a_i} \left( 2\pi + \sum_{j \neq i} D_{ij} a_j \right). \quad (7.16)$$

The other common definition is due to Tomasi \textit{et al.}\textsuperscript{198}, and is derived for a surface
element residing on a sphere of radius $R_I$:

$$D_{ii} = -\frac{S_{ii}}{2R_I}.$$  \hspace{1cm} (7.17)

Here, $S_{ii}$ is a diagonal element of the $S$ matrix, and $i \in I$. Previous reports investigating symmetric versus asymmetric $K$ matrices\textsuperscript{3,6,214} have only used Eq. (7.16). Here, we investigate both definitions.

In our previous work\textsuperscript{213}, we discussed how the $D_{ii}$ elements affect the negative definiteness of the response matrix, $Q$. We showed that Eq. (7.16) often leads to violations of the negative-definiteness requirement for $Q$ when SWIG or ISWIG discretization is employed, because the switching function allows some discretization points to reside inside of the cavity, as they are being attenuated. We then appealed to Eq. (7.17) to define $D_{ii}$, after removing the switching function from $S_{ii}$ in Eq. (7.17). This definition preserves the negative-definiteness of $Q$\textsuperscript{213}.

In regard to the current work, we will show that Eq. (7.16) and Eq. (7.17) also affect the extent to which solvation energies for $X = \text{DAS}$ differ from those predicted when $X = \text{SAD}^\dagger$. For consistency with our previous work\textsuperscript{213}, the SWIG/ISWIG notation will imply the use of Eq. (7.17), and alternatively, we will refer to the substitution of Eq. (7.16), in place of Eq. (7.17), as “subSWIG/subISWIG” discretization. The PC, GBO, and FIXPVA calculations use Eq. (7.17), except where it is explicitly stated that the sum rule in Eq. (7.16) is used instead.

187
7.3 Computational details

Previous assessments of symmetric versus asymmetric $K$ matrices have been limited to data sets consisting of $\text{H}_2\text{O}$, $\text{C}_2\text{H}_5\text{OH}$, $(\text{CH}_3)_2\text{NH}$, $\text{CH}_3\text{C(O)}\text{NH}_2$, $\text{NO}^+$, and $\text{CN}^-$\textsuperscript{3,6,214}. Here, we take the twenty amino acids that constitute the universal genetic code as a representative test set of molecules, since the side chains in these molecules exhibit a variety of chemical properties (polar versus non-polar, charged versus neutral, aliphatic versus aromatic, etc.). The geometries of the amino acids were generated, in their zwitterionic forms, using the TINKER program (version 4.2), and were not further optimized. IEF-PCM calculations were carried out in water, $\varepsilon = 78.39$.

The solute molecules in our calculations are described using either the Hartree-Fock/6-31+G* method, or else using the AMBER99 force field\textsuperscript{56}. In the former case, the solute cavity is constructed from a union of atom-centered spheres, using van der Waals radii taken from Bondi\textsuperscript{188} (except for hydrogen\textsuperscript{199}). These radii are scaled by a factor of 1.2 for use in cavity construction\textsuperscript{207}. For AMBER99 solutes, we construct a solvent-accessible surface\textsuperscript{220} by adding a solvent probe radius of 1.4 Å to the unscaled Lennard-Jones radius of each atom. All atomic radii and nuclear coordinates are provided in the Supplementary Data.\textsuperscript{221}

Calculations were performed using a locally-modified version of Q-Chem\textsuperscript{53}, in which we have recently implemented hybrid quantum mechanics/molecular mechanics methods\textsuperscript{222} and also PCM methods\textsuperscript{186,213}. In addition to SWIG and ISWIG
discretization, we have implemented the FIXPVA discretization algorithm as described in Ref. 5, except that our version uses Lebedev grids to discretize the atomic spheres, rather than the GEPOL algorithm. (We also use somewhat different atomic radii than those used in Ref. 5.) We present a comparison of our implementation of FIXPVA to the implementation in the most recent release of GAMESS 223,224 (version 1 Oct., 2010) for C-PCM calculations in Supplementary Data (Appendix C).

For comparisons of Lebedev grids versus GEPOL grids in IEF-PCM calculations, we generated GEPOL grids with GAMESS 223,224 and read them into Q-Chem. In order to use GBO on the GEPOL grids, we needed to determine an optimal value of the parameter \( \zeta \) that controls the width of the Gaussian surface charges, via

\[
\zeta_i = \frac{\zeta}{\sqrt{w_i R_i}} .
\]

(7.18)

Here, \( w_i \) is the quadrature weight of the \( i \)th grid point and \( R_i \) is the radius of the sphere on which the \( i \)th point resides. Scalmani and Frisch 4 assert that optimal parameters for GEPOL grids in the range of \( 4.5 \leq \zeta \leq 4.8 \), but they do not explicitly report their values. By numerical optimization, we arrived at the values of \( \zeta = 4.7477177485 \) and \( \zeta = 4.7396506415 \) for GEPOL grids of 60 and 240 points per atom, respectively. In C-PCM calculations, these values reproduce Born ion energies within \( \sim 10^{-6} \) kcal/mol across a variety of dielectric constants, ion charges, and cavity radii. Note that these \( \zeta \) values fall within the range specified by Scalmani and Frisch.
7.4 Results

7.4.1 Solvation energies with Lebedev grids

Figure 7.1 shows the differences in the total energy, $W$, among the three different forms of $K$ that are considered here. We report the energies obtained using the asymmetric $X = \text{SAD}^\dagger$ and $X = \text{DAS}$ forms of $K$, relative to those obtained using the symmetric form, for the GBO, SWIG, and ISWIG discretization schemes. In these calculations, the solute molecules are described using either the AMBER99 force field [Fig. 7.1(a)] or else at the Hartree-Fock (HF) level [Fig. 7.1(b)], and $N = 590$ Lebedev grid points are used per atomic sphere.

For the neutral amino acids described at the AMBER99 level, the energy differences amongst the various forms of $K$ are $\lesssim 1$ kcal/mol, which is smaller than the intrinsic error in solvation energies computed at the IEF-PCM level$^{200,225}$. At the HF/6-31+G* level, however, the discrepancies for the neutral amino acids approach 7 kcal/mol in several cases. For the charged amino acids, which exhibit the largest polarization energies, the discrepancies between different forms of $K$ approach 8 kcal/mol for AMBER99 solutes and 24 kcal/mol for HF solutes. The effects seen here for neutral versus charged solutes are further examined in Section 7.5.2.

The SWIG and ISWIG methods, which are compared in Fig. 7.1, differ only in the choice of switching function. Compared to SWIG, ISWIG does not allow the surface charges to penetrate as deeply into the interior of the cavity$^{213}$, and we observe that ISWIG affords much smaller (although still significant) differences between various forms of $K$. The GBO method, wherein Gaussian blurring is used
Figure 7.1: Relative energies of the amino acids in water, obtained at (a) the AMBER99 level and (b) the HF/6-31+G* level. The cavity surface is discretized using 590 Lebedev points per atomic sphere, and solution-phase energies obtained using $X = \text{DAS}$ or $X = \text{SAD}^\dagger$ in Eq. (7.11) are reported relative to the energy obtained using a symmetrized form of $X$. The GBO discretization uses Gaussian blurring only, whereas SWIG and ISWIG use a switching function in conjunction with Gaussian blurring.
but where grid points cannot penetrate into the cavity at all, affords slightly smaller variations than are observed using ISWIG. This suggests a trend in which differences amongst the $K$ matrices are magnified as the extent to which the grid points can penetrate into the cavity becomes larger. These results appear to favor ISWIG over SWIG, as the former affords energies in reasonably good agreement with GBO results. Nevertheless, substantial variations among the different $K$ matrices still exist for GBO discretization, which demonstrates that the switching function is not entirely to blame for these variations.

Consistent with results reported in Ref.\textsuperscript{8}, the non-symmetric form of $K$ with $X = DAS$ tends to agree best with the symmetric form. However, previous comparisons of these alternate $K$ matrices\textsuperscript{6,8,214}, using a point-charge discretization and no switching function, have found that the differences in solvation energies amongst them are extremely small ($\lesssim 0.01$ kcal/mol\textsuperscript{214}), which is clearly not the case here. One possible reason for the larger differences observed here could be due to the more complicated topography of the solute cavities in our calculations. (Recall that $DAS = SAD^\dagger$ for a spherical cavity.) Previous numerical comparisons of different $K$ matrices have employed either spherical cavities\textsuperscript{8} or else “united atom” cavities\textsuperscript{162} that, due to the fairly small size of the molecules that were considered, consist of only a few spheres\textsuperscript{6,214}. The definition of $D_{ii}$ also contributes to the lack of asymmetry amongst the $K$-matrix variants, as we shall show in Section 7.4.5.
The results in Fig. 7.1 demonstrate that the switching function alone cannot explain large energy differences among the different forms of K. Therefore, it is interesting to compare results obtained using FIXPVA discretization, where a switching function is used but Gaussian blurring is not. Using FIXPVA, we find that the various choices for K are in excellent agreement with one another, using either Eq. (7.16) or Eq. (7.17) to define $D_{ii}$. Variations in solvation energies amongst the two asymmetric forms of K are no larger than 0.2 kcal/mol for AMBER99 solutes and 0.5 kcal/mol for HF solutes, even for the charged amino acids. However, the fact that the dependence on K is small does not necessarily imply that solvation energies or surface charge distributions are accurate.

For AMBER99 solutes, where there is no escaped charge, one can compute the *exact* electrostatic solvation energy (for a given solute cavity) by solving Poisson’s equation. We do this using the adaptive Poisson-Boltzmann solver (APBS)\textsuperscript{226}, and we have carefully converged the results with respect to the three-dimensional integration grid.\textsuperscript{7,2} The energy computed with APBS (and also the PCMs considered in this work) is only the *electrostatic* part of the solvation energy, hence these results are not necessarily expected to agree with experimental solvation energies, which include non-electrostatic contributions as well. The more important consideration is whether the APBS and PCM results agree with one another, and in fact we find that converged

\textsuperscript{7,2} The APBS grid parameters we have chosen reproduce Born ion energies within $\sim0.2$ kcal/mol for several different solvent dielectric constants and ion charges. In addition, our grid parameters are comparable to tests of APBS convergence performed by other authors\textsuperscript{227–229}, who conclude that the accuracy of APBS, as compared to analytical or other highly accurate models, systematically increases with increasing grid density.
APBS energies lie within 0.13 kcal/mol of the energy that is obtained using IEF-PCM with \( X = \text{DAS} \) and GBO discretization, for each of the amino acids. (See the Supplementary Data [Appendix C]) Exact IEF-PCM calculations should afford an exact solution for \( E_{pol} \), in the absence of escaped charge, and these results demonstrate that Gaussian blurring does not change this fact, at least not for \( X = \text{DAS} \). As such, we take GBO discretization as the benchmark result, for comparison to other discretization schemes.

Figure 7.2 compares these benchmarks to energies computed using alternative discretization schemes. For the \( X = \text{DAS} \) form of \( K \) [Fig. 7.2(a)], the PC, SWIG, and ISWIG procedures all agree with GBO results to within \( \sim 0.1 \) kcal/mol for AMBER99 solutes and to within \( \sim 1 \) kcal/mol for HF solutes. FIXPVA results are the outliers and differ by \( > 1 \) kcal/mol, both for AMBER99 and HF solutes.

For the \( X = \text{SAD}^\dagger \) variant of \( K \) [Fig. 7.2(b)], there is far more scatter amongst the data, with ISWIG and FIXPVA in best agreement with the GBO results. (Note, however, that for \( X = \text{SAD}^\dagger \) the GBO results differ by as much as 4.5 kcal/mol from converged APBS energies, as shown in the Supplementary Data. [Appendix C]) The SWIG method is in poor agreement with GBO results for the charged amino acids, and the PC method is extremely irregular.

All together, ISWIG exhibits the best agreement with GBO results. These results seem to favor the use of \( X = \text{DAS} \), for which the ISWIG, SWIG, PC, and GBO results are all in excellent agreement. The data also suggest that FIXPVA does not provide solvation energies in good agreement with other discretization methods,
Figure 7.2: Energies of the amino acids in water, relative to results obtained using GBO discretization, for (a) the $X = \text{DAS}$ and (b) the $X = \text{SAD}^\dagger$ form of $K$. Cavity surfaces were discretized using 590 Lebedev points per atomic sphere.
7.4.2 Convergence with respect to the surface grid

In a previous study\textsuperscript{213}, we found that $N = 110$ affords negligible violations of Gauss’ Law, and solvation energies that are converged within $\sim 0.1 \text{ kcal/mol}$ of the $N \to \infty$ limit, yet discrepancies among the various forms of $K$ persist even at $N = 590$. In fact, these discrepancies increase with $N$: the maximum variation observed at the HF/6-31+G* level is 15 kcal/mol for $N = 50$, 16 kcal/mol for $N = 110$, and 24 kcal/mol for $N = 590$. (Results for $N = 50$ and $N = 110$ can be found in the Supplementary Data. [Appendix C])

In Fig. 7.3, we test the convergence of the discretization by computing $E_{pol}$ as a function of $N$. For this test, we use the histidine molecule described at the AMBER99 level, since histidine exhibits some of the largest discrepancies amongst the various choices for $K$. For the $X = \text{DAS}$ form of $K$, $E_{pol}$ converges smoothly (and fairly rapidly) as a function of $N$, but the same cannot be said for $X = \text{SAD}^\dagger$, where $E_{pol}$ does not appear to have converged even for $N = 1202$. In addition, the norm of the matrix $M = \text{DAS} - \text{SAD}^\dagger$ is not significantly different at $N = 1202$ than it is at $N = 50$. So long as $||M||$ remains large, the various forms of $K$ will continue to afford significantly different results.

Results for the symmetric form of $K$ are also shown in Fig. 7.3. This is the form used by Chipman in SS(V)PE calculations\textsuperscript{8,160,215}, and we note that it inherits
Figure 7.3: Convergence of the polarization energy (axis at left), as a function of the number of Lebedev grid points per atomic sphere, for histidine described at the AMBER99 level with SWIG discretization. Results for three alternative $K$ matrices are shown, along with the norm of the matrix $M = \text{DAS} - \text{SAD}^\dagger$ (axis at right).
the oscillations in $E_{pol}$ that arise from $X = \text{SAD}^\dagger$, although these oscillations are somewhat damped by the smooth convergence of the $X = \text{DAS}$ form.

### 7.4.3 Cavity surface area

We next examine the total cavity surface area predicted by each method. The total surface area, $\text{tr}(A)$, is a well-defined quantity in PCMs, and it plays an important role in all PCMs because individual surface areas appear in definitions of the $S$ and $D$ matrices. The PC and GBO surface area is defined by the Lebedev weights $\{w_i\}$ and the atomic radii $\{R_i\}$,

$$\sum_i a_i = \sum_l R_i^2 \sum_{i \in l} w_i,$$

(7.19)

and is exact in the limit $N \to \infty$. For smooth PCMs, $w_i$ in Eq. (7.19) is replaced by $w_i F_i$, where $F_i$ is the switching function.

In Table 7.1, we report errors in the total surface area, relative to PC results, for each discretization method. The PC areas are comparable to GEPO areas, having $< 1\%$ error. Both SWIG and ISWIG also exhibit errors of $< 1\%$, but FIXPVA exhibits alarmingly large errors, and severely underestimates the surface area. The fact that FIXPVA underestimates the surface area has been noted previously, but the errors that we observe are significantly larger than in previous studies. An explanation for this observation is provided in Section 7.5.1.

### 7.4.4 Induced surface charge

Gauss’ Law provides a diagnostic for gauging the accuracy of the discretization. For a cavity with unit dielectric inside, and dielectric constant $\varepsilon$ outside, Gauss’ Law states
Table 7.1: Errors in the total cavity surface area, for the amino acid data set using $N = 590$ Lebedev points per atom (SWIG, ISWIG, and FIXPVA) as well as using $N = 60$ GEPOL points per atom. The mean signed error (MSE), root mean square error (RMSE), and maximum signed error (Max) are listed, in percent, taking PC results with Lebedev grids as the benchmark.
that the total induced charge on the cavity surface, \( Q_{\text{surf}} \), is proportional to the total charge contained within the cavity, \( Q_{\text{in}} \), according to

\[
Q_{\text{surf}} = -\left( \frac{\varepsilon - 1}{\varepsilon} \right) Q_{\text{in}}.
\] (7.20)

For solutes described by a force field, \( Q_{\text{in}} \) is simply the overall solute charge. This is not the case for HF solutes, owing to the presence of escaped charge, but we can still apply Eq. (7.20) easily if we take \( Q_{\text{in}} \) to be the total nuclear charge. In this case, \( Q_{\text{surf}} \) is that part of the ASC that is induced by the electrostatic potential arising from the nuclei.

Deviations from Eq. (7.20) for the aqueous amino acids are summarized in Table 7.2. Previously\textsuperscript{213}, we showed that \( N = 590 \) Lebedev points per atom is sufficient to reduce the Gauss’ Law error below 0.001\( e \) in C-PCM calculations on this same set of molecules. The results presented here show that this same level of accuracy is achievable in IEF-PCM calculations, provided that one chooses \( X = \text{DAS} \). Although the \( X = \text{SAD}^\dagger \) variant is only slightly inferior for \textsc{amber}99 solutes, for HF solutes it affords errors as large as 55\( e \) in the nuclear part of \( Q_{\text{surf}} \).

Despite these tremendous errors in \( Q_{\text{surf}} \) for \( X = \text{SAD}^\dagger \), this method does not appear to produce any individual charges \( q_i \) that are anomalously large, at least not for any of the Gaussian-blurred discretization schemes considered here. In all cases, \(|q_i| \lesssim 0.1e\). (We have previously shown that artificially large surface charges can sometimes appear in the absence of Gaussian blurring\textsuperscript{186}.) On the other hand, Table 7.2 shows that the surface charge induced with FIXPVA discretization seems
### Total charge errors (AMBER99)

<table>
<thead>
<tr>
<th>Method</th>
<th>MAE</th>
<th>RMSE</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>GBO</td>
<td>$7.0 \times 10^{-5}$</td>
<td>$6.5 \times 10^{-5}$</td>
<td>$2.1 \times 10^{-4}$</td>
</tr>
<tr>
<td>SWIG</td>
<td>$2.9 \times 10^{-4}$</td>
<td>$3.7 \times 10^{-4}$</td>
<td>$1.3 \times 10^{-3}$</td>
</tr>
<tr>
<td>ISWIG</td>
<td>$2.0 \times 10^{-4}$</td>
<td>$2.6 \times 10^{-4}$</td>
<td>$9.1 \times 10^{-4}$</td>
</tr>
<tr>
<td>FIXPVA</td>
<td>$1.5 \times 10^{-2}$</td>
<td>$7.6 \times 10^{-3}$</td>
<td>$3.4 \times 10^{-2}$</td>
</tr>
</tbody>
</table>

### Nuclear charge errors (HF/6-31+G*)

<table>
<thead>
<tr>
<th>Method</th>
<th>MAE</th>
<th>RMSE</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>GBO</td>
<td>$7.9 \times 10^{-3}$</td>
<td>$5.6 \times 10^{-3}$</td>
<td>$2.2 \times 10^{-2}$</td>
</tr>
<tr>
<td>SWIG</td>
<td>$1.8 \times 10^{-2}$</td>
<td>$3.7 \times 10^{-3}$</td>
<td>$2.7 \times 10^{-2}$</td>
</tr>
<tr>
<td>ISWIG</td>
<td>$1.3 \times 10^{-2}$</td>
<td>$3.1 \times 10^{-3}$</td>
<td>$2.1 \times 10^{-2}$</td>
</tr>
<tr>
<td>FIXPVA</td>
<td>$9.6 \times 10^{-1}$</td>
<td>$2.4 \times 10^{-1}$</td>
<td>$1.4 \times 10^{0}$</td>
</tr>
</tbody>
</table>

Table 7.2: Gauss’ Law error statistics for IEF-PCM. Listed are the mean absolute error (MAE), the root mean square error (RMSE), and the maximum error (Max), evaluated over the amino acid data set, with $\varepsilon = 78.39$ and $N = 590$. All values are given in atomic charge units.
to be less sensitive to the $X$ variants, yet it is less accurate, as compared to the other discretization methods, for $X = \text{DAS}$.

### 7.4.5 Solvation energies with GEPOL grids

The above results were confined to the use of Lebedev grids, and one may question whether some intrinsic aspect of Lebedev grids is causing the observed $K$-matrix asymmetry. To this end, we investigate GEPOL grids here. Since many of the above discrepancies observed with the AMBER99 solutes are very similar to those with the HF/6-31+G* solutes, we restrict our investigation of GEPOL grids to the HF solutes to avoid redundancy.

Results of PC-GEPOL and GBO-GEPOL calculations are presented in Fig. 7.4. As in the case of Lebedev grids, we find large variations amongst the two asymmetric forms of $K$; for some amino acids, these variations are quite large when PC-GEPOL discretization is employed. (One outlier has been omitted from Fig. 7.4, for clarity.) Such large variations are likely caused by point charges in close proximity, and as expected, the use of Gaussian blurring sidesteps this problem, as seen in the GBO-GEPOL data of Fig. 7.4. Similar to the case of Lebedev grids in Fig. 7.1, $X = \text{DAS}$ tends to better agree with the symmetric $X$, as compared to $X = \text{SAD}^\dagger$.

These results show that GEPOL grids are susceptible to $K$-matrix asymmetry to a similar extent as the Lebedev grids are. For GBO discretization, the differences in Fig. 7.4 (60 points per sphere) are not quite as large as those in Fig. 7.1 (590 points per sphere), but this is partly because the GEPOL grids are roughly ten times less dense. As shown in Supplementary Data (Appendix C), the differences between $K$
Figure 7.4: Relative energies of the amino acids in water, obtained at the HF/6-31+G* level. The cavity surface is discretized using 60 GEPOL points per atomic sphere, and solution-phase energies obtained using $X = \text{DAS}$ or $X = \text{SAD}^\dagger$ in Eq. (7.11) are reported relative to the energy obtained using a symmetrized form of $X$. For comparison, the dashed line at -10 kcal/mol indicates the lower bound in Fig. 7.1. The PC-GEPOL $X = \text{SAD}^\dagger$ data point for amino acid H lies at -2544 kcal/mol, far outside the range of this figure.
variants tends to increase with increasing grid density. Indeed, we find in tests that using GEPOL grids of 240 points per atom tend to have less agreement among $K$ variants than that in Fig. 7.4.

### 7.4.6 Definition of $D_{ii}$

Previous investigations of discrepancies between different forms of $K$ have exclusively used the sum rule in Eq. (7.16) to define $D_{ii}$, whereas all of the calculations presented above employ Eq. (7.17). In Table 7.3, we report results of PC-GEPOL and GBO-GEPOL calculations using the sum rule of Eq. (7.16) to define $D_{ii}$. Remarkably, the variations among the different forms of $K$ disappear, and we obtain results that are consistent with previous studies finding only small discrepancies between different forms of $K^{3,6}$. It is especially surprising that the sum rule manages to fix the wild disagreement for the PC-GEPOL calculations (cf. Fig. 7.4).

In light of these results, we are compelled to see if the asymmetry in SWIG disappears in subSWIG. Table 7.3 reports the results of subSWIG, subISWIG, and GBO and PC with the sum rule for the HF amino acids. Note that the only difference between the data in Table 7.3 and Fig. 7.1 is the definition of $D_{ii}$. Like in the GEPOL calculations with the sum rule, the matrix asymmetry disappears almost completely for the GBO and PC calculations. Some asymmetry remains in subSWIG and subISWIG, probably due to grid points penetrating the cavity interior, yet it is greatly reduced compared to that of Fig. 7.1.
Table 7.3: Statistics for relative energies of the amino acids in water, obtained at the HF/6-31+G* level, when using the sum rule of Eq. (7.16). The mean absolute difference (MAD), root mean square difference (RMSD), and the maximum absolute difference (Max) in the solvation energies (in kcal/mol) are tabulated for the two asymmetric \( X \) matrices, relative to the symmetric version.


Figure 7.5: Geometrical definitions used to define the FIXPVA switching functions.

7.5 Analysis

7.5.1 FIXPVA

In Section 7.4.3 we noted that the FIXPVA algorithm significantly underestimates cavity surface areas. Wang and Li\textsuperscript{179} suggest an \textit{ad hoc} re-parameterization of the FIXPVA switching function, for the purpose of surface area calculations. Here, we show that the problem is intrinsic to the FIXPVA switching procedure itself.

The FIXPVA algorithm ensures a smooth potential surface for the solute by means of two different switching functions, $f_1$ and $f_2$. Consider two intersecting spheres, A and B, with centers at points $P_A$ and $P_B$ as depicted in Fig. 7.5(a). Then, for a
discretization point $P_1$ on the surface of sphere A, $f_1 = f_1(d_{12})$ where $d_{12}$ is the distance from $P_1$ to the point $P_2$, which is defined by the intersection of spheres A and B and the $(P_1, P_A, P_B)$ plane. The other switching function is $f_2 = f_2(d_{13})$, where $d_{13}$ represents the distance from $P_1$ to the point $P_3$ where sphere B intersects the line that connects $P_A$ and $P_B$. As $d_{12}$ and/or $d_{13}$ becomes small, the product $f_1(d_{12}) f_2(d_{13})$ is used to scale down the area associated with the grid point $P_1$.

The FIXPVA method appears to have been designed with an eye toward intersections of the type depicted in Fig. 7.5(a), where $P_3$ lies between $P_A$ and $P_B$. Other types of intersections are possible, however, including the one depicted in Fig. 7.5(b). In this case, the distance $d_{AB}$ between $P_A$ and $P_B$ is smaller than the radius of sphere B, a scenario that occurs readily if sphere A represents a hydrogen atom. Whereas in Fig. 7.5(a), it seems reasonable that $P_1$ might be somewhat attenuated, since this point would soon pass into the interior of the cavity if $d_{AB}$ were to decrease, in Fig. 7.5(b) the point $P_1$ clearly lies on the exterior of the cavity, and is necessary for the description of sphere A. Because $d_{12}$ is small, however, this point (and every other point on the surface of sphere A) may nevertheless be substantially attenuated by the switching function $f_1$.

However, a close examination of the FIXPVA code as implemented in GAMESS\textsuperscript{223,224} reveals that FIXPVA averts the situation depicted in Fig. 7.5 by arbitrarily setting $f_1 = 1$ when $d_{AB} < R_A$, the radius of sphere A. This aspect of FIXPVA is not discussed in Ref.\textsuperscript{5}, and, as a result, our previously-reported implementation of
FIXPVA\textsuperscript{186,213} lacked this feature. We have since modified our FIXPVA implementation to reflect the undiscussed aspect of $f_1$. (In the Supplementary Data [Appendix C], we show that results for energies, surface areas, and induced surface charges obtained with our current implementation can replicate those of the GAMESS implementation when using identical GEPOL grids and differ negligibly between Lebedev and GEPOL grids of similar grid density.)

Unfortunately, this undiscussed aspect of $f_1$ is a significant flaw that limits the applicability of FIXPVA because it introduces a discontinuity in $f_1$. The discontinuity exists whenever $R_A$ is less than the parameter $n_2 = 1.5$ Å\textsuperscript{5}, which controls the upper bound of the switching region of $f_1$. Referring to Fig. 7.5, the discontinuity occurs for atom A (with $R_A < n_2$) at the point when $P_A$ and $P_3$ coincide because $n_1 < d_{13} < n_2$ (i.e. it is within the switching region of $f_1$). Once $P_A$ is inside sphere B, however, $f_1$ will be immediately set to unity. We provide an example of this discontinuity in the Supplementary Data (Appendix C) using the GAMESS implementation of FIXPVA, which is also reproducible in our implementation of FIXPVA.

In Ref.\textsuperscript{5}, where the FIXPVA algorithm was introduced, such artifacts were avoided by setting the hydrogen atom radii to zero and setting all other radii to be larger than $n_2$. One could reparametrize FIXPVA, but $n_2$ will always set a lower bound on the atomic radii that can be safely used in FIXPVA, in any parametrization. This suggests that the FIXPVA approach is potentially useful in the context of united-atom cavities\textsuperscript{162}, but less so for all-atom approaches or for any sort of cavity with radii below $n_2$. Our FIXPVA comparisons with the HF solutes unfortunately use hydrogen radii.
of 1.32 Å and are subject to the $f_1$ discontinuity, which may be contributing to the observed errors in energy, surface area, and induced surface charge.

In addition, the underestimation of surface areas (Table 7.1) is also largely due to FIXPVA scaling away areas along the seams of intersecting spheres, whether their radii are above $n_2$ or not, in order to prevent the surface point charges from approaching too closely. This results in “holes” where spheres intersect on the cavity surface. Without any grid points to gather surface charge in these regions, the solvent electrostatics nearby these holes will be poorly described, contributing to errors in energy and in fulfilling Gauss’ law, as seen in Fig. 7.2 and Table 7.2.

### 7.5.2 Dependence of the solvation energy on the electrostatic potential

That the neutral amino acids afford better agreement amongst the various alternative K matrices, as compared to the charged amino acids, is ultimately a consequence of the quadratic dependence of $E_{pol}$ on the electrostatic potential; see Eq. (7.12).

Suppose that we were to scale the electrostatic potential by a factor $\lambda$, at a fixed cavity geometry and therefore a fixed response matrix, $Q$. Then we could write

$$E_{pol}(\lambda) = \frac{1}{2}\lambda^2 v^\dagger Qv .$$  \hspace{1cm} (7.21)

Suppose next that we compute $E_{pol}$ for the two asymmetric K matrices, using the same solute and cavity geometry in each case. Then the difference between the two solvation energies would be

$$\Delta E_{pol}(\lambda) = \frac{1}{2}\lambda^2 v^\dagger (Q^{DAS} - Q^{SAD})v .$$  \hspace{1cm} (7.22)
As such, any difference between the $Q$ matrices is magnified by $\lambda^2$. For AMBER99 solutes, one can verify numerically that $\Delta E_{pol}$ increases by a factor of $\lambda^2$ if the atomic point charges are scaled by $\lambda$.

For HF/6-31+G* solutes, it is not so straightforward to scale $v$ by a factor, so we instead consider a sequence of ionized histidine solutes, His$^{n+}$ ($n = 0, 1, 2, 3$). We find that the energy difference $\Delta W = W_{\text{DAS}} - W_{\text{SAD}^\dagger}$ is fit very well by a quadratic function of $n$. (See the Supplementary Data. [Appendix C]) For neutral histidine, $\Delta W = 6.5$ kcal/mol using SWIG discretization, which is not too much different from the discrepancies observed for neutral amino acids using the same PCM parameters [cf. Fig. 7.1(b)], but $\Delta W$ increases rapidly as the solute charge increases, with $\Delta W = 191.7$ kcal/mol obtained for His$^{3+}$.

In view of these results, one should expect molecules that elicit a smaller electrostatic potential to afford better agreement among the various forms of $K$. This explains the relatively large discrepancies that are observed for the charged amino acids.

7.5.3 Dependence of the surface charge on the electrostatic potential

As noted in Section 7.4.4, the $X = \text{SAD}^\dagger$ form of $K$ exhibits large violations of Gauss’ Law when Gaussian blurring is used. Since the induced surface charges depend linearly on the surface electrostatic potential, $q = Qv$, one might expect larger deviations from Gauss’ Law as the magnitude of $v$ increases for a fixed cavity geometry.
We investigate this trend by means of HF/6-31+G* calculations on a series of homonuclear diatomic molecules. The solute cavity consists of just two spheres, which represents the simplest non-trivial case (since $\text{DAS} = \text{SAD}^\dagger$ for a spherical cavity$^{160,214}$). We fix the internuclear distance at 1.0 Å and use a radius of 1.4 Å for both atomic spheres, with $N = 590$ Lebedev points per sphere. As such, the surface grid (and therefore the matrix $Q$) is held fixed as the atomic number, $Z$, is varied. The magnitude of the nuclear contribution to $v$ increases linearly with $Z$, hence the nuclear contribution to $Q_{\text{surf}}$ should increase linearly with $Z$ as well. The nuclear charge is entirely contained within the cavity, so Eq. (7.20) should hold for this part of $Q_{\text{surf}}$.

Deviations from Gauss’ Law, as a function of $Z$, are shown in Fig. 7.6. For the nuclear contribution to $Q_{\text{surf}}$, numerical errors in Gauss’ Law depend linearly on $Z$ (i.e., on the magnitude of the electrostatic potential), for both asymmetric variants of $K$, which seems reasonable in view of the fact that the total nuclear contribution to $Q_{\text{surf}}$ also increases linearly with $Z$. The magnitude of the error is quite small ($<0.01e$) for $X = \text{DAS}$, whereas for $X = \text{SAD}^\dagger$ the error reaches $0.75e$ at $Z = 10$.

Interestingly, deviations from Gauss’ Law for the total induced surface charge (considering both the nuclear and electronic contributions to the electrostatic potential) are nearly equivalent for the two asymmetric versions of IEF-PCM. These deviations are not necessarily errors, since for HF solutes the total surface charge need not obey Eq. (7.20), owing to the presence of escaped charge. [In fact, the $Z$-dependent trend in the total deviation from Gauss’ Law correlates with atomic electronegativities; the
Figure 7.6: Deviations from Gauss’ Law for the nuclear contribution to the induced surface charge, and for the total (nuclear + electronic) induced surface charge, for two different versions of IEF-PCM. (The inset shows an enlarged view of the nuclear charge error for $X = \text{DAS}$.) The solutes are a series of homonuclear diatomic molecules described at the HF/6-31+G* level. The bond lengths, solute cavities, and discretization points are the same for each molecule. SWIG, ISWIG, and GBO discretization produce essentially identical results, so only SWIG results are shown here.
electronegativity is smallest for \( Z = 3 \), leading to the largest escaped charge and also the largest deviation from Eq. (7.20).\] In the \( X = \text{SAD}^\dagger \) case, this escaped charge somehow cancels the nuclear charge error, and the total deviation from Eq. (7.20) is the same as it is for \( X = \text{DAS} \).

This analysis explains the results in Table 7.2. The Gauss’ Law error for the amino acids at the AMBER99 level is small because the atomic partial charges are small, and thus the magnitude of \( \mathbf{v} \) is relatively small. The total nuclear charge is much larger, and the \( X = \text{SAD}^\dagger \) version of IEF-PCM exhibits large deviations from Gauss’ Law, for the nuclear part of \( Q_{\text{surf}} \).

### 7.5.4 Definition of \( D_{ii} \)

The results in Section 7.4.6 demonstrate that enforcing the sum rule in Eq. (7.16) somehow removes the discrepancies in solvation energies obtained using different forms of \( \mathbf{K} \). This is a perplexing result, given that we are only modifying the diagonal elements of \( \mathbf{D} \). That is, the off-diagonal elements of \( \mathbf{D} \) remain unchanged whether using Eq. (7.16) or Eq. (7.17). Yet, it is obvious that the sum rule makes \( D_{ii} \) depend on off-diagonal elements, whereas Eq. (7.17) does not. Although we do not yet have a completely satisfactory explanation for how the sum rule accomplishes this feat, we suggest that it is very likely related to an increase in the diagonal dominance of the matrix \( \mathbf{D}_\text{A} \), as detailed below.

In Fig. 7.7, we plot every value of \( D_{ii}a_i \) across the entire amino acid data set for the discretization methods with Lebedev grids. One can clearly see that the sum rule inflates many of the magnitudes of \( D_{ii}a_i \) relative to the definition in Eq. (7.17). When
Figure 7.7: Histogram of $D_{ij}a_i$ values (unitless) collected from the Lebedev grids (590 points per atom) across all twenty of the amino acids in the HF calculations. These values are plotted (a) using Eq. (7.17) and (b) using Eq. (7.16). The inset in panel (a) shows a magnified view of the narrow range in which all the values fall.
using Eq. (7.17), all $D_{ii}a_i$ values are negative and are sharply peaked within the range of $-0.4 \leq D_{ii}a_i \leq 0$. The sum rule, on the other hand, has many significantly larger values over a wide range that even spans some positive values.

Let us consider the effect of the sum rule. We find that inflation of the diagonal element tends to be most pronounced for rows of $DA$ containing large off-diagonal elements, such that $|\sum_{j \neq i} D_{ij}a_j| > 2\pi$. In such a case, $|D_{ii}a_i| \sim |\sum_{j \neq i} D_{ij}a_j|$, and $DA$ gains approximate diagonally dominant character. In addition, the magnitude of off-diagonal elements of $DA$ tend to rapidly decay away from the diagonal elements because $(DA)_{ij} \propto r_{ij}^{-3}$. Combining these two characteristics together via the sum rule, one can have that $DA \approx AD^\dagger$, which makes $K$ approximately symmetric for any $X$ variant. In contrast, the use of Eq. (7.17) does not let $DA$ gain diagonally-dominant character, and, in that case, large off-diagonal elements have a more prominent bearing on the asymmetry.

Although we cannot firmly state that inflated diagonal elements are necessarily incorrect, we have pointed out in previous work\textsuperscript{213} that when using the sum rule, one runs the risk of compromising the negative-definiteness of the response matrix, $Q$, making the PCM energy non-variational. From another point of view, it also seems somewhat unphysical that the diagonal elements $D_{ii}$, which are related to the self-field interaction of the $i$th surface element\textsuperscript{4}, should depend on any surface element other than its own. That is, Eq. (7.16) has the strange property that the self-field interaction over a given surface element is a function of the positions of all other surface elements. Another odd property is that the sum rule produces positive values
for $D_{ii}a_i$ (Fig. 7.7), which is problematic because $DA$ is supposed to be the matrix analog of the negative-definite integral operator $\hat{D}$, and a negative-definite matrix must necessarily have all negative entries along its diagonal. Values of $D_{ii}a_i > 0$ correspond to the unphysical interpretation that the self-field interaction over the $i$th surface element is attractive$^{213}$. Thus, even though the sum rule removes variations amongst the different $K$ matrices, there is no guarantee that the energy and/or surface charges produced with it are physically accurate. Therefore, we do not appeal to the sum rule to alleviate $K$-matrix asymmetry in SWIG/ISWIG.

### 7.5.5 Conductor limit

If there is no escaped charge, then the exact relation

$$\hat{S}\sigma(\vec{s}) = -\phi_0(\vec{s})$$  \hspace{1cm} (7.23)

is valid in limit $\varepsilon \to \infty$.$^{160}$ On the other hand, Eq. (7.1) reduces to Eq. (7.23) in this limit, regardless of whether there is any escaped charge. As such, the reaction field is inexact outside of the cavity, if there is escaped charge.$^{160}$

Although the limiting form of the C-PCM and IEF-PCM equations has been derived previously$^{160,198,203,213}$, less attention has been given to the discretized matrix equations in this limit. The discretized form of the PCM working equations can be written as $-Q^{-1}q = -v$. For C-PCM, $Q = -[(\varepsilon - 1)/\varepsilon]S^{-1}$.$^{213}$ and this equation becomes

$$\left(\frac{\varepsilon}{\varepsilon - 1}\right) S q = -v .$$  \hspace{1cm} (7.24)
For the $X = \text{DAS}$ form of IEF-PCM, the corresponding equation is

$$Y^{-1} \left( \frac{1}{f_\varepsilon} I - \frac{1}{2\pi} DA \right) S q = -v , \quad (7.25)$$

whereas for the $X = \text{SAD}^\dagger$ form one has

$$Y^{-1} S \left( \frac{1}{f_\varepsilon} I - \frac{1}{2\pi} AD^\dagger \right) q = -v . \quad (7.26)$$

For a conductor, we simply have $S q = -v$, and it is immediately clear that Eqs. (7.24) and (7.25) reduce to this form in the limit $\varepsilon \to \infty$. Furthermore, the correct limit is obtained from Eq. (7.25) regardless of whether $\text{DAS} = \text{SAD}^\dagger$ or not. This feature of $X = \text{DAS}$ has also been pointed out elsewhere$^{203,214}$.

In the limit $\varepsilon \to \infty$, Eq. (7.26) can be rearranged to yield

$$S q - \frac{1}{2\pi} S \text{SAD}^\dagger q = -v + \frac{1}{2\pi} D A v . \quad (7.27)$$

To obtain the conductor limit ($S q = -v$) from this equation, it must be the case that

$$-S \text{SAD}^\dagger q = D A v = -\text{DAS} q , \quad (7.28)$$

or in other words, $\text{SAD}^\dagger = \text{DAS}$. As such, we expect that both the $X = \text{SAD}^\dagger$ and the symmetrized version of IEF-PCM will afford incorrect limits as $\varepsilon \to \infty$.

To test this presumption, we investigated the conductor limit numerically, using histidine described at the AMBER99 level. Figure 7.8 shows the solvation energy as a function of $\varepsilon$, computed using C-PCM and also both asymmetric forms of IEF-PCM. The conductor-like model tracks the $X = \text{DAS}$ model fairly closely (with small discrepancies when $\varepsilon$ is small, as expected), and by the time $\varepsilon \approx 80$, the two methods
Figure 7.8: Approach to the conductor limit, for histidine described at the AMBER99 level using SWIG discretization. The same cavity is used in each case and is discretized using 110 Lebedev points per atomic sphere.

are indistinguishable. In contrast, the SAD† variant of IEF-PCM is in fair agreement with the other two models when $\varepsilon = 2$, but already exhibits a large discrepancy by the time $\varepsilon = 10$, which only increases as $\varepsilon \to \infty$. This discrepancy is entirely the result of Eq. (7.27) failing to yield the correct conductor limit when $\text{DAS} \neq \text{SAD}†$.

To understand the agreement between C-PCM and the $X = \text{DAS}$ form of IEF-PCM, we rearrange Eq. (7.1) for IEF-PCM to obtain

$$\left[\hat{I} + \frac{1}{\varepsilon} \left(2\hat{\gamma}^{-1} - \hat{I}\right)\right]\hat{S}\sigma = -\left(\frac{\varepsilon - 1}{\varepsilon}\right)\phi_0.$$  \hspace{1cm} (7.29)

The only difference between this equation and the corresponding C-PCM equation is the term on the left that includes the factor of $1/\varepsilon$ [cf. Eq. (7.24)]. In other words, C-PCM introduces an error on the order of $\varepsilon^{-1}$ in the calculation of the polarization
energy, and is therefore increasingly accurate as $\varepsilon$ increases.$^{165,181,182}$

### 7.6 Conclusions

For IEF-PCM/SS(V)PE calculations, we recommend the use of the asymmetric “DAS” form of the $K$ matrix,

$$
K = S - \frac{1}{2\pi} \left( \frac{\varepsilon - 1}{\varepsilon + 1} \right) \text{DAS}.
$$

With this choice for $K$, the IEF-PCM method exhibits the following desirable properties.

- The electrostatic solvation energy, $E_{\text{pol}}$, converges smoothly and rapidly as the surface grid density increases.
- In the absence of escaped charge, numerical errors in $E_{\text{pol}}$ are $< 0.2$ kcal/mol, and deviations from Gauss’ Law are $\lesssim 0.02e$.
- The method has the correct limiting behavior as $\varepsilon \to \infty$.
- All of the aforementioned features are preserved when IEF-PCM is implemented using the ISWIG smooth discretization scheme.

The last point serves to underline our previous conclusion$^{213}$ that the ISWIG method represents a stable, accurate, and robust approach to cavity discretization.

Although previous studies report little difference between IEF-PCM results obtained using Eq. (7.30) versus its transpose,$^{3,6,203,214}$ we observe numerous instances where the transposed (“SAD$^\dagger$”) form of the equations leads to large errors and/or erratic behavior. The symmetrized version of the IEF-PCM equations, which has been
employed in some previous work\textsuperscript{8,160,215}, is also subjected to these errors, since this form inherits some of the ill behavior of the SAD\textsuperscript{f} form, including an incorrect limit as $\varepsilon \to \infty$. We can attribute the small differences in these previous studies to the use of the sum rule in Eq. (7.16) for diagonal elements of matrix $D$. We observe that the sum rule removes errors associated with $K$-matrix asymmetry, seemingly by inflating diagonal elements of the $D$ matrix to compensate for off-diagonal asymmetry. However, the sum rule can lead to unphysical energies and/or surface charges, and we do not recommend it as a means to alleviate discrepancies amongst the various forms of $K$. 
CHAPTER 8

A simple polarizable continuum solvation model for electrolyte solutions

8.1 Introduction

Implicit solvent models play an important role in both biomolecular simulations and quantum chemistry calculations. These models dramatically enhance conformational sampling and also reduce the complexity of free energy calculations, as compared to simulations using explicit, atomistic representations of the solvent. A wide variety of implicit solvent models have been introduced over the years, ranging from microscopic to semi-microscopic to macroscopic, and varying widely in their level of sophistication. The merits of—and problems associated with—many of these approaches have been discussed in recent reviews. The present work introduces a new implicit solvent model for electrolyte solutions that allows for variation of the solute geometry and provides smooth forces for molecular dynamics simulations, for arbitrary cavity shapes.

Specifically, we focus on solution of the classical electrostatic continuum problem (Poisson’s equation), modified by the presence of a thermal distribution of mobile

8,1 This chapter appeared as a full article in the Journal of Chemical Physics, in 2011, volume 134, pages 204110:1–15.
ions. The electrostatic interaction between two different dielectrics, or the interaction between a dielectric continuum and an atomistic region, is then described by the Poisson-Boltzmann equation.\textsuperscript{235,237,238} Truncating Boltzmann factors at first order, one obtains the \textit{linearized} Poisson-Boltzmann equation (LPBE),\textsuperscript{249}

\[
\left( \hat{\nabla}^2 - \kappa^2 \right) U(\mathbf{r}) = 0 ,
\] (8.1)

which is also known as the \textit{Debye-Hückel equation}. The quantity \( U(\mathbf{r}) \) represents the electrostatic potential. Within the LPBE, the solvent is characterized by a dielectric constant (relative permittivity), \( \varepsilon \), and a Debye length \( \lambda = \kappa^{-1} \), where

\[
\kappa = \frac{1}{\lambda} = \left( \frac{8\pi e^2 \mathcal{J}}{\varepsilon k_B T} \right)^{1/2}.
\] (8.2)

The quantity \( \lambda \) characterizes the length scale on which the mobile ions screen the Coulomb potential. This length scale depends upon the \textit{ionic strength} of the solution,

\[
\mathcal{J} = \frac{1}{2} \sum_i z_i^2 c_i.
\] (8.3)

The sum in this equation ranges over dissolved ionic species with concentrations \( c_i \) and dimensionless charge numbers \( z_i \).

Equation (8.1) forms the basis of the electrostatic solvation effects that we wish to describe. We consider an atomistic representation of the solute molecule(s) and a continuum description of the solvent, the latter parameterized by \( \varepsilon \) and \( \kappa \). In the present work, the solute is described by a set of point charges located at atomic sites, although in principle a more sophisticated description (such as the protein dipoles/Langevin dipoles of Warshel and co-workers,\textsuperscript{247,250,251} a polarizable force field,\textsuperscript{252} or
a quantum-mechanical charge distribution) could be used to describe the solute. A “solute cavity”, constructed from atom-centered spheres, will represent the interface between these two regions. Because we employ an atomistic description of the matter inside of the cavity, we set $\varepsilon = 1$ in this region.

An analytical solution of the LPBE, Eq. (8.1), is available in the case of disjoint spherical cavities with no escaped charge.\(^\text{10}\) (This analytical solution forms the basis of a new semi-analytical model for arbitrary cavity shapes.\(^\text{229}\)) More often, however, the LPBE or its nonlinear precursor has been solved numerically,\(^\text{253–261}\) but unfortunately these numerical approaches are vulnerable to discontinuities in the forces that are required for molecular dynamics simulations. As such, these numerical approaches are typically used in conjunction with a frozen internal geometry for the solute molecule(s). Although some progress has been made toward reducing the magnitude of the discontinuities, in order to provide energy-conserving forces,\(^\text{262}\) it is worth exploring alternative numerical methods that are specifically constructed so that the energy and its gradients are smooth functions of the nuclear coordinates of the solute molecule(s), and which are therefore immediately suitable for molecular dynamics applications. Such methods are introduced here.

Specifically, we consider a family of boundary-element approaches that is known in the quantum chemistry literature as either “apparent surface charge” models, “reaction-field” models, or (most often) “polarizable continuum” models (PCMs).\(^\text{159}\) These models are easily incorporated into self-consistent field procedures, and are
widely used in quantum chemistry. A variety of different PCMs are available, including the conductor-like screening model (COSMO);\textsuperscript{165} a slightly-modified variant known as generalized COSMO (GCOSMO),\textsuperscript{181,182} which has also been called the conductor-like PCM (C-PCM);\textsuperscript{166} the “integral equation formalism” (IEF-PCM),\textsuperscript{167,183,198,210,212} and the “surface and simulation of volume polarization for electrostatics” [SS(V)PE] method.\textsuperscript{160} The latter is formally equivalent to IEF-PCM at the level of integral equations,\textsuperscript{184} although some differences exist, relative to IEF-PCM, in how SS(V)PE has typically been implemented numerically.\textsuperscript{8,213,221}

For $J = 0$, and provided that none of the solute’s charge distribution “escapes” beyond the cavity that defines the solute/continuum interface, the IEF-PCM/SS(V)PE method affords the same electrostatic solvation energy as Poisson’s equation, up to discretization errors.\textsuperscript{249} Furthermore, we have recently shown how numerical IEF-PCM/SS(V)PE calculations can be performed in a manner that provides a continuous potential energy surface for the solute, and also avoids spurious numerical fluctuations due to close approach of surface discretization points.\textsuperscript{186,213} These properties make the IEF-PCM/SS(V)PE approach attractive for problems where forces are required.

As compared to IEF-PCM/SS(V)PE, the GCOSMO/C-PCM method is approximate but is formally and computationally simpler. Moreover, these methods are identical in the limit $\varepsilon \rightarrow \infty$.\textsuperscript{8,221} In the absence of escaped charge, GCOSMO solvation energies are essentially identical to IEF-PCM results for $\varepsilon \gtrsim 10$.\textsuperscript{221}
In this work, we are concerned with electrolyte solutions rather than neat liquid solvents. The IEF-PCM/SS(V)PE method has been adapted for solution of the LPBE rather than Poisson’s equation, and we will refer to these adaptations as “screened PCMs” because they involve a screened Coulomb operator (Yukawa potential), $e^{-\kappa r}/r$. To the best of our knowledge, no analytic implementation of the screened IEF-PCM or SS(V)PE models has yet been reported. Rather, numerical quadrature has been used to evaluate integrals involving the screened Coulomb potential, which is non-trivial because the integrand is singular when these integrals involve a single surface element. (This is actually a problem in many different finite-element solvation models.) In the present work, we show that our recently-developed switching/Gaussian (SWIG) approach to smooth surface integration has an additional advantage in the context of screened PCMs, in that it leads to analytic expressions for the matrix elements required in these models. These expressions greatly simplify the computations and facilitate the development of analytic gradient expressions.

While the generalization of IEF-PCM/SS(V)PE to electrolyte solutions (within the Debye-Hückel limit) has been known for some time, no analogous generalization of the conductor-like models has been reported. The reason for this is unclear to us, but we suspect that it may be due to the ad hoc nature of traditional derivation of COSMO, which can obscure the underlying physics of the model. In this work, we present an alternative derivation for COSMO that we think is more intuitive, and which immediately suggests a screened variant of this model. Because this
new model extends the “conductor-like screening model” (COSMO) to electrolytes in the Debye-Hückel limit, we refer to it as the Debye-Hückel-like screening model (DESMO).

The remainder of this work is organized as follows. Section 8.2 presents a review of the relevant PCM theory, followed by a derivation of DESMO and a mathematical comparison to screened IEF-PCM/SS(V)PE. Section 8.3 provides the details of the SWIG implementation of these models. In Section 8.4, we present several numerical tests on molecular mechanics (MM) solutes, for which there is no escaped charge, in order to evaluate the accuracy of DESMO as well as our analytic implementation of screened IEF-PCM/SS(V)PE. Finally, Section 8.5 presents a summary.

8.2 Theory

8.2.1 The Debye-Hückel model problem

Debye-Hückel theory\textsuperscript{14,263} is developed by considering the electrostatics of a fixed spherical ion immersed in a continuum dielectric solvent, within which mobile ions of a dissociated salt are present (see Fig. 8.1). The positions of the mobile ions are governed by Boltzmann statistics at a specified temperature and, in combination with the dielectric medium, give rise to an average screened Coulomb potential that interacts with the fixed ion, \textit{i.e.}, a screened reaction-field potential. The solvent is thus characterized by a dielectric constant, $\varepsilon$, and a Debye length $\lambda = \kappa^{-1}$.

For later use, let us review the analytical solution to the LPBE [Eq. (8.1)] for a single point charge, $q$, centered in a spherical cavity of radius $b$ (Fig. 8.1). The
mobile ions are assumed to be hard spheres of radius $c$ that form a layer of width $c$ around the solute cavity. Mobile ions cannot penetrate into this layer, which is therefore known as the ion exclusion layer (IEL). Let $a = b + c$ be the distance from the cavity-centered point charge to the bulk solvent boundary. The interior of the cavity ($r < b$) is assumed to have a dielectric constant of unity, whereas the IEL and bulk solvent have dielectric constant $\varepsilon$. Under these assumptions, Eq. (8.1) can be solved exactly for all space. With the point charge $q$ located at the origin, the solution is

$$U(r) = \begin{cases} 
  u_0 & \text{for } 0 < r < b \\
  u_1(r) & \text{for } b < r < a \\
  u_2(r) & \text{for } r > a
\end{cases} \quad (8.4)$$
where

\[
    u_0 = \frac{q[1 + \kappa(a - b)]}{\varepsilon b(1 + \kappa a)} 
    \quad \text{ (8.5a)}
\]

\[
    u_1(r) = \frac{q[1 + \kappa(a - r)]}{\varepsilon r(1 + \kappa a)} 
    \quad \text{ (8.5b)}
\]

\[
    u_2(r) = \frac{q e^{\kappa(a-r)}}{\varepsilon r(1 + \kappa a)} 
    \quad \text{ (8.5c)}
\]

For the region \( r > a \), note in particular the presence of the screened Coulomb potential, \( e^{-\kappa r}/r \), and the factor

\[
    \gamma = \frac{e^{\kappa a}}{1 + \kappa a} 
    \quad \text{ (8.6)}
\]

in Eq. (8.5c). The factor \( \gamma \) alone accounts for the finite radius of the ions, and we therefore refer to \( \gamma \) as the IEL factor. The total electrostatic energy of the Debye-Hückel model system is\textsuperscript{13,264}

\[
    W = -\frac{q^2}{2b} \left[ 1 - \frac{1}{\varepsilon} + \frac{\kappa b}{\varepsilon(1 + \kappa a)} \right]. 
    \quad \text{ (8.7)}
\]

For \( \kappa = 0 \), this expression reduces to the Born ion model.\textsuperscript{12}

In the remainder of this work, we assume that \( c = 0 \) (i.e., that the mobile solvent ions are point charges), so that the IEL coincides with the solute cavity. (The IEL still functions to prevent mobile solvent ions from penetrating the solute cavity interior.) This assumption is tacitly adopted in the derivations of the screened IEF-PCM and screened SS(V)PE models as well.\textsuperscript{249} The simplification that this assumption provides, in the context of PCMs, is that only one surface is present.
8.2.2 Boundary conditions and integral operators

Next, let us establish some notation and also briefly review the boundary conditions and relevant integral operators that appear in various PCMs. We use the notation in Ref. 249 as much as possible. For more details on the integrals operators, see Refs. 249 and 212.

The total electrostatic potential, $U(\vec{r})$, of the solute/solvent system can be separated into two parts,

$$U(\vec{r}) = \phi_0^p(\vec{r}) + U_{\text{rxn}}(\vec{r}) ,$$

(8.8)

where $\phi_0^p(\vec{r})$ is the vacuum electrostatic potential produced by the solute’s charge density, $\rho(\vec{r})$, and $U_{\text{rxn}}(\vec{r})$ is the reaction-field potential generated by the continuum, in response to $\rho(\vec{r})$. The solute is enclosed by surface, $\Gamma$, that constitutes a boundary between the interior volume, $V^{\text{int}}$ (where the dielectric constant is unity), and the exterior volume, $V^{\text{ext}}$, which is characterized by a dielectric constant, $\varepsilon$, and an inverse Debye length, $\kappa$. Points constrained to lie on $\Gamma$ are denoted by $\vec{s}$, whereas $\vec{r}$ can range over all space.

The total electrostatic potential is continuous across $\Gamma$. If $\vec{s}^{\text{int}}$ and $\vec{s}^{\text{ext}}$ represent points that reside on the interior and exterior faces of $\Gamma$, respectively, then

$$U(\vec{s}^{\text{int}}) = U(\vec{s}^{\text{ext}}) .$$

(8.9)

Given Eq. (8.8), this also implies that the reaction-field potential must be continuous across $\Gamma$:

$$U_{\text{rxn}}(\vec{s}^{\text{int}}) = U_{\text{rxn}}(\vec{s}^{\text{ext}}) .$$

(8.10)
The gradient of \( U \), on the other hand, is subject to a “jump” boundary condition. Specifically, the outward-pointing normal electric field generated by \( U \) is discontinuous across \( \Gamma \):

\[
\partial_\vec{s} U(\vec{s}^{\text{int}}) = \varepsilon \partial_\vec{s} U(\vec{s}^{\text{ext}}) .
\] (8.11)

We use the symbol \( \partial_\vec{s} \) as an abbreviation for the outward-pointing normal derivative, \( \partial/\partial \vec{n}_s \). Because \( \partial_\vec{s} \phi_0^\rho(\vec{r}) \) is continuous across \( \Gamma \), the reaction field is solely responsible for the jump boundary condition, and we have

\[
\partial_\vec{s} U^{\text{rxn}}(\vec{s}^{\text{ext}}) = -\left( \frac{\varepsilon - 1}{\varepsilon} \right) \partial_\vec{s} \phi_0^\rho(\vec{s}) + \frac{1}{\varepsilon} \partial_\vec{s} U^{\text{rxn}}(\vec{s}^{\text{int}}) .
\] (8.12)

The screened electrostatic potential produced by a solute charge density \( \rho \) is

\[
\phi^\rho_\kappa(\vec{r}) = \int \rho(\vec{r}') \frac{e^{-\kappa|\vec{r} - \vec{r}'|}}{|\vec{r} - \vec{r}'|} d^3\vec{r}' .
\] (8.13)

Similarly, the screened electrostatic potential produced by a single-layer surface charge density, \( \sigma(\vec{s}) \), is

\[
\chi^\sigma_\kappa(\vec{r}) = \int \sigma(\vec{s}) \frac{e^{-\kappa|\vec{r} - \vec{s}|}}{|\vec{r} - \vec{s}|} d^2\vec{s} ,
\] (8.14)

In the limit \( \kappa \to 0 \), these potentials reduce to the corresponding unscreened (vacuum) electrostatic potentials.

The self-adjoint integral operator \( \hat{S}_\kappa \) is defined such that the action of \( \hat{S}_\kappa \) on \( \sigma(\vec{s}) \) produces the screened potential at the surface point \( \vec{s} \):

\[
\hat{S}_\kappa \sigma(\vec{s}) = \chi^\sigma_\kappa(\vec{s}) .
\] (8.15)

Note that \( \chi^\sigma_\kappa(\vec{s}) \) therefore depends on the surface charge across all of \( \Gamma \). The quantity \( \partial_\vec{s} \chi^\sigma_\kappa(\vec{s}) \) represents the negative normal component of the screened electric field.
produced by $\sigma(s)$, evaluated at the point $s$. The integral operator $\hat{D}_\kappa$ is defined such that it generates this screened negative normal electric field, according to

$$\hat{D}_\kappa \sigma(s) = \partial_s \chi_\kappa^\sigma(s).$$

The discontinuity inherent in the jump boundary condition can then be expressed as

$$\left[ \hat{D}_\kappa + 2\pi I \right] \sigma(s) = \partial_s \chi_\kappa^\sigma(s) \bigg|^{s_{\text{int}}} \quad (8.17a)$$

and

$$\left[ \hat{D}_\kappa - 2\pi I \right] \sigma(s) = \partial_s \chi_\kappa^\sigma(s) \bigg|^{s_{\text{ext}}}.$$ 

The adjoint of $\hat{D}_\kappa$ is $\hat{D}_\kappa^\dagger$, and these operators obey the relation $\hat{D}_\kappa \hat{S}_\kappa = \hat{S}_\kappa \hat{D}_\kappa^\dagger$. Like the electrostatic potentials, these integral operators reduce to their unscreened forms in the limit that $\kappa \to 0$.

The total energy of the solute/solvent supersystem is

$$W = E_0 + E_{\text{pol}}, \quad (8.18)$$

where $E_0$ is the energy of $\rho(\vec{r})$ in vacuum. The second term, $E_{\text{pol}}$, is the reaction field energy (also known as the electrostatic solvation energy) that arises from polarization of the dielectric continuum. This quantity is given by

$$E_{\text{pol}} = \frac{1}{2} \int \rho(\vec{r}) U_{\text{rxn}}(\vec{r}) \, d^3 \vec{r}, \quad (8.19)$$

where the factor of $1/2$ accounts for polarization work. As pointed out by Chipman, all of the PCMs considered here invoke the ansatz

$$U_{\text{rxn}}(\vec{r}) = \chi_0^\sigma(\vec{r}), \quad (8.20)$$

for all $\vec{r}$. 

231


8.2.3 Conductor-like screening model

Traditional derivation

Conventional derivations of COSMO and the closely-related GCOSMO/C-PCM method begin by assuming that the solute resides completely inside of a cavity that is embedded in a conductor.\textsuperscript{165,166,181,182} This implies that $U(\vec{r})$ vanishes at the cavity surface, $\Gamma$, and is zero beyond this boundary:

$$U(\vec{r}) = \begin{cases} \phi_0^\rho(\vec{r}) + \chi_0^\sigma(\vec{r}) & \text{for } \vec{r} \in V^{\text{int}} \\ 0 & \text{for } \vec{r} \in V^{\text{ext}}. \end{cases} \quad (8.21)$$

To ensure that the electrostatic potential is continuous across the cavity surface [Eq. (8.9)], one must then have

$$\hat{S}_0 \sigma^{\text{conductor}}(\vec{s}) = -\phi_0^\rho(\vec{s}). \quad (8.22)$$

To obtain a model that works for a continuum with a finite dielectric constant, rather than a conductor, the standard approach is to scale $\sigma^{\text{conductor}}(\vec{s})$ by an \textit{ad hoc} factor, $f_\varepsilon$. The equation to be solved is then

$$\hat{S}_0 \sigma^{\text{COSMO}}(\vec{s}) = -f_\varepsilon \phi_0^\rho(\vec{s}). \quad (8.23)$$

In the original derivation of COSMO,\textsuperscript{165} Klamt and Schüürman propose the scaling factor $f_\varepsilon = (\varepsilon - 1)/(\varepsilon + 0.5)$. The GCOSMO/C-PCM method also uses Eq. (8.23) but with $f_\varepsilon = (\varepsilon - 1)/\varepsilon$. As compared to the original COSMO, the latter choice more closely approximates Gauss’ Law for the surface charge density.\textsuperscript{181,182}

Alternative derivation

The difficulty with this derivation, in the context of electrolyte solutions, is that the scaling factor $f_\varepsilon$ is applicable only for $\kappa = 0$. For $\kappa > 0$, no analogous scaling
factor exists for arbitrary cavity shapes. As such, the derivation of COSMO presented above cannot be extended to electrolytic solvents. Here, we reconsider the derivation of COSMO.

Rather than starting from the conductor limit and later scaling the surface charge, one can derive COSMO starting from an \textit{ansatz} that actually resembles a cavity immersed in a dielectric \cite{cf. Eq. (8.21)}:

\begin{equation}
U(\vec{r}) = \begin{cases} 
\phi_0^\rho(\vec{r}) + \chi_0^\sigma(\vec{r}) & \text{for } \vec{r} \in V^\text{int} \\
\phi_0^\rho(\vec{r})/\varepsilon & \text{for } \vec{r} \in V^\text{ext}
\end{cases}, \tag{8.24}
\end{equation}

Enforcing the continuity condition [Eq. (8.9)] upon this \textit{ansatz} immediately affords Eq. (8.23), with \( f_\varepsilon = (\varepsilon - 1)/\varepsilon \) as in GCOSMO/C-PCM. At no point do we appeal to the conductor limit. In view of this result, the solution of Eq. (8.23) can be interpreted as that particular surface charge density, \( \sigma^{\text{COSMO}}(\vec{s}) \), that makes \( U(\vec{r}) = \phi_0^\rho(\vec{r})/\varepsilon \) for all points \( \vec{r} \in V^\text{ext} \). This interpretation is not obvious from the conventional derivation of COSMO.

Although the surface charge density \( \sigma^{\text{COSMO}}(\vec{s}) \) that is obtained from Eq. (8.24) satisfies the continuity condition, it does not exactly satisfy the jump boundary condition except in special cases. This is made immediately clear by taking the normal derivative of \( U^\text{rxn}(\vec{s}^\text{ext}) \) within the \textit{ansatz} of Eq. (8.24),

\begin{equation}
\partial_\vec{s} U^\text{rxn}(\vec{s}^\text{ext}) = - \left( \frac{\varepsilon - 1}{\varepsilon} \right) \partial_\vec{s} \phi_0^\rho(\vec{s}), \tag{8.25}
\end{equation}

which is \textit{not} the correct jump boundary condition [Eq. (8.12)] because it lacks the term \( \varepsilon^{-1} \partial_\vec{s} U^\text{rxn}(\vec{s}^\text{int}) \). That is, GCOSMO/C-PCM only approximately satisfies the jump boundary condition, and Eq. (8.25) should be regarded as the fundamental
approximation in this model. On the other hand, the missing term in this boundary condition is proportional to $\varepsilon^{-1}$, so Eq. (8.25) is exact in the conductor limit, $\varepsilon \rightarrow \infty$. It is also exact if $\partial_{\vec{s}} U^\text{rxn}(\vec{s}^\text{int}) = 0$, in which case $U^\text{rxn}(\vec{s}^\text{int})$ is constant. Examples where the latter situation is realized include the trivial case of $\varepsilon = 1$ (i.e. the gas phase), wherein $U^\text{rxn}(\vec{s}^\text{int}) = 0$, or for certain highly symmetrical systems, such as a uniform surface charge distribution over a spherical cavity, as in the Born ion model.

Interestingly, Eq. (8.25) also suggests an equivalent secondary equation that could be solved to obtain $\sigma^\text{COSMO}(\vec{s})$,

$$
\left( \hat{D}^\dagger_0 - 2\pi \hat{I} \right) \sigma^\text{COSMO}(\vec{s}) = - \left( \frac{\varepsilon - 1}{\varepsilon} \right) \partial_{\vec{s}} \phi^\rho_0(\vec{s}),
$$

which follows from Eq. (8.17b). To illustrate the equivalence of Eq. (8.26) with Eq. (8.23), we multiply Eq. (8.26) with $\hat{S}_0$ and use the identity $\hat{D}_0 \hat{S}_0 = \hat{S}_0 \hat{D}_0^\dagger$ to arrive at

$$
\left( \hat{D}_0 - 2\pi \hat{I} \right) \hat{S}_0 \sigma^\text{COSMO}(\vec{s}) = - \left( \frac{\varepsilon - 1}{\varepsilon} \right) \hat{S}_0 \partial_{\vec{s}} \phi^\rho_0(\vec{s}).
$$

Next, we can make use of an identity derived by Chipman,

$$
- \left( \hat{D}_\kappa - 2\pi \hat{I} \right) \phi^\rho_\kappa(\vec{s}) + \hat{S}_\kappa \partial_{\vec{s}} \phi^\rho_\kappa(\vec{s}) = 4\pi \varphi^\rho_\kappa(\vec{s}),
$$

where $\varphi^\rho_\kappa(\vec{s})$ is the screened electrostatic potential due to that portion of solute’s charge density, $\rho(\vec{r})$, that penetrates beyond the solute cavity. Then, for the salt-free case where the solute charge density is completely contained inside the cavity, such that $\kappa = 0$ and $\varphi^\rho_\kappa(\vec{s}) = 0$, Eq. (8.28) can be substituted into Eq. (8.27) to obtain Eq. (8.23). Indeed, we have confirmed in numerical tests that the $\sigma^\text{COSMO}(\vec{s})$ obtained
from Eq. (8.26) is equivalent to Eq. (8.23). However, Eq. (8.23) is much easier to solve in practice because it does not involve surface normal derivatives.

8.2.4 Debye-Hückel–like screening model

Following the logic of our alternative derivation of GCOSMO/C-PCM, it is now clear how electrolytic screening effects can be incorporated within this model. We choose an \textit{ansatz} for \( U(\vec{r}) \) such that any point outside of the cavity has a screened potential [cf. Eq. (8.24)]:

\[
U(\vec{r}) = \begin{cases} 
\phi_0(\vec{r}) + \chi_0(\vec{r}) & \text{for } \vec{r} \in V^{\text{int}} \\
\phi_\kappa(\vec{r})/\varepsilon & \text{for } \vec{r} \in V^{\text{ext}} 
\end{cases}
\]  

(8.29)

Applying the continuity condition to this \textit{ansatz} leads immediately to an equation for the surface charge,

\[
\hat{S}_0 \sigma^{\text{DESMO}}(\vec{s}) = \frac{1}{\varepsilon} \phi_\kappa(\vec{s}) - \phi_0(\vec{s}).
\]  

(8.30)

Equation (8.30) is the primary equation for our Debye-Hückel–like screening model (DESMO). This model involves the same integral operator, \( \hat{S}_0 \), that is used in COSMO and therefore retains a great deal of COSMO’s simplicity, relative to the screened versions of SS(V)PE and IEF-PCM. In the limit \( \kappa \to 0 \), Eq. (8.30) reduces to the GCOSMO/C-PCM equation, as expected.

Like GCOSMO/C-PCM, DESMO only approximately satisfies the jump boundary condition [Eq. (8.12)]. Starting from Eq. (8.28) with \( \phi_\kappa(\vec{s}) = 0 \), and substituting
Eq. (8.30), we arrive at a secondary equation for $\sigma^{\text{DESMO}}(\vec{s})$:

$$
\left( \hat{I} - \frac{1}{2\pi} \hat{D}_\kappa \right) \hat{S}_0 \sigma^{\text{DESMO}}(\vec{s}) = - \left( \hat{I} - \frac{1}{2\pi} \hat{D}_\kappa \right) \phi_0^\rho(\vec{s}) - \frac{1}{2\pi \varepsilon} \hat{S}_\kappa \partial_\kappa \phi_0^\rho(\vec{s}) .
$$

(8.31)

At this point, Eq. (8.31) is not a completely transparent result, but its importance will become apparent below.

### 8.2.5 Screened SS(V)PE and screened IEF-PCM

It is useful to summarize Chipman’s derivation of the screened SS(V)PE model,\textsuperscript{249} for comparison to DESMO. This derivation begins by invoking a pair of separate surface charge distributions, $\sigma^\text{int}$ and $\sigma^\text{ext}$, which reside on the interior and exterior faces of $\Gamma$, respectively. The total electrostatic potential within this ansatz is

$$
U(\vec{r}) = \begin{cases} 
\phi_0^\rho(\vec{r}) + \chi^\text{int}_0(\vec{r}) & \text{for } \vec{r} \in V^\text{int} \\
\left[ \phi_\kappa^\rho(\vec{r}) + \chi^\text{ext}_\kappa(\vec{r}) \right] / \varepsilon & \text{for } \vec{r} \in V^\text{ext}
\end{cases} .
$$

(8.32)

The continuity condition for this ansatz takes the form

$$
\hat{S}_0 \sigma^\text{int}(\vec{s}) - \frac{1}{\varepsilon} \hat{S}_\kappa \sigma^\text{ext}(\vec{s}) = \frac{1}{\varepsilon} \phi_\kappa^\rho(\vec{s}) - \phi_0^\rho(\vec{s})
$$

(8.33)

whereas the jump boundary condition is

$$
\left( \hat{I} + \frac{1}{2\pi} \hat{D}_0^\dagger \right) \sigma^\text{int}(\vec{s}) + \left( \hat{I} - \frac{1}{2\pi} \hat{D}_\kappa^\dagger \right) \sigma^\text{ext}(\vec{s}) = \frac{1}{2\pi} \left[ \partial_\kappa \phi_\kappa^\rho(\vec{s}) - \partial_\kappa \phi_0^\rho(\vec{s}) \right].
$$

(8.34)
As written, Eqs. (8.33) and (8.34) are coupled, but they can be manipulated so as to uncouple $\sigma^{\text{int}}$ and $\sigma^{\text{ext}}$. The result is the primary equation for screened SS(V)PE,\(^{249}\)

\[
\left[\left(\hat{I} - \frac{1}{2\pi} \hat{D}_\kappa\right) \hat{S}_0 + \frac{1}{\varepsilon} \hat{S}_\kappa \left(\hat{I} + \frac{1}{2\pi} \hat{D}_0^\dagger\right)\right] \sigma^{\text{int}}(\vec{s})
= \left(\hat{I} - \frac{1}{2\pi} \hat{D}_\kappa\right) \left[\frac{1}{\varepsilon} \phi_\kappa^\rho(\vec{s}^-) - \phi_0^\rho(\vec{s})\right]
+ \frac{1}{2\pi\varepsilon} \hat{S}_\kappa \left[\partial_\vec{s} \phi_\kappa^\rho(\vec{s}) - \partial_\vec{s} \phi_0^\rho(\vec{s})\right].
\] (8.35)

Chipman then sets $\sigma^{\text{int}}(\vec{s}) = \sigma^{\text{SS(V)PE}}(\vec{s})$, which is used in Eq. (8.20) to obtain the reaction-field potential $U^{\text{rxn}}(\vec{r})$ that is used to evaluate $E_{\text{pol}}$.

The screened IEF-PCM method is derived using a somewhat similar approach,\(^{198}\) considering interior and exterior faces of $\Gamma$, but using only a single surface charge distribution, $\sigma^{\text{IEF-PCM}}(\vec{s})$. The resulting screened IEF-PCM equation is\(^{198,212}\)

\[
\left[\left(\hat{I} - \frac{1}{2\pi} \hat{D}_\kappa\right) \hat{S}_0 + \frac{1}{\varepsilon} \hat{S}_\kappa \left(\hat{I} + \frac{1}{2\pi} \hat{D}_0^\dagger\right)\right] \sigma^{\text{IEF-PCM}}(\vec{s})
= -\left(\hat{I} - \frac{1}{2\pi} \hat{D}_\kappa\right) \phi_0^\rho(\vec{s}) - \frac{1}{2\pi\varepsilon} \hat{S}_\kappa \partial_\vec{s} \phi_0^\rho(\vec{s}).
\] (8.36)

The identity in Eq. (8.28) can be rearranged to yield

\[
\left(\hat{I} - \frac{1}{2\pi} \hat{D}_\kappa\right) \phi_\kappa^\rho(\vec{s}) + \frac{1}{2\pi} \hat{S}_\kappa \partial_\vec{s} \phi_\kappa^\rho(\vec{s}) = 2 \varphi_\kappa^\rho(\vec{s}).
\] (8.37)

Then, as noted in Ref. 249, Eq. (8.37) can be used to recast the right side of Eq. (8.35) into a form analogous to that of Eq. (8.36). Following this manipulation, the basic equation for screened SS(V)PE [Eq. (8.35)] can be written

\[
\left[\left(\hat{I} - \frac{1}{2\pi} \hat{D}_\kappa\right) \hat{S}_0 + \frac{1}{\varepsilon} \hat{S}_\kappa \left(\hat{I} + \frac{1}{2\pi} \hat{D}_0^\dagger\right)\right] \sigma^{\text{SS(V)PE}}(\vec{s})
= -\left(\hat{I} - \frac{1}{2\pi} \hat{D}_\kappa\right) \phi_0^\rho(\vec{s}) - \frac{1}{2\pi\varepsilon} \hat{S}_\kappa \partial_\vec{s} \phi_0^\rho(\vec{s}) + \frac{2}{\varepsilon} \varphi_\kappa^\rho(\vec{s}).
\] (8.38)
Comparing this to Eq. (8.36), it is apparent that the screened versions of SS(V)PE and IEF-PCM are identical except for the term involving $\varphi_\kappa^\rho(\vec{s})$, which is absent in the latter. As such, these two models are equivalent in the absence of escaped charge. The two methods also become equivalent in the limit $\varepsilon \to \infty$.

Let us now compare DESMO to screened SS(V)PE and IEF-PCM, continuing the discussion left off at the end of Section 8.2.4. The secondary form of the DESMO equation, Eq. (8.31), is remarkably similar to Eq. (8.36) but it lacks the second term in square brackets in Eq. (8.36). However, this missing term is proportional to $\varepsilon^{-1}$, and Eq. (8.31) becomes exact in the limit $\varepsilon \to \infty$. DESMO is also exact if

$$\left(\hat{I} + \frac{1}{2\pi} \hat{D}^\dagger_0\right) \sigma^{\text{DESMO}}(\vec{s}) = \frac{1}{2\pi} \left[ \partial_\sigma \varphi^\rho(\vec{s}) - \partial_\sigma \varphi^\rho_0(\vec{s}) \right],$$

(8.39)

which can be seen by substituting Eq. (8.39) into Eq. (8.35). The left side of Eq. (8.39) is simply $\partial_\sigma \chi^\sigma_\kappa(\vec{s}^{\text{int}})$ scaled by a factor of $2\pi$, and it expresses the screened version of the limit in which GCOSMO/C-PCM is exact, $\partial_\sigma U^{\text{rxn}}(\vec{s}^{\text{int}}) = 0$. Clearly, Eq. (8.39) reduces to the condition $\partial_\sigma \chi^\sigma_\kappa(\vec{s}^{\text{int}}) = 0$ [i.e. $\partial_\sigma U^{\text{rxn}}(\vec{s}^{\text{int}}) = 0$] when $\kappa = 0$.

### 8.3 Discretization

In this section, we discretize the integral equations that define the DESMO, screened SS(V)PE, and screened IEF-PCM methods, to obtain finite-dimensional matrix equations. This discretization is accomplished using the SWIG procedure that we have recently introduced for the unscreened SS(V)PE/IEF-PCM and GCOSMO/C-PCM methods. Complete details of the SWIG procedure can be found in Ref. 213; the essential aspects are summarized here.
8.3.1 Overview of the SWIG approach

The solute cavity surface is constructed from a union of atom-centered spheres whose radii are parameters of the model. Although this approach has occasionally been criticized,\textsuperscript{265,266} and more sophisticated approaches have been explored,\textsuperscript{3,164,185,267} the use of parameterized atomic radii remains ubiquitous in the PCM literature, and provides a simple definition for the cavity surface that will facilitate comparison with exact results.

Lebedev grids\textsuperscript{191} are used to discretize the surface of each sphere into a finite set of points, \{\vec{s}_i\}. Each grid point is associated with a surface area \( \tilde{a}_i = w_i R_i^2 \), where \( R_i \) is the radius of the atomic sphere on which the point \( \vec{s}_i \) resides, and \( w_i \) is the Lebedev quadrature weight associated with this grid point. A straightforward way to discretize the integral equations in Section 8.2 would be to place a point charge, \( q_i \), at each grid point \( \vec{s}_i \), thus replacing \( \sigma(\vec{s}) \) with a vector of point charges, \( \mathbf{q} \).

The problem with such an approach is that surface grid points may emerge from—or vanish into—the interior of the cavity as the nuclei move, leading to discontinuities in the potential energy surface. To avoid these, we introduce a switching function, \( F_i \), that smoothly attenuates \( \vec{s}_i \)'s contribution to the PCM equations, as the point \( \vec{s}_i \) passes through a buffer region surrounding the cavity surface. (Precisely how \( F_i \) is incorporated into the matrix elements is discussed in Section 8.3.3.) Insofar as \( F_i \) is a smooth function of the nuclear coordinates, this guarantees a smooth potential energy surface. Note also that the surface area associated with \( \vec{s}_i \) is\textsuperscript{186}

\[
a_i = \tilde{a}_i F_i = w_i F_i R_i^2 .
\]  

(8.40)
We have observed that potential energy surfaces that are rigorously smooth (in the mathematical sense of possessing continuous derivatives) may still suffer from spurious oscillations in the energy or gradient, which arise due to the singular nature of the Coulomb potential between the surface point charges. Introduction of a switching function actually exacerbates this problem, as it allows grid points to approach one another more closely than would be the case if we simply discarded all interior grid points. This problem can be overcome via “Gaussian blurring” of the point charges $q_i$. To wit, we introduce a set of spherical Gaussian functions, $\{g_i\}$, that are centered at the points $\vec{s}_i$:

$$g_i(\vec{r}) = q_i \left( \frac{\zeta_i^2}{\pi} \right)^{3/2} e^{-\zeta_i^2 |\vec{r} - \vec{s}_i|^2}.$$  \hspace{1cm} (8.41)

The exponent $\zeta_i$, which controls the width of $g_i$, depends upon the number of Lebedev grid points used to discretize the spheres. Values of $\zeta_i$ were taken from Ref. 197, where they were optimized to obtain accurate solvation energies for the Born ion model across a range of dielectric constants. The Coulomb interaction between $g_i(\vec{r})$ and $g_j(\vec{r})$ is finite even as $|\vec{s}_i - \vec{s}_j| \to 0$, and in our experience this is sufficient to remove spurious fluctuations in the energy and gradient.

### 8.3.2 Electrostatic solvation energy and gradient

Upon discretization, the electrostatic potential $\phi^\rho_\kappa$ and the normal electric field $\partial_\vec{s} \phi^\rho_\kappa$ are replaced by vectors $\mathbf{v}_\kappa$ and $\mathbf{v}^\perp_\kappa$, respectively, whose elements are

$$v_{\kappa,i} = \phi^\rho_\kappa(\vec{s}_i)$$  \hspace{1cm} (8.42a)

$$v^\perp_{\kappa,i} = \partial_\vec{s} \phi^\rho_\kappa(\vec{s}_i).$$  \hspace{1cm} (8.42b)
Note that the unscreened electrostatic potential, \( v_0 \), can be viewed as a special case of \( v_\kappa \). The point charges \( q_i \) are collected in a vector \( q \) that, together with the surface areas \( a_i \), represent the discretization of \( \sigma(\vec{s}) \).

For consistency with the general PCM framework,\(^8,213\) we wish to cast the integral equations of Section 8.2 into finite-dimensional matrix equations of the form

\[
Kq = Rv_0
\]  

whose formal solution is

\[
q = Qv_0,
\]

where \( Q = K^{-1}R \) is known as the solvent response matrix. Table 8.1 defines the \( K \) and \( R \) matrices for DESMO and screened SS(V)PE in terms of \( S_\kappa \) and \( D_\kappa \), which are the discretized forms of the operators \( \hat{S}_\kappa \) and \( \hat{D}_\kappa \) from Section 8.2. In addition, we have introduced a diagonal matrix \( A \) that contains the surface element areas, \( A_{ij} = a_i \delta_{ij} \), as well as auxiliary matrices \( L \), \( M \), and \( N \). The latter are each diagonal, with matrix elements

\[
L_{ij} = \delta_{ij} v_{0,i}^\perp / v_{0,i}
\]

\[
M_{ij} = \delta_{ij} v_{\kappa,i} / v_{0,i}
\]

\[
N_{ij} = \delta_{ij} v_{\perp\kappa,i} / v_{0,i}.
\]

Following discretization, the reaction-field energy can be written in a variety of equivalent forms:

\[
E_{pol} = \frac{1}{2} v_0^\dagger q = \frac{1}{2} q^\dagger Q^{-1} q = \frac{1}{2} v_0^\dagger Q v_0.
\]
### Table 8.1: Definitions of $\mathbf{K}$ and $\mathbf{R}$ for screened PCM.

<table>
<thead>
<tr>
<th>Method</th>
<th>Matrix $\mathbf{K}$</th>
<th>Matrix $\mathbf{R}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>DESMO</td>
<td>$\mathbf{S}_0$</td>
<td>$-(\mathbf{I} - \frac{1}{\varepsilon}\mathbf{M})$</td>
</tr>
<tr>
<td>IEF-PCM</td>
<td>$\left(\mathbf{I} - \frac{1}{2\pi}\mathbf{D}_\kappa\mathbf{A}\right)\mathbf{S}<em>0 + \frac{1}{\varepsilon}\mathbf{S}</em>\kappa \left(\mathbf{I} + \frac{1}{2\pi}\mathbf{A}\mathbf{D}_0^\dagger\right)$</td>
<td>$\left(\mathbf{I} - \frac{1}{2\pi}\mathbf{D}<em>\kappa\mathbf{A}\right) - \frac{1}{2\pi\varepsilon}\mathbf{S}'</em>\kappa\mathbf{L}$</td>
</tr>
<tr>
<td>SS(V)PE</td>
<td>$\left(\mathbf{I} - \frac{1}{2\pi}\mathbf{D}_\kappa\mathbf{A}\right)\mathbf{S}<em>0 + \frac{1}{\varepsilon}\mathbf{S}</em>\kappa \left(\mathbf{I} + \frac{1}{2\pi}\mathbf{A}\mathbf{D}_0^\dagger\right)$</td>
<td>$\left(\mathbf{I} - \frac{1}{2\pi}\mathbf{D}<em>\kappa\mathbf{A}\right) \left(\frac{1}{\varepsilon}\mathbf{M} - \mathbf{I}\right) + \frac{1}{2\pi\varepsilon}\mathbf{S}'</em>\kappa \left(\mathbf{N} - \mathbf{L}\right)$</td>
</tr>
</tbody>
</table>
Although $Q$ is generally not symmetric, it is easily shown that $E_{pol}$ is invariant to symmetrization of $Q$, i.e., we could replace $Q$ in Eq. (8.46) with the symmetric matrix $\tilde{Q} = (Q + Q^\dagger)/2$. For solutes described using quantum mechanics (at the self-consistent field level), this symmetrization is important because if $Q = Q^\dagger$, then one can evaluate the analytic gradient of $E_{pol}$ without solving the coupled-perturbed equations.\(^6\) In practice, symmetrization of $Q$ requires solving Eq. (8.44) twice, the second time with $Q$ replaced by $Q^\dagger$, as discussed in Ref. 213. This technique will be exploited in Section 8.3.5, when we introduce the DESMO analytic gradient.

So long as $Q$ is symmetric, we can write the derivative of $E_{pol}$ with respect to a perturbation $x$ as

$$E_{pol}'^x = \frac{1}{2} v_0^\dagger Q^x v_0 + q^x v_0^x . \quad (8.47)$$

Expanding the first term,

$$v_0^\dagger Q^x v_0 = v_0^\dagger K^{-1} (R^x - K^x K^{-1} R) v_0 , \quad (8.48)$$

we have a general expression for the derivative of the reaction-field energy. The matrices $K^x$ and $R^x$ are given in Table 8.2.

### 8.3.3 Matrix elements for SWIG discretization

Let us first introduce the matrix elements of the unscreened operators $\hat{S}_0$ and $\hat{D}_0$, which are the same as in our previous work.\(^{213}\) Off-diagonal elements of $S_0$, $S_{0,ij}$, are equal to the unscreened Coulomb interaction between the functions $g_i(\vec{r})$ and $g_j(\vec{r})$. The appropriate Coulomb integral can be written in terms of the error function,

$$S_{0,ij} = \frac{\text{erf}(\zeta_{ij}s_{ij})}{s_{ij}} , \quad (8.49)$$

243
<table>
<thead>
<tr>
<th>Method</th>
<th>Matrix $K^x$</th>
<th>Matrix $R^x$</th>
</tr>
</thead>
<tbody>
<tr>
<td>DESMO</td>
<td>$S_0^x$</td>
<td>$\frac{1}{\varepsilon}M^x$</td>
</tr>
</tbody>
</table>
| IEF-PCM  | $-\frac{1}{2\pi}(D_\kappa^x A + D_\kappa A^x)S_0 + \left(I - \frac{1}{2\pi}D_\kappa A\right)S_0^x$  
$+ \frac{1}{\varepsilon}S_\kappa^x\left(I + \frac{1}{2\pi}AD_0^\dagger\right) + \frac{1}{2\pi\varepsilon}S_\kappa\left[A^xD_0^\dagger + A\left(D_0^\dagger\right)^x\right]$  
$- \frac{1}{2\pi}(D_\kappa^x A + D_\kappa A^x) - \frac{1}{2\pi\varepsilon}(S_\kappa^x L + S_L^x)$ | $-\frac{1}{2\pi}(D_\kappa^x A + D_\kappa A^x)\left(\frac{1}{\varepsilon}M - I\right) + \frac{1}{\varepsilon}\left(I - \frac{1}{2\pi}D_0 A\right)M^x$  
$- \frac{1}{2\pi\varepsilon}S_\kappa^x\left(N - L\right) - \frac{1}{2\pi\varepsilon}S_L^x\left(N^x - L^x\right)$ |
| SS(V)PE  | $-\frac{1}{2\pi}(D_\kappa^x A + D_\kappa A^x)S_0 + \left(I - \frac{1}{2\pi}D_\kappa A\right)S_0^x$  
$+ \frac{1}{\varepsilon}S_\kappa^x\left(I + \frac{1}{2\pi}AD_0^\dagger\right) + \frac{1}{2\pi\varepsilon}S_\kappa\left[A^xD_0^\dagger + A\left(D_0^\dagger\right)^x\right]$ | $-\frac{1}{2\pi}(D_\kappa^x A + D_\kappa A^x)\left(\frac{1}{\varepsilon}M - I\right) + \frac{1}{\varepsilon}\left(I - \frac{1}{2\pi}D_0 A\right)M^x$  
$- \frac{1}{2\pi\varepsilon}S_\kappa^x\left(N - L\right) - \frac{1}{2\pi\varepsilon}S_L^x\left(N^x - L^x\right)$ |

Table 8.2: Definitions of $K^x$ and $R^x$ for screened PCMs.
where \( s_{ij} = |\vec{s}_i - \vec{s}_j| \) and \( \zeta_{ij} = \zeta_i \zeta_j / (\zeta_i^2 + \zeta_j^2)^{1/2} \). The off-diagonal element \( D_{0,ij} \) is related to \( S_{0,ij} \) according to the relation \(^{159}\)

\[
D_{0,ij} = \vec{n}_j \cdot \frac{\partial S_{0,ij}}{\partial \vec{s}_j},
\]

(8.50)

where \( \vec{n}_j \) represents the outward pointing unit vector normal to the cavity surface at the point \( \vec{s}_j \). Using Eqs. (8.49) and (8.50), one obtains

\[
D_{0,ij} = \left( \text{erf}(\zeta_{ij} s_{ij}) - \frac{2 \zeta_{ij} s_{ij} e^{-\zeta_{ij}^2 s_{ij}^2}}{\sqrt{\pi} s_{ij}^3} \right) \frac{\vec{n}_j \cdot \vec{s}_{ij}}{s_{ij}^3}. \]

(8.51)

In the SWIG approach, diagonal matrix elements \( S_{0,ii} \) are obtained by taking the limit \( s_{ij} \to 0 \) in Eq. (8.49), and furthermore introducing a switching function in the denominator. The resulting definition is

\[
S_{0,ii} = \frac{\zeta_i \sqrt{2/\pi}}{F_i}. \]

(8.52)

The precise form of the function \( F_i \) can be found in Ref. 213, along with a proof that placing \( F_i \) in the denominator guarantees that the potential energy surface will be smooth, provided that the function \( F_i \) is smooth. The diagonal elements \( D_{0,ii} \) in SWIG are defined by

\[
D_{0,ii} = -\frac{S_{0,ii} F_i}{2 R_I},
\]

(8.53)

where \( R_I \) is the radius of the sphere that contains the point \( \vec{s}_i \).

Ten-no\(^{268,269}\) has discussed the evaluation of Gaussian integrals over Yukawa-type potentials, and his formulas can be used to evaluate the screened Coulomb interaction.
between $g_i(\vec{r})$ and $g_j(\vec{r})$. First, we introduce some definitions:

$$T_{ij} = \zeta_{ij} s_{ij} \quad (8.54a)$$
$$U_{ij} = \frac{\kappa}{2\zeta_{ij}} \quad (8.54b)$$
$$\alpha_{ij} = e^{-\kappa s_{ij}} \text{erfc}(U_{ij} - T_{ij}) \quad (8.54c)$$
$$\beta_{ij} = e^{\kappa s_{ij}} \text{erfc}(U_{ij} + T_{ij}) \quad (8.54d)$$
$$\alpha'_{ij} = -\kappa \alpha_{ij} + e^{-\kappa s_{ij}} \left( \frac{2\zeta_{ij}}{\sqrt{\pi}} e^{-(U_{ij} - T_{ij})^2} \right) \quad (8.54e)$$
$$\beta'_{ij} = \kappa \beta_{ij} - e^{\kappa s_{ij}} \left( \frac{2\zeta_{ij}}{\sqrt{\pi}} e^{-(U_{ij} + T_{ij})^2} \right) \quad (8.54f)$$

Note that $\text{erfc}(x) = 1 - \text{erf}(x)$. The notation in Eq. (8.54) facilitates a succinct expression for the off-diagonal matrix elements of $S_\kappa$:

$$S_{\kappa,ij} = \frac{e^{U_{ij}^2}}{2s_{ij}} (\alpha_{ij} - \beta_{ij}) \quad (8.55)$$

We can use a relation analogous to Eq. (8.50), but replacing $S_{0,ij}$ with $S_{\kappa,ij}$ and $D_{0,ij}$ with $D_{\kappa,ij}$, to obtain the off-diagonal elements of $D_\kappa$. The result is

$$D_{\kappa,ij} = \left[ \frac{e^{U_{ij}^2}}{2} \left( \frac{\alpha'_{ij} - \beta'_{ij}}{s_{ij}} \right) - \frac{S_{\kappa,ij}}{s_{ij}} \right] \vec{n}_j \cdot \vec{s}_{ij} s_{ij}^3 \quad (8.56)$$

Analogous to the salt-free case, diagonal elements of $S_\kappa$ are derived from $S_{\kappa,ij}$ by taking the limit $s_{ij} \to 0$ and introducing $F_i$ in the denominator:

$$S_{\kappa,ii} = F_i^{-1} \left[ \zeta_i \sqrt{2/\pi} - \kappa \text{erfc}(U_{ii}) \exp(U_{ii}^2) \right] \quad (8.57)$$

For $D_{\kappa,ii}$, we use a definition analogous to Eq. (8.53):

$$D_{\kappa,ii} = -\frac{S_{\kappa,ii} F_i}{2R_I} \quad (8.58)$$
Finally, we note that the diagonal elements of $R$ should not be scaled by $F_i$, because this would cancel the factors of $F_i$ contained in the $K$ matrix. Thus, we define an alternative version of $S_\kappa$, denoted $S'_\kappa$, where the switching function is absent. We can express this as

$$
S'_{\kappa,ij} = \begin{cases} 
S_{\kappa,ij} & \text{for } i \neq j \\
S_{\kappa,ii}F_i & \text{for } i = j.
\end{cases}
$$

(8.59)

The matrix $S'_\kappa$ is used in place of $S_\kappa$ in constructing the $R$ matrix (see Tables 8.1 and 8.2).

### 8.3.4 Electrostatic potentials and electric fields

To complete the implementation of the screened PCMs, we need to be able to evaluate the screened and unscreened electrostatic potentials and normal electric fields, at each discretization point $\vec{s}_i$. For solutes described using quantum mechanics, evaluation of the screened electrostatic potential, $v_\kappa$, requires evaluation of Gaussian integrals over a Yukawa-type potential. We have not implemented such integrals in general, owing to the complexity of the formulas involved.\textsuperscript{268,269} Therefore we confine our attention to solutes described at the MM level, for which $\rho(\vec{r})$ consists of a set of atom-centered point charges, $\{\rho_J\}$. In this case, only $s$-type Gaussian integrals are required. Analogous to Eq. (8.54), we make the following definitions related to the screened interaction between the point charge $\rho_J$ (located at position $\vec{r}_J$) and the $i$th
surface Gaussian, \( g_i(\vec{r}) \):

\[
T_{iJ} = \zeta_i r_{iJ} \quad (8.60a)
\]

\[
U_i = \frac{\kappa}{2\zeta_i} \quad (8.60b)
\]

\[
\alpha_{iJ} = e^{-\kappa r_{iJ}} \operatorname{erfc}(U_i - T_{iJ}) \quad (8.60c)
\]

\[
\beta_{iJ} = e^{\kappa r_{iJ}} \operatorname{erfc}(U_i + T_{iJ}) \quad (8.60d)
\]

\[
\alpha'_{iJ} = -\kappa \alpha_{iJ} + e^{-\kappa r_{iJ}} \left( \frac{2\zeta_i}{\sqrt{\pi}} e^{-(U_i - T_{iJ})^2} \right) \quad (8.60e)
\]

\[
\beta'_{iJ} = \kappa \beta_{iJ} - e^{\kappa r_{iJ}} \left( \frac{2\zeta_i}{\sqrt{\pi}} e^{-(U_i + T_{iJ})^2} \right) . \quad (8.60f)
\]

Here \( r_{iJ} = |\vec{s}_i - \vec{r}_J| \).

The unscreened electrostatic potential due to the point charges \( \{\rho_J\} \), evaluated at the point \( \vec{s}_i \), is

\[
v_{0,i} = \sum_J \rho_J \frac{\operatorname{erf}(T_{iJ})}{r_{iJ}} . \quad (8.61)
\]

The corresponding electric field at \( \vec{s}_i \) is

\[
\vec{E}_{0,i} = -\sum_J \rho_J \left( \operatorname{erf}(T_{iJ}) - \frac{2T_{iJ}}{\sqrt{\pi}} e^{-T_{iJ}^2} \right) \frac{\vec{r}_{iJ}}{r_{iJ}^3} \quad (8.62)
\]

and the normal component of this field is

\[
v_{0,i}^\perp = \vec{n}_i \cdot \vec{E}_{0,i} . \quad (8.63)
\]

We next consider how to evaluate the screened electrostatic potential, \( v_\kappa \), in practical calculations. Unlike \( S_{\kappa,ij} \) (the interaction between surface charges, subject to the screened Coulomb potential), the screened potential \( v_\kappa,i \) should account for the existence of an IEL. For the Debye-Hückel model system, the IEL factor \([\gamma, \text{in Eq. (8.6)}]\)
fulfills this role, but a simple scaling factor of this type does not exist for an arbitrary cavity shape and distribution of point charges. Therefore, we must make an approximation when computing $v_\kappa$.

Our approach is based upon two observations: that the solute cavity is constructed from a union of atom-centered spheres, and that the MM solute consists of point charges located at the centers of these spheres. For an arbitrary MM solute, a crude approximation to $\phi^\rho(\mathbf{s})$ could be constructed by supposing that the MM charge $\rho_J$ only makes a contribution to $\phi^\rho(\mathbf{s})$ for discretization points that lie on the $J$th atomic sphere. Under this assumption, $\phi^\rho(\mathbf{s}_i)$ can always be expressed using Eq. (8.4), for each discretization point $\mathbf{s}_i$, and in effect each atom gets its own IEL factor.

In reality, of course, all of the solute charges should contribute to $\phi^\rho$ at each discretization point, but then it is unclear how to account for the IEL. In the spirit of simplicity, we therefore use atomic IEL factors, yet we allow all solute charges to contribute to $\phi^\rho$ across the entire cavity surface. We refer to this as the “local IEL approximation”, and this approximation can be expressed mathematically as

$$\phi^\rho(\mathbf{s}_i) \approx \frac{e^{\kappa R_i}}{1 + \kappa R_i} \int \rho(\mathbf{r}) \frac{e^{-\kappa|\mathbf{s}_i - \mathbf{r}|}}{|\mathbf{s}_i - \mathbf{r}|} d^3\mathbf{r} \quad (8.64)$$

for all points $\mathbf{s}_i$ located on the $I$th atomic sphere.

When $\rho(\mathbf{r})$ consists of a set of point charges $\{\rho_J\}$, the local IEL approximation affords

$$v_{\kappa,i} = \frac{\gamma_i}{2} \frac{e^{U^2}}{U^2} \sum_J \rho_J \frac{\alpha_{iJ} - \beta_{iJ}}{r_{iJ}} \quad (8.65)$$
for the screened electrostatic potential at \( \vec{s}_i \). In this equation, we have defined a “local IEL factor” for the \( i \)th grid point [cf. Eq. (8.6)],

\[
\gamma_i = \frac{e^{\kappa R_I}}{1 + \kappa R_I}.
\]  

(8.66)

The screened electric field at \( \vec{s}_i \) can now be evaluated,

\[
\vec{E}_{\kappa,i} = -\gamma_i \frac{eU_i^2}{2} \sum_J \rho_J [(\alpha'_{iJ} - \beta'_{iJ}) r_{iJ} - (\alpha_{iJ} - \beta_{iJ})] \vec{r}_{iJ}^{\perp} \cdot \vec{r}_{iJ},
\]  

(8.67)

and its normal component is

\[
v^+_{\kappa,i} = \vec{n}_i \cdot \vec{E}_{\kappa,i}.
\]  

(8.68)

At this point, the accuracy of the local IEL approximation is unclear, but a number of physical arguments can be made in favor of it. In terms of its physical interpretation, the factor \( \gamma_i \) approximately removes the screening effects in the electrostatic potential, \( v_{\kappa,i} \), due to mobile solvent ions over a distance of \( R_I \). That is, mobile solvent ions are not allowed to penetrate between surface point \( \vec{s}_i \) and the source charge at \( \vec{r}_J \) for a portion \( R_I \) of the total distance \( r_{iJ} \). Obviously, this is exact for a single point charge centered in a spherical cavity (i.e., the Debye-Hückel model system), and it makes some sense that the nearest and most significant contribution to \( \phi^\rho_{\kappa}(\vec{s}_i) \) would come from the point charge \( \rho_I \) at the center of the sphere on which the point \( \vec{s}_i \) is located, for which \( \gamma_i \) is accurate given the above physical interpretation. Furthermore, the expression for \( \phi^\rho_{\kappa}(\vec{s}_i) \) in Eq. (8.64), while approximate for \( \kappa > 0 \), has the correct limit [namely, \( \phi^0_{\kappa}(\vec{s}_i) \)] as \( \kappa \to 0 \). Moreover, if \( \kappa R_I \gg 1 \) for all atomic spheres, then the integrand in Eq. (8.64) will have significant magnitude only on sphere \( I \), owing to the factor of \( \exp(-\kappa|\vec{s}_i - \vec{r}|) \) in the integrand.
On the other hand, if $\kappa R_i \ll 1$ then $\gamma_i \approx 1$. Since the use of the LPBE is most appropriate when $\kappa$ is small, one might question whether these local IEL factors make any difference at all. Numerical tests in Section 8.4 will demonstrate not only that the IEL approximation is satisfactory (at least for MM solutes), but also that setting $\gamma_i = 1$ for each $i$ affords results that are much less satisfactory.

### 8.3.5 DESMO analytic gradient

In Section 8.3.2 we presented generic expressions for the PCM gradient, in terms of the matrix derivatives $K^x$ and $R^x$. At present, we have only implemented these expressions in the DESMO case, so in the interest of brevity we will not present detailed formulas for the the screened IEF-PCM/SS(V)PE analytic gradients. [It should be clear from Table 8.2 that the IEF-PCM/SS(V)PE gradient is considerably more complicated than is the DESMO gradient.] Detailed formulas for the DESMO gradient are presented below.

According to Eq. (8.47) and Table 8.2, the necessary components for the DESMO gradient are $S_0^x$ and $M^x$ as well as the electrostatic potential gradients $v_0^x$ and $v_\kappa^x$. We will use the symbol $\hat{\nabla}_J$ to denote the derivative with respect to displacement of atom $J$. Then $\hat{\nabla}_J v_{0,i}$ is essentially just the negative of the electric field in Eq. (8.62):

$$\hat{\nabla}_J v_{0,i} = \rho_J \left( \text{erf}(T_{i,J}) - \frac{2T_{i,J}}{\sqrt{\pi}} e^{-T_{i,J}^2} \right) \frac{\vec{r}_{i,J}}{v_0^x} \theta_{i,J} .$$

(8.69)

Here, $\vec{r}_{i,J} = \vec{r}_i - \vec{r}_J$ and

$$\theta_{i,J} = \begin{cases} 0 & \text{if } i \in J \\ 1 & \text{if } i \notin J \end{cases} .$$

(8.70)
Similarly, from the electric field in Eq. (8.67) we obtain

$$\hat{\nabla}_J v_{\kappa,i} = \gamma_i \frac{e^{U_i^2}}{2} \rho_J \left[ (\alpha_{i,J} - \beta_{i,J}^\prime) r_{i,J} - (\alpha_{i,J} - \beta_{i,J}) \right] \vec{r}_{i,J} \theta_{i,J}. \quad (8.71)$$

The gradient $S_{0,i}^x$, for off-diagonal elements $i \neq j$ is

$$\hat{\nabla}_J S_{ij} = - \left( \text{erf}(T_{ij}) - \frac{2 T_{ij}^2}{\sqrt{\pi}} e^{-T_{ij}^2} \right) \frac{s_{ij}}{s_{ij}^x} (\theta_{i,J} - \theta_{j,J}), \quad (8.72)$$

and for diagonal elements it is

$$\hat{\nabla}_J S_{ii} = - \frac{S_{ii}}{F_i} \hat{\nabla}_J F_i. \quad (8.73)$$

The gradient of the switching function, $\hat{\nabla}_J F_i$, has been reported previously. Notice also that the area derivatives, $A^x$ (see Table 8.2) depend only on the switching function gradient, according to Eq. (8.40). Finally, we do not compute $M^x$ in practice but instead compute $M^x v_0$, owing to the simple formula

$$\left( \hat{\nabla}_J M_{ii} \right) v_{0,i} = \hat{\nabla}_J v_{\kappa_i} - M_{ii} \hat{\nabla}_J v_{0,i}^x. \quad (8.74)$$

The symmetrization of $Q$ as well as the gradient for DESMO can be simplified by defining several sets of charges. First, we define a set of conductor charges

$$q_\infty = -S_0^{-1} v_0 \quad (8.75)$$

that constitute a discretized solution to Eq. (8.22). We also define

$$q' = S_0^{-1} \left( \frac{1}{\varepsilon} M - I \right) v_0 \quad (8.76)$$

and

$$q'' = \left( I - \frac{1}{\varepsilon} M \right) q_\infty. \quad (8.77)$$
Using surface charges $q = (q' + q'')/2$ is then equivalent to solving Eq. (8.44) with a symmetrized $Q$ matrix. In practice, one can obtain the charges $q_\infty$, $q'$, and $q''$ either by explicitly constructing $S_1^{-1}$, or else by solving Eqs. (8.75) and (8.76) separately, using some sort of iterative algorithm. Once $q_\infty$ is determined, it is trivial to compute $q''$.

Given the above charges, one can simplify $E_{\text{pol}}^x$ for the DESMO case to afford

$$E_{\text{pol}}^x = \frac{1}{2} \left( q_\infty S_0^x q' - \frac{1}{\varepsilon} q_\infty M^x v_0^x \right) + q^\dagger v_0^x$$  \hspace{1cm} (8.78)

For MM solutes, there is no reason to symmetrize $Q$, and the charges $q'$ from Eq. (8.76) can be used in place of $q$ to compute both $E_{\text{pol}}$ and $E_{\text{pol}}^x$. This somewhat reduces the computational effort. Similarly, if only the energy (but not the gradient) is required for a DESMO calculation, then it suffices to compute only $q''$, since $E_{\text{pol}} = v_0^\dagger q/2 = v_0^\dagger q''/2$.

### 8.4 Numerical Tests

We have implemented the DESMO and screened IEF-PCM/SS(V)PE models, for MM solutes only, within a locally-modified version of the Q-Chem software package.\(^53\) This implementation builds upon our recent implementation of smooth PCMs\(^ {186,213,221}\) and hybrid quantum mechanics/molecular mechanics methods,\(^ {222}\) within this same software. In this section, we present numerical tests of these screened PCMs, both for some simple test cases where an exact analytical solution of the LPBE is available, but also for more realistic MM solutes, where three-dimensional numerical solution of the LPBE provides a benchmark.
Table 8.3: Ionic strengths (in mol/L) for each pair of parameters ε and λ explored here, assuming $T = 298$ K. Note that physiological ionic strengths are $\sim 0.1-0.2$ mol/L.\textsuperscript{9}

<table>
<thead>
<tr>
<th>ε</th>
<th>λ / Å</th>
<th>3</th>
<th>5</th>
<th>25</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>0.0524</td>
<td>0.0189</td>
<td>7.54 x 10^-4</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>0.261</td>
<td>0.0943</td>
<td>3.77 x 10^-3</td>
<td></td>
</tr>
<tr>
<td>80</td>
<td>1.05</td>
<td>0.377</td>
<td>1.51 x 10^-2</td>
<td></td>
</tr>
</tbody>
</table>

For these tests, we employ three different values of the dielectric constant ($\varepsilon = 4$, 20, and 80) and four different Debye lengths ($\lambda / \text{Å} = 3$, 5, 25, and $\infty$). The dielectric constants are chosen to represent three common categories of solvents in chemical applications: $\varepsilon = 4$ for non-polar organic solvents, (benzene, alkanes, ethers), $\varepsilon = 20$ for polar solvents (alcohols, ketones), and $\varepsilon = 80$ for water. The three Debye lengths are chosen to sample a variety of ionic strengths and to test the local IEL approximation. If the ratio $R_I/\lambda$ is small, for all atomic spheres, then we expect the local IEL approximation to be satisfactory, but if $R_I/\lambda \sim 1$, as it is when $\lambda = 3$ Å, then one might expect the local IEL approximation to break down. Note that when $\lambda = \infty$, DESMO reduces to GCOSMO/C-PCM and screened IEF-PCM/SS(V)PE reduces to its unscreened form. In Table 8.3, we provide the molar ionic strengths that correspond to each pair of $\varepsilon$ and $\lambda$ values used in these tests, at $T = 298$ K. For $\lambda = \infty$, the ionic strength is zero.
For MM solutes, the screened IEF-PCM and screened SS(V)PE methods are equivalent, at the level of integral equations, and we choose to discretize Eq. (8.35) for IEF-PCM/SS(V)PE. We also test a variant of DESMO that we call DESMO-0, which is identical except that we set $\gamma_i = 1$ in Eq. (8.66), for each $i$. In effect, DESMO-0 ignores the IEL and allows mobile solvent ions to penetrate the interior of the solute cavity. Setting $\gamma_i = 1$ in IEF-PCM/SS(V)PE calculations should not have a significant effect, because $v_\kappa$ is absent in this method. More precisely, although $\phi_\kappa^\rho$ appears (formally) in Eqs. (8.35) and (8.36), the quantity $\varphi_\kappa = 0$ for MM solutes, hence the terms in these equations that involve $\phi_\kappa^\rho$ cancel, by virtue of the identity in Eq. (8.37). This explains why, in previous work on screened IEF-PCM and screened SS(V)PE, relatively little attention has been paid to the treatment of the IEL. We will see that this issue is more important in the context of DESMO.

For all of the examples presented here, the solute cavity is constructed using SWIG discretization with 110 Lebedev grid points per atomic sphere. This is sufficient to converge GCOSMO/C-PCM solvation energies to within $\sim 0.1$ kcal/mol, relative to results obtained using much denser grids.\textsuperscript{213}

\subsection{Comparison to the Kirkwood model}

An analytical solution to the LPBE, for an arbitrary electrostatic multipole centered in a spherical solute cavity, was derived long ago by Kirkwood,\textsuperscript{13} and this model can be regarded as a generalization of the Debye-Hückel model problem presented in Section 8.2.1. Here, we compare exact results for the Kirkwood model to numerical results for various PCM.
In our first example, we take $\rho(\vec{r})$ to consist of a single point charge centered in a spherical cavity of radius $a$, equivalent to the model in Section 8.2.1 with $c = 0$. Numerical results for this model are shown in Tables 8.4 and 8.5 for an ion charge of $+e$ and a cavity radius $a = 2$ Å. The DESMO and IEF-PCM/SS(V)PE methods are essentially exact for all $\varepsilon$ and $\lambda$, but the same cannot be said for DESMO-0. Although the latter model is accurate $\lambda = \infty$ (the salt-free limit), its accuracy deteriorates as $\lambda$ decreases, most notably when $\varepsilon$ is low. Already, this indicates the importance of accounting for the IEL within DESMO. In fact, the same trends observed for this simplest of models will be seen in many of our subsequent tests.

<table>
<thead>
<tr>
<th>$\varepsilon$</th>
<th>$\lambda$ (Å)</th>
<th>$E_{\text{pol}}$ (kcal/mol)</th>
<th>error (kcal/mol)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>DESMO</td>
</tr>
<tr>
<td>4</td>
<td>$\infty$</td>
<td>-62.262</td>
<td>0.000</td>
</tr>
<tr>
<td>20</td>
<td>$\infty$</td>
<td>-78.865</td>
<td>0.000</td>
</tr>
<tr>
<td>80</td>
<td>$\infty$</td>
<td>-81.978</td>
<td>0.000</td>
</tr>
<tr>
<td>4</td>
<td>25</td>
<td>-63.799</td>
<td>0.000</td>
</tr>
<tr>
<td>20</td>
<td>25</td>
<td>-79.173</td>
<td>0.000</td>
</tr>
<tr>
<td>80</td>
<td>25</td>
<td>-82.055</td>
<td>0.000</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>-68.192</td>
<td>0.003</td>
</tr>
<tr>
<td>20</td>
<td>5</td>
<td>-80.051</td>
<td>0.001</td>
</tr>
<tr>
<td>80</td>
<td>5</td>
<td>-82.275</td>
<td>0.000</td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>-76.098</td>
<td>0.007</td>
</tr>
<tr>
<td>20</td>
<td>3</td>
<td>-81.632</td>
<td>0.001</td>
</tr>
<tr>
<td>80</td>
<td>3</td>
<td>-82.670</td>
<td>0.000</td>
</tr>
</tbody>
</table>

Table 8.4: Reaction field energies, $E_{\text{pol}}$, for a $+e$ point charge centered in a cavity of radius 2 Å, along with errors in various PCMs.
We next consider a simple model of a highly polar solute. The model consists of two point charges, $+2e$ and $-2e$, separated by a distance of 1 Å, resulting in a dipole moment of $\mu = 9.6$ debye. We place this system inside of a spherical cavity of radius 3.0 Å. The exact energy of this system can be calculated via a multipole expansion of the charge distribution, using Kirkwood’s exact formulas for each multipole order.\textsuperscript{13} For this example, an expansion up to $\ell = 7$ is sufficient to converge the energy to within $10^{-6}$ kcal/mol. Exact results, and errors for various PCMs, are listed in Table 8.6.

As in the previous example, the screened IEF-PCM/SS(V)PE method is extremely accurate for all values of $\varepsilon$ and $\lambda$. (The $\approx 0.02$ kcal/mol errors that are listed in Table 8.6 are eliminated if we use 590 Lebedev points to discretize each atomic sphere.)

<table>
<thead>
<tr>
<th>$\varepsilon$ (Å)</th>
<th>$\lambda$</th>
<th>Kirkwood (exact)</th>
<th>DESMO</th>
<th>DESMO-0</th>
<th>IEF-PCM/SS(V)PE</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>25</td>
<td>-1.54</td>
<td>-1.54</td>
<td>-1.60</td>
<td>-1.54</td>
</tr>
<tr>
<td>20</td>
<td>25</td>
<td>-0.31</td>
<td>-0.31</td>
<td>-0.32</td>
<td>-0.31</td>
</tr>
<tr>
<td>80</td>
<td>25</td>
<td>-0.08</td>
<td>-0.08</td>
<td>-0.08</td>
<td>-0.08</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>-5.93</td>
<td>-5.93</td>
<td>-6.84</td>
<td>-5.93</td>
</tr>
<tr>
<td>20</td>
<td>5</td>
<td>-1.19</td>
<td>-1.19</td>
<td>-1.37</td>
<td>-1.19</td>
</tr>
<tr>
<td>80</td>
<td>5</td>
<td>-0.30</td>
<td>-0.30</td>
<td>-0.34</td>
<td>-0.30</td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>-8.30</td>
<td>-8.30</td>
<td>-10.09</td>
<td>-8.30</td>
</tr>
<tr>
<td>20</td>
<td>3</td>
<td>-1.66</td>
<td>-1.66</td>
<td>-2.02</td>
<td>-1.66</td>
</tr>
<tr>
<td>80</td>
<td>3</td>
<td>-0.42</td>
<td>-0.42</td>
<td>-0.51</td>
<td>-0.42</td>
</tr>
</tbody>
</table>

Table 8.5: Salt shifts, $E_{pol}(\kappa) - E_{pol}(\kappa = 0)$, for a $+e$ point charge centered in a cavity of radius 2 Å.
Table 8.6: Reaction field energies, $E_{\text{pol}}$, for a Kirkwood model of a dipolar solute ($\mu = 9.6$ debye inside a cavity of radius 3 Å), along with errors in various PCMs.

<table>
<thead>
<tr>
<th>$\varepsilon$</th>
<th>$\lambda$ (Å)</th>
<th>$E_{\text{pol}}$ (kcal/mol)</th>
<th>DESMO</th>
<th>DESMO-0</th>
<th>IEF-PCM/SS(V)PE</th>
<th>error (kcal/mol)</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>$\infty$</td>
<td>-16.41</td>
<td>-2.03</td>
<td>-2.03</td>
<td>0.02</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>$\infty$</td>
<td>-22.81</td>
<td>-0.55</td>
<td>-0.55</td>
<td>0.02</td>
<td></td>
</tr>
<tr>
<td>80</td>
<td>$\infty$</td>
<td>-24.16</td>
<td>-0.12</td>
<td>-0.12</td>
<td>0.03</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>25</td>
<td>-16.46</td>
<td>-1.98</td>
<td>-2.03</td>
<td>0.02</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>25</td>
<td>-22.83</td>
<td>-0.53</td>
<td>-0.54</td>
<td>0.02</td>
<td></td>
</tr>
<tr>
<td>80</td>
<td>25</td>
<td>-24.16</td>
<td>-0.12</td>
<td>-0.12</td>
<td>0.03</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>-17.16</td>
<td>-1.28</td>
<td>-2.03</td>
<td>0.02</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>5</td>
<td>-22.99</td>
<td>-0.37</td>
<td>-0.52</td>
<td>0.03</td>
<td></td>
</tr>
<tr>
<td>80</td>
<td>5</td>
<td>-24.20</td>
<td>-0.08</td>
<td>-0.12</td>
<td>0.03</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>-17.90</td>
<td>-0.52</td>
<td>-2.15</td>
<td>0.02</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>3</td>
<td>-23.17</td>
<td>-0.19</td>
<td>-0.51</td>
<td>0.03</td>
<td></td>
</tr>
<tr>
<td>80</td>
<td>3</td>
<td>-24.25</td>
<td>-0.03</td>
<td>-0.11</td>
<td>0.03</td>
<td></td>
</tr>
</tbody>
</table>
Table 8.7: Salt shifts, $E_{\text{pol}}(\kappa) - E_{\text{pol}}(\kappa = 0)$, for a Kirkwood model of a zwitterion ($\mu = 9.6$ debye inside a cavity of radius 3 Å).

<table>
<thead>
<tr>
<th>$\varepsilon$</th>
<th>$\lambda$ (Å)</th>
<th>Kirkwood (exact)</th>
<th>DESMO</th>
<th>DESMO-0</th>
<th>IEF-PCM/SS(V)PE</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>25</td>
<td>-0.05</td>
<td>0.00</td>
<td>-0.04</td>
<td>-0.05</td>
</tr>
<tr>
<td>20</td>
<td>25</td>
<td>-0.01</td>
<td>0.00</td>
<td>-0.01</td>
<td>-0.01</td>
</tr>
<tr>
<td>80</td>
<td>25</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>-0.70</td>
<td>0.01</td>
<td>-0.70</td>
<td>-0.69</td>
</tr>
<tr>
<td>20</td>
<td>5</td>
<td>-0.18</td>
<td>0.00</td>
<td>-0.15</td>
<td>-0.18</td>
</tr>
<tr>
<td>80</td>
<td>5</td>
<td>-0.05</td>
<td>0.00</td>
<td>-0.04</td>
<td>-0.05</td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>-1.49</td>
<td>0.03</td>
<td>-1.61</td>
<td>-1.49</td>
</tr>
<tr>
<td>20</td>
<td>3</td>
<td>-0.35</td>
<td>0.01</td>
<td>-0.32</td>
<td>-0.35</td>
</tr>
<tr>
<td>80</td>
<td>3</td>
<td>-0.09</td>
<td>0.00</td>
<td>-0.08</td>
<td>-0.09</td>
</tr>
</tbody>
</table>

DESMO is less accurate than in the previous test, although the largest errors are for $\varepsilon = 4$, and for $\varepsilon = 80$ the errors are $\leq 0.12$ kcal/mol. However, DESMO does not afford an accurate trend for the “salt shift”, i.e., the difference $E_{\text{pol}}(\kappa) - E_{\text{pol}}(\kappa = 0)$ at fixed $\varepsilon$ (see Table 8.7). In fact, DESMO fails to predict a significant change in $E_{\text{pol}}$ as the ionic strength changes. In contrast, DESMO-0 affords a correct trend in the salt shift, despite the fact that the errors in $E_{\text{pol}}$, relative to exact Kirkwood results, are larger for this method than they are for DESMO.

The failure of DESMO to predict the salt shift in this example indicates that the local IEL approximation may be deficient in cases where solute charges are not located at the centers of the spheres that are used to construct the cavity surface. On the other hand, the salt shifts in this example are quite small ($< 0.4$ kcal/mol in
polar solvents) and are comparable to the errors in the solvation energies themselves. Absolute errors in $E_{pol}$ are no larger than 0.6 kcal/mol for $\varepsilon = 20$, and the errors in $E_{pol}$ are even smaller for test problems with smaller dipole moments. These results seem to indicate that DESMO (and GCOSMO, for $\kappa = 0$) is slightly less accurate than IEF-PCM/SS(V)PE when the atomic charge distribution is not spherically symmetric. As expected, however, the difference between these methods vanishes as $\varepsilon$ increases.

### 8.4.2 Comparison to the model of Lotan and Head-Gordon

The second analytical model to which we compare the screened PCMs is one derived by Lotan and Head-Gordon (LHG). The LHG model generalizes the Kirkwood model and provides an analytic solution to the LPBE for any number of non-overlapping spherical cavities, each of which contains an arbitrary charge density so long as there is no escaped charge. The analytical solution for the solvation energy is lengthy to explicate, and we refer the reader to Ref. 10 for details.

Here, we consider a set of disjoint spherical cavities whose locations are chosen at random within a 30 Å radius of the coordinate origin. The radius of each sphere is selected at random within the range 1.5–2.5 Å, and at the center of each sphere we place a single point charge, selected at random from between $-e$ and $+e$. (All of the parameters necessary to define this test system are provided in the Supplementary Material.) We calculate the electrostatic energy of this system using the formulas in Ref. 10 (with $\varepsilon = 1$ inside of each sphere), and compare electrostatic solvation energies (Table 8.8) and salt shifts (Table 8.9) to results computed using various PCMs.
<table>
<thead>
<tr>
<th>ε</th>
<th>λ</th>
<th>$E_{pol}$ (kcal/mol)</th>
<th>error (kcal/mol)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>DESMO</td>
</tr>
<tr>
<td>4</td>
<td>∞</td>
<td>-2899.49</td>
<td>-0.01</td>
</tr>
<tr>
<td>20</td>
<td>∞</td>
<td>-3672.69</td>
<td>0.00</td>
</tr>
<tr>
<td>80</td>
<td>∞</td>
<td>-3817.67</td>
<td>0.00</td>
</tr>
<tr>
<td>4</td>
<td>25</td>
<td>-2914.09</td>
<td>-0.26</td>
</tr>
<tr>
<td>20</td>
<td>25</td>
<td>-3675.61</td>
<td>-0.05</td>
</tr>
<tr>
<td>80</td>
<td>25</td>
<td>-3818.40</td>
<td>-0.01</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>-3035.50</td>
<td>-0.41</td>
</tr>
<tr>
<td>20</td>
<td>5</td>
<td>-3699.90</td>
<td>-0.08</td>
</tr>
<tr>
<td>80</td>
<td>5</td>
<td>-3824.47</td>
<td>-0.02</td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>-3122.65</td>
<td>-0.22</td>
</tr>
<tr>
<td>20</td>
<td>3</td>
<td>-3717.32</td>
<td>-0.05</td>
</tr>
<tr>
<td>80</td>
<td>3</td>
<td>-3828.83</td>
<td>-0.01</td>
</tr>
</tbody>
</table>

Table 8.8: Reaction field energies, $E_{pol}$, obtained from the LHG model,\(^{10}\) for a set of twenty randomly-positioned charges in spherical cavities (see the text for details). PCM results are given in terms of the error relative to the LHG result.
<table>
<thead>
<tr>
<th>$\varepsilon$</th>
<th>$\lambda$ (Å)</th>
<th>salt shift (kcal/mol)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>LHG</td>
<td>DESMO</td>
</tr>
<tr>
<td>20</td>
<td>25</td>
<td>-2.92</td>
</tr>
<tr>
<td>80</td>
<td>25</td>
<td>-0.73</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>-136.01</td>
</tr>
<tr>
<td>20</td>
<td>5</td>
<td>-27.21</td>
</tr>
<tr>
<td>80</td>
<td>5</td>
<td>-6.80</td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>-223.16</td>
</tr>
<tr>
<td>20</td>
<td>3</td>
<td>-44.63</td>
</tr>
<tr>
<td>80</td>
<td>3</td>
<td>-11.16</td>
</tr>
</tbody>
</table>

Table 8.9: Salt shifts, $E_{pol}(\kappa) - E_{pol}(\kappa = 0)$, obtained from the LHG model and from various PCMs, for a set of twenty randomly-positioned charges in spherical cavities (see the text for details).
Somewhat surprisingly, we find differences of up to 1.0 kcal/mol between IEF-PCM/SS(V)PE solvation energies and analytic LHG results. (Recall that in the case of the Kirkwood model problems, the IEF-PCM solvation energy differed by no more than 0.03 kcal/mol from the analytical result, and this difference provides a measure of the discretization error.) For the multi-sphere LHG problem, this disagreement persists even when very dense integration grids (up to 5294 points per sphere) are employed. For $\kappa = 0$, it is possible to turn off both the switching function and the Gaussian blurring in IEF-PCM calculations, but neither of these changes eliminates the discrepancies between IEF-PCM and LHG results. Transposition or symmetrization of the $K$ matrix in Table 8.1, as is done in some implementations of IEF-PCM, also fails to eliminate the discrepancies with respect to LHG results.

These results indicate that it is the IEF-PCM/SS(V)PE ansatz for the reaction-field potential, Eq. (8.20), rather than any numerical issue, that is responsible for the diminished accuracy in cases where the cavity consists of disjoint spheres. Chipman has noted that the IEF-PCM/SS(V)PE reaction field is incorrect outside of the cavity, if some of the solute charge lies outside of the cavity. It follows that when the solute cavity is disjoint, the mutual polarization of the various surface charge distributions may not be described correctly.

To explore this issue further, we consider a system composed of two spheres, one with a radius of 1.29 Å and a point charge of $-2e$ located at its center, and another with a radius of 1.70 Å and a point charge of $+1.5e$ located at its center. We take
Figure 8.2: Difference between the PCM solvation energy and that predicted by the analytic LHG formulas, for two point charges centered in spherical cavities with $\varepsilon = 4$ and $\kappa = 0$. The energy difference is plotted as a function of the center-to-center distance between the two spheres, the sum of whose radii is 2.99 Å. (Note that both the horizontal and vertical scales are logarithmic.)

For this calculation, and discretize the two spheres using 194 Lebedev points per sphere, with the SWIG switching function turned off. Figure 8.2 plots the difference between the IEF-PCM solvation energy and the exact LHG result, as a function of the center-to-center distance between the two spheres. The difference between the two solvation energies disappears quite slowly as a function of distance, and exceeds 0.01 kcal/mol even at 10 Å separation. At the same time, this disagreement is typically two or more orders of magnitude smaller than the total reaction-field energy, and from a pragmatic chemical standpoint this difference is fairly negligible.

As compared to IEF-PCM/SS(V)PE results, DESMO solvation energies are in much better agreement with analytical LHG results (see Table 8.8), and differ by $< 0.5$ kcal/mol for all values of $\varepsilon$ and $\lambda$. When $\kappa = 0$, DESMO is equivalent
to GCOSMO, and this model reproduces the LHG results (for $\kappa = 0$) to within 0.01 kcal/mol for all values of $\varepsilon$. Moreover, for the two-sphere example in Fig. 8.2, the GCOSMO/C-PCM result is virtually identical to the LHG result at all ion–ion distances. Only when the two spheres are nearly in contact with one another does the difference exceed 0.01 kcal/mol.

The accuracy of the DESMO-0 model, in contrast, deteriorates as $\lambda$ decreases, and for $\varepsilon = 4$ the errors are quite large (see Table 8.8). This underscores the need to account for the IEL in DESMO calculations, but also indicates that the local IEL approximation appears to be accurate even for cavities composed of disjoint spheres.

### 8.4.3 Comparison to results from the adaptive Poisson-Boltzmann solver

We next consider more realistic MM solutes and cavity shapes, for which no analytic solution is available. PCM results here will be compared to results obtained by three-dimensional numerical solution of the LPBE, using the so-called adaptive Poisson-Boltzmann solver (APBS),\textsuperscript{226} a numerical finite-difference code that performs three-dimensional volumetric integration of the LPBE. Two different solutes are considered: alanine dipeptide and thymine dinucleotide. Atomic charges and radii (i.e., Lennard-Jones $\sigma$ parameters) are taken from the AMBER99 force field.\textsuperscript{56} For the dipeptide example, we construct the cavity using unmodified Lennard-Jones radii, but for the dinucleotide we add a solvent probe radius of 1.4 Å to each Lennard-Jones radius, so that the cavity corresponds to a “solvent accessible surface”.\textsuperscript{220} The geometries of the solutes were obtained from the TINKER program,\textsuperscript{58} and are provided in the
Table 8.10: Reaction field energies, $E_{pol}$, for alanine dipeptide, obtained using the APBS software. PCM results are given in terms of the error relative to the APBS result.

Supplementary Material,\textsuperscript{270} along with the atomic charges and radii.

The APBS calculation should provide an accurate result, provided that the integration grid is very dense, and we use a $193 \times 193 \times 193$ grid, with a grid spacing of 0.05 Å for the dipeptide and 0.10 Å for the dinucleotide. For the Kirkwood model problems in Section 8.4.1, these specifications reproduce the total energies to within 0.1 kcal/mol. However, we cannot expect perfect agreement between APBS and the PCM results, even if the latter were exact, because the discretization of the solute cavity is inherently different in these two approaches.
Table 8.11: Salt shifts, $E_{pol}(\kappa) - E_{pol}(\kappa = 0)$, for alanine dipeptide, obtained using the APBS software and various PCMs.

<table>
<thead>
<tr>
<th>$\varepsilon$</th>
<th>$\lambda$ (Å)</th>
<th>APBS</th>
<th>DESMO</th>
<th>DESMO-0</th>
<th>IEF-PCM/SS(V)PE</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>25</td>
<td>-0.18</td>
<td>-0.15</td>
<td>-0.21</td>
<td>-0.20</td>
</tr>
<tr>
<td>20</td>
<td>25</td>
<td>-0.04</td>
<td>-0.03</td>
<td>-0.04</td>
<td>-0.04</td>
</tr>
<tr>
<td>80</td>
<td>25</td>
<td>-0.01</td>
<td>-0.01</td>
<td>-0.01</td>
<td>-0.01</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>-2.61</td>
<td>-2.28</td>
<td>-3.30</td>
<td>-2.87</td>
</tr>
<tr>
<td>20</td>
<td>5</td>
<td>-0.54</td>
<td>-0.46</td>
<td>-0.66</td>
<td>-0.63</td>
</tr>
<tr>
<td>80</td>
<td>5</td>
<td>-0.14</td>
<td>-0.11</td>
<td>-0.17</td>
<td>-0.16</td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>-5.00</td>
<td>-4.38</td>
<td>-6.58</td>
<td>-5.42</td>
</tr>
<tr>
<td>20</td>
<td>3</td>
<td>-1.05</td>
<td>-0.88</td>
<td>-1.32</td>
<td>-1.18</td>
</tr>
<tr>
<td>80</td>
<td>3</td>
<td>-0.27</td>
<td>-0.22</td>
<td>-0.33</td>
<td>-0.30</td>
</tr>
</tbody>
</table>

The results of these tests are shown in Tables 8.10–8.13, and reflect the same trends seen in the previous tests. The agreement between APBS and the PCMs is remarkably good ($\sim 1$ kcal/mol or better) considering that they are significantly different approaches. For $\varepsilon = 4$, the DESMO results for $E_{pol}$ differ from APBS by 1–2 kcal/mol, but the differences are much smaller in the higher-dielectric environments. This indicates that the local IEL approximation fares well even for complicated cavity shapes and non-trivial solute charge distributions. Curiously, for $\varepsilon = 80$, the DESMO and DESMO-0 models are in excellent agreement here, which was not the case in previous tests.
Table 8.12: Reaction field energies, $E_{pol}$, for thymine dinucleotide, obtained using the APBS software. PCM results are given in terms of the error relative to the APBS result.

<table>
<thead>
<tr>
<th>$\varepsilon$</th>
<th>$\lambda$ (Å)</th>
<th>$E_{pol}$ (kcal/mol)</th>
<th>relative energy (kcal/mol)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>DESMO</td>
</tr>
<tr>
<td>4</td>
<td>$\infty$</td>
<td>-27.88</td>
<td>-0.91</td>
</tr>
<tr>
<td>20</td>
<td>$\infty$</td>
<td>-36.64</td>
<td>0.17</td>
</tr>
<tr>
<td>80</td>
<td>$\infty$</td>
<td>-38.41</td>
<td>0.51</td>
</tr>
<tr>
<td>4</td>
<td>25</td>
<td>-29.25</td>
<td>-0.95</td>
</tr>
<tr>
<td>20</td>
<td>25</td>
<td>-36.92</td>
<td>0.17</td>
</tr>
<tr>
<td>80</td>
<td>25</td>
<td>-38.48</td>
<td>0.51</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>-32.00</td>
<td>-1.11</td>
</tr>
<tr>
<td>20</td>
<td>5</td>
<td>-37.48</td>
<td>0.15</td>
</tr>
<tr>
<td>80</td>
<td>5</td>
<td>-38.62</td>
<td>0.50</td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>-33.19</td>
<td>-0.98</td>
</tr>
<tr>
<td>20</td>
<td>3</td>
<td>-37.74</td>
<td>0.20</td>
</tr>
<tr>
<td>80</td>
<td>3</td>
<td>-38.69</td>
<td>0.52</td>
</tr>
</tbody>
</table>

Table 8.12: Reaction field energies, $E_{pol}$, for thymine dinucleotide, obtained using the APBS software. PCM results are given in terms of the error relative to the APBS result.
Table 8.13: Salt shifts, $E_{\text{pol}}(\kappa) - E_{\text{pol}}(\kappa = 0)$, for thymine dinucleotide, obtained using the APBS software and various PCMs.

8.5 Summary

We have formulated a new boundary-element PCM for electrolytic solvents that we call the Debye-Hückel–like screening model, or DESMO. This model generalizes COSMO to solutions with non-zero ionic strength, within the Debye-Hückel (linearized Poisson-Boltzmann) limit. DESMO is closely related to “screened” versions of the PCMs known as IEF-PCM and SS(V)PE that have previously been developed for electrolytes.\textsuperscript{167,198,249} Like COSMO in the salt-free case, however, DESMO is significantly simpler to implement and much less expensive to evaluate, as compared to these alternative PCMs.

We have also presented the first fully analytic implementation of the screened
IEF-PCM/SS(V)PE model. This implementation builds upon our “SWIG” method for smooth discretization of the integral equations that appear in boundary-element PCMs, and our implementation of these screened PCMs is the first that is amenable to analytic gradient techniques. For single-cavity problems, the IEF-PCM/SS(V)PE approach affords the same solvation energy as does the LPBE, but for systems composed of disjoint cavities, we observe small discrepancies between the screened IEF-PCM/SS(V)PE solvation energy and that obtained by exact solution of the LPBE. We ascribe these differences to a fundamental limitation of the PCM reaction-field ansatz.

For solutes described by molecular mechanics point charges, we have shown that DESMO accurately approximates the exact solution of the LPBE, in cases where the exact solution is available, and accurately approximates other numerical solutions in cases where no analytic solution is available. For polar solvents (modeled here using $\varepsilon = 20$ and $\varepsilon = 80$), the total energies predicted by DESMO lie within $\sim 0.5$ kcal/mol of benchmark results, across a broad range of ionic strengths. To obtain this level of accuracy, an approximate treatment of the ion exclusion layer appears to be mandatory, and use of the “DESMO-0” model (which neglects the IEL, and was introduced here solely for comparative purposes) is not recommended. Further refinements to the treatment of the IEL, beyond the “local IEL approximation” that is introduced here, may improve the accuracy of DESMO.

Given the accuracy, simplicity, and computational efficiency of DESMO, we think that this model represents an important addition to the PCM toolbox.
CHAPTER 9

Conclusion

This work has set out to model a considerably complicated quantum chemistry problem, the excited electronic states of aqueous DNA. Because the system is too large to model in full with high accuracy quantum chemistry, we must make intelligent approximations in order to make a computationally feasible model. This has led us to develop a QM/MM/PCM multi-layer approach. We have shown that doing so, though, is highly non-trivial and requires special attention to the details of methodology being employed in order to avoid contamination with artifacts.

For the electronic excitation, we have shown that standard hybrid functional TDDFT fails in the condensed phase, incorrectly predicting many spurious CT states. A QM/MM approach partially alleviates this issue, but to fully rectify the intrinsic error, one must employ LRC-TDDFT, which allows us to explore the excited electronic states of DNA with much more confidence in the accuracy of the CT state energies.

Modeling solvation is another issue, though, and can become very cumbersome when expensive excited state quantum chemistry calculations are involved. PCMs provide a means to avoid this computational burden by assuming a “pre-averaged”
continuum solvent charge density (i.e., an implicit solvent). However, the details of PCM discretization is very important, and if one is not careful, an implementation of PCM can be numerically unstable and yield very poor results. The SWIG method mitigates these issues, allowing us to utilize PCMs in new ways.

We have combined all of the above into a robust multi-layer QM/MM/PCM methodology that will be available as a part of the Q-Chem v4.0 software package. This will enable other researchers in the scientific community to carry out similar types of calculations as done in this work as well as many other types.

A number of future extensions of this work are currently underway and will be reported in the near future. We summarize them here. We have parallelized the PCM code to enable rapid calculations of over $10^6$ MM atoms, involving $10^8$ surface grid points. We have also shown important theoretical connections between PCM models and the Generalized Born model for implicit solvation. Furthermore, we have implemented algorithms for more advanced molecular cavity surface definitions, and we have implemented the DESMO PCM for QM/MM by using linear combinations of $\text{erfc}(r)/r$ Gaussian integrals, which are readily available in Q-Chem, to approximate the necessary Yukawa-type integrals. The research described here in this dissertation along with these continuing efforts is helping to advance the capabilities of quantum chemistry in the condensed phase, providing us the tools to better understand the intricacies of complex chemical systems.
APPENDIX A

Supporting Information for: “Both intra- and interstrand charge-transfer excited states in aqueous B-DNA are present at energies comparable to, or just above, the \( ^1\pi\pi^* \) excitonic bright states” A.1

A.1 TD-DFT basis set dependence

We examine basis set dependence of the vertical excitation energies of \( \pi\pi^* \) and charge transfer (CT) states at the LRC-\( \omega \)PBE level in Table A.1. For A:T and A2, the \( \pi\pi^* \) excitation energies all tend to decrease, by roughly the same amount, as basis set size increases. As such, the order and energy spacing of the \( \pi\pi^* \) states remain nearly constant for all basis sets. The CT states of A2 show a similar trend as the \( \pi\pi^* \) states except for the 6-311(2+,2+)G** basis. The \( \pi\pi^* \) states of A2 do not exhibit much stabilization in moving from the 6-311+G* basis to the 6-311(2+,2+)G** basis, whereas the CT states are more sensitive to this move. On the other hand, the CT state of A:T appears to experience only a small decrease in energy as a result of increasing basis set.

A.1 This chapter appeared as the Supporting Information to the full article in the Journal of the American Chemical Society, in 2009, volume 131, page 3913.
<table>
<thead>
<tr>
<th>Excited state</th>
<th>6-31G*</th>
<th>6-311G*</th>
<th>6-31+G*</th>
<th>6-311+G*</th>
<th>6-311(2+,2+)G**</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>A:T base pair</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thy $\pi\pi^*$</td>
<td>5.47 (0.17)</td>
<td>5.41 (0.17)</td>
<td>5.33 (0.18)</td>
<td>5.30 (0.18)</td>
<td>5.28 (0.18)</td>
</tr>
<tr>
<td>Ade $\pi\pi^*$ (W)</td>
<td>5.67 (0.02)</td>
<td>5.64 (0.02)</td>
<td>5.60 (0.08)</td>
<td>5.57 (0.09)</td>
<td>5.56 (0.09)</td>
</tr>
<tr>
<td>Ade $\pi\pi^*$ (B)</td>
<td>5.91 (0.37)</td>
<td>5.85 (0.38)</td>
<td>5.74 (0.33)</td>
<td>5.71 (0.32)</td>
<td>5.70 (0.31)</td>
</tr>
<tr>
<td>Ade $\rightarrow$ Thy CT</td>
<td>6.50 (0.03)</td>
<td>6.45 (0.02)</td>
<td>6.46 (0.04)</td>
<td>6.44 (0.04)</td>
<td>6.43 (0.05)</td>
</tr>
<tr>
<td><strong>A$_2$ $\pi$-stacked dimer</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\pi\pi^*$ (W$^-$)</td>
<td>5.70 (0.00)</td>
<td>5.66 (0.01)</td>
<td>5.54 (0.03)</td>
<td>5.51 (0.05)</td>
<td>5.49 (0.03)</td>
</tr>
<tr>
<td>$\pi\pi^*$ (W$^+$)</td>
<td>5.72 (0.00)</td>
<td>5.68 (0.00)</td>
<td>5.63 (0.04)</td>
<td>5.61 (0.24)</td>
<td>5.59 (0.05)</td>
</tr>
<tr>
<td>$\pi\pi^*$ (B$^-$)</td>
<td>5.82 (0.04)</td>
<td>5.74 (0.04)</td>
<td>5.66 (0.01)</td>
<td>5.64 (0.20)</td>
<td>5.63 (0.01)</td>
</tr>
<tr>
<td>$\pi\pi^*$ (B$^+$)</td>
<td>6.00 (0.44)</td>
<td>5.93 (0.43)</td>
<td>5.79 (0.42)</td>
<td>5.77 (0.03)</td>
<td>5.75 (0.40)</td>
</tr>
<tr>
<td>3$'$-Ade $\rightarrow$ 5$'$-Ade CT</td>
<td>6.35 (0.01)</td>
<td>6.29 (0.01)</td>
<td>6.13 (0.01)</td>
<td>6.12 (0.00)</td>
<td>6.01 (0.03)</td>
</tr>
<tr>
<td>5$'$-Ade $\rightarrow$ 3$'$-Ade CT</td>
<td>6.55 (0.00)</td>
<td>6.48 (0.00)</td>
<td>6.37 (0.00)</td>
<td>6.35 (0.04)</td>
<td>6.16 (0.00)</td>
</tr>
<tr>
<td><strong>A$_2$ $\pi$-stacked dimer in a water cluster</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\pi\pi^*$ (W$^-$)</td>
<td>5.60 (0.05)</td>
<td>5.53 (0.07)</td>
<td>5.45 (0.07)</td>
<td>5.43 (0.07)</td>
<td></td>
</tr>
<tr>
<td>$\pi\pi^*$ (W$^+$)</td>
<td>5.63 (0.01)</td>
<td>5.58 (0.01)</td>
<td>5.55 (0.03)</td>
<td>5.53 (0.03)</td>
<td></td>
</tr>
<tr>
<td>$\pi\pi^*$ (B$^-$)</td>
<td>5.68 (0.01)</td>
<td>5.62 (0.01)</td>
<td>5.59 (0.00)</td>
<td>5.57 (0.00)</td>
<td></td>
</tr>
<tr>
<td>$\pi\pi^*$ (B$^+$)</td>
<td>5.85 (0.41)</td>
<td>5.77 (0.47)</td>
<td>5.67 (0.49)</td>
<td>5.65 (0.47)</td>
<td></td>
</tr>
<tr>
<td>3$'$-Ade $\rightarrow$ 5$'$-Ade CT</td>
<td>6.62 (0.00)</td>
<td>6.60 (0.02)</td>
<td>6.53 (0.06)</td>
<td>6.52 (0.07)</td>
<td></td>
</tr>
<tr>
<td>5$'$-Ade $\rightarrow$ 3$'$-Ade CT</td>
<td>5.89 (0.09)</td>
<td>5.85 (0.02)</td>
<td>5.83 (0.01)</td>
<td>5.81 (0.01)</td>
<td></td>
</tr>
</tbody>
</table>

Table A.1: Basis set dependence of vertical excitation energies (in eV) of simple nucleobase systems, calculated at the LRC-ωPBE level. Oscillator strengths are shown in parentheses.
Table A.1 also presents excitation energies for A₂ embedded in a cluster of DFT water molecules representing its first solvation shell. (This geometry is one particular snapshot from a molecular dynamics simulation.) The \( \pi\pi^* \) exciton states of A₂ all appear to experience a solvent red shift of \( \sim 0.1 \) eV, independent of basis set. For the CT states, the basis-set dependence of the solvent shift is only slightly more pronounced, amounting to a 0.3–0.4 eV blue shift for the 3'-adenine \( \rightarrow \) 5'-adenine CT, and a 0.4–0.6 eV red shift in the the 5'-adenine \( \rightarrow \) 3'-adenine CT state. (As noted in the manuscript, the configurationally-averaged solvent shift is toward lower CT energies in aqueous solution, but individual solvent configurations can exhibit blue shifts, as is seen for the one of the two CT states in the present example.)

Proceeding from the smallest to largest basis sets in Table A.1, the energy gap between the brightest \( \pi\pi^* \) state and the adenine \( \rightarrow \) thymine CT state in A:T changes by only 0.1 eV. The bright state/CT energy gap in A₂ shows a similar dependence on basis set, both in the gas phase and in the water cluster.

The 6-31G* basis set thus yields energy difference between excited states that are comparable to those obtained with much larger basis sets. Since our main focus is the relative ordering of the \( \pi\pi^* \) and CT excitation energies, in rather large systems with explicit solvent molecules, 6-31G* is our basis of choice for most of the calculations described in the manuscript.
<table>
<thead>
<tr>
<th>Molecule</th>
<th>Excitation</th>
<th>Type</th>
<th>Benchmark(^a)</th>
<th>LRC-(\omega)PBE(^b)</th>
<th>LRC-(\omega)PBEh(^c)</th>
</tr>
</thead>
<tbody>
<tr>
<td>dipeptide</td>
<td>(n_1 \rightarrow \pi_2^*)</td>
<td>CT</td>
<td>8.07</td>
<td>8.02</td>
<td>7.75</td>
</tr>
<tr>
<td></td>
<td>(\pi_1 \rightarrow \pi_2)</td>
<td>CT</td>
<td>7.18</td>
<td>7.03</td>
<td>6.93</td>
</tr>
<tr>
<td></td>
<td>(n_1 \rightarrow \pi_1^*)</td>
<td>L</td>
<td>5.62</td>
<td>5.57</td>
<td>5.65</td>
</tr>
<tr>
<td></td>
<td>(n_2 \rightarrow \pi_2^*)</td>
<td>L</td>
<td>5.79</td>
<td>5.83</td>
<td>5.91</td>
</tr>
<tr>
<td>(\beta)-dipeptide</td>
<td>(n_1 \rightarrow \pi_2^*)</td>
<td>CT</td>
<td>9.13</td>
<td>8.83</td>
<td>8.45</td>
</tr>
<tr>
<td></td>
<td>(\pi_1 \rightarrow \pi_2^*)</td>
<td>CT</td>
<td>7.99</td>
<td>8.33</td>
<td>8.01</td>
</tr>
<tr>
<td></td>
<td>(n_1 \rightarrow \pi_1^*)</td>
<td>L</td>
<td>5.40</td>
<td>5.55</td>
<td>5.64</td>
</tr>
<tr>
<td></td>
<td>(n_2 \rightarrow \pi_2^*)</td>
<td>L</td>
<td>5.10</td>
<td>5.69</td>
<td>5.77</td>
</tr>
<tr>
<td>tripeptide</td>
<td>(\pi_1 \rightarrow \pi_2^*)</td>
<td>CT</td>
<td>7.01</td>
<td>7.02</td>
<td>6.92</td>
</tr>
<tr>
<td></td>
<td>(\pi_2 \rightarrow \pi_3^*)</td>
<td>CT</td>
<td>7.39</td>
<td>7.23</td>
<td>7.15</td>
</tr>
<tr>
<td></td>
<td>(\pi_1 \rightarrow \pi_3^*)</td>
<td>CT</td>
<td>8.74</td>
<td>9.22</td>
<td>8.72</td>
</tr>
<tr>
<td></td>
<td>(n_1 \rightarrow \pi_3^*)</td>
<td>CT</td>
<td>9.30</td>
<td>9.31</td>
<td>8.88</td>
</tr>
<tr>
<td></td>
<td>(n_2 \rightarrow \pi_3^*)</td>
<td>CT</td>
<td>8.33</td>
<td>8.50</td>
<td>8.22</td>
</tr>
<tr>
<td></td>
<td>(n_1 \rightarrow \pi_2^*)</td>
<td>CT</td>
<td>8.12</td>
<td>7.97</td>
<td>7.69</td>
</tr>
<tr>
<td></td>
<td>(n_1 \rightarrow \pi_1^*)</td>
<td>L</td>
<td>5.74</td>
<td>5.58</td>
<td>5.67</td>
</tr>
<tr>
<td></td>
<td>(n_2 \rightarrow \pi_2^*)</td>
<td>L</td>
<td>5.61</td>
<td>5.84</td>
<td>5.91</td>
</tr>
<tr>
<td></td>
<td>(n_3 \rightarrow \pi_3^*)</td>
<td>L</td>
<td>5.91</td>
<td>5.92</td>
<td>6.00</td>
</tr>
<tr>
<td>N-phenyl</td>
<td>1 (^1)B_2</td>
<td>L</td>
<td>4.85</td>
<td>5.14</td>
<td>5.14</td>
</tr>
<tr>
<td>pyrrole</td>
<td>2 (^1)A_1</td>
<td>L</td>
<td>5.13</td>
<td>5.30</td>
<td>5.25</td>
</tr>
<tr>
<td></td>
<td>2 (^1)B_2</td>
<td>CT</td>
<td>5.47</td>
<td>5.55</td>
<td>5.36</td>
</tr>
<tr>
<td></td>
<td>3 (^1)A_1</td>
<td>CT</td>
<td>5.94</td>
<td>6.47</td>
<td>6.05</td>
</tr>
<tr>
<td>DMABN</td>
<td>(^1)B</td>
<td>L</td>
<td>4.25</td>
<td>4.75</td>
<td>4.75</td>
</tr>
<tr>
<td></td>
<td>(^1)A</td>
<td>CT</td>
<td>4.56</td>
<td>5.02</td>
<td>4.99</td>
</tr>
</tbody>
</table>

\(^a\)From Peach et al.\(^\text{271}\)
\(^b\)\(C_{HF} = 0, \omega = 0.3 \, a_0^{-1}\)
\(^c\)\(C_{HF} = 0.2, \omega = 0.2 \, a_0^{-1}\)

Table A.2: Vertical excitation energies (in eV) for a set of localized (L) and charge-transfer (CT) excitations, including root-mean-square errors (RMSEs) and mean signed errors (MSEs).
A.2 Benchmarking the LRC functionals

Rohrdanz et al.\textsuperscript{121} have recently analyzed the performance of various LRC density functionals for TD-DFT excitation energies. In particular, they examined a set of benchmark (CC2 and CASPT2) excitation energies\textsuperscript{271} that includes both localized ($n\pi^*$ and $\pi\pi^*$) and CT excitation energies. Table A.2 here is analogous to Table II of Rohrdanz et al., and presents excitation energies for this set of molecules, calculated at the TD-LRC-$\omega$PBE/cc-pVDZ and TD-LRC-$\omega$PBEh/cc-pVDZ levels. In the notation of Rohrdanz et al., the functional that we call LRC-$\omega$PBEh corresponds to the parameters $C_{HF} = 0$ (\textit{i.e.}, 20\% short-range Hartree–Fock exchange) and $\omega = 0.3 \, a_0^{-1}$; this is the recommended parameter set from that study. The functional that we call LRC-$\omega$PBE corresponds to $C_{HF} = 0$ and $\omega = 0.3 \, a_0^{-1}$, a parameter set that was not considered in detail by Rohrdanz et al. As such, for $C_{HF} = 0$ we present $\omega$-dependent excitation energy curves in Fig. A.1, for each of the molecules in Table A.2.

Both functionals afford root-mean-square errors (RMSEs) of 0.3 eV, for both localized and CT excitation energies. (For comparison, TD-PBE0 affords RMSEs of 0.3 eV and 3.0 eV, respectively, for the localized and CT excitation energies in this data set.\textsuperscript{121,271}) When the signs of the errors are taken into account, however, a more nuanced picture emerges.

For localized excitations, the two LRC functionals afford excitation energies that are in reasonable agreement with one another, and with the benchmarks. In most cases, the LRC functionals overestimate the excitation energies (corresponding to positive errors, in our sign convention), and the mean \textit{signed} error (MSE) for either
Figure A.1: TD-LRC-\(\omega\)PBE excitation energies (as a function of \(\omega\), with \(C_{\text{HF}} = 0\)) for the benchmark set of molecules in Table A.2. All calculations were performed with the cc-pVDZ basis set and SG-1 quadrature grid.
functional is $+0.2$ eV for the localized excitations.

For the CT states, the LRC-$\omega$PBE functional affords errors of either sign, with about equal frequency, and has a MSE of $+0.1$ eV (corresponding to overestimation, in the mean). The LRC-$\omega$PBEh functional, on the other hand, underestimates the CT excitation energies in 10/13 cases. Moreover, in two of the cases where LRC-$\omega$PBEh overestimates a CT excitation energy. The only case where LRC-$\omega$PBEh significantly overestimates a CT excitation energy is the $1A^1$ state of 4-($N,N$-dimethylamino)benzonitrile (DMABN), a state that actually exhibits very little CT character in the gas phase. The MSE for CT states in the case of LRC-$\omega$PBEh is $+0.2$ eV, i.e., the signed error is in the opposite direction as it is in the case of LRC-$\omega$PBE. (To one significant digit, the MSEs for CT states do not change if we throw out the questionable DMABN data point.)

As we have pointed out previously, CT excitation energies are especially sensitive to the value of $\omega$. Therefore we wish to examine the $\omega$-dependence of the CT excitation energies in a prototype DNA system, the $\pi$-stacked adenine dimer, $A_2$. Figures A.2 and A.3 plot $\omega$-dependent excitation energy profiles for the LRC-$\omega$PBE and LRC-$\omega$PBEh functionals, respectively. (In this context, we take LRC-$\omega$PBE to mean that $C_{HF} = 0$ and LRC-$\omega$PBEh to mean that $C_{HF} = 0.2$.) In the former case, reasonable agreement with CC2/TZVPP benchmarks, for both valence and CT states, is obtained using $\omega = 0.3 \ a_0^{-1}$. Smaller values of $\omega$, while they provide slightly better accuracy for the $n\pi^*$ and $\pi\pi^*$ states, but significantly underestimate CT excitation energies. Hence, $\omega = 0.3 \ a_0^{-1}$ is the value used for the LRC-$\omega$PBE
Figure A.2: Vertical excitation energies for the $^{1}\pi\pi^{*}$ and $^{1}\text{CT}$ states of $A_2$, as computed at the LRC-$\omega$PBE/6-311+G* level ($C_{\text{HF}} = 0$) as a function of the LRC range parameter, $\omega$. CC2/TZVPP results are shown as horizontal dotted lines.

Figure A.3: Vertical excitation energies for the $^{1}\pi\pi^{*}$ and $^{1}\text{CT}$ states of $A_2$, as computed at the LRC-$\omega$PBEh/6-311+G* level (with $C_{\text{HF}} = 0.2$) as a function of the LRC range parameter, $\omega$. CC2/TZVPP results are shown as horizontal dotted lines.
functional in this work. Note that the overestimation of valence excitation energies, which is statistically apparent from Table A.2, is also evident for \( A_2 \) in Fig. A.2, at or near our preferred value of \( \omega \). In the case of LRC-\( \omega \)PBEh the value \( \omega = 0.2 \, a_0^{-1} \) employed in this work (which was chosen for its good statistical performance in the study of Rohrdanz et al.\(^{121}\)) slightly underestimates the CT excitation energies while slightly overestimating the valence excitation energies. Again, this is in accord with the benchmark results in Table A.2. Together, the comparisons in this section justify the excitation energy shifts that we apply to calculated TD-LRC-\( \omega \)PBE and TD-LRC-\( \omega \)PBEh excitation energies.

In what follows (as in 4), the LRC-\( \omega \)PBE functional uses the parameters \( C_{HF} = 0 \) and \( \omega = 0.3 \, a_0^{-1} \), while the LRC-\( \omega \)PBEh functional employs the parameters \( C_{HF} = 0.2 \) and \( \omega = 0.2 \, a_0^{-1} \).

### A.3 Supplementary gas-phase data

Included below are the vertical excitation energies and oscillator strengths of some additional nucleobase assemblies. These include simple monomers and dimers (Table A.3, which is analogous to Table 4.1 in the manuscript but also provides the oscillator strengths); the single-stranded ATATA system (Table A.4), and the double-stranded ATA:TAT system (Table A.5).

Figure A.4 presents gas-phase stick spectra that show how the \( \pi \pi^* \) states and the CT states in \( A_n, T_n, \) and \( A_n:T_n \) evolve from \( n = 1 \) to \( n = 4 \). (The nucleobases in these calculations are arranged in their canonical B-DNA geometries.)
### Table A.3: Vertical excitation energies (in eV) and oscillator strengths for low-lying singlet excited states of simple nucleobase systems. All methods used the 6-311+G* basis set except CC2, where the TZVP basis was used. Geometries correspond to canonical B-DNA. Some of the CC2 relaxed oscillator strengths failed to converge and are therefore not reported here. For the CIS(D) and SCS-CIS(D) excitation energies, the oscillator strengths are CIS values.

<table>
<thead>
<tr>
<th>Excited State</th>
<th>Method</th>
<th>PBE0</th>
<th>LRC-ωPBE</th>
<th>CIS(D)</th>
<th>SCS-CIS(D)</th>
<th>CC2</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Adenine Monomer</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>nπ*</td>
<td>5.19 (0.00)</td>
<td>5.26 (0.00)</td>
<td>5.87 (0.01)</td>
<td>5.80 (0.01)</td>
<td>5.34 (0.00)</td>
<td></td>
</tr>
<tr>
<td>ππ* (W)</td>
<td>5.46 (0.19)</td>
<td>5.62 (0.06)</td>
<td>5.55 (0.36)</td>
<td>5.35 (0.36)</td>
<td>5.46 (0.00)</td>
<td></td>
</tr>
<tr>
<td>ππ* (B)</td>
<td>5.56 (0.10)</td>
<td>5.75 (0.30)</td>
<td>5.67 (0.07)</td>
<td>5.36 (0.07)</td>
<td>5.58 (0.32)</td>
<td></td>
</tr>
<tr>
<td><strong>Thymine Monomer</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>nπ*</td>
<td>5.26 (0.00)</td>
<td>4.93 (0.00)</td>
<td>4.89 (0.00)</td>
<td>4.74 (0.00)</td>
<td>4.84 (0.00)</td>
<td></td>
</tr>
<tr>
<td>ππ*</td>
<td>5.59 (0.27)</td>
<td>5.34 (0.22)</td>
<td>5.43 (0.41)</td>
<td>5.30 (0.41)</td>
<td>5.31 (0.20)</td>
<td></td>
</tr>
<tr>
<td><strong>A:T Base Pair</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thy nπ*</td>
<td>4.87 (0.00)</td>
<td>5.10 (0.00)</td>
<td>5.27 (0.00)</td>
<td>5.18 (0.00)</td>
<td>4.94</td>
<td></td>
</tr>
<tr>
<td>Ade nπ*</td>
<td>5.35 (0.00)</td>
<td>5.51 (0.00)</td>
<td>5.54 (0.00)</td>
<td>5.50 (0.00)</td>
<td>5.54</td>
<td></td>
</tr>
<tr>
<td>Thy ππ*</td>
<td>5.14 (0.15)</td>
<td>5.30 (0.18)</td>
<td>5.37 (0.35)</td>
<td>5.23 (0.35)</td>
<td>5.21</td>
<td></td>
</tr>
<tr>
<td>Ade ππ* (W)</td>
<td>5.41 (0.20)</td>
<td>5.57 (0.09)</td>
<td>5.46 (0.37)</td>
<td>5.25 (0.37)</td>
<td>5.40</td>
<td></td>
</tr>
<tr>
<td>Ade ππ* (B)</td>
<td>5.52 (0.12)</td>
<td>5.71 (0.32)</td>
<td>5.66 (0.14)</td>
<td>5.39 (0.14)</td>
<td>5.47</td>
<td></td>
</tr>
<tr>
<td>Ade → Thy CT</td>
<td>4.45 (0.00)</td>
<td>6.44 (0.04)</td>
<td>6.98 (0.00)</td>
<td>6.96 (0.00)</td>
<td>6.04</td>
<td></td>
</tr>
<tr>
<td><strong>A2 π-Stacked Dimer</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5’ nπ*</td>
<td>5.15 (0.00)</td>
<td>5.22 (0.00)</td>
<td>5.42 (0.00)</td>
<td>5.33 (0.00)</td>
<td>5.26 (0.00)</td>
<td></td>
</tr>
<tr>
<td>3’ nπ*</td>
<td>5.16 (0.00)</td>
<td>5.24 (0.00)</td>
<td>5.42 (0.00)</td>
<td>5.34 (0.00)</td>
<td>5.27 (0.00)</td>
<td></td>
</tr>
<tr>
<td>ππ*(W−)</td>
<td>5.30 (0.04)</td>
<td>5.51 (0.03)</td>
<td>5.40 (0.05)</td>
<td>5.22 (0.05)</td>
<td>5.39 (0.05)</td>
<td></td>
</tr>
<tr>
<td>ππ*(W+)</td>
<td>5.43 (0.17)</td>
<td>5.61 (0.04)</td>
<td>5.57 (0.24)</td>
<td>5.46 (0.24)</td>
<td>5.41 (0.00)</td>
<td></td>
</tr>
<tr>
<td>ππ*(B−)</td>
<td>5.51 (0.03)</td>
<td>5.64 (0.01)</td>
<td>5.68 (0.20)</td>
<td>5.58 (0.20)</td>
<td>5.42 (0.00)</td>
<td></td>
</tr>
<tr>
<td>ππ*(B+)</td>
<td>5.56 (0.16)</td>
<td>5.77 (0.41)</td>
<td>5.76 (0.03)</td>
<td>5.74 (0.03)</td>
<td>5.55 (0.40)</td>
<td></td>
</tr>
<tr>
<td>3’-Ade → 5’-Ade CT</td>
<td>4.95 (0.00)</td>
<td>6.12 (0.01)</td>
<td>6.28 (0.00)</td>
<td>6.28 (0.00)</td>
<td>6.19</td>
<td></td>
</tr>
<tr>
<td>5’-Ade → 3’-Ade CT</td>
<td>5.09 (0.00)</td>
<td>6.33 (0.01)</td>
<td>6.50 (0.04)</td>
<td>6.48 (0.04)</td>
<td>6.32</td>
<td></td>
</tr>
<tr>
<td>Excited State</td>
<td>VEE/ eV</td>
<td>Oscillator Strength</td>
<td>Character</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------------</td>
<td>---------</td>
<td>---------------------</td>
<td>-----------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>4.92</td>
<td>0.00</td>
<td>Thy 2 nπ*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>4.93</td>
<td>0.00</td>
<td>Thy 4 nπ*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>5.25</td>
<td>0.00</td>
<td>Ade 3 nπ*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>5.25</td>
<td>0.00</td>
<td>Ade 5 nπ*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>5.26</td>
<td>0.00</td>
<td>Ade 1 nπ*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>5.41</td>
<td>0.09</td>
<td>Thy 4 ππ*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>5.42</td>
<td>0.06</td>
<td>Thy 2 ππ*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>5.65</td>
<td>0.00</td>
<td>Ade 3 →Thy4 CT</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>5.67</td>
<td>0.00</td>
<td>Ade 1 ππ*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>5.71</td>
<td>0.01</td>
<td>Ade 5 ππ*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>5.84</td>
<td>0.19</td>
<td>Ade 3 ππ*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>5.86</td>
<td>0.01</td>
<td>Ade 3 →T4 CT</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>5.87</td>
<td>0.00</td>
<td>Ade 1 nπ*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>5.87</td>
<td>0.00</td>
<td>Ade 5 nπ*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>5.89</td>
<td>0.02</td>
<td>Ade 3 →Thy4 CT</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>5.89</td>
<td>0.28</td>
<td>Ade 1 ππ*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>5.92</td>
<td>0.27</td>
<td>Ade 3 →Thy4 CT (+ Ade5 ππ*)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>5.97</td>
<td>0.01</td>
<td>Ade 1 →Thy2 CT</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>6.18</td>
<td>0.00</td>
<td>Ade 1 nπ*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>6.19</td>
<td>0.00</td>
<td>Ade 3 nπ*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>6.21</td>
<td>0.00</td>
<td>Ade 5 nπ*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>6.21</td>
<td>0.00</td>
<td>Ade 3 →Thy2 CT</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table A.4: LRC-ωPBE/6-31G* vertical excitation energies (VEEs) and oscillator strengths of ATATA. The monomers are labeled by number in 5’ to 3’ order.
<table>
<thead>
<tr>
<th>Excited State</th>
<th>VEE/ eV</th>
<th>Oscillator Strength</th>
<th>Character</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5.12</td>
<td>0.00</td>
<td>Thy 3β nπ*</td>
</tr>
<tr>
<td>2</td>
<td>5.13</td>
<td>0.00</td>
<td>Thy 1β nπ*</td>
</tr>
<tr>
<td>3</td>
<td>5.19</td>
<td>0.00</td>
<td>Thy 2α nπ*</td>
</tr>
<tr>
<td>4</td>
<td>5.35</td>
<td>0.03</td>
<td>Thy 2α ππ*(+ Ade 2β → Thy 3β CT)</td>
</tr>
<tr>
<td>5</td>
<td>5.38</td>
<td>0.15</td>
<td>Ade 2β → Thy 3β CT</td>
</tr>
<tr>
<td>6</td>
<td>5.45</td>
<td>0.06</td>
<td>Thy 1β ππ*</td>
</tr>
<tr>
<td>7</td>
<td>5.51</td>
<td>0.00</td>
<td>Ade 1α nπ*</td>
</tr>
<tr>
<td>8</td>
<td>5.52</td>
<td>0.00</td>
<td>Ade 2β nπ*</td>
</tr>
<tr>
<td>9</td>
<td>5.54</td>
<td>0.00</td>
<td>Ade 3α nπ*</td>
</tr>
<tr>
<td>10</td>
<td>5.57</td>
<td>0.01</td>
<td>Ade 2β → Thy 3β CT</td>
</tr>
<tr>
<td>11</td>
<td>5.63</td>
<td>0.01</td>
<td>Ade 1α → Thy 2α CT</td>
</tr>
<tr>
<td>12</td>
<td>5.64</td>
<td>0.01</td>
<td>Ade 3α ππ*</td>
</tr>
<tr>
<td>13</td>
<td>5.75</td>
<td>0.11</td>
<td>Ade 2β → Thy 3β CT</td>
</tr>
<tr>
<td>14</td>
<td>5.77</td>
<td>0.39</td>
<td>Ade 2β ππ*(+ Ade 3α ππ*)</td>
</tr>
<tr>
<td>15</td>
<td>5.83</td>
<td>0.20</td>
<td>Ade 3α ππ*(+ Ade 1α ππ*)</td>
</tr>
<tr>
<td>16</td>
<td>5.87</td>
<td>0.15</td>
<td>Ade 1α ππ*(+ Ade 2β ππ*)</td>
</tr>
<tr>
<td>17</td>
<td>5.92</td>
<td>0.03</td>
<td>Ade 1α → Thy 2α CT</td>
</tr>
<tr>
<td>18</td>
<td>5.97</td>
<td>0.00</td>
<td>Ade 2β nπ*(+ Ade 2α nπ*)</td>
</tr>
<tr>
<td>19</td>
<td>5.98</td>
<td>0.00</td>
<td>Ade 3α nπ*</td>
</tr>
<tr>
<td>20</td>
<td>5.98</td>
<td>0.00</td>
<td>Ade 1β nπ*(+ Ade 2α nπ*)</td>
</tr>
<tr>
<td>21</td>
<td>6.04</td>
<td>0.00</td>
<td>Ade 2β → Thy 1β CT</td>
</tr>
<tr>
<td>22</td>
<td>6.12</td>
<td>0.01</td>
<td>Ade 3α → Thy 1α CT</td>
</tr>
<tr>
<td>23</td>
<td>6.22</td>
<td>0.00</td>
<td>Thy 3β nπ*</td>
</tr>
<tr>
<td>24</td>
<td>6.22</td>
<td>0.00</td>
<td>Thy 2α nπ*</td>
</tr>
<tr>
<td>25</td>
<td>6.31</td>
<td>0.00</td>
<td>Thy 1β nπ*</td>
</tr>
<tr>
<td>26</td>
<td>6.32</td>
<td>0.00</td>
<td>Ade 1α nπ*</td>
</tr>
<tr>
<td>27</td>
<td>6.34</td>
<td>0.00</td>
<td>Ade 2β nπ*</td>
</tr>
<tr>
<td>28</td>
<td>6.40</td>
<td>0.02</td>
<td>Ade 3α → Thy 1β CT</td>
</tr>
<tr>
<td>29</td>
<td>6.41</td>
<td>0.00</td>
<td>Ade 3α nπ*</td>
</tr>
<tr>
<td>30</td>
<td>6.50</td>
<td>0.02</td>
<td>Ade 2β → Thy 2α CT</td>
</tr>
</tbody>
</table>

Table A.5: LRC-ωPBE/6-31G* vertical excitation energies (VEEs) and oscillator strengths of ATA:TAT. The monomers are labeled by number in 5′ to 3′ order and additionally according to their strand (α = ATA, β = TAT).
Figure A.4: Stick spectra of various gas-phase Aₙ, Tₙ, and Aₙ:Tₙ multimers composed of nucleobases arranged in the canonical B-DNA geometries. All calculations are performed at the LRC-ωPBE/6-31G* level. To avoid congestion, optically-weak nπ* excitations are omitted from these spectra, as are Rydberg excitations (which are largely absent in this energy range, due to the omission of diffuse basis functions). The 1ππ* states in the multimers (n ≥ 2) are delocalized excitons. The lowest adenine → thymine CT states of A:T appear at 6.5 eV, out of the range depicted here.
A.4 MD simulations and QM/MM models

We model solvation through a mixed quantum mechanics/molecular mechanics (QM/MM) approach with water clusters derived from molecular dynamics (MD) simulations. B-DNA oligomers (whose canonical geometries are generated by the nucleic program, which is part of the TINKER software package\textsuperscript{58}) were added to a pre-equilibrated, cubic box ($L = 49.323$ Å) of water molecules at a density of 0.99 g/cm\textsuperscript{3}. MD simulations were run at constant temperature of 298 K under periodic boundary conditions with the AMBER99\textsuperscript{176,177} and TIP3P\textsuperscript{57} force fields, as implemented in the TINKER program.\textsuperscript{58} In most cases, the nucleobases and/or backbone were rigidly constrained during the MD simulations to retain the same B-DNA conformation as in the gas-phase calculations. (The except is the adenine dinucleotide calculation discussed in the manuscript, in which the adenine monomers—but not the backbone—were allowed to relax.) Following equilibration, water molecules within a specified radius (see below) of any nucleobase atom were retained as explicit QM waters in the excited-state calculation. Ten configurations, each separated by at least 5 ps, were extracted from the MD simulations for QM/MM calculations.

Calculations using exciton model Hamiltonians reveal that the extent of exciton delocalization is strongly affected by intramolecular dynamics of the DNA molecule,\textsuperscript{68,272} whereas we use frozen monomer geometries. In larger systems, it is therefore absolutely crucial to consider the effect of DNA fluctuations, as perfectly helical DNA supports infinitely-delocalized $\pi\pi^*$ excitons. In the small systems consider here (no more than two stacked bases, in aqueous solution) this is probably not so important,
however, at least for configurations in which the bases are well stacked. (Unstacked geometries of a dinucleotide will not exhibit delocalized excited states.) We plan to address the issue of structural fluctuations in future simulations.

Initially, we considered a cluster solvation model and two different QM/MM models, before settling on the description of aqueous solvation that is used in the manuscript. These models are defined as follows.

1. All water molecules within 2.5 Å of any nucleobase atom are included in the QM region, and all remaining water molecules are completely omitted. We call this the “Cluster Model”.

2. A second model includes the aforementioned cluster as the QM region, but also includes TIP3P point charges for water molecules within 10 Å of the QM region. We call this “QM/MM Model 1”.

3. QM/MM Model 2 is analogous to Model 1, but the QM region extends to 4.0 Å around the nucleobases and the MM region extends out another 14.0 Å.

For the aqueous A:T and A$_2$ systems discussed below, both the cluster model and QM/MM Model 1 include about 15 QM water molecules, whereas QM/MM Model 2 includes approximately 50 QM water molecules. QM/MM Models 1 and 2 include $\sim$550 and $\sim$700 MM water molecules, respectively.

Figure A.5 displays absorption spectra for hydrated A:T and A$_2$, calculated using each of these three solvation models. The Cluster Model actually performs reasonably well, although the CT states tend to appear at slightly lower energies as compared
Figure A.5: Absorption spectra for aqueous A:T and A₂ at the LRC-ωPBE/6-31G* level using three different models of aqueous solvation. (To avoid congestion, the optically-weak nπ⁺ states are omitted.) Gaussian distributions are obtained from averages over solvent configuration; for the CT states, the stick spectra are shown as well.
to the more complete QM/MM solvation models. Comparing QM/MM Models 1 and 2, however, we see that the significantly larger QM region in the latter scarcely changes the absorption spectrum. QM/MM Model 1 is also small enough that we can perform CIS(D)/6-311+G* calculations using this solvation model. Therefore, QM/MM Model 1 is used for all of the solution-phase calculations in the manuscript. Henceforth (as in the manuscript), “QM/MM” refers to Model 1.

A.5 Supplementary solution-phase data

Figure A.6 compares CIS(D) and SCS-CIS(D) absorption spectra based on QM/MM calculations of aqueous A:T and A₂. These two methods predict virtually identical excitation energies for the CT states, but the $\pi\pi^*$ excitation energies are lower at the SCS-CIS(D) level, by $\sim 0.2$ eV.

As discussed in the manuscript, the TD-LRC-$\omega$PBE method appears to overestimate CT states by $\sim 0.1$ eV while the TD-LRC-$\omega$PBEh method underestimates the same states by $\sim 0.2$ eV. Both methods overestimate $n\pi^*$ and $\pi\pi^*$ excitation energies by $\sim 0.2$ eV, as compared to CC2 and CASPT2 benchmarks. If we take these mean errors are empirical shifts to be applied to the solution-phase absorption spectra obtained from TD-DFT (i.e., the spectra that appear in Fig. A.5), the result is the set of spectra appearing in Fig. A.7. These shifts (which can also be justified by appeal to the $\omega$-dependent excitation energy curves, Figs. A.2 and A.3) bring the TD-LRC-$\omega$PBE and TD-LRC-$\omega$PBEh excitation energies into excellent agreement. Note, however, that the oscillator strengths are no longer reliable, since these were
Figure A.6: Absorption spectra for hydrated A:T (left panels) and hydrated A₂ (right panels), obtained at the CIS(D)/6-311+G* level using QM/MM Model 1. Gaussian distributions were obtained from averages over solvent con figuration (using CIS oscillator strengths), though for the CT states the stick spectra are shown as well. In A:T there is considerable mixing between the second ππ* band and the CT states, lending significant oscillator strength to the latter.
Figure A.7: TD-LRC-DFT absorption spectra for aqueous A:T and A₂, in which individual excitation energies have been shifted according to estimated errors as described in the manuscript. The LRC-ωPBE functional is used in (a) and (b), whereas the LRC-ωPBEh functional is used in (c) and (d). Oscillator strengths for the CT states are not reliable, as a result of intensity borrowing at the original, calculated excitation energies.
calculated at excitation energies that in some cases allow for intensity borrowing that would not be possible at the shifted energies. This is especially evident in the case of CT states calculated at the TD-LRC-ωPBEh level; in the original calculation, these states appear right on top of the brightest part of the $\pi\pi^*$ exciton band, from which they borrow substantial intensity. Such intensity borrowing does not occur at the higher, shifted excitation energies. (Because the TD-LRC-ωPBE method puts the CT excitations above the $\pi\pi^*$ excitations in the first place, the oscillator strengths calculated with this method are probably more reliable.)

The largest aqueous system that we consider is $A_2:T_2$, an absorption spectrum for which is shown in Fig. A.8, which was calculated using the LRC-ωPBE functional.
On average, the $\pi\pi^*$ exciton energies for this system are within 0.05 eV of those obtained for the analogous QM/MM model of A$_2$. (This close similarity arises in part because we constrain the nucleobases to their canonical B-DNA geometries during the molecular dynamics simulation; only the water degrees of freedom are subject to configurational averaging. As a result, the spectrum in Fig. A.8 is very nearly a superposition of the solution-phase A:T and A$_2$ spectra that are presented in the manuscript.) Compared to gas-phase A$_2$:T$_2$, the exciton states of both adenine and thymine exhibit solvatochromatic red shifts of $\sim$0.1 eV.

Both intra- and interstrand CT states in this system are stabilized by 0.1–0.2 eV relative to the gas phase. As in aqueous A$_2$, the low-energy tail of the adenine $\rightarrow$ adenine CT band has a slight overlap with the most intense $\pi\pi^*$ band. Although the interstrand CT states are higher in energy, on average, than the intrastrand CT states (both adenine $\rightarrow$ adenine and thymine $\rightarrow$ thymine), there is significant overlap between the intra- and interstrand CT bands. Moreover, the interstrand CT states in aqueous A$_2$:T$_2$ are about 0.2 eV lower than those in aqueous A:T. In the larger system, delocalization of the virtual orbitals over adjacent $\pi$-stacked bases stabilizes the interstrand CT states.

Interstrand CT states between non-hydrogen–bonded nucleobases were observed in only half of the cluster configurations and are thus omitted from Fig. A.8. When such states are observed, they appear in the range of 6.3–7.0 eV, and in each case there exists a lower-energy adenine $\rightarrow$ thymine CT state localized on a hydrogen-bonded base pair. It is possible that—as observed in the gas phase—these delocalized
interstrand CT states are stabilized in larger oligomers. We plan to investigate this possibility in future work.
APPENDIX B

Supporting information for “Polarizable continuum reaction-field solvation models affording smooth potential energy surfaces”  

B.1 Additional information on the SWIG formalism

In Section B.1.1 of this Supporting Information, we present a formal derivation of the SWIG methodology and a rigorous proof that this method affords smooth potential energy surfaces. These arguments generalize the S-COSMO model of York and Karplus\(^{168}\) to more a more general class of PCMs.\(^{3,8,160}\) In slightly modified form, these arguments can be used to extend the FIXPVA method\(^{5}\) to these more general PCMs, although the FIXPVA method as described in Ref. 5 uses a different switching function that the one employed here. Our switching function is described in detail in Section B.1.2.

B.1.1 Derivation

Chipman\(^{3,8,160}\) provides a detailed discussion of the integral-equation formalism for apparent surface charge PCMs, which includes the C-PCM/COSMO, IEF-PCM, and

\(^{B.1}\)This chapter appeared as the Supporting Information to the full article in the *Journal of Physical Chemistry Letters*, in 2010, volume 1, page 556.
SS(V)PE methods, and the reader is referred to Refs. 160–3 for details. For our purposes, it suffices to note that the PCM equations, in the continuous representation, are expressed in terms of a pair of integral operators, ˆ\(S\) and ˆ\(D\). The operator ˆ\(S\) acts on the surface charge density, \(\sigma(s)\), to generate the corresponding surface potential at the surface point \(s\),

\[
\hat{S}\sigma(s) = \int_{\text{cavity surface}} ds' \sigma(s') \frac{1}{|s - s'|} .
\]

(B.1)

whereas the action of ˆ\(D\) on \(\sigma(s)\) generates the surface dipole potential at \(s\),

\[
\hat{D}\sigma(s) = \int_{\text{cavity surface}} ds' \sigma(s') \frac{\partial}{\partial n_{s'}} \frac{1}{|s - s'|} .
\]

(B.2)

Here \(n_{s'}\) indicates the normalized, outward-pointing surface normal vector at the surface point \(s'\).

The matrices \(S\) and \(D\) are discretized representations of the operators ˆ\(S\) and ˆ\(D\). Within the SWIG approach, we blur the point charge \(q_i\), which is located at the \(i\)th Lebedev grid point, \(r_i\), using a normalized gaussian function

\[
g_i(r) = \left(\frac{\zeta_i^2}{\pi}\right)^{3/2} e^{-\zeta_i^2 |r - r_i|^2} .
\]

(B.3)

The exponents of these gaussians are treated as fixed parameters that characterize the model.

The matrix element \(S_{ij}\) of \(S\) represents the projection of ˆ\(Sg_j\) onto \(g_i\). For \(i \neq j\), we determine \(S_{ij}\) by using Eq. (B.1), projecting onto \(g_i\), and evaluating surface integrals over \(s\) and \(s'\). The result is

\[
S_{ij} = \frac{\text{erf}(\zeta_{ij} r_{ij})}{r_{ij}} \quad \text{for} \ i \neq j ,
\]

(B.4)
where $\zeta_{ij} = \zeta_i \zeta_j / (\zeta_i^2 + \zeta_j^2)^{1/2}$ and $r_{ij} = |\mathbf{r}_i - \mathbf{r}_j|$. Diagonal elements of $\mathbf{S}$ are obtained according to $S_{ii} = \lim_{r_{ij} \to 0} S_{ij}$. Recognizing that $\zeta_{ii} = \zeta_i / \sqrt{2}$, this limit is evaluated to obtain

$$S_{ii} = \zeta_i \sqrt{2/\pi}.$$  \hfill (B.5)

Equations (B.4) and (B.5) are nearly the same as Eq. (3) in the text, except for a factor of $F_i^{-1}$ that appears in $S_{ii}$ as given in Eq. (3). We next explain the origin of this factor.

According to Eq. (1) in the text, the surface charges $q_i$ are determined according to $q_i = K^{-1} \mathbf{Rv}$. If a certain switching function $F_k = 0$, however, we wish to ensure that $q_k = 0$, so that the $k$th surface grid point does not contribute to the PCM solvation energy. However, the electrostatic potential $v_k$ is generally nonzero, even if $F_k = 0$, and furthermore $\mathbf{R} \propto \mathbf{I}$ for C-PCM/COSMO. Hence the $k$th element of $\mathbf{Rv}$ is generally nonzero. As such, the only way to ensure that the “switched-off” grid points do not contribute to the solvation energy is to require that $K^{-1}$ has a null space that corresponds precisely to those surface grid points for which $F_k = 0$.

Partitioning $\mathbf{K}$ into diagonal and off-diagonal parts,

$$\mathbf{K} = \mathbf{K}_{\text{diag}} + \mathbf{K}_{\text{off}},$$ \hfill (B.6)

in which the matrix $\mathbf{K}_{\text{diag}}$ is diagonal, we may write

$$\mathbf{K} = \mathbf{K}_{\text{diag}}^{1/2} \left( \mathbf{I} + \mathbf{K}_{\text{diag}}^{-1/2} \mathbf{K}_{\text{off}} \mathbf{K}_{\text{diag}}^{-1/2} \right) \mathbf{K}_{\text{diag}}^{1/2},$$ \hfill (B.7)

whence

$$\mathbf{K}^{-1} = \mathbf{K}_{\text{diag}}^{-1/2} \left( \mathbf{I} + \mathbf{K}_{\text{diag}}^{-1/2} \mathbf{K}_{\text{off}} \mathbf{K}_{\text{diag}}^{-1/2} \right)^{-1} \mathbf{K}_{\text{diag}}^{-1/2}.$$ \hfill (B.8)
Now, if the matrix elements of diagonal matrix $K_{\text{diag}}^{-1}$ were zero whenever $F_k = 0$, the effect of the right and left multiplication by $K_{\text{diag}}^{-1/2}$ in Eq. (B.8) would be to annihilate any rows and columns of $(I + K_{\text{diag}}^{-1/2} K_{\text{off}} K_{\text{diag}}^{-1/2})^{-1}$ for which $F_k = 0$, or in other words, to create the desired null space for $K^{-1}$. Similarly, the matrix $I + K_{\text{diag}}^{-1/2} K_{\text{off}} K_{\text{diag}}^{-1/2}$ would become block diagonal and equal to the unit matrix within the null space of $K^{-1}$, and the dimension of $q = K^{-1} R v$ could be reduced without approximation to just the number of grid points whose switching functions are non-zero.

These arguments suggest that we should replace the expression for $S_{ii}$ in Eq. (B.5) with

$$S_{ii} = \zeta_i \sqrt{2/\pi} F_i,$$  \hspace{1cm} (B.9)

so that $S_{ii} \to \infty$ as $F_i \to 0$. This equation, together with Eq. (B.4) for $S_{ij}$, defines the $S$-matrix used in this work. Note that

$$(\text{DAS})_{ii} = \sum_{j \neq i} D_{ij} a_j S_{ji},$$  \hspace{1cm} (B.10)

owing to our choice $D_{ii} = 0$, as discussed in the text. Since $S_{ij}$ is finite, as is $D_{ij} = -\hat{n}_j \cdot (\partial S_{ij}/\partial r_j)$ [Eq. (6)], it follows that $K_{ii} \to \infty$ as $F_i \to 0$, for both C-PCM/COSMO and SS(V)PE. As York and Karplus point out, the quantity $S_{ii}$ is the self-energy of the $i$th surface tessera, so the definition in Eq. (B.9) has the effect of imposing an increasing steep penalty for polarizing the $i$th tessera, as the $i$th Lebedev point enters and passes through the switching region from the cavity surface, until finally $S_{ii} = \infty$ within the cavity interior, and the $i$th tesserae can no longer be polarized.
Although the dimension of the null space of $K^{-1}$ can and does change as the atoms are moved, by construction (Section B.1.2), $F_i \rightarrow 0$ smoothly, and therefore the block structure of $K^{-1}$ changes smoothly as well. This proves that the SWIG potential energy surfaces are rigorously smooth, using exact arithmetic. In practice, of course, we must employ a finite drop tolerance, and in the present work we remove any Lebedev point for which $F_i < 10^{-8}$. (In practice, this is no different than the finite drop tolerances used to discard shell pairs and to screen AO integrals in standard electronic structure calculations. Potential energy surfaces are only as smooth as these finite thresholds allow, but the thresholds can be made as small as machine precision, if desired.) Figure B.1 shows how the number of non-negligible points (and thus

Figure B.1: Number of non-negligible ($F_i > 10^{-8}$) Lebedev grid points for NaCl, as a function of the Na–Cl distance.
the dimension of \( K \) changes as a function of bond length, for the NaCl calculation described in the paper. The dimension of \( K \) changes numerous times as the atoms are pulled apart, yet the SWIG potential surface remains smooth.

**B.1.2 Switching function**

The switching function \( F_i \) used in the current work is equivalent to that used by York and Karplus.\(^{168}\) For a given surface grid point, \( r_i \), the function \( F_i \) is a product of elementary switching functions, \( f \):

\[
F_i = \prod_{J, i \notin J} f(\hat{r}_{i,J}). \tag{B.11}
\]

The notation “\( i \notin J \)” indicates that the product over \( J \) omits the atom associated with the \( i \)th Lebedev point. The elementary switching function used in Eq. (B.11) is

\[
f(x) = \begin{cases} 
0 & \text{if } x < 0 \\
x^3(10 - 15x + 6x^2) & \text{if } 0 \leq x \leq 1 \\
1 & \text{if } x > 1
\end{cases} \tag{B.12}
\]

The argument of \( f \) in Eq. (B.11) is a dimensionless number, \( \hat{r}_{i,J} \), that quantifies the extent to which the \( i \)th discretization point has penetrated into the buffer region surrounding the \( J \)th atom:

\[
\hat{r}_{i,J} = \frac{r_{i,J} - R_{\text{in},J}}{R_{\text{sw},J}}. \tag{B.13}
\]

Here, \( r_{i,J} = |r_i - R_{J}| \) is the distance from the \( i \)th discretization point to the \( J \)th atom, \( R_{\text{in},J} < R_J \) is the radius of the inner limit of the switching region for atom \( J \), and \( R_{\text{sw},J} \) is the width of this switching region. The parameters \( R_{\text{in},J} \) and \( R_{\text{sw},J} \) are determined by \( R_J \), the radius of the \( J \)th atomic sphere, and \( N_J \), the number of
Lebedev points used to discretize this sphere. Specifically,\(^\text{168}\)

\[ R_{\text{sw},J} = \gamma_J R_J \]  

(B.14)

and

\[ R_{\text{in},J} = R_J - \alpha_J R_{\text{sw},J} \, , \]  

(B.15)

where

\[ \alpha_J = \frac{1}{2} + \frac{1}{\gamma_J} - \sqrt{\frac{1}{\gamma_J^2} - \frac{1}{28}} \]  

(B.16)

and

\[ \gamma_J = \sqrt{\frac{14}{N_J}} \, . \]  

(B.17)

B.2 Computational details

Solute cavities were discretized using atom-centered Lebedev grids.\(^\text{273}\) Grid points were discarded whenever \( F_i < 10^{-8} \); tighter thresholds had little discernible effect on the calculations. The rest of this section provides details regarding the (adenine)(H\(_2\)O)\(_{52}\) and NaCl calculations discussed in this work.

B.2.1 (Adenine)(H\(_2\)O)\(_{52}\) optimization and frequencies

An (adenine)(H\(_2\)O)\(_{52}\) cluster was carved out of an equilibrated molecular dynamics simulation of adenine in bulk water under ambient conditions, by selecting all H\(_2\)O molecules within 5 Å of any adenine atom. This cluster was optimized in the gas phase (using the AMBER99 force field) for 50 optimization cycles, in order to eliminate any artifacts in the gradient caused by extracting the cluster from the bulk. The
resulting structure was then used as a starting point for the VTN-, FIXPVA-, and SWIG-COSMO optimizations.

Solute cavities for these optimizations were constructed from atom-centered spheres whose radii were chosen by adding 1.4 Å to the AMBER99 Lennard–Jones radii. This surface was discretized using 50 Lebedev points per atomic sphere.

Q-Chem default convergence criteria were used for the geometry optimizations (maximum step size = 0.3 Å; maximum gradient component = $3.0 \times 10^{-4}$ a.u.; maximum atomic displacement = $1.2 \times 10^{-3}$ a.u.; maximum energy change = $1.0 \times 10^{-6}$ a.u.). Harmonic frequencies were calculated by finite difference of analytic energy gradients. For the SWIG procedure, the finite-difference step size was taken to be the Q-Chem default, $3.0 \times 10^{-4}$ Å. The step size used for FIXPVA was reduced to $3.0 \times 10^{-5}$ Å, in order to avoid imaginary frequencies that were encountered when larger displacements were used.

**B.2.2 NaCl dissociation**

Dissociation curves for NaCl [computed at the HF/6-31+G*//SS(V)PE level] consist of single-point calculations spaced 0.005 Å apart. The total energy includes the SS(V)PE electrostatic interaction energy as well as non-electrostatic cavitation, dispersion, and repulsion interactions.

Each non-electrostatic component of the solvation energy is defined by a fairly simple function of the cavity surface area. The cavitation energy was computed using the formula given in Ref. 161, whereas dispersion and repulsion energies were computed as described in Ref. 180 Lennard–Jones parameters for the dispersion/repulsion...
energy were taken from the AMBER99 force field\textsuperscript{56} for Na\textsuperscript{+} and Cl\textsuperscript{−}, and from the TIP3P force field\textsuperscript{57} for H\textsubscript{2}O.

As suggested in Ref. 161, the non-electrostatic interaction terms utilize a somewhat different cavity surface than that used for the electrostatic interactions. For the electrostatic [SS(V)PE] terms, we use Bondi’s atomic van der Waals radii,\textsuperscript{188} each scaled by a factor of 1.2. For the dispersion and repulsion interaction energies, a solvent-accessible surface was constructed by adding the Lennard–Jones minimum-energy distance for each solvent atom to the scaled Bondi radius. In either case, each atomic sphere was discretized using 110 Lebedev points.
APPENDIX C

Supplementary data for “Symmetric versus asymmetric discretization of the integral equations in polarizable continuum solvation models”

C.1 Additional data

In Section II A of the paper, we discussed the general matrix formulation of PCMs in terms of Eq. 7. We present the corresponding $K$ and $R$ matrices for C-PCM and IEF-PCM in Table C.1, where the $K$-matrix variants considered in this work are explicitly shown.

C.1 This chapter appeared as supplementary data to the full article in Chemical Physics Letters, in 2011, volume 509, pages 77–87.

<table>
<thead>
<tr>
<th>Method</th>
<th>Matrix $K$</th>
<th>Matrix $R$</th>
</tr>
</thead>
<tbody>
<tr>
<td>C-PCM</td>
<td>$S$</td>
<td>$-[(\varepsilon - 1)/\varepsilon] I$</td>
</tr>
<tr>
<td>IEF-PCM ($X = DAS$ form)</td>
<td>$S - (f_\varepsilon/2\pi)DAS$</td>
<td>$-f_\varepsilon Y$</td>
</tr>
<tr>
<td>IEF-PCM ($X = SAD^\dagger$ form)</td>
<td>$S - (f_\varepsilon/2\pi)SAD^\dagger$</td>
<td>$-f_\varepsilon Y$</td>
</tr>
<tr>
<td>IEF-PCM (symmetric form)</td>
<td>$S - (f_\varepsilon/4\pi)(SAD + SAD^\dagger)$</td>
<td>$-f_\varepsilon Y$</td>
</tr>
</tbody>
</table>

Table C.1: Possible definitions for the matrices $K$ and $R$ in Eq. 7.
In Section IV A of the paper, we discuss the differences in the total energy, $W$, between the symmetric and asymmetric forms of $K$ for the amino acids in water. For the data shown in Fig. 1 of the paper, cavity surfaces were discretized using $N = 590$ Lebedev points per atomic sphere, but we have performed the same calculations using $N = 50$ and $N = 110$ as well. Figure C.1 below shows the energy differences that we obtain using $N = 110$. (This figure is the analogue of Fig. 1 in the paper, using a less dense grid.) We find that the denser grid can lead to even larger variations among the different forms of $K$. We tests three different grid densities ($N = 50, 110,$ and $590$), and Table C.2 shows the largest variations in energy, amongst the different forms of $K$, that are obtained in each case, for a variety of discretization procedures as described in the paper.

In Fig. 2 of the text, we plot the $W$ for a variety of different discretization schemes, relative to the energy obtained using Gaussian blurring only (“GBO”). Effectively, we are taking GBO discretization as the benchmark, and here we assess the reliability of this benchmark. Figure C.2 plots the energy difference $W_{\text{GBO}} - W_{\text{APBS}}$ for the aqueous amino acids, where $W_{\text{GBO}}$ represents the IEF-PCM solvation energy computed using GBO discretization and $W_{\text{APBS}}$ is the energy obtained using the adaptive Poisson-Boltzmann (APBS) software to solve the three-dimensional Poisson-Boltzmann equation. The same solute cavity is used for both the APBS and the IEF-PCM calculations, and the solute is described at the AMBER99 level so that there is no escaped charge. The APBS calculations use a $193 \times 193 \times 193$ grid with a resolution of 0.1 Å, and appear to be converged with respect to the grid parameters.
Figure C.1: Relative energies of the amino acids in water \((\varepsilon = 78.39)\), obtained at (a) the AMBER99 level and (b) the HF/6-31+G* level. The cavity surface was discretized using 110 Lebedev points per atomic sphere, and solution-phase energies are reported relative to the energy obtained using the symmetric form of \(K\).
Table C.2: Maximum (signed) difference, in kcal/mol, between the IEF-PCM energy computed using $X = \text{DAS}$ (or $X = \text{SAD}^{\dagger}$) and that computed using the symmetric form of $X$. The data set is the amino acids in water. Three different values of $N$, the number of Lebedev points per atomic sphere, and three different discretization procedures are compared.

<table>
<thead>
<tr>
<th>$N$</th>
<th>$X = \text{DAS}$</th>
<th>$X = \text{SAD}^{\dagger}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>GBO</td>
<td>ISWIG</td>
</tr>
<tr>
<td>50</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>110</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>590</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td>110</td>
<td>110</td>
</tr>
<tr>
<td>50</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>110</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>590</td>
<td>0.0</td>
<td>0.0</td>
</tr>
</tbody>
</table>

Table C.2: Maximum (signed) difference, in kcal/mol, between the IEF-PCM energy computed using $X = \text{DAS}$ (or $X = \text{SAD}^{\dagger}$) and that computed using the symmetric form of $X$. The data set is the amino acids in water. Three different values of $N$, the number of Lebedev points per atomic sphere, and three different discretization procedures are compared.
As such, $W_{\text{APBS}}$ should be the exact solution-phase energy, since we consider only electrostatic contributions to the solvation energy. Using $X = \text{DAS}$, we find that $|W_{\text{GBO}} - W_{\text{APBS}}| < 0.13$ kcal/mol for the entire amino acid data set. This means that IEF-PCM with $X = \text{DAS}$ and GBO discretization is essentially exact (in the absence of escaped charge), which justifies the use of this method as a benchmark in Fig. 2 of the paper.

![Graph showing comparison of total energies obtained for AMBER99 solutes](image)

Figure C.2: Comparison of total energies obtained for AMBER99 solutes, computed by numerical solution of the Poisson-Boltzmann equation (using the APBS software) to those obtained from two different forms of IEF-PCM with GBO discretization. The APBS and IEF-PCM solute cavities are identical. APBS calculations used a $193 \times 193 \times 193$ grid with a grid resolution of 0.1 Å, whereas IEF-PCM calculations used $N = 590$ Lebedev points per atomic sphere.

Section V B of the paper discusses the quadratic scaling of $E_{\text{pol}}$ with respect to
the electrostatic potential. This relationship is exact when a force field is used to represent the solute, but the same need not be true for a quantum-mechanical solute. We have tested this relationship using HF/6-31+G* calculations on a series of ionized histidine solutes, His$^{n+}$. Assuming that the solute’s internal energy, $E_0$, is approximately the same regardless of whether the $X = \text{DAS}$ or $X = \text{SAD}^\dagger$ form of IEF-PCM is employed, and assuming furthermore that the electrostatic potential grows in proportion to the molecular charge, $n$, then based upon the analysis in Section V B one should expect the energy difference $W_{\text{DAS}} - W_{\text{SAD}^\dagger}$ to grow quadratically with $n$. Calculated values of $W_{\text{DAS}} - W_{\text{SAD}^\dagger}$ are plotted versus $n$ in Fig. C.3, and the quadratic dependence is evident. This demonstrates that the analysis in Section V B is likely to be valid even for quantum-mechanical solutes.

**C.2 Comparison of FIXPVA Implementations**

We compare the energy, surface area, and Gauss’ law error for nuclear charges produced by our implementation of FIXPVA with Lebedev grids to the implementation in GAMESS$^{223,224}$ (ver. 1 Oct. 2010) with GEPOL grids. We have confirmed the equivalence of the two implementations by reading FIXPVA grids generated by GAMESS into our code, which for HF/6-31+G*/C-PCM calculations replicate the total energy to within $10^{-9}$ hartrees. A more interesting comparison is between the GEPOL and Lebedev grids, and we present such results on HF/6-31+G*/C-PCM calculations for the amino acid set (with Bondi radii scaled by 1.2 and $\varepsilon = 78.39$) in Table C.3.

Given that Su and Li$^5$ report a rotational variance of 0.9 and 0.2 kcal/mol for
Figure C.3: Difference in total energy between the $X = \text{SAD}^\dagger$ and $X = \text{DAS}$ forms of $K$, for successive ionization of histidine described at the HF/6-31+G* level. (The dashed curve is a quadratic fit to the four data points.) The cavity surface is discretized with SWIG using $N = 590$ Lebedev points per atomic sphere, and $\varepsilon = 78.39$. 

---

310
FIXPVA with GEPOL grids of 60 and 240 points per sphere, respectively, there is little substantial energetic difference between the GEPOL and Lebedev grids, even at these moderate grid densities. On average, the surface areas agree within 1 Å², and the Gauss’ law error for nuclear charges agrees within 0.1e. The use of denser grids should yield even better agreement between GEPOL and Lebedev grids. Nonetheless, the differences in Table C.3 are much smaller than the errors in K-matrix asymmetry reported elsewhere in this work.

C.3 FIXPVA discontinuity

We provide an example of the FIXPVA discontinuity. To do so, we plot the electrostatic solvation energy, $E_{pol}$ and the surface area along the dissociation of NaCl at the HF/6-31G* level. In order to observe the discontinuity, we use a radius of 1.2 Å for the Na atom and 1.4 Å for Cl atom, such that both radii are less than $n_2 = 1.5$ Å. The calculation is carried out with GAMESS, and results are plotted in Fig. C.4. It is clear that $E_{pol}$ and the surface area experience discontinuities at interatomic distances which are equal to the atomic radii.

Fig. C.4 serves only as an example in which the discontinuity can be realized, and it is not meant to be chemically relevant. Of course, one may argue that the atomic radii in this example are too small, but one can easily come up with many chemically relevant situations where atomic radii, especially those belonging to hydrogen atoms, will be below the $n_2$ threshold, such as proton transfer reactions, O–H vibrational motions, cationic organic molecules, cationic transition metal atoms, etc.
<table>
<thead>
<tr>
<th>Amino Acid</th>
<th>GEPOL $W$ (a.u.)</th>
<th>SA ($\text{Å}^2$)</th>
<th>$\Delta_{\text{Gauss}}$</th>
<th>Lebedev $W$ (a.u.)</th>
<th>SA ($\text{Å}^2$)</th>
<th>$\Delta_{\text{Gauss}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>-321.88334</td>
<td>83.82</td>
<td>0.71</td>
<td>-321.88320</td>
<td>83.00</td>
<td>0.76</td>
</tr>
<tr>
<td>C</td>
<td>-719.38547</td>
<td>102.72</td>
<td>0.76</td>
<td>-719.38496</td>
<td>102.87</td>
<td>0.82</td>
</tr>
<tr>
<td>F</td>
<td>-551.43051</td>
<td>120.57</td>
<td>1.18</td>
<td>-551.43027</td>
<td>121.31</td>
<td>1.28</td>
</tr>
<tr>
<td>G</td>
<td>-282.84839</td>
<td>74.18</td>
<td>0.57</td>
<td>-282.84827</td>
<td>75.07</td>
<td>0.57</td>
</tr>
<tr>
<td>I</td>
<td>-438.96596</td>
<td>102.45</td>
<td>1.12</td>
<td>-438.96556</td>
<td>101.11</td>
<td>1.27</td>
</tr>
<tr>
<td>L</td>
<td>-438.90410</td>
<td>102.37</td>
<td>1.15</td>
<td>-438.90354</td>
<td>100.68</td>
<td>1.35</td>
</tr>
<tr>
<td>M</td>
<td>-797.45500</td>
<td>121.21</td>
<td>1.03</td>
<td>-797.45490</td>
<td>121.77</td>
<td>1.07</td>
</tr>
<tr>
<td>N</td>
<td>-489.62292</td>
<td>101.99</td>
<td>0.92</td>
<td>-489.62244</td>
<td>101.07</td>
<td>1.02</td>
</tr>
<tr>
<td>P</td>
<td>-398.77831</td>
<td>92.77</td>
<td>0.98</td>
<td>-398.77769</td>
<td>91.92</td>
<td>1.11</td>
</tr>
<tr>
<td>Q</td>
<td>-528.69266</td>
<td>113.56</td>
<td>1.06</td>
<td>-528.69229</td>
<td>113.95</td>
<td>1.10</td>
</tr>
<tr>
<td>S</td>
<td>-396.73413</td>
<td>91.53</td>
<td>0.74</td>
<td>-396.73352</td>
<td>90.59</td>
<td>0.85</td>
</tr>
<tr>
<td>T</td>
<td>-435.76094</td>
<td>96.16</td>
<td>0.86</td>
<td>-435.75998</td>
<td>95.21</td>
<td>1.00</td>
</tr>
<tr>
<td>V</td>
<td>-399.94481</td>
<td>98.27</td>
<td>1.02</td>
<td>-399.94464</td>
<td>98.15</td>
<td>1.09</td>
</tr>
<tr>
<td>W</td>
<td>-682.17613</td>
<td>134.52</td>
<td>1.34</td>
<td>-682.17580</td>
<td>134.13</td>
<td>1.46</td>
</tr>
<tr>
<td>Y</td>
<td>-626.29219</td>
<td>130.31</td>
<td>1.22</td>
<td>-626.29207</td>
<td>131.13</td>
<td>1.28</td>
</tr>
<tr>
<td>H</td>
<td>-603.39912</td>
<td>123.95</td>
<td>1.46</td>
<td>-603.39848</td>
<td>123.88</td>
<td>1.57</td>
</tr>
<tr>
<td>K</td>
<td>-546.00669</td>
<td>112.63</td>
<td>1.15</td>
<td>-546.00601</td>
<td>112.00</td>
<td>1.30</td>
</tr>
<tr>
<td>R</td>
<td>-494.47461</td>
<td>108.82</td>
<td>1.35</td>
<td>-494.47405</td>
<td>110.25</td>
<td>1.41</td>
</tr>
<tr>
<td>E</td>
<td>-548.07604</td>
<td>115.23</td>
<td>0.97</td>
<td>-548.07566</td>
<td>116.19</td>
<td>1.01</td>
</tr>
<tr>
<td>D</td>
<td>-509.01129</td>
<td>105.27</td>
<td>0.80</td>
<td>-509.01070</td>
<td>103.79</td>
<td>0.91</td>
</tr>
</tbody>
</table>

Table C.3: FIXPVA on GEPOL grids with 240 points per atom versus Lebedev grids with 194 points per atom. Total energies ($W$), surface area (SA), and Gauss’ law error for nuclear charges in a.u. ($\Delta_{\text{Gauss}}$) are compared. Calculations are run at HF/6-31+G*//C-PCM with Bondi radii scaled by 1.2 and $\varepsilon = 78.39$. GEPOL calculations are run with GAMESS, and Lebedev calculations are run with our locally modified version of Q-Chem.
Figure C.4: Illustration of FIXPVA discontinuity along NaCl dissociation at HF/6-
31+G* run with GAMESS. Electrostatic solvation energy ($E_{pol}$) in kcal/mol is plotted
with the solid line, and the surface area in $\text{Å}^2$ is plotted with the dashed line. A
GEPOL grid of 60 points per sphere is used. The atomic radii are Na = 1.2 Å and Cl
= 1.4 Å, which are marked with a light vertical line and coincide with the observed
discontinuities. Data points are plotted at 0.01 Å intervals.
Bibliography


[245] C. J. Cramer and D. G. Truhlar, SMx continuum models for condensed phases, in Trends and Perspectives in Modern Computational Science, edited


[263] P. Debye and E. Hückel, On the theory of electrolytes. i. freezing point depres-

sion and related phenomena, in Collected Papers of Peter J. W. Debye, pages


