Monash University

Doctoral Thesis

Large-scale calculations of Ionic Liquids

Author:
Jason Rigby

Supervisors:
Dr. Ekaterina I. Izgorodina
Dr. Santiago Barrera Acevedo

A thesis submitted in fulfilment of the requirements
for the degree of Doctor of Philosophy

in the

Monash Computational Chemistry Group
School of Chemistry

October 2014

Errata and Addendum

Page 6, line 14: delete "a Taylor series expansion of the Fock operator truncated the second order" and replace with "a perturbative approach involving a second-order Taylor series expansion of the eigenfunctons and eigenvalues of the exact wavefunction, where the zeroth order perturbation is represented by the HF Hamiltonian"

Page 6, line 19: delete "which are at the pinnacle of accuracy"

Page 6, para 2: Comment: Coupled cluster approaches are always superior to the CI equivalent in cases where the CI expansion is truncated (*i.e.* all but Full CI).

Page 6, line 34: delete "assume that electrons are non-interacting" and replace with "include terms that are exact for non-interacting electrons"

Page 78, lines 1-4: Comment: While Becke was the first one to highlight the concept of a "density functional approximation" (DFA) in literature, this was done in response to a longstanding argument by Levy that such a distinction should be made. The term DFA cannot be solely attributed to Becke, and is rather the product of a longstanding dialogue between scholars in the field of Density Functional Theory.
See reference 8 of chapter 5: Axel D. Becke. Perspective: Fifty years of density-functional theory in chemical physics. *The Journal of Chemical Physics*, 140(18), 2014.

Page 80, para 2: delete "Nedler" and replace with "Nelder"

Page 105, line 6: insert "DFT-D3" between "utilising" and "type" such that it reads "…by utilising DFT-D3 type empirical dispersion corrections."

Notice

General declaration

In accordance with Monash University Doctorate Regulation 17.2 Doctor of Philosophy and Research Master's regulations the following declarations are made:

I hereby declare that this thesis contains no material which has been accepted for the award of any other degree or diploma at any university or equivalent institution and that, to the best of my knowledge and belief, this thesis contains no material previously published or written by another person, except where due reference is made in the text of the thesis.

This thesis includes three original papers published in peer reviewed journals and one unpublished publication. The core theme of the thesis is developing methodology to perform *ab initio* quantum chemical calculations on a large scale for ionic liquids. The ideas, development and writing up of all the papers in the thesis were the principal responsibility of myself, the candidate, working within the School of Chemistry under the supervision of Dr. Ekaterina I. Izgorodina and Dr. Santiago Barrera Acevedo.

The inclusion of co-authors reflects the fact that the work came from active collaboration between researchers and acknowledges input into team-based research.

In the case of Chapters 2, 3 and 4 my contribution to the work involved the following:

Thesis chapter	Publication title	Publication status	Nature and extent of candidate's contribution
2	Assessment of atomic partial charge schemes for polarisation and charge transfer in ionic liquids	Published	90%
3	Large-scale *ab initio* calculations of archetypical ionic liquids	Published	80%
4	New SCS- and SOS-MP2 coefficients fitted to semi-Coulombic systems	Published	90%
4	SCS-IL-MP2 produces accurate interaction energies for ionic liquid clusters	Submitted	90%

I have not renumbered sections of submitted or published papers in order to generate a consistent presentation within the thesis.

Signed: ██████████

Date: 17/10/2014

MONASH UNIVERSITY

Abstract

Faculty of Science
School of Chemistry

Doctor of Philosophy

Large-scale calculations of Ionic Liquids

by Jason Rigby

Ionic liquids (ILs) hold great promise in many fields including energy storage and generation, mechanical, pharmaceutical, synthetic and separation applications to name just a few. For any given application, the desired physical properties of the ideal IL may differ substantially from others and no widely applicable patterns or trends to facilitate intuitive design. The origins of physical properties lie in the characteristics of the intermolecular energetics, which consist of a complex interplay between electrostatic and dispersion forces. This thesis investigates and develops computational methodologies for calculating a reliable description of the intermolecular interactions for this challenging class of solvents and electrolytes of the future.

The electrostatic approximation used in classical molecular dynamics (MD) where atomic partial charges are assigned was investigated in terms of methods based on density matrix partitioning, and the restrained electrostatic potentials (RESP) approaches. It was found that the "geodesic" atomic partial charge scheme, part of the RESP family, produced the most accurate charges. This was measured in terms of (a) charge convergence with increasing basis set size; (b) charge invariance with changes to the coordinate system; (c) insensitivity to minor structural changes on the resulting charges; (d) adequate capture of charge transfer effects; and (e) the preservation of symmetric of charges in symmetric molecules. Although charges can vary dramatically depending on the scheme used, the careful use of atomic partial charge schemes may still produce reliable force-fields, or at least serve as a rapid diagnostic tool to quantify electrostatic interactions and charge transfer.

In moving towards unbiased *a priori* descriptions of IL intermolecular interactions, second-order Møller Plesset perturbation theory (MP2) was used with the linear-scaling fragment molecular orbital (FMO) framework to assess the extent to which dispersion forces play a role in the intermolecular energetics. ILs of increasing size were examined such that the many-body effects may be captured. The dispersion energy contribution

formed up to 20% of the total interaction energy. Furthermore, the interaction energy as produced by FMO was within 1 kJ mol^{-1} of the full-wavefunction MP2 interaction energy when three-body effects were included. As the dispersion interaction is purely a quantum mechanical phenomenon, correlated quantum mechanical methods, such as MP2 or coupled-cluster approaches, are required to provide an unbiased account of these effects.

Ab initio methods such as MP2 and CCSD(T) scale formally as N^5 an N^7, respectively, with chemical system size. While the FMO approach provides a marked improvement in efficiency, the counterpoise (CP) approach to correcting the basis set superposition error (BSSE) is not amenable to fragmented approaches and requires each ion in the cluster to be calculated in the basis set of the entire system. In order to remove this bottleneck, the spin-component scaled second-order Møller Plesset perturbation theory (SCS-MP2) methodology was refined by fitting 174 non-CP corrected interaction energies at the MP2/cc-pVTZ level of theory to CP corrected CCSD(T)/CBS benchmark energies. This has resulted in an implicit BSSE correction that may be used within the highly efficient FMO framework, and is shown to yield results on par with or exceeding the accuracy of MP2/cc-pVQZ for clusters of two and four ion pairs (IPs). This new approach as been termed SCS-IL-MP2.

An alternative dispersion corrected density functional theory (DFT) approach, DFT-D3, was assessed and refined in view of producing accurate interaction energies at the same CCSD(T)/CBS quality. The same test set of 174 ILs was used to fit the SCS-IL-MP2 approach was used to refit the DFT-D3 approach for both the Hartree-Fock (HF) wavefunction and the PBE and BLYP density functionals (DFs). In most cases, the selection of the DF and associated DFT-D version 3 (DFT-D3) parameters differed negligibly with all reaching within 1 to 2 kJ mol^{-1} per IP. HF-D3 parameters, on the other hand, showed a substantial improvement, particularly when used with the Becke-Johnson (BJ) damping function. Refitted HF-D3 and the BJ damping function was able to consistently provide interaction energy errors below 5 kJ mol^{-1} per IP. It would be worthwhile further investigating the application of the refined HF-D3 in its ability to produce reliable energies and geometries over a more diverse set of ILs.

Both the SCS-IL-MP2 and new DFT-D3 approaches may be applied to the highly scalable FMO framework. These form the core elements of the quantum chemistry toolbox for the study and understanding of the physicochemical properties of ILs that have so far been only superficially characterised by electronic structure theory. From this starting point, a rigorous and unbiased *a priori* understanding of the intermolecular interactions and resulting physicochemical properties may be predicted by means of efficient *ab initio* molecular dynamics (AIMD) techniques.

Acknowledgements

Contents

List of Figures

List of Tables

List of Abbreviations

[BF$_4$]$^-$ tetrafluoroborate.

[C$_1$mim]$^+$ 1,3-dimethylimidazolium.

[C$_1$mim]Br 1,3-dimethylimidazolium bromide.

[C$_1$mim]Cl 1,3-dimethylimidazolium chloride.

[C$_1$mim][BF$_4$] 1,3-dimethylimidazolium tetrafluoroborate.

[C$_1$mim][DCA] 1,3-dimethylimidazolium dicyanamide.

[C$_1$mpyr]$^+$ *N,N*-dimethylpyrrolidinium.

[C$_1$mpyr]Br *N,N*-dimethylpyrrolidinium bromide.

[C$_1$mpyr]Cl *N,N*-dimethylpyrrolidinium chloride.

[C$_1$mpyr][BF$_4$] *N,N*-dimethylpyrrolidinium tetrafluoroborate.

[C$_1$mpyr][DCA] *N,N*-dimethylpyrrolidinium dicyanamide.

[C$_2$mpyr]$^+$ *N*-methyl-*N*-ethylpyrrolidinium.

[C$_2$mpyr]Br *N*-methyl-*N*-ethylpyrrolidinium bromide.

[C$_2$mpyr]Cl *N*-methyl-*N*-ethylpyrrolidinium chloride.

[C$_2$mpyr][BF$_4$] *N*-methyl-*N*-ethylpyrrolidinium tetrafluoroborate.

[C$_2$mpyr][DCA] *N*-methyl-*N*-ethylpyrrolidinium dicyanamide.

[C$_2$py]$^+$ *N*-ethylpyridinium.

[C$_2$py]Br *N*-ethylpyridinium bromide.

[C$_2$py]Cl *N*-ethylpyridinium chloride.

[C$_2$py][BF$_4$] *N*-ethylpyridinium tetrafluoroborate.

[C$_2$py][DCA] *N*-ethylpyridinium dicyanamide.

[C$_3$m$_2$im]$^+$ 1-propyl-2,3-dimethylimidazolium.

[C$_3$m$_2$im]I 1-propyl-2,3-dimethylimidazolium iodide.

[C$_3$mim]$^+$ 1-propyl-3-methylimidazolium.

[C$_3$mim]I 1-propyl-3-methylimidazolium iodide.

[C$_4$mim][BF$_4$] 1-butyl-3-methylimidazolium tetrafluoroborate.

[C$_n$mim][BF$_4$] 1-alkyl-3-methylimidazolium tetrafluoroborate.

[C$_n$mpyr]$^+$ *N*-alkyl-3-methylpyridinium.

[DCA]$^-$ dicyanamide.

[NTf$_2$]$^-$ bis(trifluoromethylsulfonyl)amide.

[PF$_6$]$^-$ hexafluorophosphate.

[mes]$^-$ mesylate.

[tos]$^-$ tosylate.

AIM atoms-in-molecules.

AIMD *ab initio* molecular dynamics.

BJ Becke-Johnson.

BLYP Becke, Lee, Yang and Parr.

BO-AIMD Born-Oppenheimer AIMD.

BSSE basis set superposition error.

CBS complete basis set limit.

CCSD(T) coupled-cluster with single, double and perturbative triple excitations.

CHARMM Chemistry at HARvard Molecular Mechanics.

CI Configuration Interaction.

COMPASS Condensed-phase Optimized Molecular Potentials for Atomistic Simulation Studies.

CP counterpoise.

CP-AIMD Car-Parrinello AIMD.

CPU central processing unit.

DF density functional.

DFT density functional theory. **DFT-**

D dispersion corrected DFT. **DFT-**

D3 DFT-D version 3.

FMO fragment molecular orbital.

FMO2 two-body FMO.

FMO3 three-body FMO.

GGA generalised gradient approximation.

HF Hartree-Fock.

IL ionic liquid.

IP ion pair.

MAE mean absolute error.

MD molecular dynamics.

MP2 second-order Møller Plesset perturbation theory.

PBE Perdew, Burke and Ernzerhof.

PES potential energy surface.

RESP restrained electrostatic potentials.

RTIL room-temperature ionic liquid.

SAPT0 zeroth-order symmetry-adapted perturbation theory.

SCS-IL-MP2 ionic liquid specific SCS-MP2.

SCS-MP2 spin-component scaled second-order Møller Plesset perturbation theory.

To mum and dad

Interactive features

The 2D page is inherently limited when conveying 3D concepts. To this end, this thesis takes advantage of mobile technology to augment the experience by allowing the reader to access 3D representations of selected molecular structures and manipulate them in 3D space. This is achieved using QR code technology, which allows mobile devices to access web content by scanning a 2D barcode that contains the encoded URL. The QR codes are shown in the outer margin as shown in this example.[QR1] All QR codes are referenced in text with a superscript number prefixed with "QR" in square brackets, which can be seen in the previous sentence. If you do not have a compatible mobile device, the QR code caption is a hyperlink that will take you directly to the interactive content.

QR1

The 3D content has been optimised for mobile devices, including most mobile phones and tablet devices. In most cases, your device will require additional software to read the QR codes. No recommendations for specific software will be made here, however one may simply search for QR code readers on the appropriate platform (e.g. Google Play for Android, or the iTunes store for iOS devices) in order to find a suitable application. To test your device, please feel free to scan [QR1]. If you have a compatible device with the correct software installed, you should see a 3D model of caffeine, without which this thesis would not be possible.

It should be noted that the mobile content is designed to complement the concepts described in this thesis, and is not integral to understanding the content. Further, the structures are shown only as a qualitatively accurate visual guide and do not necessarily reflect the precise bond lengths and angles obtained from theory or experiment.

Interactive content was made possible using the Google App Engine and JSmol, which is part of the open-source Jmol package.

Chapter 1

Introduction

1.1 What are ionic liquids?

Ionic liquids (ILs) are a highly customisable and versatile class of organic compound composed entirely of ions. ILs are loosely defined as having a melting point of less than 100 $^\circ$C, but may be more specifically categorised as room-temperature ionic liquids (RTILs).[1] As illustrated by Figure 1.1,[QR1] ILs are uniquely positioned with, at a first approximation, significant contributions from both dispersion and electrostatic forces to the overall interaction energy.[2–4] Additionally, ILs tend to be bulky and asymmetric, which impedes packing ability and consequently their ability to crystallise.[5, 6]

QR1
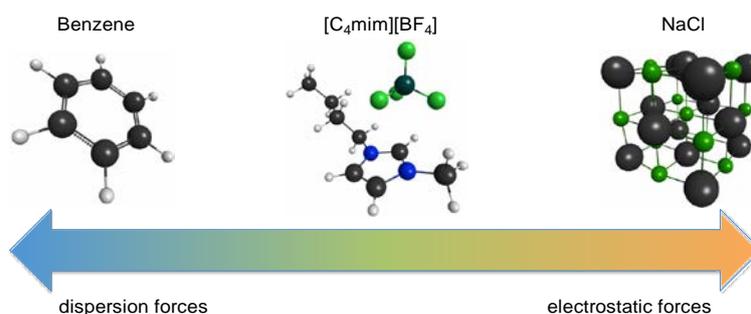
Figure 1.1: Ionic liquids (e.g. 1-butyl-3-methylimidazolium tetrafluoroborate, 1-butyl-3-methylimidazolium tetrafluoroborate ([C$_4$mim][BF$_4$])) compared with molecular organic solvents (e.g. benzene) and crystalline salts (e.g. sodium chloride)

A unique array of properties arise from the combination of non-negligible dispersion and electrostatic properties, and poor crystal packing. In general, ILs are electrically conductive,[7, 8] have low vapour pressure[9] and high thermal stability.[10] For these reasons, ILs are particularly well-suited to electrochemical applications including energy storage and solar cells,[11] and for use in place of petroleum-based mechanical

lubricants.[12, 13] Often referred to as "green solvents,"[14] ILs are typically of low toxicity,[15, 16] and may be used to catalyse a large number of reactions[17] with the possibility of being recycled[18–20] Further to this, ILs may be used in molecular separations,[21] sample storage and preservation,[22] or even as an alternative form for active pharmaceutical ingredients.[23] This is far from an exhaustive list, however it can be seen that ILs are applicable to a wide range of niche applications.
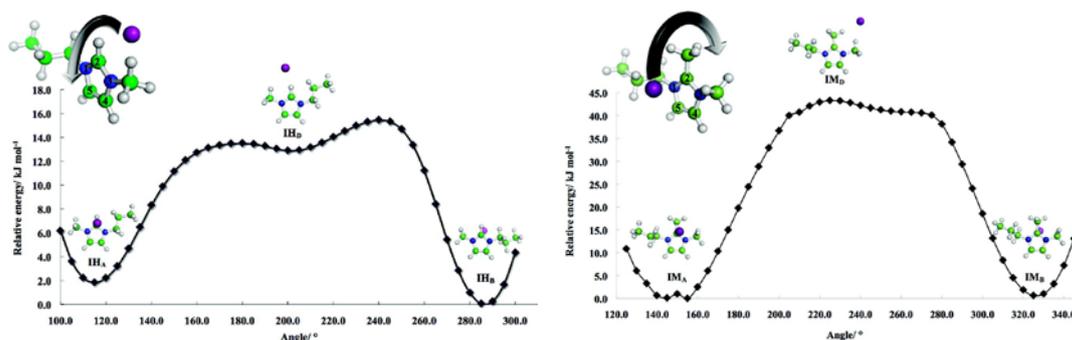
(a) 1-propyl-3-methylimidazolium iodide ([C$_3$mim]I) potential energy surface

(b) 1-propyl-2,3-dimethylimidazolium iodide ([C$_3$m$_2$im]I) potential energy surface

Figure 1.2: Comparison of potential energy surface scans between [C$_3$mim]I and [C$_3$m$_2$im]I. Reprinted with permission from Izgorodina *et al.*[24] Copyright 2011 American Chemical Society.

While the partial list of applications given above is remarkably diverse, the number of ILs that possess the ideal set of properties of each is not unlike finding a needle in a haystack. ILs are said to number in excess of 10^{18} when considering binary and ternary ion combinations,[14] where it is reasonable to assume that only a small proportion of these will satisfy the requirements for any particular application. Consequently, the need for the development of an accurate framework with which one may rationally design task-specific ILs is of paramount importance. It is not yet precisely known how the interactions between the constituent ions correlate with the physical properties for a given application, which is likely the result of a fine interplay between the fundamental intermolecular forces. An example to illustrate this is the increase in viscosity observed in ILs upon methylation of the 1-propyl-3-methylimidazolium ([C$_3$mim]$^+$) cation at the C2 position to form 1-propyl-2,3-dimethylimidazolium ([C$_3$m$_2$im]$^+$) cation.[QR2] It would seem logical to see the opposite trend; the ability for the acidic C2 proton to form hydrogen bonds is eliminated, thus interaction between ions would decrease and so too the viscosity. In this case, rather counter-intuitively, there is a marked difference in viscosity with only a subtle change in structure; measured at 85 °C, an almost six-fold increase in viscosity of 35 to 195 cP is observed.[24] This introduction does not intend to give a thorough account of the origins of this specific observation, although it is important to note that the explanation remains unclear. There have been a number of propositions given to rationalise this observation,[25–28] however debate appears to

QR2

focus on whether the lack of hydrogen bonding causes a loss of entropy due to a reduced conformational space[25] or whether ion mobility is reduced[24] given that the respective potential energy surface minima differing negligibly (see Figure 1.2).

Viscosity notwithstanding, melting points in general are notoriously difficult to predict. ILs tend to show only local trends that are either non-transferable to other systems or do not extrapolate generally. For example, for 1-alkyl-3-methylimidazolium tetrafluoroborate ([$C_n$mim][BF_4]) ILs, melting points generally decrease as the length of the alkyl chain increases due to increasing disorder and disrupted crystal packing.[29] However, as the alkyl chain length extends beyond 10 carbon atoms, the melting point quite abruptly increases as attractive van der Waals forces become more significant.[30] Contributions from the anion further complicate matters; N-alkyl-3-methylpyridinium ([$C_n$mpyr]$^+$) cations, for example, similarly show a general decrease in melting point with increasing chain length with the tosylate anion giving quite low melting points that steadily decrease to a cation alkyl chain length of at least four carbon atoms. Conversely, when the mesylate and tetrafluoroborate anions are present, an increase in melting point is observed at an alkyl chain length of as little as two and three carbon atoms, respectively.[3]

In cases such as these, the only tools available that can resolve the intermolecular behaviour at the atomic level are computational in nature, although the validity of the computational methods available vary substantially. In the following section, key methods of computational analysis will be discussed with specific reference to the strengths, weaknesses and applications to ILs.

1.2 Current methods of modelling

1.2.1 Classical molecular dynamics

1.2.1.1 Theoretical overview

Classical molecular dynamics (MD) is a time-resolved chemical simulation technique where the intramolecular potential is typically parameterised in terms of bond stretching, angle bending and torsional components, and the intermolecular interaction is governed by the 12-6 Lennard-Jones potential and Coulomb force, shown in Equation 1.1.[31] Collectively, these parameterisations are known as a forcefield. The classical Newtonian equations of motion are integrated using this potential to generate the simulation trajectory.

$$U(\text{non-covalent})_{\alpha\beta} = \sum_{i}\sum_{j} 4c\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right] + \frac{1}{4\pi c_0}\frac{q_i q_j}{r_{ij}} \qquad (1.1)$$

It can be seen that the non-covalent interaction potential is a function of monomer pairs (α and β) with dispersion being represented as a 12-6 Lennard-Jones potential, and electrostatics as the Coulombic interaction between non-bonding sites (i and j) of each monomer of the pair. The total non-bonded energy is simply a sum of all pair-wise interactions.

1.2.1.2   Implications and criticisms

Classical force-field calculations are an attractive choice simply since they generally have a computational complexity of $O(N^2)$ which can be reduced to $O(N \log N)$ [32] as opposed to a quantum chemical method such as second-order Møller Plesset perturbation theory (MP2) calculations which have a complexity of $O(N^5)$.[33] Both small-scale *ab initio* calculations and classical forcefield based methods have major drawbacks: the former, while capturing effects related to electron correlation, neglects significant contributions from longer-ranged electrostatic interactions as well as more local anion-anion/cation-cation dispersive interactions since the molecules that make up this interaction are simply not considered. The latter includes crude approximations for correlative effects that are either based on empirical data from spectroscopic measurements of model compounds (for example, the CHARMM forcefield optimised for DNA [34]), or parameterised based on system-specific small-scale *ab initio* calculations (for example, the COMPASS forcefield optimised for common organic molecules [35]) and operate on a purely pair-wise basis thereby neglecting many-body effects. That is, the total energy of the entire system is not equal to the sum of each interacting pair's energy. Many-body effects are investigated in detail in Chapter 2, where it is shown that many-body effects form a substantial part of the interactions found in IL. Consequently, this two-body approximation has severe deficiencies regardless of from where parameters were sourced.

Classical forcefields are generally fitted to reproduce some macroscopic property – typically density – and may not be easily transferred to new systems; the forcefield may need to be reparameterised for previously unaccounted for interactions and/or the nature of the atom-atom interactions may change as the simulation progresses.[36] Moreover, these forcefields do not allow for bond breakage or formation, and consequently are completely inappropriate for protic ionic liquids with labile protons. These complexities limit the use of classical dynamics to a more retrospective approach; they are a step forward in

explaining existing experimental observations, but are insufficient for exploring new and poorly understood systems and have questionable predictive power. The complex interplay between anion-anion, cation-cation and cation-anion interactive forces and their constituent components (such as dispersive, inductive and electrostatic forces) responsible for the unique physicochemical properties found in ionic liquids quickly put property prediction beyond the realm of classical physics.

1.2.2 Small-scale *ab initio* calaculations

1.2.2.1 Theoretical overview

The following is a brief overview of *ab initio* and density functional theory. For readers interested in the finer details, introductory textbooks such as Computational Chemistry by Jensen[37] or Molecular Electronic Structure Theory by Helgaker *et al.*[38] are highly recommended.

Ab initio calculations, small-scale or otherwise, all attempt to solve a form of the Schrödinger equation, a partial differential equation that describes the quantum state of a set of a physical system. Usually this is the time-independent Schrödinger equation, shown in Equation 1.2.

$$E\Psi = \hat{H}\Psi \tag{1.2}$$

The time-independent Schrödinger equation is an eigenvalue problem where the Hamiltonian operator, $\hat{H}$, operates on the wavefunction, Ψ, to produce the identical wavefunction scaled by the eigenvalue, E, which is the energy of the system.

For the vast majority of *ab initio* computational chemistry applications, the Schrödinger equation is restricted to only the electronic wavefunction with the nuclei treated as fixed point charges in order to reduce computational complexity. This is possible due to the Born-Oppenheimer approximation, which asserts that the mass of the nuclei is high enough that the Heisenberg uncertainty principle may be ignored; i.e. their wave-like nature is negligible. This allows the Hamiltonian to take the following form:

$$\hat{H}_{elec} = \hat{T}_e + \hat{V}_{en} + \hat{V}_{ee} + \hat{V}_{nn} \tag{1.3}$$

with $\hat{T}_e$ representing the kinetic energy of the electrons, $\hat{V}_{en}$ the electron-nucleus attraction, $\hat{V}_{ee}$ the electron-electron repulsion, and $\hat{V}_{nn}$ the nucleus-nucleus repulsion. The

electron-electron repulsion term is the most challenging to calculate as the repulsion experienced by any given electron is affected by all other electrons – a term that becomes intractable as the number of electrons increases. This is known as "electron correlation."

As a first approximation to solving the electronic Schrödinger equation, the Hartree-Fock (HF) method considers the electron-electron repulsion term for each electron acting only in the average field of all other electrons. An "exchange" term is also included to account for the interaction between electrons of the same spin, which is in addition to the Coulombic force. Although the exchange energy is a form of electron correlation, the HF method is typically considered an uncorrelated method. Thus, correlation energy is defined as the difference between the HF energy and any method accounting for Coulombic correlation.

The HF method is the foundation of many *ab initio* methods, all of which are categorised (rather unimaginatively) as "post-Hartree-Fock" methods. These include: MP2,[39] which is a Taylor series expansion of the Fock operator truncated the second order; coupled-cluster methods, which use the HF wavefunction as a reference to which excitation operators are applied up to an arbitrary order, accounting for all possible excitations to the given limit (e.g. coupled-cluster with single, double and perturbative triple excitations (CCSD(T)), truncated at triple-excitations);[40, 41] and Configuration Interaction (CI) methods,[42] which are at the pinnacle of accuracy where the assumption that the wavefunction is a single Slater determinant is disregarded in favour of multiple weighted Slater determinants. CI methods can be extended to include all possible electron excitations yielding the true solution to the Schrödinger equation; this is known as Full CI. There are many more *ab initio* methods than listed above, however apart from Full CI, these are arguably the most common in contemporary usage. *Ab initio* methods beyond CCSD(T) are seldom used owing to their computational complexity and limited computable chemical systems. However, the important point to note about *ab initio* methods is that they are systematically improvable to approach the true solution to whichever level of accuracy the computational power of the day provides.

Another approach to calculating electronic structure are the density functional theory (DFT) methods, which attempt to circumvent the electron correlation problem by asserting that there is a direct, albeit unknown, relationship between electron *density* and energy.[43] In Kohn-Sham DFT, approximate kinetic energy and electron correlation terms may be determined, however these approximations inadequately describe the electronic structure as they assume that electrons are *non-interacting*. Therefore, corrections need to be added in order to account for wavefunction antisymmetry, i.e. electron exchange energy, and additional electron-electron correlation beyond the average field as in HF theory – the *exchange-correlaton energy*. These corrections are themselves

unknown, however attempts to formulate such corrections are known as "density functionals" or "exchange-correlation functionals." In the GAUSSIAN 09 software package, there are over 200 density functionals available.[44]

In expressing the energy of a system in terms of electron density over three-dimensional space instead of the wavefunction, the dimensionality is substantially reduced. Whereas the wavefunction has dimensionality of $4N$, including the three-dimensional spatial and spin coordinates for each electron, electron density is expressed only in three spatial dimensions. This makes DFT an inexpensive alternative to wavefunction-based methods. However in contrast to wavefunction-based methods, a major criticism of DFT methods is that they are not systematically improvable. While for canonical implementations of wavefunction-based methods it can be said in general that with increased accuracy comes increased cost, the same cannot be said for DFT; there is no consistent hierarchy of DFT density functionals, although a notional "Jacob's ladder" description has been proposed.[45]

DFT is not completely without its merits, however; it is widely used to produce minimum energy structures. Energetics, on the other hand, can vary substantially.[46, 47] The origins of this poor behaviour stem from the assumptions made by the respective DFT functional which cannot be guaranteed to extend to all systems. Those for which the functional was designed generally perform exceptionally well, however as more novel chemicals are explored, the validity of the functional is not known. As a consequence, the best energetics are obtained from the wavefunction methods, which make only systematically improvable approximations.

1.2.2.2   Implications and criticisms

The majority of *ab initio* calculations of ionic liquids has been limited to ion pairs (IPs)[3, 48, 49] or small ion clusters.[50, 51] This is necessitated by most conventional wavefunction or DFT methods, which scale poorly with system size. The value of IP calculations is limited as it is far removed from the physical reality; ions are not in a gas phase environment, but rather surrounded by other ions of like or opposite charge and in constant motion. Supposing that ionic liquids did not have a significant contribution from many-body effects, this would be largely unproblematic, however a number of properties are affected by this.

In Chapter 3, it is shown that many-body effects are considerable in terms of the overall interaction energy, including in electron correlation effects. On a per IP basis, the interaction energy increases substantially from a single IP to clusters consisting of eight

IPs.[2] This is in part due to the polarisability of ionic liquids which allow for reinforced dipole moments and thus an increase in the induction contribution of the overall interaction,[52–54] much like the dipole moments in gas phase water compared to the liquid phase.

The behaviour of IL in the bulk compared with isolated ion pars or small clusters has significant implications on the classical approach detailed in Section 1.2.1. One of the key components of the classical intermolecular interaction forces are the electrostatic interactions between ions. Due to the many-body effects, this is a not a linearly additive quantity; that is, the charges of the ions change with the system size. In classical MD, this is represented by "charge transfer," or a reduction of the net charge on the constituent ions; instead of unity charges, which result in artificially slower dynamics, net charges of $0.8e$ and $-0.8e$ may be used.[55] In classical MD, polarisation is achieved using *atomic partial charges*, but this is fundamentally flawed in non-polarisable forcefields as these charges cannot change during the simulation and thus the true polarisability of the ions is not properly reflected. It should also be noted that the use of atomic partial charges in classical MD is simply a tool to model electrostatic interactions, but atomic partial charges themselves are artificial and without concrete definition.

Small-scale *ab initio* calculations have demonstrated by means of IPs and small clusters that many-body effects are important, but still do not truly reflect the environment as in experiment. Despite this deficiency, they have shown that classical approaches are inappropriate due to the polarisability of the ions and resulting induction forces and therefore indicate that best efforts should be placed on overcoming the scalability problems preventing the application of *ab initio* techniques on large systems. As ionic liquids are poorly characterised and often lack an intuitive structure-property relationship, models that assume such a relationship are likely to fail. Thus, wavefunction based methods are presently the most reliable way forward for accurate *a priori* characterisation.

1.2.3 *Ab initio* molecular dynamics

1.2.3.1 Theoretical overview

Ab initio molecular dynamics (AIMD) recognises the shortcomings of the classical approach and uses electronic structure methods to determine atomic forces that would otherwise be determined using ball-and-spring potentials, and two-body intermolecular potentials as in Section 1.2.1. AIMD is typically performed in one of two ways; the Carr-Parrinello method[56] or the Born-Oppenheimer method.

In Born-Oppenheimer AIMD (BO-AIMD), the forces are obtained by a standard DFT or wavefunction based method as the negative gradient on each nuclei. The standard Newtonian equations of motion are integrated using an algorithm such as Velocity Verlet or Leapfrog in a manner identical to the classical case. This is the most simple approach to AIMD, however as at each timestep a full wavefunction optimisation is required, it has not been widely adopted, with Car-Parrinello AIMD (CP-AIMD) gaining popularity as a more efficient method.

CP-AIMD involves only a single electronic structure optimisation as an initial step, as opposed to BO-AIMD, which requires optimisation at each step. This is typically done using a DFT approach. By introducing the electrons as an additional degree of freedom, both the orbitals and the nuclei are evolved as the simulation progresses. The consequence of this, however, is that in order to prevent the electrons from leaving the ground state, the electrons must be given an artificially increased mass, known as the "fictitious mass," which is set to a default of 200 a.u. in the CP-AIMD program.[57] Selection is critical to keep the electrons following the Born-Oppenheimer ground state surface, however a fictitious mass too large may lead to unconverged electronic properties and inaccurate dynamics.[58, 59]

1.2.3.2   Implications and criticisms

AIMD is currently at the forefront of IL simulation; it the least biased approach that requires no fitting to experimental data. It is intrinsically capable of accounting for charge transfer, induction and polarisation, and bond breakage and formation. Recent studies of ILs utilising AIMD include gas solubility studies[60–62] and analysis of IL mixtures,[63–65] to name just two areas of active research.

The overwhelming majority of AIMD simulations are calculated over a DFT potential energy surface, and thus contemporary AIMD may be argued somewhat of a misnomer; the DFT functional itself may be parameterised empirically and therefore is strictly speaking not a first-principles approach. With this in mind, while the current state of AIMD is a notable improvement over the classical approach, it still suffers from an unpredictable treatment of exchange and correlation energies.

With MD already sensitive to parameters such as the time step and simulation length, CP-AIMD further complicates the situation by introducing an artificial electronic mass, the selection of which is critical to physically realistic dynamics. Consequently the preferred AIMD technique is BO-AIMD, which does not have this parameter, however this is largely out of the realm of that which is considered computationally feasible. A central theme to this thesis involves overcoming computational bottlenecks involved in

ab initio electronic structure theory, and as such is it envisaged that BO-AIMD will increase in relevance in the near future.

1.3   Aims and overview

This thesis focuses on the computational bottlenecks that restrict the application of unbiased *ab initio* methods to large systems of ILs. It is structured in the following way:

Chapter 2 focuses on the calculation of atomic partial charges that are routinely used in classical forcefields. ILs are used as the subject of this study as they have been shown to be highly polarisable and subject to charge transfer. A variety of partial charge schemes are assessed to (a) highlight the variety of non-unique solutions to finding atomic partial charges, and (b) to use atomic partial charge schemes as a diagnostic tool for evaluating polarisability and charge transfer in large ionic clusters.

Chapter 3 introduces the fragment molecular orbital (FMO) approach[66] as applied to IL clusters. An evaluation of the error of FMO with respect to the MP2 level of theory is provided, which paves the way for widespread adoption of this large-scale technique for the calculation of accurate energetics reflecting the bulk properties of ILs.

Chapter 4 examines spin-component scaled second-order Møller Plesset perturbation theory (SCS-MP2) approaches, culminating in revised SCS-MP2 scaling coefficients that form the basis for the new ionic liquid specific SCS-MP2 (SCS-IL-MP2) approach proposed in this thesis. The SCS-IL-MP2 approach is then applied to larger clusters with remarkable improvements in energy calculations.

Chapter 5 considers non-equilibrium geometries and compares benchmark CCSD(T)/CBS correlation energies with MP2, SCS-MP2, SCS-IL-MP2 and DFT-D-type empirical dispersion corrections.[67, 68] The approach is then refined, with new empirical dispersion coefficients derived to reflect the dispersion energy of ionic liquid clusters resulting in a protocol for achieving high accuracy.

A core feature of this work is a focus on rigorous, systematic analysis. As such, the thesis structure outline above is designed to form the foundation of an electronic structure framework with which one may reliably form hypotheses based on theory with a well-founded understanding of the advantages and limitations.

References

[1] John S. Wilkes. A short history of ionic liquids-from molten salts to neoteric solvents. *Green Chemistry*, 4:73–80, 2002.

[2] Ekaterina I. Izgorodina, Jason Rigby, and Douglas R. MacFarlane. Large-scale ab initio calculations of archetypical ionic liquids. *Chemical Communications*, 48:1493–1495, 2012.

[3] Uditha L. Bernard, Ekaterina I. Izgorodina, and Douglas R. MacFarlane. New insights into the relationship between ion-pair binding energy and thermodynamic and transport properties of ionic liquids. *The Journal of Physical Chemistry C*, 114(48):20472–20478, 2010.

[4] Karina Shimizu, Mohammad Tariq, Margarida F. Costa Gomes, Luís P. N. Rebelo, and José N. Canongia Lopes. Assessing the dispersive and electrostatic components of the cohesive energy of ionic liquids using molecular dynamics simulations and molar refraction data. *The Journal of Physical Chemistry B*, 114(17):5831–5834, 2010.

[5] Ekaterina I. Izgorodina, Uditha L. Bernard, Pamela M. Dean, Jennifer M. Pringle, and Douglas R. MacFarlane. The madelung constant of organic salts. *Crystal Growth & Design*, 9(11):4834–4839, 2009.

[6] Pamela M. Dean, Jennifer M. Pringle, and Douglas R. MacFarlane. Structural analysis of low melting organic salts: perspectives on ionic liquids. *Physical Chemistry Chemical Physics*, 12:9144–9153, 2010.

[7] Douglas R. MacFarlane, Maria Forsyth, Ekaterina I. Izgorodina, Andrew P. Abbott, Gary Annat, and Kevin Fraser. On the concept of ionicity in ionic liquids. *Physical Chemistry Chemical Physics*, 11:4962–4967, 2009.

[8] Oldamur Hollóczki, Friedrich Malberg, Tom Welton, and Barbara Kirchner. On the origin of ionicity in ionic liquids. ion pairing versus charge transfer. *Physical Chemistry Chemical Physics*, 16(32):16880–16890, 2014.

[9] Luis P. N. Rebelo, José N. Canongia Lopes, José M. S. S. Esperança, and Eduardo Filipe. On the critical temperature, normal boiling point, and vapor pressure of ionic liquids. *The Journal of Physical Chemistry B*, 109(13):6040–6043, 2005.

[10] Helen L Ngo, Karen LeCompte, Liesl Hargens, and Alan B McEwen. Thermal properties of imidazolium ionic liquids. *Thermochimica Acta*, 357–358:97 − 102, 2000.

[11] Douglas R. MacFarlane, Naoki Tachikawa, Maria Forsyth, Jennifer M. Pringle, Patrick C. Howlett, Gloria D. Elliott, James H. Davis, Masayoshi Watanabe, Patrice Simon, and C. Austen Angell. Energy applications of ionic liquids. *Energy & Environmental Science*, 7:232–250, 2014.

[12] Zhuo Zeng, Benjamin S. Phillips, Ji-Chang Xiao, and Jean'ne M. Shreeve. Polyfluoroalkyl, polyethylene glycol, 1,4-bismethylenebenzene, or 1,4-bismethylene-2,3,5,6-tetrafluorobenzene bridged functionalized dicationic ionic liquids: Synthesis and properties as high temperature lubricants. *Chemistry of Materials*, 20(8):2719–2726, 2008.

[13] Chengfeng Ye, Weimin Liu, Yunxia Chen, and Laigui Yu. Room-temperature ionic liquids: a novel versatile lubricant. *Chemical Communications*, pages 2244–2245, 2001.

[14] Robin D. Rogers and Kenneth R. Seddon. Ionic liquids–solvents of the future? *Science*, 302(5646):792–793, 2003.

[15] Kathryn M. Docherty and Charles F. Kulpa, Jr. Toxicity and antimicrobial activity of imidazolium and pyridinium ionic liquids. *Green Chemistry*, 7:185–189, 2005.

[16] David J. Couling, Randall J. Bernot, Kathryn M. Docherty, JaNeille K. Dixon, and Edward J. Maginn. Assessing the factors responsible for ionic liquid toxicity to aquatic organisms via quantitative structure-property relationship modeling. *Green Chemistry*, 8:82–90, 2006.

[17] Thomas Welton. Room-temperature ionic liquids. solvents for synthesis and catalysis. *Chemical Reviews*, 99(8):2071–2084, 1999.

[18] Martyn J. Earle, Paul B. McCormac, and Kenneth R. Seddon. Diels-Alder reactions in ionic liquids . a safe recyclable alternative to lithium perchlorate-diethyl ether mixtures. *Green Chemistry*, 1:23–25, 1999.

[19] Hua-Ping Zhu, Fan Yang, Jie Tang, and Ming-Yuan He. Bronsted acidic ionic liquid 1-methylimidazolium tetrafluoroborate: a green catalyst and recyclable medium for esterification. *Green Chemistry*, 5:38–39, 2003.

[20] Nicolas Audic, Hervé Clavier, Marc Mauduit, and Jean-Claude Guillemin. An ionic liquid-supported ruthenium carbene complex: A robust and recyclable catalyst for ring-closing olefin metathesis in ionic liquids. *Journal of the American Chemical Society*, 125(31):9248–9249, 2003.

[21] Xinxin Han and Daniel W. Armstrong. Ionic liquids in separations. *Accounts of Chemical Research*, 40(11):1079–1086, 2007.

[22] Nolene Byrne, Li-Min Wang, Jean-Philippe Belieres, and C. Austen Angell. Reversible folding-unfolding, aggregation protection, and multi-year stabilization, in high concentration protein solutions, using ionic liquids. *Chemical Communications*, pages 2714–2716, 2007.

[23] Jelena Stoimenovski, Douglas R. MacFarlane, Katharina Bica, and Robin D. Rogers. Crystalline vs. ionic liquid salt forms of active pharmaceutical ingredients: A position paper. *Pharmaceutical Research*, 27(4):521–526, 2010.

[24] Ekaterina I. Izgorodina, Radha Maganti, Vanessa Armel, Pamela M. Dean, Jennifer M. Pringle, Kenneth R. Seddon, and Douglas R. MacFarlane. Understanding the effect of the C2 proton in promoting low viscosities and high conductivities in imidazolium-based ionic liquids: Part i. weakly coordinating anions. *The Journal of Physical Chemistry B*, 115(49):14688–14697, 2011.

[25] Patricia A. Hunt. Why does a reduction in hydrogen bonding lead to an increase in viscosity for the 1-butyl-2,3-dimethyl-imidazolium-based ionic liquids? *The Journal of Physical Chemistry B*, 111(18):4844–4853, 2007. PMID: 17388550.

[26] Koichi Fumino, Alexander Wulf, and Ralf Ludwig. Strong, localized, and directional hydrogen bonds fluidize ionic liquids. *Angewandte Chemie International Edition*, 47(45):8731–8734, 2008.

[27] Hualin Li, Murvat Ibrahim, Ismail Agberemi, and Mark N. Kobrak. The relationship between ionic structure and viscosity in room-temperature ionic liquids. *The Journal of Chemical Physics*, 129(12), 2008.

[28] Yong Zhang and Edward J. Maginn. The effect of C2 substitution on melting point and liquid phase dynamics of imidazolium based-ionic liquids: insights from molecular dynamics simulations. *Physical Chemistry Chemical Physics*, 14:12157–12164, 2012.

[29] Hiroyuki Tokuda, Kikuko Hayamizu, Kunikazu Ishii, Md. Abu Bin Hasan Susan, and Masayoshi Watanabe. Physicochemical properties and structures of room temperature ionic liquids. 2. variation of alkyl chain length in imidazolium cation. *The Journal of Physical Chemistry B*, 109(13):6103–6110, 2005.

[30] John D. Holbrey and Kenneth R. Seddon. The phase behaviour of 1-alkyl-3-methylimidazolium tetrafluoroborates; ionic liquids and ionic liquid crystals. *Journal of the Chemical Society, Dalton Transactions*, pages 2133–2140, 1999.

[31] Wendy D. Cornell, Piotr Cieplak, Christopher I. Bayly, Ian R. Gould, Kenneth M. Merz, David M. Ferguson, David C. Spellmeyer, Thomas Fox, James W. Caldwell,

and Peter A. Kollman. A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *Journal of the American Chemical Society*, 117(19):5179–5197, 1995.

[32] Blake G. Fitch, Aleksandr Rayshubskiy, Maria Eleftheriou, T. J. Christopher Ward, Mark Giampapa, Michael C. Pitman, and Robert S. Germain. Blue matter: approaching the limits of concurrency for classical molecular dynamics. In *Proceedings of the 2006 ACM/IEEE conference on Supercomputing*, SC '06, New York, NY, USA, 2006. ACM.

[33] David E. Bernholdt. Scalability of correlated electronic structure calculations on parallel computers: A case study of the RI-MP2 method. *Parallel Computing*, 26(7-8):945 – 963, 2000.

[34] Nicolas Foloppe and Alexander D. MacKerell, Jr. All-atom empirical force field for nucleic acids: I. parameter optimization based on small molecule and condensed phase macromolecular target data. *Journal of Computational Chemistry*, 21(2):86–104, 2000.

[35] H. Sun. COMPASS: An ab initio force-field optimized for condensed-phase applications overview with details on alkane and benzene compounds. *The Journal of Physical Chemistry B*, 102(38):7338–7364, 1998.

[36] D. Marx and J. Hutter. *Ab initio molecular dynamics: basic theory and advanced methods.* Cambridge University Press, 2009.

[37] F. Jensen. *Introduction to Computational Chemistry*. Wiley, 2013.

[38] T. Helgaker, J. Olsen, and P. Jorgensen. *Molecular Electronic-Structure Theory*. Wiley, 2013.

[39] Chr. Møller and M. S. Plesset. Note on an approximation treatment for many-electron systems. *Physical Review*, 46:618–622, 1934.

[40] Jiří Čížek. On the correlation problem in atomic and molecular systems. calculation of wavefunction components in urselltype expansion using quantumfield theoretical methods. *The Journal of Chemical Physics*, 45(11):4256–4266, 1966.

[41] George D. Purvis and Rodney J. Bartlett. A full coupledcluster singles and doubles model: The inclusion of disconnected triples. *The Journal of Chemical Physics*, 76(4):1910–1918, 1982.

[42] C. David Sherrill and Henry F. Schaefer III. The configuration interaction method: Advances in highly correlated approaches. volume 34 of *Advances in Quantum Chemistry*, pages 143 – 269. Academic Press, 1999.

[43] W. Kohn and L. J. Sham. Self-consistent equations including exchange and corre-lation effects. *Physical Review*, 140:A1133–A1138, 1965.

[44] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, Ö. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, and D. J. Fox. Gaussian09 Revision D.01. Gaussian Inc. Wallingford CT 2009.

[45] John P. Perdew and Karla Schmidt. Jacob's ladder of density functional approxima-tions for the exchange-correlation energy. *AIP Conference Proceedings*, 577(1):1–20, 2001.

[46] Ekaterina I. Izgorodina, Uditha L. Bernard, and Douglas R. MacFarlane. Ion-pair binding energies of ionic liquids: Can DFT compete with ab initio-based methods? *The Journal of Physical Chemistry A*, 113(25):7064–7072, 2009. PMID: 19462960.

[47] Stefan Zahn, Douglas R. MacFarlane, and Ekaterina I. Izgorodina. Assessment of Kohn-Sham density functional theory and Møller-Plesset perturbation theory for ionic liquids. *Physical Chemistry Chemical Physics*, 15:13664–13675, 2013.

[48] Matthew T. Clough, Karolin Geyer, Patricia A. Hunt, Jurgen Mertes, and Tom Welton. Thermal decomposition of carboxylate ionic liquids: trends and mecha-nisms. *Physical Chemistry Chemical Physics*, 15:20480–20495, 2013.

[49] Ekaterina I. Izgorodina, Dorothea Golze, Radha Maganti, Vanessa Armel, Maria Taige, Thomas J. S. Schubert, and Douglas R. MacFarlane. Importance of dis-persion forces for prediction of thermodynamic and transport properties of some common ionic liquids. *Physical Chemistry Chemical Physics*, 16:7209–7221, 2014.

[50] Richard P. Matthews, Tom Welton, and Patricia A. Hunt. Competitive pi interac-tions and hydrogen bonding within imidazolium ionic liquids. *Physical Chemistry Chemical Physics*, 16:3238–3253, 2014.

[51] Richard P Matthews, Claire Ashworth, Tom Welton, and Patricia A Hunt. The impact of anion electronic structure: similarities and differences in imidazolium based ionic liquids. *Journal of Physics: Condensed Matter*, 26(28):284112, 2014.

[52] Stefan Zahn, Frank Uhlig, Jens Thar, Christian Spickermann, and Barbara Kirchner. Intermolecular forces in an ionic liquid ([Mmim][Cl]) versus those in a typical salt (NaCl). *Angewandte Chemie International Edition*, 47(19):3639–3641, 2008.

[53] Katharina Wendler, Stefan Zahn, Florian Dommert, Robert Berger, Christian Holm, Barbara Kirchner, and Luigi Delle Site. Locality and fluctuations: Trends in imidazolium-based ionic liquids and beyond. *Journal of Chemical Theory and Computation*, 7(10):3040–3044, 2011.

[54] Jason Rigby and Ekaterina I. Izgorodina. Assessment of atomic partial charge schemes for polarisation and charge transfer effects in ionic liquids. *Physical Chemistry Chemical Physics*, 15:1632–1646, 2013.

[55] Oleg Borodin. Polarizable force field development and molecular dynamics simulations of ionic liquids. *The Journal of Physical Chemistry B*, 113(33):11463–11478, 2009. PMID: 19637900.

[56] R. Car and M. Parrinello. Unified approach for molecular dynamics and density-functional theory. *Physical Review Letters*, 55:2471–2474, 1985.

[57] Wanda Andreoni and Alessandro Curioni. New advances in chemistry and materials science with CPMD and parallel computing. *Parallel Computing*, 26(7–8):819 – 842, 2000.

[58] Sergei Izvekov and Gregory A. Voth. Ab initio molecular-dynamics simulation of aqueous proton solvation and transport revisited. *The Journal of Chemical Physics*, 123(4), 2005.

[59] Paul Tangney. On the theory underlying the Car-Parrinello method and the role of the fictitious mass parameter. *The Journal of Chemical Physics*, 124(4), 2006.

[60] Dzmitry S. Firaha and Barbara Kirchner. CO_2 absorption in the protic ionic liquid ethylammonium nitrate. *Journal of Chemical & Engineering Data*, Article ASAP, 2014.

[61] Bo Han, Yubao Sun, Maohong Fan, and Hansong Cheng. On the CO_2 capture in water-free monoethanolamine solution: An ab initio molecular dynamics study. *The Journal of Physical Chemistry B*, 117(19):5971–5977, 2013.

[62] Oldamur Hollóczki, Dzmitry S. Firaha, Joachim Friedrich, Martin Brehm, Richard Cybik, Martin Wild, Annegret Stark, and Barbara Kirchner. Carbene formation in ionic liquids: Spontaneous, induced, or prohibited? *The Journal of Physical Chemistry B*, 117(19):5898–5907, 2013.

[63] Martin Brehm, Henry Weber, Alfonso S. Pensado, Annegret Stark, and Barbara Kirchner. Liquid structure and cluster formation in ionic liquid/water mixtures – an extensive ab initio molecular dynamics study on 1-ethyl-3-methylimidazolium acetate/water mixtures – part 2. *Zeitschrift für Physikalische Chemie International*, 227:177–203, 2013.

[64] Martin Thomas, Martin Brehm, Oldamur Hollóczki, Zsolt Kelemen, László Nyulászi, Tibor Pasinszki, and Barbara Kirchner. Simulating the vibrational spectra of ionic liquid systems: 1-ethyl-3-methylimidazolium acetate and its mixtures. *The Journal of Chemical Physics*, 141(2), 2014.

[65] Mohammad Hadi Ghatee and Amin Reza Zolghadr. Local depolarization in hydrophobic and hydrophilic ionic liquids/water mixtures: Car–Parrinello and classical molecular dynamics simulation. *The Journal of Physical Chemistry C*, 117(5):2066–2077, 2013.

[66] Kazuo Kitaura, Eiji Ikeo, Toshio Asada, Tatsuya Nakano, and Masami Uebayasi. Fragment molecular orbital method: an approximate computational method for large molecules. *Chemical Physics Letters*, 313(3–4):701 – 706, 1999.

[67] Stefan Grimme. Accurate description of van der waals complexes by density functional theory including empirical corrections. *Journal of Computational Chemistry*, 25(12):1463–1473, 2004.

[68] Stefan Grimme, Jens Antony, Stephan Ehrlich, and Helge Krieg. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *The Journal of Chemical Physics*, 132(15), 2010.

Chapter 2

The trouble with classical mechanics:
Partial charge schemes, polarisability and charge transfer

2.1   Declaration for thesis Chapter 2

Declaration by Candidate

In the case of Chapter 2, the nature and extent of my contribution to the work was the following:

Nature of contribution	Extent of contribution(%)
Key ideas, data acquisition, data analysis, manuscript development and writing up	90

The following co-authors contributed to the work. If co-authors are students at Monash University, the extent of their contribution in percentage terms must be stated:

Name	Nature of contribution	Extent of contribution (%) for student co-authors only
Dr. Ekaterina I. Izgorodina	Key ideas, introduction write-up, proofreading and final drafting	–

The undersigned hereby certify that the above declaration correctly reflects the nature and extent of the candidates and co-authors contributions to this work.

Candidate's Signature:

Date: 17/10/2014

Supervisor's Signature:

Date: 17/10/2014

2.2   Overview

A core feature of the description of intermolecular forces in any classical molecular dynamics simulation are atomic partial charges, which are used to characterise the electrostatic component of the non-covalent interaction (see Equation 1.1). Atomic partial charges are problematic, however, as they do not correspond to any quantum mechanical operator; there is no charge operator in quantum mechanics.* This means that any number of charges may be assigned to each atom that, when summed, reflect the overall charge of the molecule but to not accurately reflect the charge distribution.

The question is, what properties must an atomic partial charge distribution have to be considered accurate? This is not a straightforward question to answer simply because there is no definitive definition of an atom in a molecule;[2] at its essence, a molecule is a collection of positively charged nuclei supported by negatively charged electrons – there is no discrete atomic unit. For this study, atomic partial charges are considered accurate if they satisfy the following criteria:

- Convergence with increasing basis set size

- Invariant with respect to coordinate system

- No major fluctuations with minor structural changes

- Ability to capture charge transfer effects

- Accurate treatment of symmetric molecules

In terms of the above, it was concluded that charges derived from the restrained electrostatic potentials (RESP) with geodesic point selection[3] were generally best suited to this. Conversely, schemes based on the wavefunction derived density matrix performed poorly. Mulliken[4] and Löwdin[5] charges suffered from severe basis set dependence, and all fluctuated drastically compared to the RESP schemes and showed low levels of charge transfer. Natural population analysis[6, 7] was relatively stable with basis set, however could not capture charge transfer. The only outstanding property of the density matrix-based methods was the treatment of charge symmetry, which was perfectly symmetrical about the mirror plane bisecting 1,3-dimethylimidazolium chloride ([C_1mim]Cl). Further to this, the dipole moment distribution was determined for each

---

*The atoms-in-molecules (AIM) approach *does* define an atom electron population operator; the "number operator."[1] While this operator has a unique solution, this does not mean that the atomic charges have a unique solution. Charges determined via the number operator are only valid within the AIM formalism, which does not reflect on its suitability in describing electrostatic interactions.

partial charge scheme in a series of ionic liquid clusters. Broad distributions were found that supported previous findings in this area.[8, 9]

This study highlights the importance of good partial charge selection to achieve the maximum possible accuracy in an classical molecular dynamics (MD) simulation. However, the broad dipole moment distribution found in ionic liquid (IL) clusters indicates that non-polarisable forcefields are fundamentally limited, whereas polarisable forcefields are time-consuming to fit and not easily transferred to structurally distinct systems. Thus, this paper is positioned to emphasise the need for *ab initio* approaches, which require no such approximation. Partial charge schemes may be useful as a diagnostic tool, but special care must be taken where their usage is fundamental to the simulation.

3D structures of IL clusters shown in Figures 9 (a), (b) and (c) in the paper that follows may be viewed by scanning [QR1], [QR2] and [QR3], respectively.

QR1

QR2

QR3

PCCP

RSC Publishing

PAPER

Assessment of atomic partial charge schemes for polarisation and charge transfer effects in ionic liquids †

Jason Rigby and Ekaterina I. Izgorodina*

In this work, we assess several popular atomic partial charge schemes with the view of accurately quantifying charge distribution, dipole moments and charge transfer in routinely used ionic liquids (ILs). We investigated a series of ion pairs of imidazolium-based ILs such as $[C_{(1-4)}mim]X$ (where X = Cl, BF_4 and NTf_2) and ionic clusters of $[NMe_4][BF_4]$, $[C_1mim][BF_4]$ and $[C_1mim]Cl$ that were composed of two, four and eight ion pairs. Assessed partial charge schemes include restrained electrostatic potentials (RESP) employing ChelpG, Connolly and Geodesic point selection algorithms, as well as density matrix partitioning schemes including Mulliken, Löwdin and Natural Population Analysis (NPA). The quality of charge distribution was analysed using the following criteria: (1) treatment of symmetry identical atoms, (2) invariance of charge in the imidazolium ring with increasing alkyl chain and (3) recalculation to dipole moments as a measure of electronic polarisation. The RESP schemes such as Connolly and Geodesic clearly outperform the ChelpG scheme as well as the density matrix-based schemes for these three criteria. Calculated partial charges reveal that dipole moments were best represented by the RESP schemes and confirmed the presence of charge transfer in ILs to a various degree. The degree of charge transfer was dependent on anions as well as cluster size. In the ion pairs, the chloride anion showed the largest charge transfer, followed by the NTf_2 and BF_4 anions. In ionic clusters the charge transfer was shown to gradually converge from two to eight ion pairs in the case of the $[NMe_4][BF_4]$ and $[C_1mim][BF_4]$ ILs to a value, close to that for corresponding ion pairs. In contrast, charge transfer in the $[C_1mim]Cl$ clusters converges to a lower value, showing an unusually strong inter-ionic bond with the chloride anion. NPA charges were found to perform poorly, with near-unity charges being retained on the anions and cations in ion pairs and ionic clusters. Mulliken and Löwdin charges were shown to be highly basis set dependent and unpredictable with marked fluctuations in partial charges and therefore their use for ILs is particularly discouraged. Ability of the partial charge schemes to capture fluctuations in the dipole moment within the ionic clusters was also examined. The Connolly and Geodesic RESP schemes were found to slightly outperform ChelpG. Evidence to suggest that chloride-based ILs might be poor model systems for ILs is also presented.

1 Introduction

Quantum chemical methods have become powerful tools for predicating a number of molecular properties including dipole moments and electronic polarisability. The reason behind this success is the availability of the well-defined operators that are used to calculate quantum chemical averages of physical quantities such as electronic energy and dipole moment. One of the most challenging concepts in quantum chemistry is that of atomic charge within a molecule. Firstly, the charge operator does not exist in quantum mechanics and secondly, since all the electrons within the molecule are treated as indistinguishable the definition of the individual atom in the molecule becomes rather subjective.

Ionic liquids (ILs) consisting entirely of ions represent an additional challenge for quantum chemistry, as the prediction of atomic charges for charged species adds to the ambiguity due to an arbitrary choice for the centre of the overall charge.

School of Chemistry, Monash University, Wellington Road, Clayton VIC 3800, Australia. E-mail:

Classical molecular dynamics (MD) simulations of ILs strongly rely on the accuracy of atomic charges due to the importance of electrostatic interactions between ions. These electrostatic interactions are treated according to Coulomb's law and the charge distribution for each ion is used to determine the overall contribution from electrostatics. Although organic-type ions constituting ILs reside at distances >3 Å, the orbital overlap is still possible to some extent,[1,2] thus resulting in a non-negligible charge transfer from the anion to the cation and hence reducing the overall charge below unity on each ion.[3] It was shown by a few research groups that downscaling of the unity charge on ions was essential for the reliable prediction of dynamic properties of ILs such as self-diffusion coefficients and viscosity, as the reduction in electrostatic interactions allowed for faster ion dynamics and better agreement with experimental data.[4–6] Further, a number of experimental studies carried out on imidazolium- and pyrrolidinium-based ionic liquids using X-ray photoelectron spectroscopy showed a shift in binding energies of electrons within the cation as a funciton of anions. The observed trends were interpreted as charge-transfer between the cation and the anion.[7,8] The situation becomes more complicated as polarisation of IL ions arising from induced dipole moments was shown to also play an important role in the prediction of bulk properties.[9] Borodin and Smith[4] identified that for classical MD simulations to be reliable, many-body polarisable force fields needed to allow for the charge fluctuation in ions during a simulation. The downside of these simulations is their cost and the need to develop new polarisable force fields for structurally different IL ions to the ones studied before.[10,11] Ab initio (DFT) MD (AIMD) simulations are paving the way towards un-biased bulk simulations of ILs as the fluctuation in charge distribution is accounted auto-matically by performing an electronic structure calculation on-the-fly for each ionic arrangement. Ab initio (DFT) MD (AIMD) simulations of small- to medium-sized clusters from 8 ion pairs to 30, 48 and 64 ion pairs of imidazolium-based ILs[12–14] coupled with the chloride, SCN and dicyanamide anions showed

a broad dipole moment distribution for both cations and anions. The charges on cations significantly deviated from the unity charge and fell in the range of 0.55 to 0.7 e. Interestingly, the results from clusters containing 8 and over 30 ion pairs appeared to be rather similar, suggesting localisation of the charge fluctuation to neighbouring ions that are in the direct contact with the reference ion.[12,13] Fluctuations of the partial atomic charges in the AIMD simulations on 30 ion pairs of the [C_1mim]Cl IL highlighted the importance of electronic polarisa-tion,[14] which classical MD simulations with the fixed charges cannot capture. The dipole moments calculated from the corres-ponding atomic charges showed similar distributions for the C_1mim and C_2mim cations in the [C_1mim]Cl and [C_2mim]Cl ILs, respectively, with the dipole moment distribution of the anion being almost identical regardless of the alkyl chain length on the cation.[12] The tremendous expense of AIMD simulations inhibits their progress for the prediction of bulk properties of condensed systems such as ILs, as MD runs on the ns scale are not computationally feasible yet. Therefore, the accurate description of the charge distribution in ions is needed to further under-stand the implications of polarisation on properties of ILs.

The practicality of the atomic charge prediction relies on two major approaches:[15] (1) separation of the one particle density matrix and (2) separation of the one-particle electron density (see Fig. 1).

The one-particle density matrix approach pioneered in quantum mechanics after the development of the Hartree–Fock theory. Mulliken population analysis was the first scheme for predicting individual atomic charges and became very popular due to its simplicity.[16,17] Within the scheme atomic charge is defined as the difference between the total number of electrons in the ground state of the free neutral atom and the gross atomic population on the atom within a molecule. The latter depends on the equally shared overlap population with the adjacent atoms, that is best described as when the boundary between atoms is placed in the middle of the bond. The scheme cannot be considered qualitative as Mulliken himself
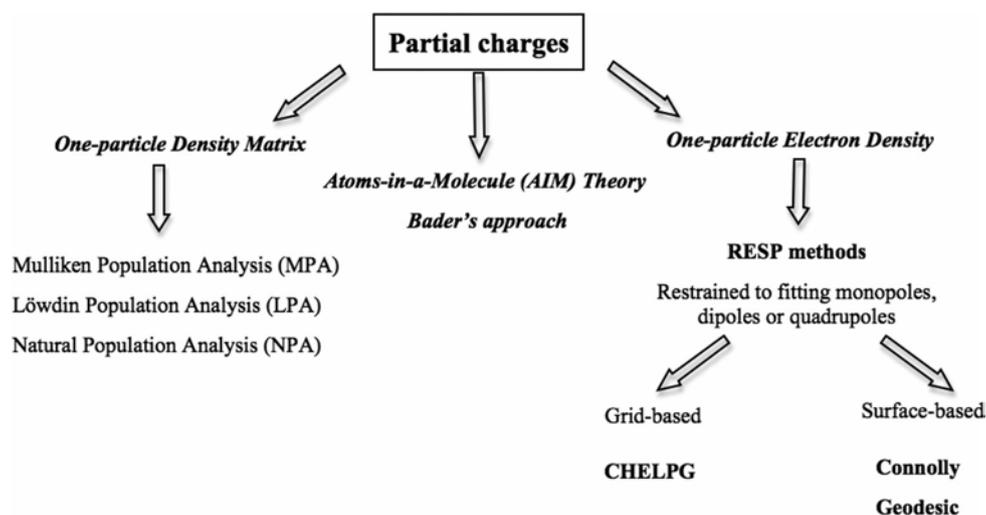
Fig. 1 Overview of common partial charge schemes.

This journal is © the Owner Societies 2013

Phys. Chem. Chem. Phys., 2013, **15**, 1632–1646 | **1633**

highlighted the fact that "fundamentally there is no such thing as an atom in a molecule except in an approximate sense".[16] The Löwdin population analysis (LPA) scheme is fundamentally identical to that of Mulliken's with the only difference arising from the use of an orthogonalised basis set.[18] This requirement makes the LPA charges more sensible with respect to the number of electrons in each atomic orbital, whereas MPA can produce gross orbital populations that are either negative or more than two. It is well accepted that both population schemes can result in unreasonable atomic charges due to incompleteness of atom-centred basis sets and the presence of diffuse functions that do not resemble atomic orbitals.[19] Fundamentally, both MPA and LPA (except for special cases[19]) schemes are invariant to a general orbital transformation, making atomic charges independent of the coordinate system. As a consequence, atomic charges of the symmetry equivalent atoms in a molecule are identical. Natural Population Analysis (NPA) proposed by Weinhold and Reed[20,21] was designed to fix the problems existing in the Mulliken scheme by constructing a set of natural atomic orbitals (NAOs) in an arbitrary atomic basis set. The construction of these natural orbitals begins with occupancy-weighted symmetric orthogonalisation of atom-centred basis functions that are separated into a highly occupied natural minimal basis and a largely unoccupied natural Rydberg basis.[22] This procedure produces an orthonomal set of atomic orbitals that retain a great degree of their atom-centred features, thus ensuring that the shape of the strongly occupied orbitals is preserved better than that of the weakly occupied ones (referred to as Rydberg orbitals) that consist of many diffuse orbitals from using extended basis sets. The diagonal elements of the density matrix made from the NAOs represent the electron occupancies of each NAO summing exactly to the total number of electrons. The natural orbitals are known to be inherent to the wavefunction, rather the quality of the basis set. The NPA scheme has been recognized as a reliable tool for calculating atomic charges and has been widely used for studying a number of chemical systems, including neutral and ionic complexes of the donor–acceptor type.[21] For more information, see a number of useful publications comparing the performance of these density matrix schemes to study atomic charges in various classes of neutral molecules.[23–25]

The one-particle electron density approach is based on fitting atomic point charges to reproduce the electrostatic potential (ESP) of a molecule. The electrostatic potential is defined via the density matrix and, therefore, is directly accessible from HF calculations. Both ab initio[26] and semi-empirical[27,28] methods have been used to generate the potentials. It has to be noted that semi-empirical methods are used less frequently due to increased computer power allowing for first principle HF calculations. When the ESP methods were first introduced, it was obvious that some of the fitted charges were poorly reproduced, strongly depending on the molecular orientation and conformation and having poor transferability between common groups of homologous molecules.[29–31] The main reason behind these observations lied in the statistical nature of the fitting process resulting in

an insufficient number of points on the ESP surface to treat atoms "buried" in the molecular structure, e.g. carbon atoms in an alkyl chain.[32,33] Although various solutions were introduced to address these issues,[26,29] the most robust technique consisted of constraining the fitting procedure to reproduce a target charge on non-hydrogen atoms.[32,33] The proposed restraints improve the quality of fitted atomic charges, especially those that are least well determined by the electro-static potential such as in non-polar functional groups.[33] The restrained ESP (RESP) methods can also be made to reproduce other molecular properties such as dipole and quadrupole moments.[29,32] There are three major algorithms used for calculating RESP atomic charges: ChelpG,[34,35] Connolly[27–30,33,36] and Geodesic.[37] The main difference stems from the way these algorithms select the points on the ESP surface for the fitting procedure. The ChelpG algorithm is grid-based and selects regularly spaced points by defining a cube of points around the molecule spaced 0.3 to 0.8 Å apart at the distance of 2.8 Å away from the nuclei. The points that fall inside the molecular van der Waals radius are excluded from the fitting procedure to avoid large distortions due to the proximity to the nuclei. In the Connolly algorithm a spherical surface of points is computed around each atom at a probe radius that is chosen as a multiple of the van der Waals radius of the atom.[27] The molecular surface is constructed by combining the individual atomic surfaces and discarding the points within the chosen multiple of the van der Waals radius of any of the atoms.[27] This algorithm generates the surface that does not change smoothly with increasing probe radius, displaying rather sharp peaks and hence, anisotropies.[37] It has to be noted that in the GAUSSIAN software the Connolly algorithm is known as the MK scheme.[38] Both ChelpG and Connolly schemes demonstrate a fair degree of dependence on molecular orientation.[37,39] The Geodesic scheme was designed to eliminate this problem by smoothing out the position of the Connolly points (for more details about the Geodesic algorithm see ref. 37). This strategy was shown to reduce the dependence of the derived atomic charges on molecular orientation.[37] A promising new family of charge schemes known as the electrostatically embedded many-body method[40] and the ESP-based screened charge model[41] have been introduced by the Truhlar group as an improvement to traditional RESP schemes, however owing to their recent development, these methods are not yet implemented in popular computational chemistry software packages and hence have been excluded from this study. Reliability of these novel methods for predicting charges in ionic liquid ions certainly needs to be performed in the future.

A relatively recent approach that borders two major approaches for calculations of charge distribution as shown in Fig. 1 is that of Richard Bader's, referred to as Atoms-in-Molecules (AIM) theory. Bader proposed to partition the electronic charge density distribution of a molecule, an experi-mentally measurable quantity, into atomic volumes whose surfaces had a zero flux of the charge density gradient.[42] It was shown that within these volumes the quantum (atomic) sub-systems obeyed a local virial relation and therefore, quantities such as atomic charges, atomic dipole moments etc. were well defined within the quantum mechanical

1634 | *Phys. Chem. Chem. Phys.*, 2013, **15**, 1632–1646

This journal is © the Owner Societies 2013

formulation.[43] For example, integration over each atomic volume generates atomic population that can be recalculated to atomic charge. For each atomic sub-system the maximum of the charge density occurs around the nucleus, decaying quite rapidly in any direction away from the centre of the nucleus. The cornerstone of the AIM theory lies in the accurate identification of critical points on the charge density surface with a zero flux gradient, which becomes a very complex procedure for condensed phases consisting of a number of molecular species.[44] It has to be noted that for ILs the organic nature of individual ions represents an additional challenge due to subtle charge transfer effects between ions. Although a significant improvement of the Bader charge density analysis has recently been proposed by Henkelman et al.[43,45] charge transfer effects have only been found in ionic systems with atom-centred anions such as chloride-based ILs[12] or NaCl.[45] For more complex ILs such as [C₄mim][BF₄] no charge transfer was present in the Bader analysis,[46] which is quite unusual as it was observed in RESP schemes for imidazolium-based ILs.[3,47,48] For these reasons the AIM theory was not considered in this study.

Among the widely used schemes for fitting atomic charges for ILs are ChelpG,[47-50] Connolly (usually erroneously referred to simply as RESP),[13,48,51-53] the Blöchl analysis,[12,13] Mulliken analysis,[54] Natural Population Analysis[48,55] and less used ones.[5,56] One of the major concerns arising from these studies is unsystematic reproducibility of atomic charges for the same cations and anions, clearly indicating strong dependence of charges on the scheme chosen as well as molecular orientation. There are only a handful of studies that compare ESP derived atomic charges using various schemes. For example, the 1-butyl-imidazolium cation atomic charges from a number of publications show similar trends but the actual charge magnitudes display a great variation.[47,50,53,56,57] An analogous situation was observed in the C₁mim cation of the [C₁mim]Cl ion pair, for which large fluctuations of the charges in the imidazolium ring were reported using a number of charge fitting schemes.[14] In the calculations of individual ion pairs of [C₁mim]Cl the charge on the chloride anion showed less fluctuation depending on the level of theory (such as MP2 and the PBE functional) as well as the scheme used for calculations of charges (such as Connolly, Bader and Blöchl).[58] Significant fluctuations of the Mulliken-derived atomic charges in the imidazolium ring was found when the alkyl chain increased from ethyl to octyl, further highlighting the molecular orientation dependence.[54] The NPA charges were found to be more pronounced (either more positively charged or more negatively charged) than those of the Connolly algorithm in the case of [C₄mim]Br.[48] One of the implications of this unsystematic behaviour of various charge schemes lies in the fact that the charge distribution could drastically affect transport properties. For example, Kohagen et al. indicated that the actual charges had a strong effect on the structural arrangement of ILs, potentially influencing the dynamics of ions in the liquid state.[48] When the position of the charge was intentionally moved off centre in model ILs consisting of univalent spherical ions, a sharp decrease

in conductivity and a rapid increase in viscosity were observed in MD simulations.[59] Due to the lack of systematic studies on the performance of various charge-fitting schemes for ILs with respect to the level of theory and dependence on the molecular orientation, the question as to which scheme is most reliable still remains widely open.[60]

In this work we performed a systematic study to compare the performance and reliability of partial charges in ionic liquid ions calculated using (a) schemes based on the density matrix approach such as MPA, LPA and NPA and (b) schemes based on the RESP approach such as ChelpG, Connolly and Geodesic. In the case of RESP schemes fitting to atomic charges as well as dipole and quadrupole moments was explored. A number of ion pairs were included consisting of the C₍₁₋₄₎mim cations coupled with the routinely used anions such as Cl, BF₄ and NTf₂. Influence of the level of theory (uncorrelated HF wavefunciton vs. correlated MP2 wavefunction) and basis sets (cc-pVDZ, aug-cc-pVDZ and aug-cc-pVTZ) on charge distribution was considered. The quality of charge distribution was analysed based on three criteria: (1) treatment of symmetry identical atoms, (2) invariance of partial charges in the imidazolium ring with increasing alkyl chain and (3) dipole moments calculated from the predicted charge distribution as a measure of electronic polarisation. Dipole moment distribution and charge transfer effects were also analysed for larger ionic clusters of the [NMe₄][BF₄], [C₁mim][BF₄] and [C₁mim]Cl ionic liquids consisting of 2, 4 and 8 ion pairs.

It should be noted that while this study focuses specifically on assessing the quality of widely used charge schemes on charge distribution in ionic liquid ions, these schemes have been applied to a much wider range of molecular systems incorporating ionic interactions. In particular, the vast majority of dynamic simulations of charged DNA- and protein-like systems include counter-ions to ensure energetic stabilisation of the resulting structures.[61-63] Thus, the outcomes of this work represent implications not only on a fast growing ionic liquid electrolyte community but also on a broader simulation community as well as experimental scientists who rely on the results of the simulations to guide their experimental work.

2 Theoretical procedures

Ion pairs of [C₍₁₋₄₎mim]X and [NMe₄]X, where X = Cl⁻, [BF₄]⁻ and [NTf₂]⁻, were fully screened for the lowest energy conformation by exploring all possible alkyl chain torsions and anion positions around the ring. The lowest energy configuration was located for all anions, in which they interact from above the imidazolium ring. For [C₁mim]Cl another energetically stable configuration, in which the chloride interacts with the C2–H bond in the plane of the imidazolium ring, was located. This configuration was not considered in the study due to lower binding energy. Clusters of [NMe₄][BF₄], [C₁mim][BF₄] and [C₁mim]Cl were constructed in a way identical to those used by Izgorodina et al.[64] where in the case of the imidazolium ILs ion pairs were systematically added in order to form 3D structures incorporating instances where the anion is both in

This journal is © the Owner Societies 2013

Phys. Chem. Chem. Phys., 2013, **15**, 1632–1646 | **1635**

the plane of the imidazolium ring, and either above or below the ring, and alkyl chains were allowed to interact. The [NMe$_4$][BF$_4$] clusters were formed in a rectangular prism such that all interionic interactions are maximised. All geometry optimisations were performed using the M06-2X DFT functional[65] and the 6-31 + G(d) basis set for all ion pairs and clusters except the eight ion pair cluster of [C$_1$mim][BF$_4$] and the [C$_1$mim]Cl clusters, for which Dunning's cc-pVDZ basis set was used to avoid orbital linear dependence issues. Atomic coordinates of all structures are available in the ESI.† Single-point calculations for the RESP, Mulliken and Löwdin charges were performed using the GAMESS-US software package,[66] whereas NPA[20] charge calculations were performed using the GAUSSIAN 09 suite of programs.[38] RESP charges were calculated using ChelpG,[34] Connolly[30] and Geodesic[37] point selection algorithms; in the case of ion pairs, dipole and quadrupole moment fitting was explored in addition to the standard charge fitting schemes. All single-point charge calculations were performed using an Ahlrichs-type triple-z doubly polarised basis set, TZVPP,[67] except for analysis of basis set dependence, for which Dunning's cc-pVDZ, aug-cc-pVDZ and aug-cc-pVTZ basis sets were used. Both correlated MP2 and uncorrelated HF wavefunctions were used for ion pair calculations, while HF wavefunctions were used exclusively for cluster calculations. Unless otherwise specified, charges and derived properties use the Hartree–Fock wavefunction with RESP charges restrained to reproducing the overall charge of the chemical system. With the exception of results shown in Section 3.4, RESP charge fitting point density parameters were selected as follows: 0.3 Å for the ChelpG grid spacing, 1 point per Å for the Connolly surface point density, and a $\{3,5+\}_{3,0}$ template for the Geodesic

Fig. 2 Imidazolium numbering scheme.

scheme. Standard numbering is used throughout this paper with respect to atoms in the imidazolium cation as detailed in Fig. 2.

3 Results and discussion

3.1 Basis set dependence

Basis set sensitivity was evaluated using the systematically augmented Dunning basis sets; namely, cc-pVDZ (CCD), aug-cc-pVDZ (ACCD) and aug-cc-pVTZ (ACCT). Table 1 shows charges calculated using the density matrix-based and the RESP-based Geodesic charge scheme for the lowest energy configuration of [C$_1$mim][BF$_4$]. ChelpG and Connolly RESP schemes behave similarly to Geodesic and are shown in the ESI.† Mulliken and Löwdin charge schemes, which are already known for being heavily basis-set dependent, show such behaviour as expected. For example, in the case of Mulliken charges the methyl carbons vary from +0.84 e to -0.99 e from ACCD to ACCT. This range of charge is concerning not only for the sheer magnitude, but also for the inversion of sign, which would impact severely on any specific intermolecular interactions derived. Considering the temptation to base conclusions on Mulliken charges, ubiquitously printed in all computational chemistry packages, these results alone highlight the risk in doing so. Large fluctuations are also observed for Löwdin charges, although over a smaller range compared to Mulliken. For example, the largest variation is observed for the boron atom, with the charge decreasing from -0.14 e to -0.98 e. Interestingly, the Löwdin charge scheme assigns positive charges to the fluorine atoms for the aug-cc-pVTZ basis set, which also occurs for the TZVPP basis set used for the remainder of the calculations performed in this study. Lastly, the NPA charges, specifically designed to minimise basis set dependence, do indeed show minimal basis set effects within the density matrix-based schemes. The largest variation of 0.23 e again occurs for the boron atom.

All RESP schemes show notably less basis set dependence when compared to the density matrix-based schemes. The maximum variation in all cases belongs to the boron atom, with
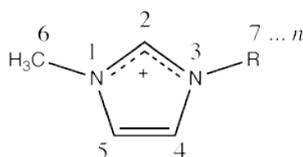
Table 1 Mulliken, Löwdin, NPA density matrix-based charges as well as the RESP Geodesic charges calculated with Dunning's cc-pVDZ (CCD), aug-cc-pVDZ (ACCD) and aug-cc-pVTZ (ACCT) basis sets for the lowest energy configuration of [C$_1$mim][BF$_4$]. All but the C2-H hydrogen atoms not shown

Atom (description)	Mulliken			Löwdin			NPA			Geodesic		
	CCD	ACCD	ACCT	CCD	ACCD	ACCT	CCD	ACCD	ACCT	CCD	ACCD	ACCT
C (C4/5)	0.11	0.63	-0.50	-0.08	0.03	0.07	0.01	0.01	-0.01	-0.23	-0.25	-0.24
N (N1/3)	-0.39	-0.45	0.24	0.17	-0.13	0.12	-0.48	-0.48	-0.37	0.23	0.28	0.28
C (C2)	0.43	0.73	-0.45	0.06	0.12	0.08	0.48	0.46	0.42	-0.22	-0.27	-0.26
N (N1/3)	-0.39	-0.49	0.20	0.17	-0.12	0.12	-0.48	-0.48	-0.37	0.26	0.31	0.31
C (C4/5)	0.11	0.64	-0.52	-0.08	0.05	0.07	0.01	0.01	-0.01	-0.27	-0.30	-0.30
C (methyl)	0.13	0.85	-0.99	-0.11	0.12	0.23	-0.31	-0.31	-0.27	-0.32	-0.36	-0.36
C (methyl)	0.13	0.85	-0.99	-0.11	0.11	0.25	-0.31	-0.31	-0.27	-0.34	-0.39	-0.39
H (C2)	0.14	-0.32	0.80	0.13	0.11	-0.02	0.27	0.27	0.26	0.31	0.31	0.31
F	-0.41	-0.87	-0.76	-0.12	-0.18	0.05	-0.65	-0.66	-0.61	-0.40	-0.37	-0.40
B	0.66	2.31	1.88	-0.49	-0.14	-0.98	1.62	1.65	1.42	0.80	0.74	0.85
F	-0.42	-0.85	-0.77	-0.12	-0.17	0.10	-0.66	-0.67	-0.61	-0.44	-0.42	-0.44
F	-0.32	-0.74	-0.71	-0.05	-0.16	0.06	-0.62	-0.63	-0.57	-0.40	-0.39	-0.42
F	-0.42	-0.85	-0.77	-0.12	-0.18	0.07	-0.66	-0.67	-0.61	-0.44	-0.42	-0.44

charges differing by o0.10 e. This is a significant improvement upon all density matrix-based schemes; the electrostatic potential varies insignificantly between these basis sets, thus allowing for a less costly calculation. As the electrostatic potential converges with increasing basis set, charges consistently "improve" (within the definition of the RESP charge scheme), showing convergence to some value. This cannot be guaranteed in the case of the density matrix-based schemes, particularly where extensive basis sets such as aug-cc-pVTZ are used.

3.2 Treatment of symmetry

The reliability of the charge schemes can be measured in terms of equal treatment of symmetric atoms and can be used as a measure of error. This aspect was investigated using the $[C_1mim]Cl$ ion pair, which possesses C_s symmetry with a mirror plane intersecting the chloride and C2–H bond. Therefore, an ideal scheme would yield identical charges for atoms either side of this plane and consequently their difference should be zero. Fig. 3 shows the difference between the methyl carbon atoms (C6 and C7), the nitrogen atoms, N1 and N3, and the carbon backbone atoms, C4 and C5, as the absolute difference between the two charges. In all instances the charge difference does not exceed 0.07 e, showing only marginal differences between charge schemes. The ChelpG scheme consistently shows the worst treatment of the C4 and C5 atoms, improving only at quadrupole fitting. The Connolly scheme shows a slight variation in symmetric charges when restrained to charge or dipole moment. Connolly charges fitted to the quadrupole moment show the largest deviation of all schemes assessed, which is likely to be the result of anisotropies in the Connolly point selection algorithm. The Geodesic scheme that is known for its highly isotropic surface[37] shows only a minor increase in charge deviation when fitted to the quadrupole moment. Out of the RESP schemes, Geodesic point selection is the most systematic at describing symmetric charges. As expected, the density matrix-based schemes, Mulliken, Löwdin and NPA, show essentially no asymmetry in predicted charges as this is an inherited property of the diagonalised density matrix, thus further reinforcing their

independence of molecular orientation. It has to be noted that using the MP2 correlated wavefunction for calculating partial charges does not introduce any appreciable changes to the trends discussed here.

3.3 Invariance with alkyl chain length

Ideally, given the peripheral nature of the alkyl chain, changes in length should impact minimally on spatially distinct regions of the molecule in question. In testing structural invariance, the charge scheme behaviour on selected atoms in the 1-alkyl-3-imidazolium cation upon modification of the alkyl chain was investigated. An ideal scheme would show negligible changes in charges for the C6, N1, C2 and N3 atoms (see inset structure in Fig. 4) as these atoms are expected to experience little influence from substituent groups beyond ethyl.

Fig. 4 shows the variation in charge with increasing alkyl chain in $[C_nmim]BF_4$ ion pairs for both density matrix-based and RESP schemes. The charges of hydrogen atoms were added to those of the bonded carbon atoms, analogous to the "united atom" approach. The Connolly and Geodesic schemes show negligible changes in charge as the alkyl chain is increased, and follow each other closely for the same cation. The grid-based ChelpG scheme shows notably different trends to those of the Connolly and Geodesic schemes, producing opposite trends for atoms within the ring (C6–N1–C2–N3–C7) and shows close agreement only in atoms that are part of the alkyl chain. Similar trends were observed for the $[C_nmim]Cl$ and $[C_nmim][NTf_2]$ ILs.

Density matrix-based schemes appear insensitive to the alkyl chain in their charge assignments; variation in charge with structural modification is insignificant, with the charge of the terminal methyl group on the alkyl chain converging perfectly to neutral in all cases. More importantly, the changes between the partial charges of the cations in ion pairs are insignificant compared to those in the isolated cations. While the alkyl chain is treated consistently, the respective schemes differ from one another for charges within the imidazolium ring. The Löwdin scheme is substantially different from the other two giving
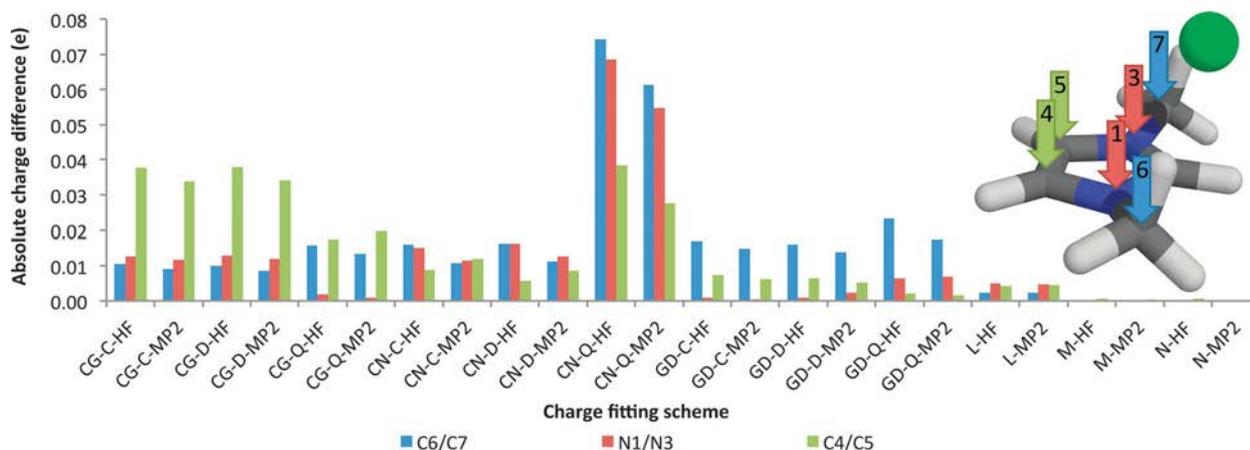
Fig. 3 Absolute difference in charges assigned to symmetrically equivalent methyl carbon atoms and imidazolium nitrogen and carbon atoms in the $[C_1mim]Cl$ ion pair. Charge fitting scheme abbreviations as follows: CG = ChelpG, CN = Connolly, GD = Geodesic, L = Löwdin, M = Mulliken and N = Natural Population Analysis; C = Charge fitting, D = Dipole fitting and Q = Quadrupole fitting.

This journal is © the Owner Societies 2013

Phys. Chem. Chem. Phys., 2013, **15**, 1632–1646 | **1637**
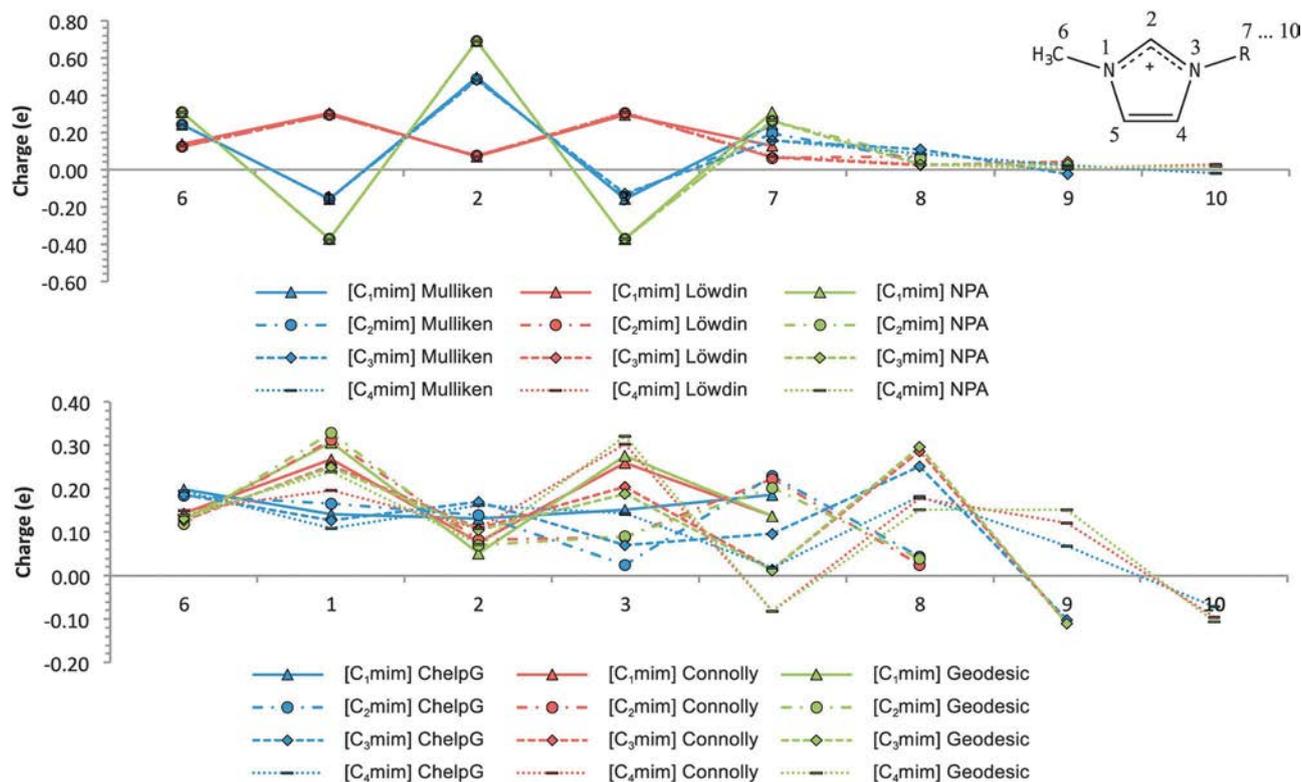
Fig. 4 Variation in charge with increasing alkyl chain for the $[C_nmim][BF_4]$ ion pair using density matrix-based schemes (top) and RESP schemes (bottom). Charges assigned to carbon atoms at positions 6, 2, and 7–10 include charges on any bonded hydrogen atoms.

charges entirely in the positive-to-neutral regime, whereas the Mulliken and NPA schemes give negative charges for the N1 and N3 atoms, which is the opposite to the trends observed for the RESP schemes. While C2 is positive for all density matrix-based and RESP schemes, Mulliken and NPA give the largest positive values. Interestingly, the Löwdin charges mirror those of the Connolly and Geodesic RESP schemes more closely than the Mulliken and NPA schemes. The NTf2- and chloride-based ion pair series show trends that are almost identical to those of the BF4 series and are shown in the ESI.† It should be noted, however, that the results for Mulliken and Löwdin are specific to the TZVPP basis set and may give strikingly different trends with other basis sets, as was shown in Section 3.1.

To summarise, out of the schemes studied, the RESP Connolly and Geodesic schemes appear to perform more consistently for the imidazolium cation with increasing alkyl chain. Density matrix-based schemes do not show variation in charges compared to isolated cations. Although Mulliken and NPA show almost no variation in partial charge on the ring, the nitrogen atoms are found to be negatively charged compared to the opposite trend for the RESP fitting schemes. Therefore, Mulliken and NPA are not recommended for the prediction of partial charges for IL ions.

3.4 RESP dependence on point density

Each of the three RESP schemes have been tested to explore the impact of point density on the resulting charges using a number of scheme-specific parameters. In the case of ChelpG, this is the grid spacing; for the Connolly scheme, surface point

density; for the Geodesic scheme, the geometric template. In this section, both the charge symmetry and distribution effects are explored.

In considering the impact on charge symmetry, the same $[C_1mim]Cl$ system used in Section 3.2 was tested by varying one of the aforementioned parameters depending on the scheme. Fig. 5 shows the differences between symmetry equivalent charges as a function of the number of fitting points for all three RESP schemes. From this it can be seen that the Geodesic scheme yields the highest accuracy charges with the fewest fitting points with considerably more stability than the Connolly scheme as fitting points are increased. Even with as few as 189 fitting points ($\{3,5+\}_{1,0}$ template‡), charges are still markedly more accurate than the ChelpG scheme with 374 fitting points (1.0 Å grid spacing). Similarly, the Connolly scheme, although slightly more erratic when compared with the Geodesic scheme, reaches appreciable accuracy after as few as 774 fitting points (0.9 points per Å²). It was also noted that across all schemes and all point densities, charge transfer remained essentially constant with a standard deviation of 0.01 e. These results clearly highlight the superiority of surface-based methods over grid-based with the latter requiring at least an order of magnitude more fitting points to achieve comparable levels of accuracy.

Additionally, the impact of ChelpG grid spacing on the charge distribution of the cation in the widely used

---

‡ Geodesic template notation follows conventions described by Coxeter.[68]

Fig. 5 Dependence of charge accuracy on the number of fitting points for [C₁mim]Cl shown as the difference between symmetry equivalent charges. Dashed lines indicate the settings used in this study.

Fig. 6 ChelpG charges as a function of grid spacing (all bonded hydrogen charges included) for [C₂mim][BF₄].

charge. From these data it is apparent that charges begin to converge only after 0.4 Å, indicating that this should be the optimal number for point density to produce statistically well-fitted charges.

In summary, it has been demonstrated that the surface-based schemes provide appreciable accuracy with significantly fewer fitting points than the grid-based ChelpG scheme. ChelpG has been shown to give dramatically varying charges if sparse grids are used, including charges with opposite signs which may bring into question the validity of dynamic properties.

3.5 Charge transfer

Charge transfer is associated with orbital overlap such that the HOMO electrons of one interacting molecule can partly populate the LUMO of the other,[69] thus allowing a fraction of charge to be transferred resulting in energetic stabilisation, similar to that for hydrogen bonding.[2] The ability of electrons to be partially transferred from one molecule to another implies a combined energetic and orbital symmetry criterion to permit the transfer. The amount of energy associated with the charge transfer is difficult to quantify, especially between charged species for fundamental reasons as discussed by Gordon et al.[70] However, in adhering to a point-charge model, the magnitude of the transferred charge may itself be quantified, albeit within the definition of the charge scheme. In practical terms, this means being able to reproduce as best as possible the true quantum mechanical electrostatic potential using atomic charges. The important distinction needs to be made between charge transfer and polarisation or induction effects, as the latter refer to changes in the overall charge distribution due to the dipole moment induced by surrounding polar or charged species.[9] It has to be noted that polarisation is a broader concept, incorporating charge transfer as a component, therefore polarisation effects in ion pairs and ionic clusters are discussed separately in Section 3.6.

[C₂mim][BF₄] IL previously analysed in Section 3.3 was investigated. Data shown in Fig. 6 show how the charges of symmetry equivalent atoms change as the grid spacing is reduced (fitting points increased). The grid spacing range shown includes the 0.8 Å–0.3 Å range recommended by Breneman et al.[34] and highlights the importance of selecting the optimum spacing for this scheme given that the default may vary between software implementations. In the most extreme case, the grid of 0.8 points per Å² gives charges of opposite sign on the N3 nitrogen, and the C2 carbon shows an unusually large positive

This journal is © the Owner Societies 2013

Phys. Chem. Chem. Phys., 2013, **15**, 1632–1646 | **1639**

The ability of charge schemes to account for and estimate charge transfer has been examined in individual ion pairs as well as ionic clusters. The importance of correlated wavefunctions, such as Møller Plesset perturbation theory (MP2), for charge transfer has also been investigated. Charge transfer was calculated as the difference between the unity charge of the isolated ion and the calculated charge of the same ion in the ion pair or the cluster (shown in eqn (1)).

$$CT = N - \sum q_{cation}^{cluster} \qquad (1)$$

where N is the number of ion pairs in the cluster. As shown in Fig. 7, the RESP schemes are largely indistinguishable from one another. The charge tansfer values fall in the range between 0.2 and 0.3 e for the NTf$_2$ and Cl ion pairs, whereas BF$_4$ displays smaller charge transfer values of o0.2 e. The RESP predicted charge transfers reflect the scaling factors of 0.6 to 0.7 that were shown to improve the prediction of transport properties in classic MD simulations.[5,6] Given that charge scaling improves the speed of ion dynamics, the RESP schemes suggest that the reason why a greater agreement with experiment is achieved is indeed the product of a more realistic description of charge transfer. In general for RESP schemes, charge transfer observed in the case of the chloride anion is comparable to that of the NTf$_2$ anion and considerably higher than that of the BF$_4$ anion. The magnitude of these charge transfer values is expected to be high for NTf$_2$, as it is a particularly bulky anion and therefore, easily polarisable,[71] but is somewhat surprising for the chloride anion, which is a much smaller and hence, "harder" anion.

By contrast, owing to their tendency to localise strongly occupied orbitals, the Mulliken and NPA schemes typically underestimate charge transfer and this is observed especially in the case of the NPA scheme that does not predict charge transfer for both NTf$_2$ and BF$_4$ ion pairs. It is only in the case of

the chloride that the NPA scheme shows a non-negligible charge transfer of 0.09 e for the Hartree–Fock wavefunction and 0.13 e for the MP2 one. This is indicative of a comparatively high degree of covalency between the anion and the cation,[72,73] reflected in several ab initio calculations.[2,74,75]

The impact of electron correlation is generally minimal, however slightly more so in the density matrix schemes. MP2 yielded an increase in charge transfer by less than 0.04 e for the RESP schemes, and by less than 0.06 e for the density matrix schemes. With changes in charge of this magnitude, it is difficult to justify the added computational expense required for the MP2 calculation.

In considering charge transfer in IL clusters, data for three ILs – [NMe$_4$][BF$_4$], [C$_1$mim][BF$_4$] and [C$_1$mim]Cl – composed of one, two, four and eight ion pairs are presented in Fig. 8. Amongst the RESP schemes, it remains difficult to identify any one particular point selection algorithm that produces the most realistic results. In the case of [C$_1$mim][BF$_4$], ChelpG shows a sharp drop in charge transfer after four ion pairs, whereas the other schemes show a modest increase from one to two ion pairs, and then a slight decrease with charge transfer for eight ion pairs, converging close to the value of the ion pair. For the [NMe$_4$][BF$_4$] clusters charge transfer increases rapidly within the ChelpG scheme, whereas Geodesic shows a gradual increase from one to eight ion pairs. For both [C$_1$mim][BF$_4$] and [NMe$_4$][BF$_4$] charge transfer in the 8-ion pair clusters converges to that of the individual ion pairs when using the Connolly scheme.

Two counteracting effects play a role in charge transfer trends with increasing cluster size. On one hand, an increasing number of interacting ions should further increase charge transfer due to increased orbital overlap. On the other hand, the distance between ions becomes longer with increasing cluster size, thus reducing degree of orbital overlap and hence, charge transfer. Therefore, it is expected that charge transfer should converge to a number with increasing cluster size and is dependent on the nature and structural arrangement of ILs. The observed trends for the Connolly and Geodesic schemes in Fig. 8 are the manifestation of these two opposite effects, suggesting that clusters of four ion pairs for [NMe$_4$][BF$_4$] and [C$_1$mim][BF$_4$] are already good models for estimating the extent of charge transfer.

In contrast to the general trends observed for the BF$_4$ ILs, the chloride system presents an unusual case as it shows the opposite trend found in [C$_1$mim][BF$_4$] and [NMe$_4$][BF$_4$]. With the exception of the Löwdin charges, there is a net decrease in charge transfer as the cluster size is increased. Since [C$_n$mim]Cl ILs are known to exhibit unusually strong inter-ionic bonds between the imidazolium ring and the chloride anion as shown by quantum chemical calculations[2] and X-ray crystal structures[73] of chloride-based salts, the reduction in charge transfer actually indicates a reduction in strength of these interactions. The fact that the NPA charges consistently show non-negligible charge transfer, otherwise not seen for the BF$_4$-based clusters, further supports this conclusion. As is the case with the ion pairs, the chloride cluster of eight ion pairs still displays charge

Fig. 7 Ion pair charge transfer based on both HF and MP2 wavefunctions for [C$_{(1-4)}$mim]X, where X = [NTf$_2$]$^-$, [BF$_4$]$^-$ and Cl$^-$.

Paper

Section 2.3. *Assessment of atomic partial charge schemes for polarisation and charge* ^{View Article Online}
transfer in ionic liquids

32   PCCP

Fig. 8 Charge transfer with increasing cluster size for (a) [NMe$_4$][BF$_4$], (b) [C$_1$mim][BF$_4$] and (c) [C$_1$mim]Cl.

poorly fitted charges. One should keep in mind that regardless of the point selection algorithm the uncertainty in fitted partial charges of atoms "buried" inside the cluster is expected to increase with increasing cluster size, thus producing ambiguous estimates of charge transfer. Therefore, the quality of the various RESP schemes cannot always be guaranteed and this issue is discussed in more detail in Section 3.6. It is suggested that charges calculated using ChelpG should generally be avoided in preference to the more recent Connolly and Geodesic schemes.

3.6 Dipole moment fluctuations

Prado et al.[9] and more recently Wendler et al.[13] have shown that dipole moments within ILs are not homogeneous. These ab initio MD simulations suggest that rather than having a fixed dipole moment, fluctuations occur giving rise to broad distributions of dipole moments of ions due to electronic polarisation. The ability of the charge schemes to reproduce these findings was tested using atomic charges with the magnitude of the dipole moment on any ion, m, given by eqn (2):

$$m = \left| \sum_i q_i \left(\bar{r}_i - \bar{r}_0 \right) \right| \qquad (2)$$

where q_i and $\bar{r}_i$ are the charge and the position of atom i, respectively, and $\bar{r}_0$ is some point of reference. Since the dipole moment requires a point of reference, typically the centre of mass, it is difficult to make comparisons between molecular systems. To remedy this, we chose to use the centre of the imidazolium ring for [C$_n$mim]$^+$, shown as the red markers in Fig. 9, and the geometric centre of [BF$_4$]$^-$. The definition adopted for the dipole moment requires that the molecular charge be distributed over atom-centred monopoles, therefore the chloride anion is undefined and not included in this study. These "pseudo" dipole moments give a measure of polarisation that can be compared between IL systems.

Approximate symmetry observed after optimisation in the clusters allows for a thorough analysis of the dipole moment distribution in a way similar to data presented in Section 3.2, whereby the absolute difference between symmetry equivalent ions is used as a measure of scheme quality. Clusters of [C$_1$mim]Cl and [C$_1$mim][BF$_4$] possess approximate C$_i$ symmetry with an inversion point at the centre of the cluster; cations numbered 1, 2, 3 and 4 have symmetry equivalent counterparts numbered 8, 7, 6 and 5, respectively. Similarly, [NMe$_4$][BF$_4$] approximates D$_2$ symmetry, for which cations numbered 2, 3, 5 and 8 are equivalent to each other, and cations 1, 4, 6 and 7 are equivalent to each other. While strictly speaking the ions identified are not perfectly equivalent – small changes in geometry such as rotations of the anion or methyl groups break symmetry – these should not affect the dipole moment to any great extent.

Fig. 10 shows the absolute difference between the dipole moments of equivalent cations. In the cases of [C$_1$mim][BF$_4$] and [C$_1$mim]Cl, this is simply the difference between equivalent pairs, whereas in the [NMe$_4$][BF$_4$] cluster the maximum difference among the four equivalent individual pairs is considered (for more

transfer of about 0.1 e higher than that of the BF$_4$ cluster. It is also important to note that all RESP schemes are in close agreement with one another, which is not observed for the other two ILs. This is likely the result of the very simple structure of the anion – a single atom – which is unlikely to suffer biased point sampling to any great extent.

To this end, it is recommended that cluster charge transfer calculated using any RESP scheme must be interpreted with caution as it is not always possible to detect results that include

This journal is © the Owner Societies 2013

Phys. Chem. Chem. Phys., 2013, **15**, 1632–1646 | **1641**

Section 2.3. *Assessment of atomic partial charge schemes for polarisation and charge* View Article Online
transfer in ionic liquids

PCCP   33   Paper

Fig. 9 Eight ion pair structures of (a) [NMe$_4$][BF$_4$], (b) [C$_1$mim][BF$_4$] and (c) [C$_1$mim]Cl. Red markers indicate centre points of imidazolium rings.

Fig. 10 Differences between dipole moments of symmetry equivalent cations in (a) [NMe$_4$][BF$_4$], (b) [C$_1$mim][BF$_4$] and (c) [C$_1$mim]Cl.

The vertical lines on the graphs show the dipole moments of isolated ions. Poorly fitted charges manifest themselves in overly broad distributions, including Connolly charges for [NMe$_4$][BF$_4$] and both ChelpG and Geodesic for [C$_1$mim][BF$_4$]. These instances of poor fitting are similarly reflected in Fig. 10.

Analysis of Fig. 11 reveals that all schemes apart from Löwdin produce rather broad distributions of dipole moments. Noticeably, in the [NMe$_4$][BF$_4$] cluster the distribution extends up to 2.0 D, especially considering that the isolated ions have zero dipole moments. Compared to the density-based schemes, the RESP schemes tend to yield broader distributions, which in some cases could be an artefact of increased uncertainty in fitting partial charges for atoms "embedded" or "buried" inside the cluster. In the vast majority of cases, the ions in the cluster exhibit a greater degree of polarisation compared to the isolated forms, with the degree of polarisation depending on the IL. For example, the C$_1$mim cation appears to be slightly more polarised in the BF$_4$-based IL than in the chloride IL. The RESP schemes produce almost identical dipole moment distributions for the [C$_1$mim]Cl cluster, again emphasising a special case of strong inter-ionic interactions between ions in Cl-based ILs. Wendler et al.[13] observed similar distributions for the [C$_1$mim]Cl cluster consisting of 30 ion pairs. In their case the dipole moment fell in the range between 0.0 and 5.0 D, with the maximum being around 3.0 D, which is in a good agreement with our results from the Connolly and Geodesic schemes for the eight ion pair cluster of [C$_1$mim]Cl.

Surprisingly, except for [C$_1$mim]Cl density matrix-based schemes yield similar trends to the RESP schemes, although being insensitive to charge transfer effects. The Löwdin scheme produces by far the narrowest distribution, and, with the

details see Fig. 9). The density matrix-based schemes show the smallest deviation from zero for the same reasons discussed in Section 3.2, whereas the RESP schemes show larger deviations. The largest deviations of 2.81 D and 3.28 D are observed between cation pairs 4 & 5 in [C$_1$mim][BF$_4$] (Geodesic scheme) and between cation pairs 4 & 7 in [NMe$_4$][BF$_4$] (Connolly scheme), respectively. The [C$_1$mim]Cl system also shows relatively poorly fitted values, although this only results in a maximum deviation observed for the Connolly scheme of 0.09 D, which is rather negligible. The reason for this substantially reduced errors for [C$_1$mim]Cl compared to [C$_1$mim][BF$_4$] might possibly be the result of a simplified inter-ionic interaction between the cation and a mono-atomic anion as opposed to the bulkier BF$_4$ anion. Observed errors are also similarly reflected in the dipole moment distributions and are detailed below.

Fig. 11 shows the dipole distributions in clusters of eight ion pairs of [NMe$_4$][BF$_4$], [C$_1$mim][BF$_4$] and [C$_1$mim]Cl (structures shown in Fig. 9). The distributions were generated by fitting dipole moments to a Gaussian curve using the mean and standard deviation (for more details see Table S1 in the ESI†).

Fig. 11 Dipole moment distributions in eight ion pair clusters of (a) [NMe$_4$][BF$_4$], (b) [C$_1$mim][BF$_4$] and (c) [C$_1$mim]Cl. Vertical lines for [C$_1$mim]$^+$ indicate the dipole moment for the isolated cation while all other isolated ions possess a zero dipole moment.

exception of [C$_1$mim]Cl, shows the least polarisation of all schemes. The Löwdin analysis was also the only density matrix scheme showing some charge transfer in ion pairs of the [C$_n$mim][BF$_4$] series. Conversely, while the NPA scheme almost entirely neglects charge transfer, it remains sensitive to polarisation effects and tends to show distributions in closer agreement to the RESP schemes. The only exception is the chloride cluster, for which both the Löwdin scheme and in particular the NPA scheme show a decrease in polarisation. This is likely due to the sensitivity of the NPA scheme towards decreased covalency between the imidazolium cation and the chloride anion.

To summarise, all the schemes, be it density matrix-based or RESP, show some degree of cationic and anionic polarisation for the BF$_4$-based ILs, with Löwdin giving the narrowest dipole moment distribution. Out of the RESP schemes it is hard to

pick and choose which scheme produces best results, with all three inevitably displaying poorly fitted charges in some cases. According to the presented data ChelpG tends to produce slightly larger errors in the "pseudo" dipole moments and therefore, the Connolly and Geodesic schemes are recommended for estimating partial charges in IL ions. Surprisingly, NPA shows an appreciable degree of polarisation of ions in the BF$_4$-based clusters similar to that of Connolly and Geodesic, although no charge transfer was observed between ions. The chloride clusters represent an exceptional case, with NPA and Löwdin slightly depolarising the cation with respect to the isolated form. This finding emphasises that polarisation effects reflect the changes in charge distribution and do not necessarily follow the trends in the charge transfer process that requires orbital overlap.

This journal is © the Owner Societies 2013

Phys. Chem. Chem. Phys., 2013, **15**, 1632–1646 | **1643**

Section 2.3. *Assessment of atomic partial charge schemes for polarisation and charge* View Article Online
transfer in ionic liquids

PCCP   35   Paper

4 Comments on the quality of charge fitting and chloride-based ionic liquids

4.1 Charge fitting

In discussing cluster charge transfer and polarisation, it has become obvious that no one particular RESP scheme provides perfect results; there is always at least one outlying value that, while not being particularly obvious, alters the trends resulting in unphysical results. This appears to be a consequence of the close proximity in which the ions sit with one another, causing the exclusion of fitting points around parts of the molecules embedded inside the cluster. The "buried atom" problem is a well-known deficiency of RESP schemes in general, but typically referred to only in an intramolecular sense.

GAMESS-US, in its RESP routines, provides a fitting diagnostic based on a w^2 distribution termed the "Estimated Standard Deviation" (ESD) to evaluate the quality of the assigned charges. As is expected, in all cases these values become large in instances where an atom is considered embedded; these are typically atoms such as methyl carbon atoms, the central nitrogen in NMe_4 or the central boron atom in BF_4. With simple molecules, this does not pose a significant problem as these numbers remain quite small, however in clusters these embedded regions extend beyond atoms that are embedded in an intramolecular sense to those that become embedded through intermolecular crowding. In these instances the ESD numbers can become staggeringly large indicating a severe deficiency in fitting points.

4.2 Chloride-based ionic liquids

The chloride ILs have presented an unusual case in terms of both their unique inter-ionic interactions as well as their performance with respect to the charge schemes analysed in this study. In terms of the former, in contrast with the BF_4 and NTf_2 ILs, the chloride series exhibited behaviour characteristic of an increased covalency between ions, which is in agreement with observations by Wang et al.,[72] Kirchner et al.[48] and our group[2] that showed a relatively high degree of orbital overlap. This type of interaction elicits a most notable response from the NPA charge scheme in the form of non-negligible charge transfer around 0.1 e, which otherwise gave almost no charge transfer for all other ILs. RESP schemes were essentially all equally as good at fitting the charge, whereas the other ILs studied each suffered poorly fitted values for at least one scheme, which in itself was unsystematic, i.e. no RESP charge scheme consistently performed well. This is attributed to the monoatomic nature of the anion which simplifies fitting procedures by limiting excluded fitting points and eliminating any "buried" atoms. While fortunate for the chloride ILs, the charge fitting quality remains problematic, as ILs typically consist of bulky anions and cations, of which chloride is far from. Thus, the chloride ILs are considered atypical and hence, poor model systems for testing the quality of quantum chemical methods and charge fitting schemes.

5 Conclusions

In this paper it has been shown that the RESP schemes are preferred when producing atomic partial charges as they can effectively capture electronic polarisation effects and show an appreciable degree of charge transfer. Among the RESP schemes, it appears that ChelpG is the least systematic for ion pairs of ionic liquids and should be used with a grid spacing of at least 0.4 Å, although instances for larger ionic clusters, for which Connolly and Geodesic schemes fail, are also present. Thus, RESP warrants a degree of caution, particularly in ionic clusters of a few ion pairs, as the charge fitting procedure may produce erroneous charges owing to insufficient fitting points for "buried" atoms or those embedded in a tightly bound cluster. Density matrix-based methods such as the Mulliken and Löwdin population analyses have shown unacceptable basis set dependence and should be disregarded despite their prevalence in most quantum chemistry packages. In particular, the Löwdin scheme was generally insensitive to polarisation effects, which are expected to dominate in the liquid state of ILs. In contrast to this, the NPA scheme was surprisingly responsive to polarisation and was largely basis set independent. However, the scheme almost entirely neglected charge transfer in ion pairs and ionic clusters, apart from the chloride-based ones. Therefore, density matrix-based schemes inadequately represent the electrostatic interactions in ILs. The presented results also show that chloride-based ILs should not be considered as model systems to test the quality and reliability of quantum chemical methods as well as describe general trends in thermodynamic and transport properties for a wider population of ILs due to exceptional inter-ionic interactions and statistically well-fitted charges for all RESP schemes studied here for mono-atomic anions such as chloride.

Acknowledgements

References

1 S. B. C. Lehmann, M. Roatsch, M. Schoppke and B. Kirchner, *Phys. Chem. Chem. Phys.*, 2010, 12, 7473–7486.
2 E. I. Izgorodina and D. R. MacFarlane, *J. Phys. Chem. B*, 2011, 115, 14659–14667.
3 F. Dommert, K. Wendler, R. Berger, L. Delle Site and C. Holm, *ChemPhysChem*, 2012, 13, 1625–1637.
4 O. Borodin and G. D. Smith, *J. Phys. Chem. B*, 2006, 110, 11481–11490.
5 R. M. Lynden-Bell and T. G. A. Youngs, *J. Phys.: Condens. Matter*, 2009, 21, 424120–7.

6 T. G. A. Youngs and C. Hardacre, ChemPhysChem, 2008, 9, 1548–1558.

7 T. Cremer, C. Kolbeck, K. Lovelock, N. Paape, R. Wölfel, P. Schulz, P. Wasserscheid, H. Weber, J. Thar, B. Kirchner, F. Maier and H.-P. Steinrück, Chem.–Eur. J., 2010, 16, 9018–9033.

8 S. Men, K. R. J. Lovelock and P. Licence, Phys. Chem. Chem. Phys., 2011, 13, 15244–15255.

9 C. E. R. Prado, M. G. D. Ppolo, T. G. A. Youngs, J. Kohanoff and R. M. Lynden-Bell, Mol. Phys., 2006, 104, 2477–2483.

10 O. Borodin, J. Phys. Chem. B, 2009, 113, 11463–11478.

11 J. B. Hooper, O. Borodin and S. Schneider, J. Phys. Chem. B, 2011, 115, 13578–13592.

12 K. Wendler, F. Dommert, Y. Y. Zhao, R. Berger, C. Holm and L. Delle Site, Faraday Discuss., 2012, 154, 111–132.

13 K. Wendler, S. Zahn, F. Dommert, R. Berger, C. Holm, B. Kirchner and L. Delle Site, J. Chem. Theory Comput., 2011, 7, 3040–3044.

14 J. Schmidt, C. Krekeler, F. Dommert, Y. Zhao, R. Berger, L. D. Site and C. Holm, J. Phys. Chem. B, 2010, 114, 6150–6155.

15 S. Saha, R. K. Roy and P. W. Ayers, Int. J. Quantum Chem., 2009, 109, 1790–1806.

16 R. Mulliken, J. Chem. Phys., 1955, 23, 1841.

17 R. Mulliken, J. Chem. Phys., 1962, 36, 3428.

18 P. Löwdin, J. Chem. Phys., 1950, 18, 365.

19 G. Bruhn, E. R. Davidson, I. Mayer and A. E. Clark, Int. J. Quantum Chem., 2006, 106, 2065–2072.

20 A. Reed, J. Chem. Phys., 1985, 83, 735.

21 A. E. Reed, L. A. Curtiss and F. Weinhold, Chem. Rev., 1988, 88, 899–926.

22 K. B. Wiberg and P. R. Rablen, J. Comput. Chem., 1993, 14, 1504–1518.

23 F. Martin and H. Zipse, J. Comput. Chem., 2005, 26, 97–105.

24 A. Clark, J. Chem. Phys., 2004, 121, 2563.

25 F. De Proft, C. Van Alsenoy, A. Peeters, W. Langenaeker and P. Geerlings, J. Comput. Chem., 2002, 23, 1198–1209.

26 C. Chipot, B. Maigret, J. L. Rivail and H. A. Scheraga, J. Phys. Chem., 1992, 96, 10276–10284.

27 B. H. Besler, K. M. Merz and P. A. Kollman, J. Comput. Chem., 1990, 11, 431–439.

28 K. M. Merz, J. Comput. Chem., 1992, 13, 749–767.

29 C. A. Reynolds, J. W. Essex and W. G. Richards, J. Am. Chem. Soc., 1992, 114, 9075–9079.

30 U. C. Singh and P. A. Kollman, J. Comput. Chem., 1984, 5, 129–145.

31 R. J. Woods and R. Chappelle, J. Mol. Struct., 2000, 527, 149–156.

32 C. I. Bayly, P. Cieplak, W. Cornell and P. A. Kollman, J. Phys. Chem., 1993, 97, 10269–10280.

33 W. D. Cornell, P. Cieplak, C. I. Bayly and P. A. Kollmann, J. Am. Chem. Soc., 1993, 115, 9620–9631.

34 C. M. Breneman and K. B. Wiberg, J. Comput. Chem., 1990, 11, 361–373.

35 L. E. Chirlian and M. M. Francl, J. Comput. Chem., 1987, 8, 894–905.

36 M. L. Connolly, J. Appl. Crystallogr., 1983, 16, 548–558.

37 M. A. Spackman, J. Comput. Chem., 1996, 17, 1–18.

38 M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, Ö. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski and D. J. Fox, Gaussian 09, Revision B.1, 2009.

39 E. Sigfridsson and U. Ryde, J. Comput. Chem., 1998, 19, 377–395.

40 H. R. Leverentz, K. A. Maerzke, S. J. Keasler, J. I. Siepmann and D. G. Truhlar, Phys. Chem. Chem. Phys., 2012, 14, 7669–7678.

41 B. Wang and D. G. Truhlar, J. Chem. Theory Comput., 2012, 8, 1989–1998.

42 R. F. W. Bader, Acc. Chem. Res., 1985, 18, 9–15.

43 E. Sanville, S. D. Kenny, R. Smith and G. Henkelman, J. Comput. Chem., 2007, 28, 899–908.

44 A. Pendás, A. Costales and V. Luaña, Phys. Rev. B: Condens. Matter Mater. Phys., 1997, 55, 4275–4284.

45 W. Tang, E. Sanville and G. Henkelman, J. Phys.: Condens. Matter, 2009, 21, 084204–084207.

46 M. J. F. Martins, A. R. Ferreira, E. Konstantinova, H. A. de Abreu, W. F. Souza, S. S. X. Chiaro, L. G. Dias and A. A. Leitão, Int. J. Quantum Chem., 2012, 112, 3234–3239.

47 T. I. Morrow and E. J. Maginn, J. Phys. Chem. B, 2002, 106, 12807–12813.

48 M. Kohagen, M. Brehm, J. Thar, W. Zhao, F. Müller-Plathe and B. Kirchner, J. Phys. Chem. B, 2010, 115, 693–702.

49 J. N. Canongia Lopes and A. A. H. Pádua, J. Phys. Chem. B, 2004, 108, 16893–16898.

50 J. K. Shah, J. F. Brennecke and E. J. Maginn, Green Chem., 2002, 4, 112–118.

51 Z. Liu, T. Chen, A. Bell and B. Smit, J. Phys. Chem. B, 2010, 114, 4572–4582.

52 Z. Liu, S. Huang and W. Wang, J. Phys. Chem. B, 2004, 108, 12978–12989.

53 J. de Andrade, E. S. Böes and H. Stassen, J. Phys. Chem. B, 2002, 106, 13344–13351.

54 S. M. Urahata and M. C. C. Ribeiro, J. Chem. Phys., 2004, 120, 1855–1863.

55 M. H. Ghatee, A. R. Zolghadr, F. Moosavi and Y. Ansari, J. Chem. Phys., 2012, 136, 124706.

56 C. G. Hanke, S. L. Price and R. M. Lynden-Bell, Mol. Phys., 2001, 99, 801–809.

This journal is © the Owner Societies 2013

Phys. Chem. Chem. Phys., 2013, 15, 1632–1646 | 1645

57 C. J. Margulis, H. A. Stern and B. J. Berne, J. Phys. Chem. B, 2002, 106, 12017–12021.

58 F. Dommert, J. Schmidt, C. Krekeler, Y. Y. Zhao, R. Berger, L. Delle Site and C. Holm, J. Mol. Liq., 2010, 152, 2–8.

59 H. V. Spohr and G. N. Patey, J. Chem. Phys., 2010, 132, 154504.

60 P. Hunt, Mol. Simul., 2006, 32, 1–10.

61 V. Babin, J. Baucom, T. A. Darden and C. Sagui, J. Phys. Chem. B, 2006, 110, 11571–11581.

62 S. Patel, A. D. Mackerell and C. L. Brooks, J. Comput. Chem., 2004, 25, 1504–1514.

63 G. T. Ibragimova and R. C. Wade, Biophys. J., 1998, 74, 2906–2911.

64 E. I. Izgorodina, J. Rigby and D. R. MacFarlane, Chem. Commun., 2012, 48, 1493–1495.

65 Y. Zhao and D. G. Truhlar, Acc. Chem. Res., 2008, 41, 157–167.

66 M. W. Schmidt, K. K. Baldridge, J. A. Boatz, S. T. Elbert, M. S. Gordon, J. H. Jensen, S. Koseki, N. Matsunaga, K. A. Nguyen, S. Su, T. L. Windus, M. Dupuis and J. A. Montgomery, J. Comput. Chem., 1993, 14, 1347–1363.

67 F. Weigend and R. Ahlrichs, Phys. Chem. Chem. Phys., 2005, 7, 3297–3305.

68 H. S. M. Coxeter, in A Spectrum of Mathematics essays presented to H. G. Forder, ed. J. C. Butcher, Auckland Auckland University Press, Wellington Oxford University Press, 1971, pp. 98–107, ch. Virus Macromolecules and Geodesic Domes.

69 A. Stone, The Theory of Intermolecular Forces, Clarendon Press, 1997.

70 H. Li, M. S. Gordon and J. H. Jensen, J. Chem. Phys., 2006, 124, 214108.

71 E. I. Izgorodina, M. Forsyth and D. R. MacFarlane, Phys. Chem. Chem. Phys., 2009, 11, 2452–2458.

72 Y. Wang, H. Li and S. Han, J. Chem. Phys., 2006, 124, 044504.

73 A. K. Abdul-Sada, S. Al-Juaid, A. M. Greenway, P. B. Hitchcock, M. J. Howells, K. R. Seddon and T. Welton, Struct. Chem., 1990, 1, 391–394.

74 E. I. Izgorodina, U. L. Bernard and D. R. MacFarlane, J. Phys. Chem. A, 2009, 113, 7064–7072.

75 E. I. Izgorodina, R. Maganti, V. Armel, P. M. Dean, J. M. Pringle, K. R. Seddon and D. R. MacFarlane, J. Phys. Chem. B, 2011, 115, 14688–14697.

Paper

Section 2.3. *Assessment of atomic partial charge schemes for polarisation and charge*
Chapter 2. *Partial charge schemes, polarisability and charge transfer*
transfer in ionic liquids

View Article Online

38
38

PCCP

References

[1] R.F.W. Bader and P.F. Zou. An atomic population as the expectation value of a quantum observable. *Chemical Physics Letters*, 191(1–2):54 – 58, 1992.

[2] R. Mulliken. Electronic population analysis on LCAO-MO molecular wave functions. ii. overlap populations, bond orders, and covalent bond energies. *The Journal of Chemical Physics*, 23:1841, 1955.

[3] Mark A. Spackman. Potential derived charges using a geodesic point selection scheme. *Journal of Computational Chemistry*, 17:1–18, 1996.

[4] R. Mulliken. Criteria for the construction of good self-consistent-field molecular orbital wave functions, and the significance of LCAO-MO population analysis. *The Journal of Chemical Physics*, 36:3428, 1962.

[5] P. Löwdin. On the non-orthogonality problem connected with the use of atomic wave functions in the theory of molecules and crystals. *The Journal of Chemical Physics*, 18:365, 1950.

[6] A. Reed. Natural population analysis. *The Journal of Chemical Physics*, 83:735, 1985.

[7] Alan E. Reed, Larry A. Curtiss, and Frank Weinhold. Intermolecular interactions from a natural bond orbital, donor-acceptor viewpoint. *Chemical Reviews*, 88:899–926, 1988.

[8] Carlos E. Resende Prado, Mario G. Del Pópolo, T. G. A. Youngs, Jorge Kohanoff, and R. M. Lynden-Bell. Molecular electrostatic properties of ions in an ionic liquid. *Molecular Physics*, 104(15):2477–2483, 2006.

[9] Katharina Wendler, Florian Dommert, Yuan Yuan Zhao, Robert Berger, Christian Holm, and Luigi Delle Site. Ionic liquids studied across different scales: A computational perspective. *Faraday Discussions*, 154:111–132, 2012.

Chapter 3

More than just an ion pair: The need for large-scale calculations

3.1   Declaration for thesis Chapter 3

Declaration by Candidate

In the case of Chapter 3, the nature and extent of my contribution to the work was the following:

Nature of contribution	Extent of contribution(%)
Key ideas, data acquisition, data analysis and manuscript development	80

The following co-authors contributed to the work. If co-authors are students at Monash University, the extent of their contribution in percentage terms must be stated:

Name	Nature of contribution	Extent of contribution (%) for student co-authors only
Dr. Ekaterina I. Izgorodina	Key ideas, data acquisition, data analysis, manuscript development and write-up	_
Prof. Douglas R. MacFarlane	consultation	–

The undersigned hereby certify that the above declaration correctly reflects the nature and extent of the candidates and co-authors contributions to this work.

Candidate's Signature:

Date: 17/10/2014

Supervisor's Signature:
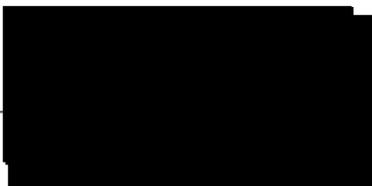
Date: 17/10/2014

3.2 Overview

Accurate energetics are critical in fully understanding thermodynamic and transport properties of ionic liquids (ILs).[1] In order to move away from classical methods, which generally account for only two-body interactions, and density functional theory (DFT) methods that are unpredictable in their ability capture electron correlation and exchange effects,[2] this paper assesses the fragment molecular orbital (FMO) approach[3] as a way forward for large-scale *ab initio* calculations of ILs.

The FMO approach is a many-body approach to calculate the electronic structure of large chemical systems most commonly using the second-order Møller Plesset perturbation theory (MP2) wavefunction.[4] The FMO many-body expansion is usually truncated at the two- (FMO2) or three-body (FMO3) term,[5] although four-body calculations have been reported.[6, 7] In this study, the MP2 level of theory was examined, as it represents the least expensive correlated level of theory and is in routine use for IL energetics.[2, 8, 9] Through this work, two important features of the energetics of ILs are shown. First, the interaction energy of the ionic liquid clusters on a per ion pair (IP) basis remains unconverged even at a size of eight IPs, particularly for the dispersion component. Second, the Hartree-Fock (HF) energy (approximated as the electrostatic energy) requires at least a three-body expansion whereas the dispersion energy can be accounted for as only a two-body term.

This paper is significant as it is the first assessment of a large-scale fully *ab initio* method for ILs in terms of both cluster size and accuracy with respect to a conventional correlated wavefunction-based method. By better understanding the many-body effects of ILs, appropriately sized clusters can be constructed that may be useful in analysing more realistic interactions, particularly in *ab initio* molecular dynamics simulations. Additionally, the rate of convergence of the HF and correlation energies suggests that FMO3 may be used for the HF component, while only using FMO2 for the correlated part, thus further accelerating the calculation. It is worth mentioning that the effects of basis set superposition error are still significant; Table S3 in the supplementary information given in Section 3.4 shows errors approaching 200 kJ mol^{-1} for the eight IP clusters.

3D structures of IL clusters shown in Figure 1 (a) and (b) in the paper that follows may be viewed by scanning [QR1], [QR2], respectively.

QR1

QR2

ChemComm

COMMUNICATION

Large-scale ab initio calculations of archetypical ionic liquids[w][z]

Ekaterina I. Izgorodina,* Jason Rigby and Douglas R. MacFarlane

Fully ab initio large-scale calculations of archetypical ionic liquids consisting of up to eight ion pairs are presented for the first time. These are used to validate the computationally efficient Fragment Molecular Orbital approach applied to these semi-Coulombic systems, paving the way towards accurate prediction of their transport properties.

Ionic liquids (ILs) represent a fascinating class of organic salts that due to their unique physical properties in some cases such as low flammability, low vapour pressure and low melting point, have attracted significant attention in a broad range of applications. For example, they offer a unique set of properties as liquid electrolytes for alternative energy devices such as metal–ion batteries,[1] thermo-cells,[2] fuel cells[3] and solar cells.[4] For enhanced performance of these electrochemical devices the electrolyte must exhibit additional desirable transport properties such as high conductivity and low viscosity and there has been an intense search for ILs that exhibit these critical properties. Computer-aided design of ionic liquids with tuned thermodynamic and transport properties has the potential to make a significant contribution to this search, however these materials represent a significant challenge for the current theoretical approaches.[5] Classic molecular dynamics (MD) simulations are limited due to their simplified description of intermolecular interactions.[6] Current force fields for ionic liquids are based on two-body interacting potentials for non-covalently bonded interactions such as Coulomb (electrostatics) and dispersion,[7] disregarding other corrections for many-body effects. Although long-range Coulomb interactions are undoubtedly the strongest interactions amongst the forces between ions in the compounds,[8] the importance of short-range dispersion interactions has recently become clear.[9] Therefore, for MD simulations to be reliable and robust a wavefunction-based ab initio (AI) method accounting for electron correlation needs to be used. Simulations based on DFT functionals are becoming more popular[10] due to robustness of the non-biased description of fundamental forces; however, a universal DFT functional performing reliably for all types of

ionic liquids is yet to be found.[11] On the other hand, while wavefunction-based AI methods are very accurate, their successful application in MD simulations of ILs is hindered because of poor scalability with molecular size, making the calculation task rapidly unwieldy.

The fragment molecular orbital (FMO) approach offers an exciting solution to this problem, by making any AI-based calculation scale linearly with molecular size.[12] The idea behind the approach lies in dividing a large molecular system into smaller fragments, which can be treated individually at a high level of AI theory. For ionic liquids the FMO approach is particularly attractive as each ion can already be treated as an individual fragment. The split of a large-sized ionic cluster into individually treated ions makes the AI calculation hugely parallelisable (one ion per CPU) and results in significantly improved computational time without any sacrifice in accuracy,[13] thus paving the way towards fully AI MD simulations of ILs.

Here we present FMO-based calculations of IL clusters using the wavefunction-based MP2 method for each ion and compare the result with large-scale MP2 calculations of the whole cluster. A triple-x doubly polarised Ahlrichs type basis set, TZVPP, was used for both calculations. This first systematic study focuses on analysing the role of many-body effects and dispersion interactions with increasing number of ion pairs in the ionic cluster. We show that the FMO approach in combination with the MP2 method provides a reliable and accurate description of the interaction energies of ionic clusters, as long as three-body corrections are included.

The ionic liquids chosen are $[NMe_4][BF_4]$, where Me stand for the methyl group, and a series of imidazolium-based ILs with increasing alkyl chain length on the cation: $[C_1mim][BF_4]$, $[C_3mim][BF_4]$ and $[C_4mim][BF_4]$. The ionic clusters studied consisted of 1, 2, 4 and 8 ion pairs. Single ion pair optimisations of the imidazolium-based series were fully screened for the lowest energy conformations, in which the BF_4 anion interacted either with the C2–H bond of the imidazolium-ring or above/below the imidazolium ring. The resulting conformations were used to build the subsequent clusters, followed by geometry optimisations. In all cases, the alkyl chains of the cation extended uniformly on the same side, allowing for van der Waals interactions between chains. The clusters constructed in this manner were ensured to converge to conformations that are close to the global minima. Examples of optimised ionic clusters consisting of 8 ion pairs are shown in Fig. 1. All calculations were performed using GAMESS-US.[14] (For more detail see ESI.[z])

School of Chemistry, Monash University, Wellington Rd, Clayton, VIC 3800, Australia. E-mail: ▮▮▮▮▮▮▮▮▮
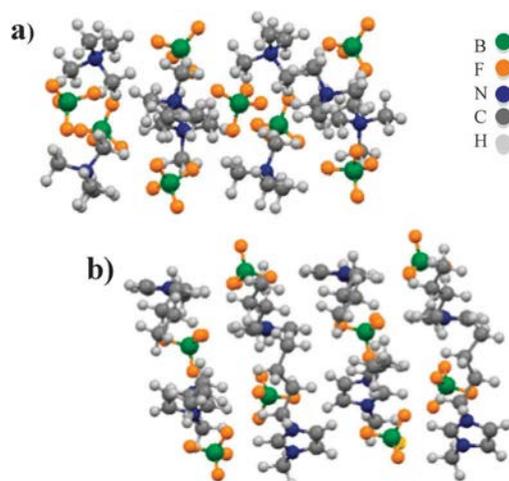
Fig. 1 Optimised structures of ionic clusters with 8 ion pairs of [NMe$_4$][BF$_4$] (a) and [C$_4$mim][BF$_4$] (b).

In the FMO approach the total electronic energy of the system can be then expressed as follows:

$$E = \sum_I E_I + \sum_{I>J} (E_{IJ} - E_I - E_J) $$
$$+ \sum_{I>J>K} [(E_{IJK} - E_I - E_J - E_K) - (E_{IJ} - E_I - E_J) $$
$$- (E_{JK} - E_J - E_K) - (E_{IK} - E_K - E_I)] \quad (1)$$

Where I, J and K represent different ions. The second sum represents a correction for the two-body inter-fragment inter-actions (abbreviated as FMO2), whereas the third sum is a correction for the three-body inter-fragment interactions (abbreviated as FMO3). If used together with the MP2 level of theory, the method becomes FMO2-MP2 and FMO3-MP2, respectively. The rest of the system is treated as a "Coulomb bath", in which intermolecular interactions beyond the cut-off point are treated as purely electrostatic.

Interaction energies of the IL clusters were calculated based on the geometries of ions in the cluster using the counterpoise approach by Boys and Bernardi[15] to account for the basis set superposition error. The energies increase linearly with increasing number of ion pairs (see ESI Fig. S2z). When the interaction energy is normalised by the number of ion pairs in the cluster (Fig. 2), a steep behaviour is observed initially, from 1 to 2 ion pairs up to 60 kJ mol^{-1}; the trend begins to level out for larger clusters, with the energy per ion pair increasing only slightly from 4 to 8 ion pairs. The observed trends are a manifestation of the many-body effects that are important in any system consisting of a number of interacting non-covalently bonded species, especially charged ones.

One of the main questions here is how far reaching are these many-body effects? Are the two-body pair-wise interactions already sufficient for accurate description of the ionic cluster energy? In order to answer this question, the FMO calculations were performed using the two-body and three-body corrections. In the FMO calculations the ion clusters were fragmented into constituent cations and anions. The FMO2 and FMO3 results are compared with the MP2 calculations of the whole
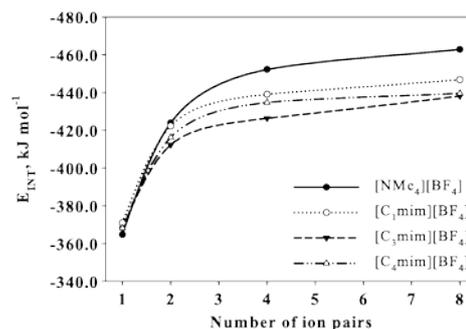
Fig. 2 Interaction energies (E_{INT}, kJ mol^{-1}) per ion pair.

clusters in Table 1. It becomes obvious that for the majority of the ionic clusters the two-body correction reproduces the MP2 interaction energy within 1 to 2 kJ mol^{-1}. However, the FMO2-MP2 errors seem to be rather unsystematic and increase slightly with increasing cluster size. Closer inspection of the FMO2 and FMO3 energies reveals that the errors arise mainly from the HF component of the interaction energy consisting of Coulomb, exchange-repulsion and induction terms (see ESI Table S2z), with the dispersion component being already treated accurately within the two-body approach. In our previous work we showed that although long-range electrostatic forces converged in ionic liquids,[16] the convergence range extended over a few Angstroms from the reference ion to achieve the accuracy of 1 kJ mol^{-1} for the electrostatic lattice energy.[16] The current FMO results highlight the fact that the two-body potential for Coulomb interactions might not be enough for ILs, requiring at least a three-body treatment for higher accuracy. On the other hand, FMO3-MP2 produces results within 0.2 kJ mol^{-1}, which is better than spectroscopic accuracy of 1 kJ mol^{-1}. More importantly, the errors seem to be independent of the cluster size, as well as the ion size. This level of accuracy is extremely desirable for accurate MD simulations of transport properties such as conductivity and viscosity of any semi-Coulombic system.

Highly accurate FMO3-MP2 calculations come with another advantage of significantly reduced computational cost. For example, a single-point FMO3-MP2 calculation of the [C$_3$mim][BF$_4$] cluster containing 8 ion pairs took 78 h to complete on 32 cores and required as little as 2 GB of RAM per core. The same cluster calculated at the full MP2 level needed 221 h (≈9 days) on 16 cores and used 29 GB of RAM per core. In this case, the job could only be completed on a 1 TB RAM machine. Memory consumption alone puts the full MP2 beyond the reach of most hardware presently used in typical High Performance Computing (HPC) environments. Combined with fast-interconnect nodes the FMO3-MP2 method can be easily performed on HPC clusters.

Table 1 Deviations (in kJ mol^{-1}) of the FMO2-MP2 and FMO3-MP2 interactions energies per ion pair from the MP2 results

	[NMe$_4$]		[C$_1$mim]		[C$_3$mim]		[C$_4$mim]	
N	FMO2	FMO3	FMO2	FMO3	FMO2	FMO3	FMO2	FMO3
2	-1.6	0.1	-0.8	-0.2	-1.3	0.0	-1.1	-0.1
4	-1.4	0.1	-1.0	0.1	-1.6	-0.3	-1.2	0.1
8	-1.9	0.2	-1.0	0.0	-0.5	-0.1	—	—
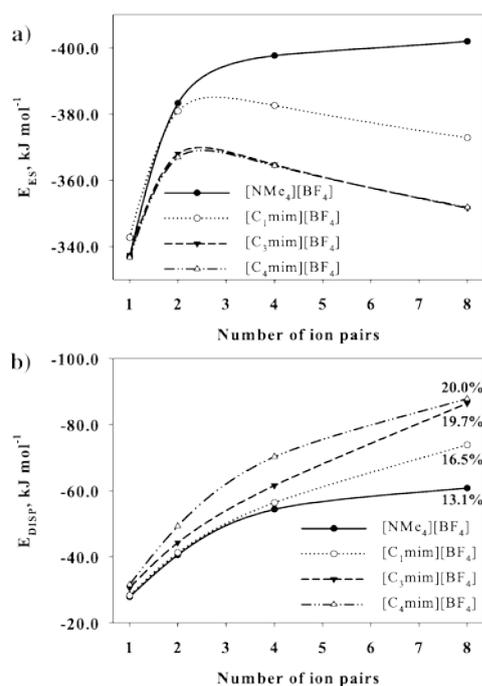
Fig. 3 (a) Total electrostatic (E_{ES}) and (b) Dispersion (E_{DISP}) components of interaction energies per ion pair.

the fundamental components of the interaction energy, accurate energy decomposition schemes are required.

In summary, fully AI large-scale calculations of archetypical ILs showed that many-body effects beyond two-body interactions were particularly important for Coulomb forces if accuracy below 1 kJ mol^{-1} is needed. Dispersion interactions increase rapidly with increasing cluster size, highlighting the need for accurate description of electron correlation in order to predict physical properties of ionic liquids. The FMO approach shows significant potential for studying these semi-Coulombic systems. In combination with the MP2 level of theory, FMO3-MP2 (incorporating the three-body correction) gives excellent accuracy of 0.2 kJ mol^{-1} at a significantly reduced computational cost compared to the MP2 method, thus paving the way towards fully AI MD simulations of ionic liquids and hence, accurate prediction of their transport properties.

The trend in the interaction energies in Fig. 2 hides a subtle interplay between the two main fundamental forces in ionic liquids: electrostatic and dispersion. In the first approximation the dispersion interactions were estimated as the difference between MP2 and HF interaction energies, whereas HF interaction energies were considered as total electrostatic energies. Fig. 3 shows the electrostatic and dispersion components of the interaction energy per ion pair. The dispersion component (Fig. 3-b) increases rapidly with increasing cluster size from about 30 kJ mol^{-1} for a single ion pair to 61 kJ mol^{-1} for [NMe$_4$][BF$_4$] and about 87 kJ mol^{-1} for [C$_3$mim][BF$_4$] and [C$_4$mim][BF$_4$] for the largest cluster size. The dispersion component begins to contribute from as little as 8% for single ion pairs up to 20% of the overall energy in the large clusters emphasising the importance of electron correlation effects as cluster size increases. This, in part, reflects the presence in the larger cluster of all of the possible inter-ion interactions, including like-ion dispersion interactions such as inter-alkyl chains, some of which may be missing in the smaller cluster. The total electrostatic component appears to slightly decrease with the cluster size in the case of the imidazolium-based ILs (Fig. 3-a). Although unusual at first glance, the trend is a result of a complex interplay between the charge transfer due to orbital overlap between ions and polarisation effects due to induced dipole moments. The latter are subject to electron correlation and are likely to be already included in the calculated dispersion component. To quantitatively evaluate

Notes and references

1 S. Seki, T. Kobayashi, N. Serizawa, Y. Kobayashi, K. Takei, H. Miyashiro, K. Hayamizu, S. Tsuzuki, T. Mitsugi, Y. Umebayashi and M. Watanabe, J. Power Sources, 2010, 195, 6207.
2 T. J. Abraham, D. R. MacFarlane and J. M. Pringle, Chem. Commun., 2011, 47, 6260.
3 S. Y. Lee, T. Yasuda and M. Watanabe, J. Power Sources, 2010, 195, 5909.
4 V. Armel, J. M. Pringle, M. Forsyth, D. R. MacFarlane, D. L. Officer and P. Wagner, Chem. Commun., 2010, 46, 3146.
5 E. I. Izgorodina, Phys. Chem. Chem. Phys., 2011, 13, 4189.
6 E. Maginn, J. Phys.: Condens. Matter, 2009, 21, 3731019.
7 (a) J. N. Canongia Lopes, J. Deschamps and A. A. H. Pá dua, J. Phys. Chem. B, 2004, 108, 11250; (b) J. N. Canongia Lopes, J. Deschamps and A. A. H. Pá dua, J. Phys. Chem. B, 2004, 108, 2038; (c) J. N. Canongia Lopes and A. A. H. Pá dua, J. Phys. Chem. B, 2004, 108, 16893.
8 S. Tsuzuki, H. Tokuda, K. Hayamizu and M. Watanabe, J. Phys. Chem. B, 2005, 109, 16474.
9 (a) P. Ballone, C. Pinilla, J. Kohanoff and M. G. Del Pó polo, J. Phys. Chem. B, 2007, 111, 4938; (b) K. Shimizu, M. Tariq, M. F. C. Gomes, L. P. N. Rebelo and J. N. C. Lopes, J. Phys. Chem. B, 2010, 114, 5831; (c) U. L. Bernard, E. I. Izgorodina and D. R. MacFarlane, J. Phys. Chem. C, 2010, 114, 20472.
10 S. Zahn and B. Kirchner, J. Phys. Chem. A, 2008, 112, 8430.
11 E. I. Izgorodina, U. L. Bernard and D. R. MacFarlane, J. Phys. Chem. A, 2009, 113, 7064.
12 (a) D. G. Fedorov and K. Kitaura, J. Chem. Phys., 2004, 120, 6832; (b) D. G. Fedorov and K. Kitaura, J. Chem. Phys., 2004, 121, 2483.
13 M. S. Gordon, J. M. Mullin, S. R. Pruitt, L. B. Roskop, L. V. Slipchenko and J. A. Boatz, J. Phys. Chem. B, 2009, 113, 9646.
14 M. W. Schmidt, K. K. Baldridge, J. A. Boatz, S. T. Elbert, M. S. Gordon, J. H. Jensen, S. Koseki, N. Matsunaga, K. A. Nguyen, S. Su, T. L. Windus, M. Dupuis and J. A. Montgomery, J. Comput. Chem., 1993, 14, 1347.
15 S. F. Boys and F. Bernardi, Mol. Phys., 1970, 19, 553.
16 E. I. Izgorodina, U. L. Bernard, P. M. Dean, J. M. Pringle and D. R. MacFarlane, Cryst. Growth Des., 2009, 9, 4834.

Large-scale calculations of archetypical ionic liquids

Ekaterina I Izgorodina,[*,a] **Jason Rigby**[a] **and Douglas R MacFarlane**[a]

School of Chemistry, Monash University, Wellington Rd, Clayton, VIC 3800, AUSTRALIA. Fax: +61 3 9905 4590; Tel: +61 3 9905 8639; E-mail: katya.izgorodina@monash.edu

SUPLLEMENTARY INFORMATION

Section 3.4. *Supporting information: Large-scale* ab initio *calculations of archetypical ionic liquids (abridged)* 46

Details of quantum chemical calculations

Clusters of ILs consisting of 1, 2, 4 and 8 ion-pairs were optimised using a meta-GGA hybrid M06-2X functional[1] with the 6-31+G(d) basis set for 1, 2 and 4 ion pairs, and 6-31G(d) for 8 ion pairs. Improved electronic energies were calculated at the MP2, FMO2-MP2 and FMO3-MP2 level of theory using the TZVPP basis set and the Resolution-of-Identity approaximation.[2] All interaction energies were counterpoised corrected using the Boys and Bernardi approach.[3]

References

1. Zhao, Y.; Truhlar, D. G., *Acc. Chem. Res.* 2008, **41**, 157.
2. Katouda, M.; Nagase, S., *Int. J. Quant. Chem.* 2009, **109**, 2121.
3. Boys, S. F.; Bernardi, F., *Mol. Phys.* 1970, **19**, 553.
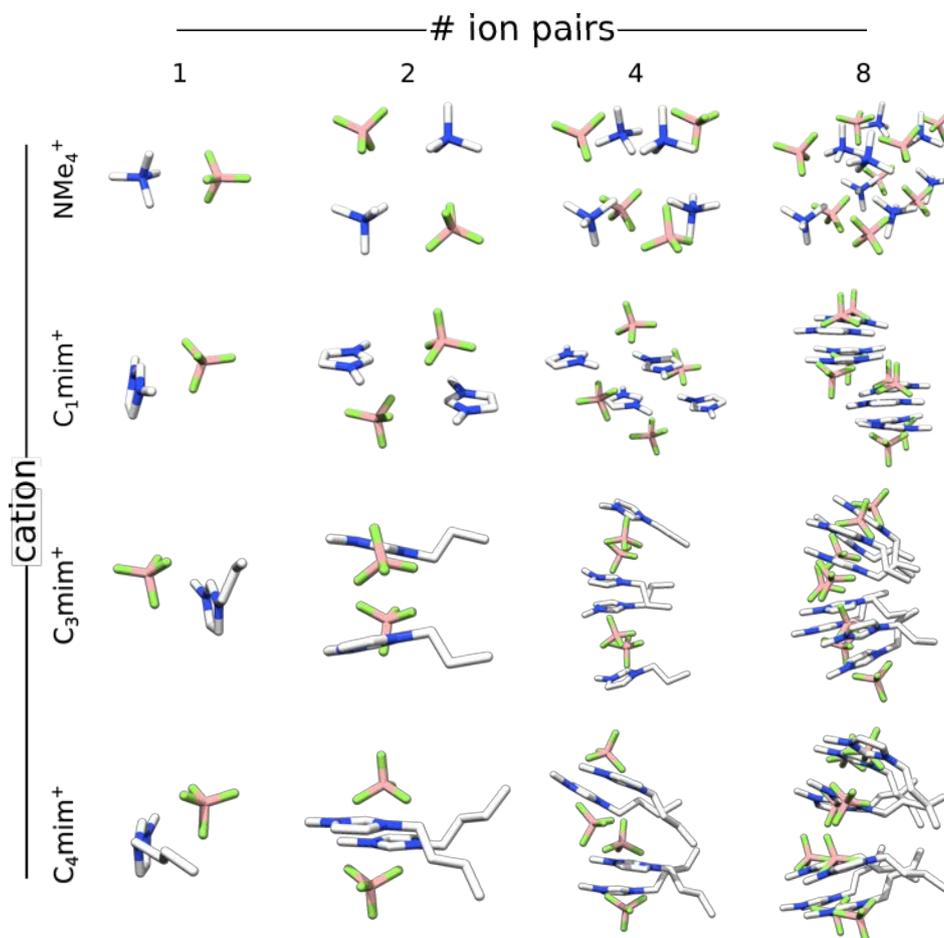
Figure S1. Optimised geometries of the ionic clusters studied.

Section 3.4. *Supporting information: Large-scale* ab initio *calculations of archetypical ionic liquids (abridged)* 48
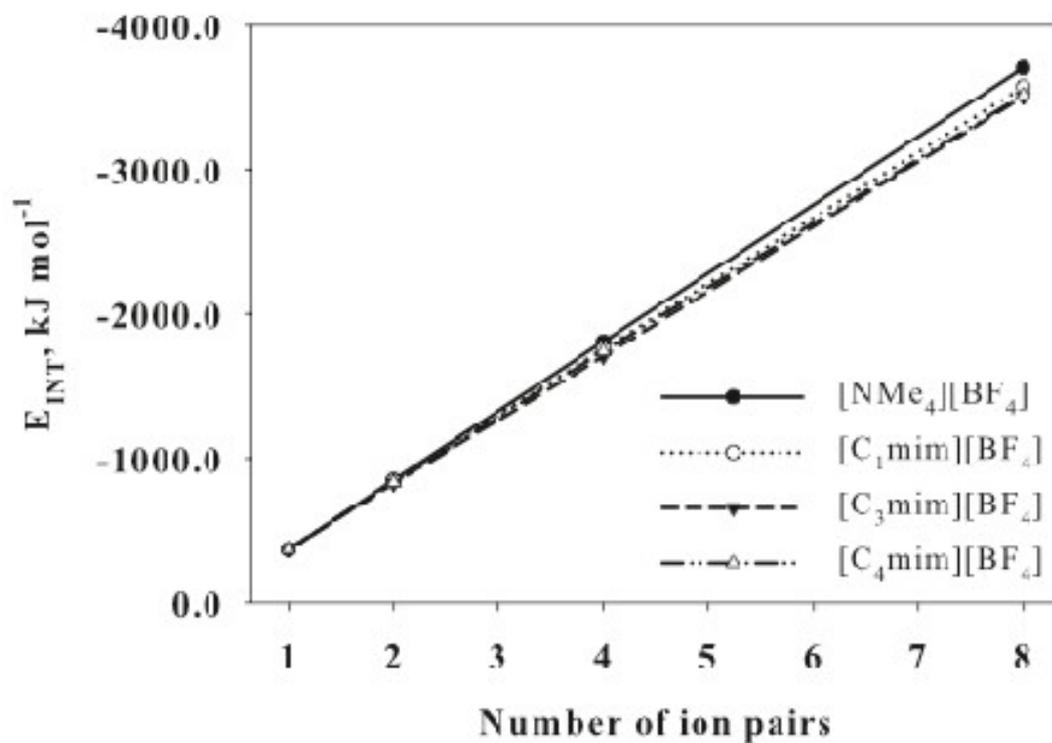
Figure S2. Total interaction energies (in kJ mol^{-1}) of ionic clusters

Table S2. Errors for the FMO2-RHF and FMO3-RHF interactions energies (in kJ mol^{-1}) per ion pair with respect to full RHF.

	[NMe$_4$][BF$_4$]		[C$_1$mim][BF$_4$]		[C$_3$mim][BF$_4$]		[C$_4$mim][BF$_4$]	
	FMO2	FMO3	FMO2	FMO3	FMO2	FMO3	FMO2	FMO3
2	-0.7	0.1	-0.8	0.0	-0.9	0.1	-1.0	0.1
4	-1.0	0.1	-1.4	0.2	-1.6	0.0	-1.5	0.3
8	-1.5	0.2	-1.4	0.2	-1.0	0.3	-	-

Table S3. RI-MP2 Basis set superposition errors, total and per ion pair (IP) in the interaction energies of the ionic clusters studied. N is the number of ion pairs in the cluster.

	[NMe$_4$][BF$_4$]		[C$_1$mim][BF$_4$]		[C$_3$mim][BF$_4$]		[C$_4$mim][BF$_4$]	
N	Total BSSE	BSSE per IP	Total BSSE	BSSE per IP	Total BSSE	BSSE per IP	Total BSSE	BSSE per IP
1	-9.6	-9.6	-11.0	-11.0	-11.8	-11.8	-12.6	-12.6
2	-26.6	-13.3	-32.1	-16.1	-33.5	-16.7	-37.5	-18.7
4	-66.1	-16.5	-74.4	-18.6	-80.5	-20.1	-84.3	-21.1
8	-143.8	-18.0	-174.0	-21.7	-193.0	-24.1	-	-

The Counterpoise method by Boys and Bernadi was used to account for the basis set superposition error. Interaction energies were calculated as follows:

$$\Delta E_{INT} = E_{\text{cluster}} - \sum E_{\text{ion}}$$

where the energies of the constituting ions were calculated within the basis set of the entire cluster.

References

[1] Ekaterina I. Izgorodina. Towards large-scale, fully ab initio calculations of ionic liquids. *Physical Chemistry Chemical Physics*, 13:4189–4207, 2011.

[2] Stefan Zahn, Douglas R. MacFarlane, and Ekaterina I. Izgorodina. Assessment of Kohn-Sham density functional theory and Møller-Plesset perturbation theory for ionic liquids. *Physical Chemistry Chemical Physics*, 15:13664–13675, 2013.

[3] Kazuo Kitaura, Eiji Ikeo, Toshio Asada, Tatsuya Nakano, and Masami Uebayasi. Fragment molecular orbital method: an approximate computational method for large molecules. *Chemical Physics Letters*, 313(3–4):701 – 706, 1999.

[4] D. Fedorov and K. Kitaura. *The Fragment Molecular Orbital Method: Practical Applications to Large Molecular Systems*. Taylor & Francis, 2009.

[5] Dmitri G. Fedorov and Kazuo Kitaura. The importance of three-body terms in the fragment molecular orbital method. *The Journal of Chemical Physics*, 120(15):6832–6840, 2004.

[6] Tatsuya Nakano, Yuji Mochizuki, Katsumi Yamashita, Chiduru Watanabe, Kaori Fukuzawa, Katsunori Segawa, Yoshio Okiyama, Takayuki Tsukamoto, and Shigenori Tanaka. Development of the four-body corrected fragment molecular orbital (FMO4) method. *Chemical Physics Letters*, 523:128 – 133, 2012.

[7] Yoshio Okiyama, Takayuki Tsukamoto, Chiduru Watanabe, Kaori Fukuzawa, Shigenori Tanaka, and Yuji Mochizuki. Modeling of peptide–silica interaction based on four-body corrected fragment molecular orbital (FMO4) calculations. *Chemical Physics Letters*, 566:25 – 31, 2013.

[8] Seiji Tsuzuki, Hiroyuki Tokuda, Kikuko Hayamizu, and Masayoshi Watanabe. Magnitude and directionality of interaction in ion pairs of ionic liquids: relationship with ionic conductivity. *The Journal of Physical Chemistry B*, 109(34):16474–16481, 2005. PMID: 16853095.

[9] Ekaterina I. Izgorodina, Uditha L. Bernard, and Douglas R. MacFarlane. Ion-pair binding energies of ionic liquids: Can DFT compete with ab initio-based methods? *The Journal of Physical Chemistry A*, 113(25):7064–7072, 2009. PMID: 19462960.

Chapter 4

Overcoming bottlenecks: Implicit BSSE correction with a modified SCS-MP2 approach

4.1 Declaration for thesis Chapter 4

Declaration by Candidate

In the case of Chapter 4, the nature and extent of my contribution to the work was the following:

Nature of contribution	Extent of contribution(%)
Key ideas, data acquisition, data analysis, manuscript development and writing up	90

The following co-authors contributed to the work. If co-authors are students at Monash University, the extent of their contribution in percentage terms must be stated:

Name	Nature of contribution	Extent of contribution (%) for student co-authors only
Dr. Ekaterina I. Izgorodina	Key ideas, proofreading and final drafting	_
Dr. Santiago Barrera Acevedo	Structure optimisation for work in Section 4.4	_

The undersigned hereby certify that the above declaration correctly reflects the nature and extent of the candidates and co-authors contributions to this work.

Candidate's Signature:

Date: 17/10/2014

Supervisor's Signature:

Date: 17/10/2014

4.2 Overview

Spin-component scaled second-order Møller Plesset perturbation theory (SCS-MP2)[1] has become a widely used computational technique owing to its ability to approach the accuracy of higher levels of theory at no increased computational expense. It has been noted previously that in addition to providing improved energies, a unique combination of SCS-MP2 and a triple-ζ basis set could yield high-quality energies without any counterpoise correction;[2] a technique that is used to remove basis set superposition error (BSSE).[3, 4]

The publication that follows extends on the SCS-MP2 approach, where the same- and opposite-spin components of the second-order Møller Plesset perturbation theory (MP2) correlation energy are fitted to CCSD(T)/CBS benchmark energies of ionic liquid (IL) ion pairs (IPs), to propose a new set of coefficients targeted to . The new method, known as ionic liquid specific SCS-MP2 (SCS-IL-MP2), is founded on a rigorous statistical analysis and is the first to unequivocally prescribe both scaling factors *and* basis set. Furthermore, this work is explicitly geared towards circumventing the need for counterpoise correction, and to this end errors of less than 1 kJ mol^{-1} with respect to CCSD(T)/CBS have been obtained.

In Section 4.4, the SCS-IL-MP2 method is applied to IL clusters to show its broader application beyond the IP. Its performance is attributed to the accurate treatment of correlation energy as a two-body term[5] and BSSE being only a short-ranged source of error. Section 4.4 is presented as a submitted manuscript.

This work removes the long-standing bottleneck of large-scale *ab initio* calculations of counterpoise correction, which requires all monomers of the IL cluster to be calculated in the basis set of the entire cluster. This represents a substantial computational expense as such a calculation cannot be performed using the fragment molecular orbital (FMO) theory since no further fragmentation can be achieved. By implicitly accounting for BSSE, the efficiency of highly parallel methods such as the FMO approach may be fully exploited.

QR1

3D structures of IL clusters shown in Figures 2 (a), (b) and (c) in the paper presented in Section 4.4 may be viewed by scanning [QR1], [QR2] and [QR3], respectively.

QR2

QR3

New SCS- and SOS-MP2 Coefficients Fitted to Semi-Coulombic Systems

Jason Rigby and Ekaterina I. Izgorodina*

School of Chemistry, Monash University, Wellington Road, Clayton Victoria 3800, Australia

⁎ Supporting Information

ABSTRACT: Spin-component scaled second-order Møller−Plesset perturbation theory (SCS-MP2) energy calculations, which independently scale the opposite- and same-spin components of the MP2 correlation energy, are known to consistently provide improved interaction energies in comparison to conventional MP2. This has led to the development of a number of SCS-MP2 derivatives that target particular classes of molecules, interactions or properties. In this study, SCS-MP2 scaling coefficients targeted to interaction energies of single ion pair semi-Coulombic ionic liquid (IL) systems are presented in view of circumventing the need for counterpoise correction to eliminate basis set superposition error (BSSE). A set of 174 IL ion pairs consisting of imidazolium ($[C_{(1-4)}mim]^+$) and pyrrolidinium ($[C_{(1-4)}mpyr]^+$) cations and routinely used anions such as Br^-, Cl^-, $[BF_4]^-$, $[PF_6]^-$, $[DCA]^-$ (dicyanamide), $[tos]^-$ (tosylate), $[mes]^-$ (mesylate), and $[NTf_2]^-$ (bis-(trifluoromethylsulfonyl)amide), each of which were arranged in multiple favorable conformations, were calculated at the MP2 level of theory with 17 popular basis sets ranging from double- to quadruple-ζ quality and at the CCSD(T)/CBS limit. For each basis set, the spin components of the IL set were scaled via least-squares multiple linear regression with respect to CCSD(T)/CBS benchmark interaction energies that were corrected for BSSE using the Boys and Bernardi approach. SCS-MP2 spin component coefficients of 1.05 and 0.68 are recommended for the opposite- and same-spin components, respectively, in conjunction with Dunning's cc-pVTZ basis set, which resulted in the most statistically reliable regression. Alternatively, a scaled opposite-spin MP2 (SOS-MP2) scaling factors of 1.64 is recommended for the opposite-spin component and should be used where the omission of the same-spin component results in a calculation speed-up. These two scaling schemes are termed SCS-IL-MP2 and SOS-IL-MP2, respectively. The SCS-IL-MP2 and SOS-IL-MP2 approaches show interaction energy errors on average less than 1.0 kJ mol^{-1} with respect to CCSD(T)/CBS benchmark results and highlights the important consideration of basis set dependence when selecting spin-component coefficients. By calculating multiple conformations for each ion pair and scaling to reproduce BSSE corrected benchmark energies, it is suggested that improved energies may be obtained for larger IL clusters beyond ion pairs without performing costly counterpoise corrections.
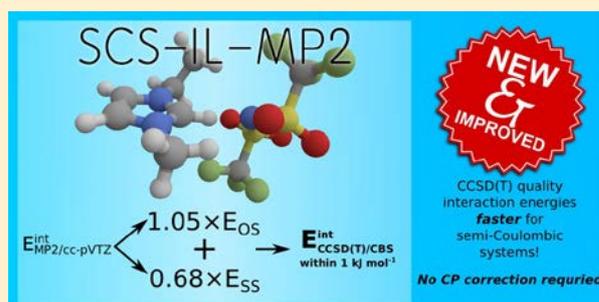
1. INTRODUCTION

Ionic liquids (ILs) have found themselves suited to a surprisingly large array of applications owing to their unique physical properties.[1] These arise from the staggering variety of anion/cation combinations焀speculated to be in excess of 10^{18}, considering binary and ternary ILs[2]焀and the unique combination of intermolecular interactions thereof. Consequently, ILs are an exceptionally challenging class of compounds to model computationally; while driven primarily by electrostatic and induction interactions, charge transfer and dispersion forces (e.g., π−π stacking, alkyl chain interactions) are non-negligible and have a strong influence on their physical properties.[3] In particular, dispersion interactions are notoriously difficult to treat with appreciable accuracy despite the ongoing development of cost-effective methods, such as Density Functional Theory (DFT),[4] and attempts to derive empirical London dispersion coefficients.[5] These approaches rely on assumptions that are unavoidably linked to the chemical systems used during their development; as more novel and unique systems are explored, the reliability of these methods is questionable. For example, DFT has been shown to generally overestimate charge transfer interactions,[6] which are particularly important for ILs in molecular dynamics simulation.[7] Similarly, the reliability of empirical dispersion coefficients comes into question as, for ionic liquids, error cancellation between the anion and cation is said to explain the "unexpectedly well" performing DFT-D3 dispersion coefficients.[8] This implies that error cancellation may not hold for all ion pair combinations or clusters of multiple ion pairs. Thus, the only truly unbiased methods remain in the realm of ab initio quantum mechanics by virtue of their systematic convergence to the exact solution of the Schrödinger equation.

The second-order Møller−Plesset perturbation theory (MP2) is the lowest level ab initio method incorporating electron correlation and is therefore an inexpensive, yet unbiased account of dispersion interactions. Its computational cost, formally $O(N^5)$ where N is the number of basis functions, has been the focus of significant optimization. Some notable examples include the now standard Resolution-of-Identity (RI) approximation, local MP2 (LMP2)[9] methods that boast linear scalability, and attempts to leverage GPU technology.[10] More broadly, the Fragment Molecular Orbital (FMO) approach[11] is routinely used by our group to explore large systems[12] and may, in theory, be combined with any of the aforementioned MP2 incarnations.

In order to improve MP2 energies, a novel approach known as "spin-component scaled MP2" (SCS-MP2) was proposed by Grimme, whereby the spin components of the MP2 correlation energy are scaled such that the energetic properties better reproduce those of higher levels of theory.[13] The originally proposed MP2 scaling factors are shown in eq 1 as α and β for the opposite-spin and same-spin, respectively.

$$E_{MP2}^{corr} = \alpha E_{OS} + \beta E_{SS} \tag{1}$$

In Grimme's scheme, the coefficients were determined under the assumption that the MP2 correlation energy is systematically underestimated at approximately 80%; thus, the opposite-spin component (numerically, the largest) should be scaled up by an additional 20%, hence α = 6/5. The same-spin component was then determined empirically in order to best reproduce a set of 51 reaction energies computed at the QCISD(T) level of theory with an augmented quadruple-ζ basis set, ultimately yielding β = 1/3.

Following Grimme's initial work on SCS-MP2, a number of derivatives have emerged that aim to further optimize these coefficients. Most notably, scaled opposite-spin MP2 (SOS-MP2) was proposed by Jung et al.[14] in an attempt to achieve comparable accuracy with respect to SCS-MP2 while neglecting entirely the same-spin component (α = 1.3, β = 0). By neglecting the same-spin component, SOS-MP2 in conjunction with the RI approximation can be shown to have an increased scalability of $O(N^4)$, while conventional and RI-MP2 both have a scalability of $O(N^5)$. Other noteworthy SCS-MP2 derivatives include the SCS(MI)-MP2 approach by Distasio and Head-Gordon,[15] where the coefficients were scaled using multivariate linear regression analysis fitted to intermolecular interaction energies of the S22 data set[16] with respect to CCSD(T) benchmark data (α = 0.40, β = 1.29); and the SCSN-MP2 approach by Hill and Platts,[17] which is optimized for nucleic acid base pair interaction energies using LMP2 and the RI approximation (α = 0, β = 1.76). Their study also presented data showing impact of basis set in the choice of coefficients with α ranging from 1.29 to 1.75, and β ranging from 0.17 to 0.40 involving double-, triple-, quadruple-ζ, and CBS extrapolated MP2 energies.

In terms of a theoretical understanding of the scaling coefficients, Szabados[18] utilized two-parameter Feenberg scaling[19] of the zeroth order Hamiltonian to minimize the third-order energy, thereby deriving α and β without fitting to benchmark levels of theory. In Szabados' study, the average values were α = 1.12 and β = 0.84 for the systems examined. Subsequently, Fink[20] has presented scaling factors where α = 1.15 and β = 0.75, the latter being almost three times that of the original SCS-MP2 method. Fink's coefficients were found for

an "S2 perturbation theory," which is an extension of Møller−Plesset perturbation theory addressing a criticism of SCS-MP2 that it is no longer systematically improvable nor a truly ab initio method, as noted by Grimme himself.[13] Fink's scaling factors may be found either by Feenberg scaling or by fitting to the full configuration interaction (FCI) wave function. It is important to note that scaling factors derived entirely from theory also demonstrate significant basis set dependence, as noted by Szabados.[18] Further, the derivation of theory-based scaling factors is nontrivial for large systems as the third-order energies required are computationally expensive for Feenberg scaling, and substantially more so for fitting to the FCI wave function. For a comprehensive review of spin-component scaled methods including those not mentioned here, see ref 21.

Apart from improving MP2 energies in terms of their agreement with higher levels of theory,[22] a well-documented phenomenon known as basis set superposition error (BSSE) must also be addressed for accurate interaction energies. BSSE manifests itself as an artificial overstabilization that arises from an unbalanced description of orbitals at the interface between interacting molecules and exists within the Linear Combination of Atomic Orbitals (LCAO) framework. The counterpoise (CP) correction scheme by Boys and Bernardi[23] is the most common method used to quantify this error and involves the calculation of the complex and monomers in their own basis set, and the basis set of the complex, as shown in eqs 2−4,

$$\Delta E_{int} = E_{complex} - \sum_i E_i \tag{2}$$

$$\Delta E_{int}^{CP} = E_{complex} - \sum_i^{All\ BFs} E_i \tag{3}$$

$$\Delta E_{BSSE} = \Delta E_{int}^{CP} - \Delta E_{int} \tag{4}$$

where E_i is monomer i in the complex and superscript "All BFs" indicates the presence of ghost atoms representing the basis functions of the entire complex, that is, all basis functions.

It becomes obvious that the removal of BSSE is nontrivial for clusters of large sizes which are necessary to account for the many-body effects in condensed systems such as ionic liquids; the number of calculations required increases as 2N + 1 where N is the number of monomers in the complex. Our group has shown that, for ionic liquids, the near-linearly scalable (with respect to the number of CPUs used) FMO approach is an extremely accurate method for the calculation of the total electronic energy, that is, $E_{complex}$.[12] In essence, the FMO method works by considering the total electronic energy as the sum of the constituent fragments (monomers) that are then corrected by many-body terms up to an arbitrary order; usually two- (dimer) or three-body (trimer) effects are considered (see ref 24 for a detailed description of FMO theory). While the total electronic energy is straightforward, the calculation of each monomer in the basis set of the complex is challenging. This is due to a limitation of the FMO method where all fragments must have occupied orbitals; fragments consisting entirely of ghost atoms, and therefore contain no electrons, are not allowed. With this in mind, scaling factors well suited to the ionic liquid (semi-Coulombic) systems our group is interested in, fitted to CP corrected interaction energies, are likely to minimize the cumulative error associated with scaling large systems and the BSSE. Interestingly, Antony and Grimme[25] have noted that with a triple-ζ quality basis set, BSSE is minimized through a cancellation of errors, with several other

Article
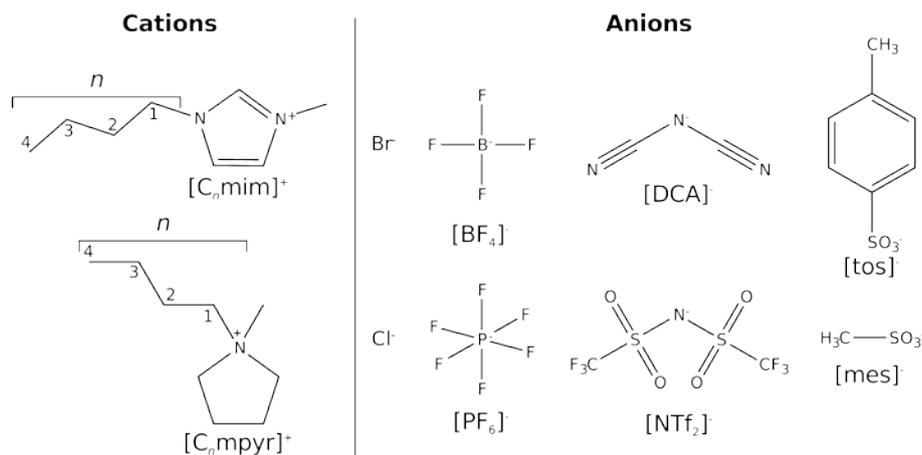
Figure 1. Summary of cations (left) and anions (right) used in this study.

studies showing that CP correction of SCS-MP2 energies has a detrimental effect.[22,26,27] Zahn et al.[22] have shown that, in the case of ionic liquids, uncorrected SCS-MP2 energies alone may produce acceptable energies, thus circumventing the need for CP correction. While this is a positive finding, the data in that study indicate that while SCS-MP2 is undoubtedly an improvement, there are still many outliers within a considerable error range, particularly for the cc-pVTZ basis set.

We present a detailed analysis of the performance of MP2 with respect to benchmark energies of a series of semi-Coulombic IL systems calculated at the CCSD(T)/CBS level of theory and new scaling factors for SCS- and SOS-MP2 approaches to spin-component scaled MP2. This includes a discussion of CBS extrapolation in section 3.2, correlation energy recovery in MP2 in section 3.3, the importance of CP correction in section 3.4, and finally a detailed discussion of newly fitted SCS- and SOS-MP2 scaling factors in section 3.5. With an exclusive focus on refining the SCS- and SOS-MP2 methods, DFT-based methods such as DFT-D3 are beyond the scope of this paper.

2. THEORETICAL PROCEDURES

A series of 186 energetically favorable IL conformations were optimized using the GAUSSIAN 09 suite of programs[28] including $[C_{(1-4)}mim]X$ (1-alkyl-3-methylimidazolium based ILs) and $[C_{(1-4)}mpyr]X$ (N-methyl-N-alkylpyrrolidinium based ILs) where $X = Br^-$, Cl^-, $[BF_4]^-$, $[PF_6]^-$, $[DCA]^-$ (dicyanamide), $[tos]^-$ (tosylate), $[mes]^-$ (mesylate), and $[NTf_2]^-$ (bis(trifluoromethylsulfonyl)amide), the structures of which are shown in Figure 1. This comprehensive set of IL structures canvasses the predominant structural motifs involved extensively in, but not limited to, synthetic catalysis and electrochemical applications.[29,30] All conformations are available for download as part of the ESI and have been published in part by our group previously;[3] a sample of these conformations is shown in Figures 2 and 3 for $[C_2mim]Br/Cl$, $[C_2mpyr]Br/Cl$, $[C_2mim][NTf_2]$, and $[C_2mpyr][NTf_2]$, which all possess an ethyl group. In these examples, a number of key interactions are highlighted; $[C_2mim]Cl/Br$ P1−P4 shows the anion interacting above (P1), below (P4) and in the plane of the imidazolium ring (P2 and P3) where the anion in P2 interacts more with the methyl group and in P3, it interacts more with the alkyl group. In the case of the pyrrolidinium cations, there are typically three key areas of interaction, where the anion interacts above the pyrrolidinium ring (P1) and on either side of the plane
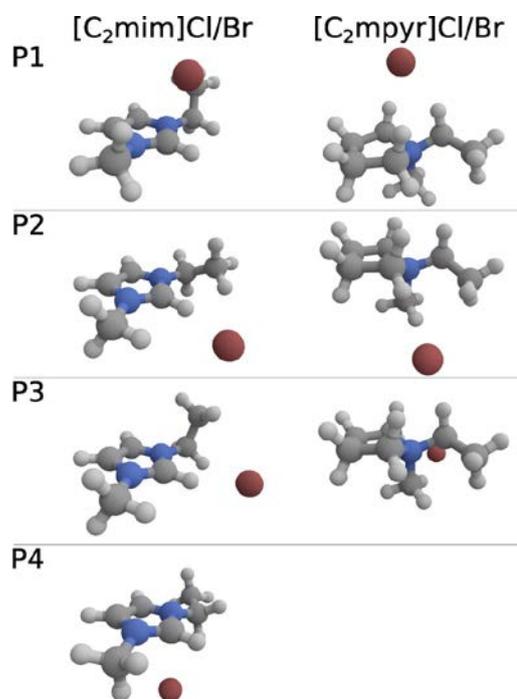
Figure 2. $[C_2mim]Br/Cl$ and $[C_2mpyr]Br/Cl$ structures, anion positions labeled as P1−P4.

described by the methyl group, nitrogen atom, and the alkyl group (P2 and P3). More complex interactions can be seen in Figure 3 for the $[NTf_2]^-$ anion, as in these cases the $[NTf_2]^-$ anion can interact either via the imide or sulfonyl groups. Similar to the structures in Figure 2, $[C_2mim][NTf_2]$ has been constructed with interactions above (P1), below (P4) and in the plane (P2) of the imidazolium ring. An additional structure (P3) shows interaction primarily via the sulfonyl group, which adopts a geometry that is between an exclusively above- and in-plane configuration. The $[C_2mpyr][NTf_2]$ set shows interactions in each of the three areas given for the Br^-/Cl^- examples, however, also include the multiple modes of interaction; P1, P3, and P4 interact primarily via the imide group, whereas P2, P5, and P6 interact primarily via the sulfonyl groups.

The selection of ion pair conformations involved systematic rotation of single bonds within each isolated ion such that the
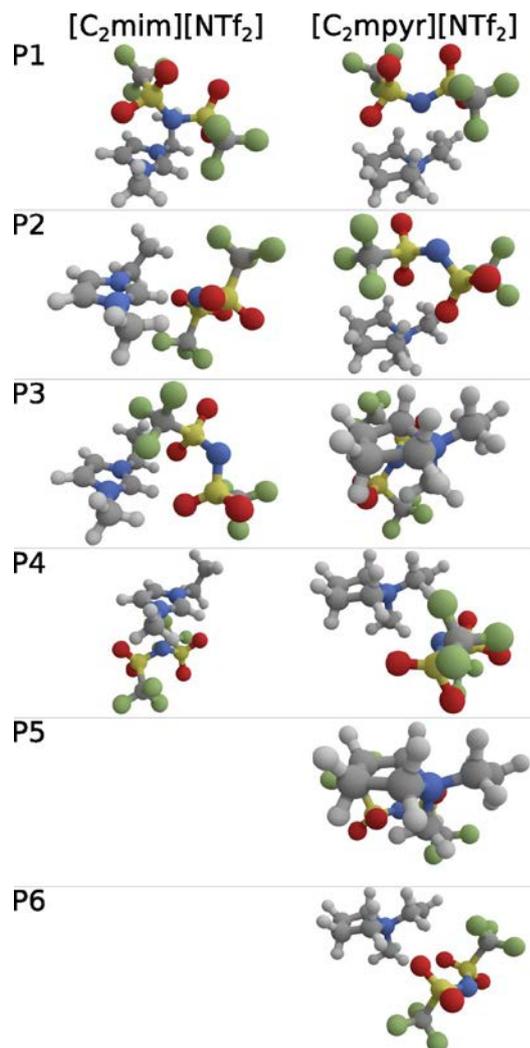
Figure 3. $[C_2mim][NTf_2]$ and $[C_2mpyr][NTf_2]$ structures, anion positions labeled as P1−P6.

global minima of each were identified. Then, ion pairs were formed by pairing the minimum energy anion and cation geometries in three to six conformations designed to reflect the diversity of the expected configurations in the condensed phase. In the case of the $[C_nmim]^+$ series, the MP2/6-31+G(d,p) level of theory was used for optimization due to the importance of dispersion forces, except for the halides which were optimized using MP2/aug-cc-pVDZ due to orbital linear dependence issues. The $[C_nmpyr]^+$ series was optimized at the B3LYP/6-31+G(d) level of theory. It should be noted that while the diversity of structures remains an important consideration, the level of theory used for optimization should have minimal impact on the subsequent single-point energy calculations. As the same structures are used for both the MP2 and benchmark energy calculations described below, our methodology remains valid despite some disparity between structure optimization methods.

Single-point energy calculations at the MP2 level of theory were performed on the entire set of 186 ILs; however, CCSD(T) calculations were performed on a smaller subset of 174 ILs, as the excluded systems were too time-consuming to calculate. The reduced set was used in all analyses except in section 3.2, since CCSD(T)/CBS benchmark energies were not required for analysis here. Excluded ILs include the $[C_3mim]$-$[NTf_2]$, $[C_4mim][NTf_2]$, and $[C_4mim][tos]$ series. All single-point energy calculations were performed using the Psi4 software package[31] with frozen-core and density-fitting approximations. In addition, the frozen natural orbital approximation for CCSD(T) was used to truncate the virtual orbital space, thus providing further acceleration with little impact on accuracy.[32] MP2 energies were calculated using 17 popular basis sets, their abbreviations given in parentheses: Ahlrichs' def2-SV(P) (KSV(P)), def2-SVP (KSVP), def2-TZVP (KTZVP), def2-TZVPP (KTZVPP) basis sets;[33] Dunning's correlation consistent cc-pVDZ (CCD), aug-cc-pVDZ (ACCD), cc-pVTZ (CCT), aug-cc-pVTZ (ACCT), cc-pVQZ (CCQ), aug-cc-pVQZ (ACCQ) basis sets;[34−36] and Dunning's correlation consistent basis sets with tight d-functions added, namely, cc-pV(D+d)Z (CC(D+d)), aug-cc-pV(D+d)Z (ACC(D+d)), cc-pV(T+d)Z (CC(T+d)), and aug-cc-pV(T+d)Z (ACC(T+d));[37] last, Truhlar's May and June "calendar" basis sets jun-cc-pVDZ (JCCD), jun-cc-pVTZ (JCCT), and may-cc-pVTZ (MCCT).[38] The calendar basis sets have varying levels of reduced augmentation when compared with the fully augmented Dunning's basis sets, prefixed with "aug". MP2 energies were also calculated at the complete basis set (CBS) limit by two-point extrapolation performed with cc-pVDZ and cc-pVTZ basis sets (D → T extrapolation), as well as cc-pVTZ and cc-pVQZ basis sets (T → Q extrapolation), including their augmented counterparts. A limited subset of ILs were also calculated with aug-cc-pVQZ and aug-cc-pV5Z basis sets for a Q → 5 extrapolation, including the $[C_1mim]X$ and $[C_1mpyr]Y$ where X = Br⁻, Cl⁻, $[BF_4]^-$, $[PF_6]^-$, $[DCA]^-$, $[tos]^-$, $[mes]^-$, and $[NTf_2]^-$, Y = Cl⁻, $[BF_4]^-$, $[DCA]^-$. All extrapolations followed a Helgaker et al.[39] scheme where the correlation energy converges as X^{-3}, as shown in eq 5,

$$\Delta E_{MP2/CBS} = \frac{X^3 \Delta E_{MP2/X} - Y^3 \Delta E_{MP2/Y}}{X^3 - Y^3} \tag{5}$$

where X and Y are the cardinal numbers of the respective basis sets used for extrapolation (X = 3, Y = 4), and MP2/X and MP2/Y are MP2 energies at basis sets X and Y, respectively.

Correlation energies at the CCSD(T)/CBS level of theory were calculated using the method by Jurecka et al. as given in eq 6,[16,40]

$$\Delta E_{CCSD(T)/CBS}^{corr} = \Delta E_{MP2/CBS}^{corr} + \underbrace{\left(\Delta E_{CCSD(T)/ACCD}^{corr} - \Delta E_{MP2/ACCD}^{corr} \right)}_{CCSD(T)\ correction} \tag{6}$$

where the difference between MP2 and CCSD(T) is considered converged with small basis sets—aug-cc-pVDZ in this case—and can therefore be added to the MP2/CBS energies to form a reasonable approximation for CCSD(T)/CBS. All components of the CCSD(T)/CBS benchmark were corrected for BSSE using the Boys and Bernardi counterpoise correction scheme[23] as described by eqs 2−4.

All errors presented in this paper, unless otherwise specified, are with respect to the CCSD(T)/CBS benchmark. Where box and whisker plots are used, the horizontal line bisecting the boxes indicate the median values, the upper and lower boundaries of the box indicate the first and third quartiles, and the whiskers indicate data extrema. Where appropriate, data have been analyzed using standard statistical measures, including mean error, mean absolute error, maximum absolute

error, and standard deviation. Mean error and mean absolute error were calculated according to eqs 7 and 8, respectively, where N is the sample size, x_i is the ith benchmark ion pair interaction energy (e.g., CCSD(T)/CBS energy), $\hat{x}_i$ is the ith estimated ion pair interation energy (e.g., SCS-MP2 energy).

$$ME = \frac{1}{N}\sum_{i=1}^{N}(x_i - \hat{x}_i) \tag{7}$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N}(|x_i - \hat{x}_i|) \tag{8}$$

Errors mentioned in-text are given as the mean error with uncertainty shown as one standard deviation either side of the mean, unless otherwise specified. Confidence intervals (CIs) for the regression coefficients have been derived using the bootstrapping technique combined with a bias corrected and accelerated approach[41] over 10 000 replicates.

3. RESULTS AND DISCUSSION

3.1. Significance of Dispersion Energy. As previously mentioned, difficulties in accurately calculating the energetic properties of ionic liquids arise primarily due to the relative significance of dispersion interactions (accounted for mainly through electron correlation) compared with the electrostatic component (included mainly in the Hartree−Fock energy). In the set of ILs ion pairs examined in this study, correlation energy at the CCSD(T)/CBS limit corresponds to a contribution on average of approximately 7−20% to the overall interaction energy, or about 27−70 kJ mol^{-1}, which is in accordance with an earlier study in our group.[3] Additionally, we have shown contributions upward of 20% in large clusters[12] and has been shown to exceed as much as 60% in a molecular dynamics study by Shimizu et al.[42] The accuracy of the dispersion component is, therefore, critical to understanding the overall energetic behavior of ionic liquids owing to both its magnitude and variability between structurally distinct systems.

3.2. MP2 Complete Basis Set Extrapolation and the Choice of Basis Sets. The behavior of the MP2/CBS extrapolation with respect to the choice of basis set was analyzed in terms of the effects of basis set augmentation (i.e., addition of diffuse functions) and in terms of the two points selected for the extrapolation (namely, the choice of double-ζ → triple-ζ (D → T) versus a triple-ζ → quadruple-ζ (T → Q) extrapolation). In the case of the former, augmentation refers to Dunning's basis sets with the "aug" prefix; specifically, aug-cc-pVDZ, aug-cc-pVTZ and aug-cc-pVQZ.

First, an analysis of the role of basis set augmentation has shown that for T → Q extrapolation, basis set augmentation does not have a significant impact on the extrapolated energies with an mean absolute difference of only 0.66 kJ mol^{-1} and a maximum absolute difference of only 1.05 kJ mol^{-1} between the extrapolation performed with and without augmented basis functions. Interestingly, while the BSSE is reduced when using augmented basis sets in general, only ILs containing the bromide anion show a marked increase in BSSE by 135% on average, which is contrary to expectations owing to the increased "completeness" of the augmented basis sets. For all other ion pairs, BSSE was reduced by an average of 34%. With this in mind, the use of nonaugmented basis sets are recommended for T → Q extrapolation because (a) the energies after CP correction differ negligibly (−0.66 ± 0.35 kJ

mol^{-1}), (b) the exclusion of augmentation reduces computation time, and (c) the likelihood for orbital linear dependence is reduced. It is also noted that for the subset of ILs calculated using Q → 5, the mean absolute error from the T → Q energies was found to be only 0.06 kJ mol^{-1}, with a maximum of 0.22 kJ mol^{-1}. While this is only a small proportion of the total ILs studied, with the exception of bromide, all cations were included. Therefore, it is reasonable to conclude that the CBS limit is sufficiently reached with a T → Q two-point extrapolation.

Second, in the case of D → T extrapolation, basis set augmentation plays a more important role. Taking the augmented T → Q energies as the benchmark, a mean and maximum absolute error of 0.43 kJ mol^{-1} and 1.49 kJ mol^{-1} were found for the augmented D → T, respectively. When augmentation is removed, errors increase with a mean and maximum absolute error of 1.26 kJ mol^{-1} and 3.02 kJ mol^{-1}, respectively. These results indicate that an augmented D → T extrapolation is sufficient in reaching the CBS limit, and without augmentation, results are still "chemically accurate" after CCSD(T) corrections (shown in eq 6) are included. For the purposes of this study, however, D → T extrapolations are not suitable for the basis of CCSD(T)/CBS benchmark energies as the unsigned CCSD(T) correction on top of MP2 is approximately 2.15 ± 1.92 kJ mol^{-1} (7.75 kJ mol^{-1} maximum), which is on the order of the MP2/CBS extrapolation errors themselves.

3.3. Recovery of MP2 Correlation Energies. The ability for CP corrected MP2 to recover electron correlation energy as a function of basis set has been assessed for the ionic liquids test set used in this study, with CCSD(T)/CBS energies used as the benchmark. Figure 4 shows that, at the CBS limit, MP2
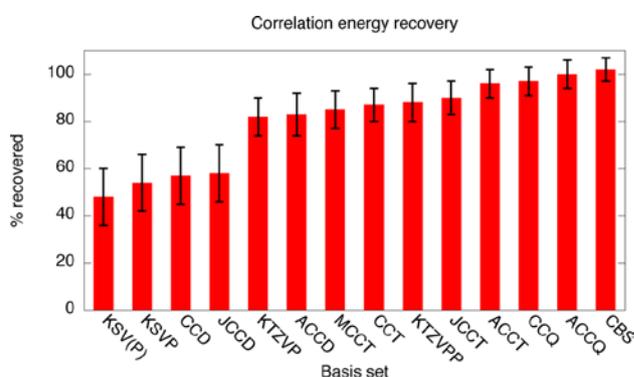
Figure 4. Average CP corrected MP2 correlation energy recovery with respect to CCSD(T)/CBS; error bars shown as one standard deviation.

overestimates correlation by 2% on average, representing a marginal increase from the aug-cc-pVQZ basis set, which recovered on average 100% of the correlation energy. In the extreme case, ILs that included the [DCA]$^-$ anion consistently showed disproportionate recovery of correlation energy; [C$_3$mim][DCA] was overestimated by as much as 13%, and still by 8% with the aug-cc-pVTZ basis set where the average recovery overall was 96%.

It has been shown that MP2, at its CBS limit, tends to overestimate the correlation energy, consistent with previous literature involving dispersion-driven interactions.[16,43,44] Difficulties arise due the unsystematic treatment of different ion-pair

interactions; notably, those involving the [DCA]$^-$ anion. To this end, CCSD(T) benchmark energies are essential when providing a reference energy to which the spin components can be fitted.

3.4. Importance of Counterpoise Correction. A complicating factor in the accurate determination of dispersion energy is the basis set superposition error (BSSE), which is an artifact of a truncated basis set that manifests as an overstabilization and converges to zero as the basis set is increased. Minimisation of BSSE effects is typically achieved in two ways: (1) a sufficiently large basis set (generally triple-ζ quality or greater) is used, and (2) counterpoise correction is performed.[23] The counterpoise approach to BSSE correction calculates the monomers in the basis set of the complex, which is believed to cancel out the BSSE energy. The magnitude of BSSE is therefore the difference between the corrected and uncorrected interaction energies. Figure 5 shows the BSSE
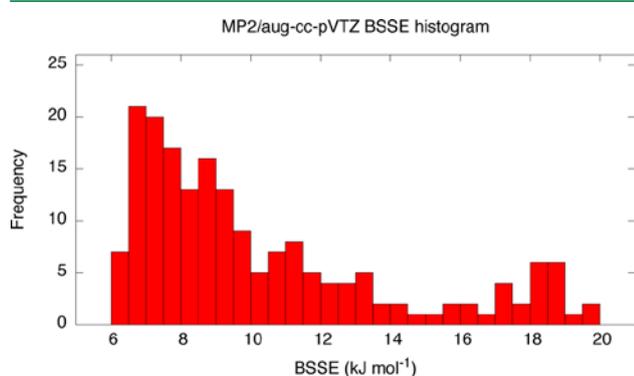
Figure 5. Histogram of BSSE for MP2/aug-cc-pVTZ interaction energies.

distribution for the IL test set using the MP2/aug-cc-pVTZ level of theory, routinely used in studies of energetic properties of ILs.[45−49] With this level of theory, 63% of the test set show

BSSE greater than 10 kJ mol^{-1} and would consequently be of little value should counterpoise correction not be performed.

It should be noted that with the exception of the above, this paper does not address BSSE exclusively; errors discussed herein are considered with respect to the CCSD(T)/CBS benchmark and therefore as an aggregate quantity composed of both insufficient correlation recovery (see section 3.3) and BSSE.

The effect of CP correction on reproducing CCSD(T)/CBS energies was investigated, with Figure 6 indicating that in order to maintain chemical accuracy, CP correction remains important even with the aug-cc-pVQZ basis set. The unscaled MP2 errors are notably unsystematic when compared to the CP corrected MP2 errors, suggesting that there is an unpredictable interplay between errors arising from basis set incompleteness (i.e., basis set superposition error), and the inherent limitation of MP2 to accurately capture correlation energy; the two opposing effects may fortuitously cancel in some instances. The result yields somewhat surprising errors where, for example, the non-CP corrected cc-pVDZ basis set gives errors of 1.01 ± 4.52 kJ mol^{-1} whereas the much more complete aug-cc-pVQZ basis set gives errors of 5.94 ± 4.36 kJ mol^{-1}. These errors become systematic and follow expected trends after CP correction with errors of -17.00 ± 4.47 kJ mol^{-1} and 0.19 ± 2.50 kJ mol^{-1} for cc-pVDZ and aug-cc-pVQZ, respectively.

Interestingly, CP correction of SCS-MP2 and SOS-MP2 appears to have a detrimental effect. In considering three sufficiently large basis sets (def2-TZVPP, aug-cc-pVTZ, and aug-cc-pVQZ), Figure 7 shows that, regardless of the basis set chosen, CP correction has a negative impact on accuracy of SCS-MP2 with mean absolute errors (MAEs) increasing significantly from the CP corrected standard MP2, even beyond that of the non-CP corrected standard MP2 energies. It can be seen that the MAEs of SCS-MP2 non-CP corrected energies for the def2-TZVPP (4.07 kJ mol^{-1}) and aug-cc-pVTZ (2.52 kJ mol^{-1}) basis sets are qualitatively indistinguishable from their CP corrected standard MP2 counterparts at 4.66 kJ mol^{-1} and 2.45 kJ mol^{-1}, respectively. The SCS-MP2 non-CP
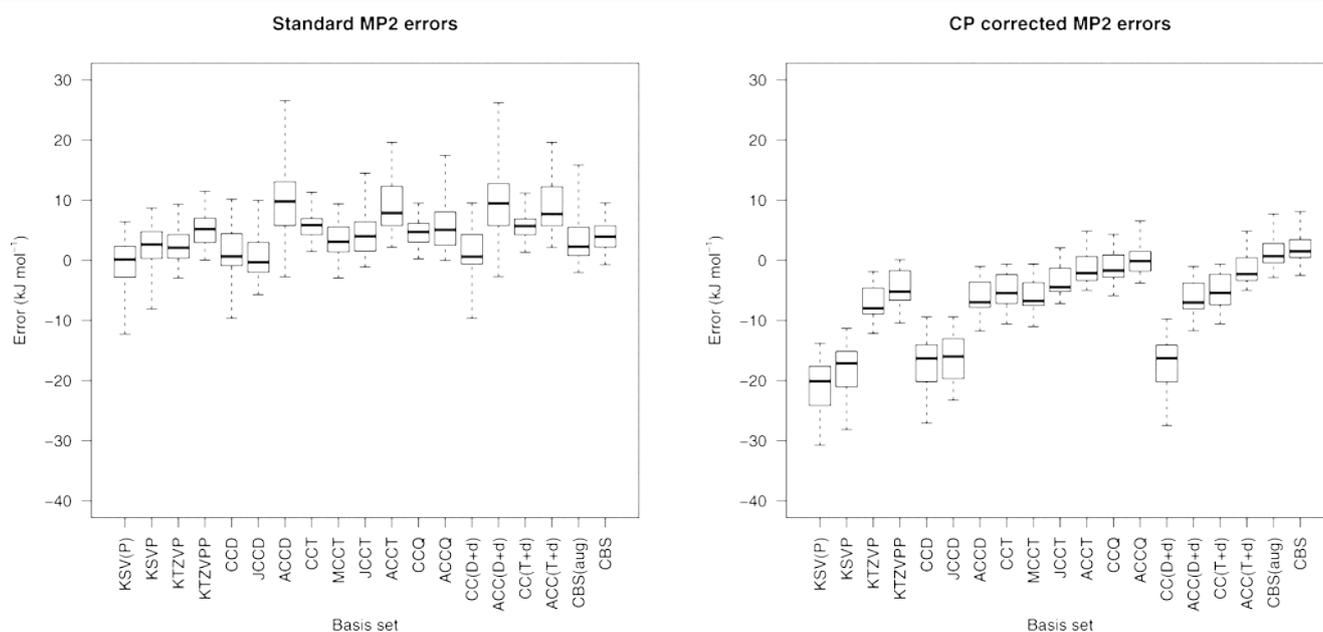
Figure 6. Comparison of errors between standard (i.e., no CP correction) (left) and counterpoise corrected (right) MP2 energies.
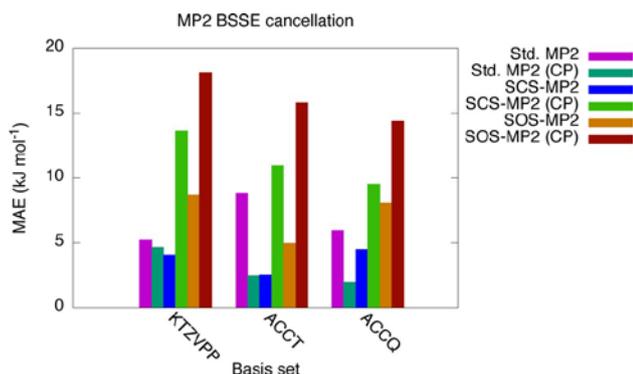
Figure 7. Comparison of errors between SCS-MP2 and standard MP2 with and without CP correction.

corrected energies using the aug-cc-pVQZ basis set do not perform as well as the triple-ζ basis sets with only a marginal improvement yielding a MAE of 4.50 kJ mol^{-1}, compared to the non-CP corrected standard MP2 error of 5.97 kJ mol^{-1}. A detailed analysis of the performance of SCS- and SOS-MP2 both with and without CP correction is shown in Figure 8, where it can be seen that when CP corrected, errors tend toward an approximate 10 kJ mol^{-1} underestimation when compared with the CCSD(T) benchmark.

These results show that, when using conventional MP2, CP correction is essential in order to achieve energies that closely reproduce CCSD(T)/CBS. While smaller basis sets may fortuitously cancel out inadequacies of MP2 with basis set incompleteness, it would not be possible to rely on these assumptions for other systems. Conversely, CP correction does not improve the errors for SCS- and SOS-MP2 methods, and it is best to use these uncorrected, preferably with a triple-ζ basis set. The convenient observation that SCS-MP2 performs best with a triple-ζ basis set (and not higher) indicates a significant basis set dependence where there currently exists no well-established convention or recommendation. The remainder of this paper is concerned with the refitting of the scaling coefficients, with particular focus on the ideal basis set combination.

3.5. Fitted Spin-Component Scaled and Spin-Opposite Scaled MP2. In this study, we present updated scaling factors for the same- and opposite-spin components for semi-Coulombic systems that take into account basis set effects, fitted to 174 ionic liquids calculated at the CCSD(T)/CBS level of theory. Coefficients for 17 popular basis sets were determined using both a standard multiple least-squares regression described by eq 9,

$$\Delta E_{\text{CCSD(T)/CBS}}^{\text{corr}} \approx \alpha(\Delta E_{\text{OS}}) + \beta(\Delta E_{\text{SS}}) \qquad (9)$$

where $\Delta E_{\text{CCSD(T)/CBS}}^{\text{corr}}$ is the CCSD(T)/CBS correlation energy and ΔE_{OS} and ΔE_{SS} represent the opposite- and same-spin contributions to the MP2 interaction correlation energy, respectively. All multiple linear regression calculations were performed using the R statistical package.[50]

The CCSD(T)/CBS correlation energies have been fitted to spin components that have not been counterpoise corrected. The rationale for this is 2-fold: first, as previously mentioned, there has already been observed an error cancellation effect with SCS-MP2 and the triple-ζ basis sets that we wish to exploit further. Second, we present scaling factors for very extensive basis sets including at the CBS limit that have BSSE

approaching zero. As the primary motivation is for computationally efficient, yet accurate, large-scale energy calculations of condensed semi-Coulombic systems, it follows that fitting to counterpoise corrected spin components is not advantageous in this regard.

Table 1 shows the coefficients determined for the SCS- and SOS-MP2 methods, respectively. In addition to coefficients determined for all basis sets, two sets of coefficients are calculated for the CBS limit; "CBS(aug)" refers to the aug-cc-pVTZ → aug-cc-pVQZ extrapolation, while "CBS" refers to the cc-pVTZ → cc-pVQZ extrapolation. The CBS limit will be referred to explicitly as one or the other in this section.

The scaling factors presented here show instances where the α coefficient is negative, and this has been argued a nonphysical quantity[15,17] as by definition both spin components themselves must be negative. Work by Hill and Platts,[17] as well as Distasio and Head-Gordon,[15] show that negative coefficients can arise from their respective means of data fitting. Distasio and Head-Gordon argue that this can be an artifact of both a biased training set and basis set deficiencies; their study shows that when fitting to a large basis set (>double-ζ quality), invalid coefficients are not present. Our data shows negative scaling factors even for basis sets such as aug-cc-pVQZ and CBS(aug), which are not only extensive, but at the limits of practical usage. This suggests that there exists no universal set of coefficients that statistically best scale MP2 while eliminating nonphysical quantities and also highlights the importance of considering MP2 in conjunction with its basis set.

It should be noted that an alternative position regarding the regression coefficients might be taken whereby the physical rationality is disregarded in preference to a purely statistical interpretation. In this case, it can be argued that by virtue of the multiple linear regression procedure, regardless of sign, the coefficients are statistically more likely to produce an improved energy. This has the advantage of exposing the reader to a wider range of valid coefficients and consequently basis sets of lower computational complexity. The authors are of the view that physically sensible coefficients are preferable, but we present all data here to stimulate further discussion on this topic.

With the success of SCS-MP2, it seems invariably the case that applying some set of coefficients improves the MP2 energies by some degree. However, with the rise of alternative schemes such as SCSN-MP2 and SCS(MI)-MP2, a one-size-fits-all solution is evidently insufficient. Therefore, the consideration of which chemical systems are similar enough to warrant their own coefficients is important; those that do not fit the model will naturally form outliers in the linear regression and may ultimately decrease the quality of the fit for the remaining data set. Further, our data shows that the choice of

basis set may significantly affect the behavior of the spin components and may do so even with only marginal effects on the overall interaction energy. We discuss one such instance of this; however, detailed diagnostic plots of all regressions performed are available in the Supporting Information.

Figure 9 shows a Q−Q plot (quantile−quantile plot) that assumes a normal residual error distribution; data lying on the diagonal indicates that this assumption holds. It can be seen that while the cc-pVTZ regression behaves predictably with very few outlying data points, the introduction of augmented basis functions causes quite a distinct deviation among the bromide series. Chemically speaking, this is unsurprising as bromide is a large anion and the introduction of diffuse

Article
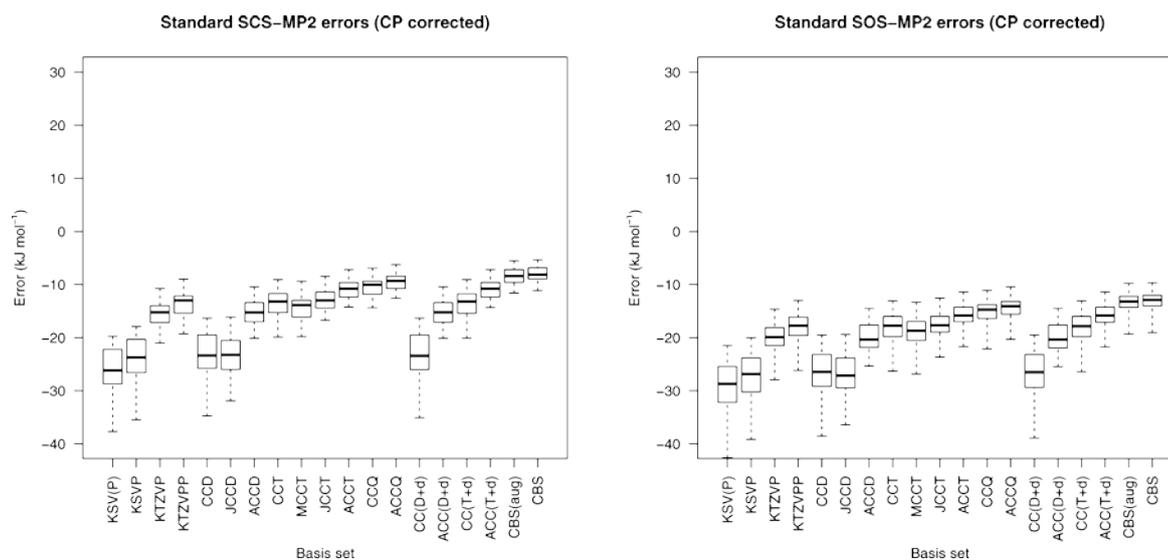
(a) Counterpoise corrected energies
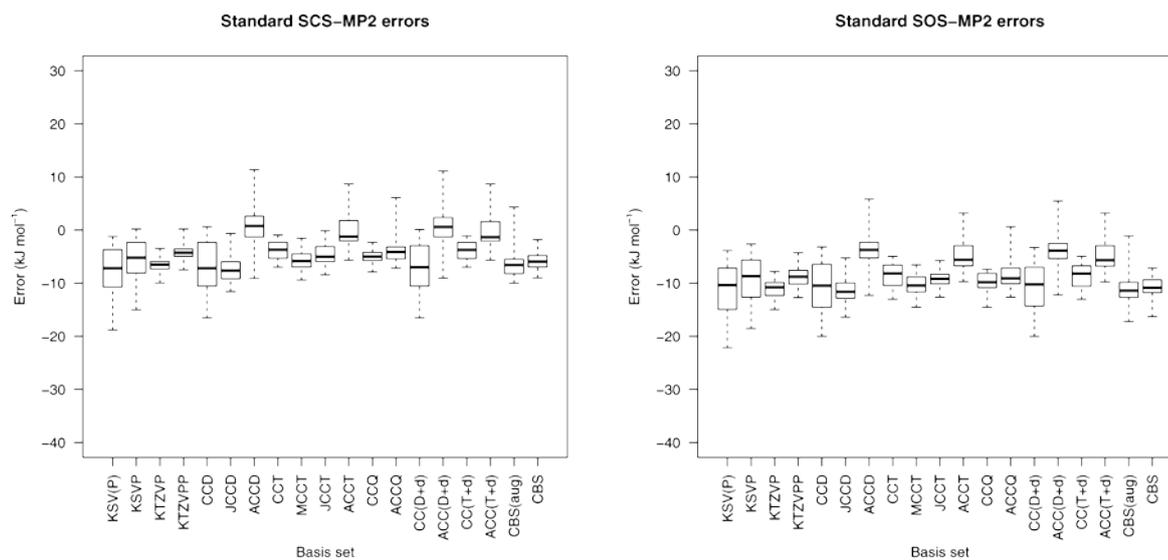
(b) Non-counterpoise corrected energies

Figure 8. Errors of CP corrected (a) and uncorrected (b) standard spin-component scaled MP2 energies.

functions is likely to allow interactions that would otherwise be insignificant in the remaining systems, thereby causing the error distribution to deviate from normality. While the specific error distributions between the various basis sets differ, as the size of the basis set increases, it generally observed that the error distributions become increasingly specific to the anion combinations. In other words, one could take steps to scale the spin components of each subset of IL; however, this would bring the utility of such coefficients into question.

In order to objectively identify the ideal combination of basis set and scaling factors, we make use of the bootstrapping technique[41] combined with a bias corrected and accelerated approach[41] for estimating the errors associated with the scaling factors. Briefly, bootstrapping is a resampling technique for estimating the errors associated with a sample population. Bootstrapping involves repeated, random resampling allowing replacement. The parameters—in this case, α and β coefficients—are estimated for each random sample population, and overall errors are derived. Bootstrapping is advanta-

Article

Table 1. Fitted SCS-MP2 and SOS-MP2 Coefficients

| basis set | SCS-MP2 | | | | | | SOS-MP2 | | | |
	α	95% CI	β	95% CI	mean error (kJ mol^{-1})	std. dev. (kJ mol^{-1})	α	95% CI	mean error (kJ mol^{-1})	Std. Dev. (kJ mol^{-1})
KSV(P)	0.35	[−0.05, 0.75]	1.88	[1.34, 2.41]	−0.48	4.40	1.76	[1.73, 1.78]	−0.47	5.01
KSVP	0.83	[0.55, 1.08]	1.10	[0.76, 1.46]	−0.22	3.72	1.67	[1.64, 1.69]	−0.22	4.08
KTZVP	0.61	[0.37, 0.86]	1.30	[1.04, 1.57]	−0.32	1.97	1.76	[1.74, 1.78]	−0.47	2.48
KTZVPP	0.98	[0.76, 1.21]	0.77	[0.51, 1.02]	−0.15	1.71	1.64	[1.63, 1.66]	−0.19	1.90
CCD	0.69	[0.41, 0.96]	1.32	[0.99, 1.66]	−0.34	4.38	1.73	[1.70, 1.76]	−0.37	4.85
JCCD	0.43	[0.14, 0.74]	1.63	[1.27, 1.97]	−0.58	3.14	1.76	[1.73, 1.78]	−0.88	3.94
ACCD	−0.29	[−0.47, −0.07]	2.23	[1.96, 2.47]	−0.32	2.45	1.40	[1.38, 1.42]	−0.72	3.98
CCT	1.05	[0.87, 1.24]	0.68	[0.47, 0.89]	−0.03	1.55	1.64	[1.63, 1.65]	−0.05	1.71
MCCT	0.65	[0.42, 0.91]	1.23	[0.94, 1.48]	−0.32	2.26	1.72	[1.70, 1.74]	−0.46	2.63
JCCT	0.17	[0.00, 0.39]	1.74	[1.50, 1.93]	−0.33	1.98	1.65	[1.63, 1.67]	−0.67	2.98
ACCT	0.07	[−0.17, 0.28]	1.77	[1.51, 2.06]	−0.32	2.60	1.45	[1.43, 1.47]	−0.42	3.41
CCQ	0.76	[0.59, 0.94]	1.04	[0.85, 1.23]	−0.15	1.37	1.70	[1.69, 1.71]	−0.23	1.72
ACCQ	−0.54	[−0.83, −0.24]	2.51	[2.18, 2.84]	−0.31	2.68	1.60	[1.57, 1.62]	−0.61	3.96
CC(D+d)	0.71	[0.44, 0.96]	1.31	[1.00, 1.63]	−0.29	4.16	1.74	[1.72, 1.77]	−0.32	4.66
ACC(D+d)	−0.25	[−0.43, −0.03]	2.19	[1.90, 2.41]	−0.31	2.36	1.41	[1.39, 1.43]	−0.71	3.90
CC(T+d)	1.02	[0.84, 1.19]	0.72	[0.53, 0.93]	−0.03	1.51	1.64	[1.63, 1.65]	−0.06	1.70
ACC(T+d)	0.07	[−0.16, 0.28]	1.78	[1.51, 2.05]	−0.32	2.60	1.45	[1.43, 1.47]	−0.43	3.42
CBS(aug)	−1.00	[−1.23, −0.76]	2.99	[2.73, 3.23]	−0.24	2.52	1.72	[1.68, 1.75]	−0.88	4.87
CBS	0.00	[−0.15, 0.15]	1.88	[1.72, 2.04]	−0.21	1.78	1.74	[1.72, 1.75]	−0.48	2.70
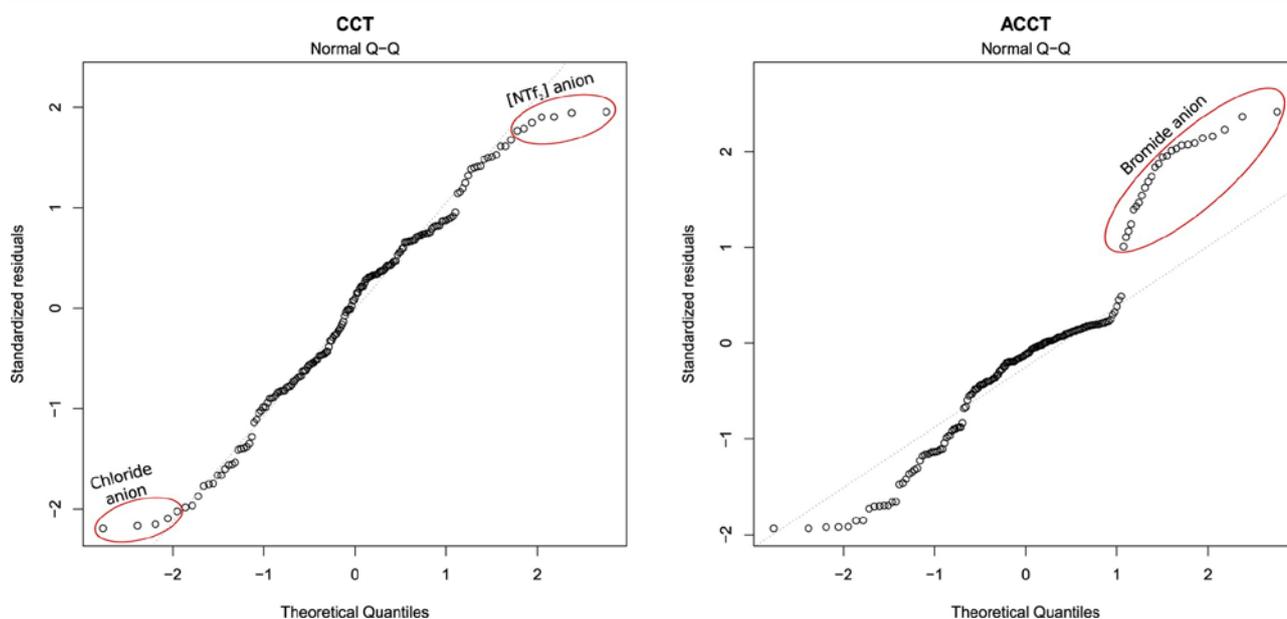
Figure 9. Q–Q plots of SCS-IL-MP2 multiple linear regression for cc-pVTZ (left) and aug-cc-pVTZ (right) basis sets.

geous because it makes no assumptions of the underlying error distribution. In this instance, the coefficients were found over 10 000 replicates with the 95th percentile confidence interval (CI) presented herein.

Figures 10 and 11 show the range of the 95% CIs shown in Table 1 for the SCS- and SOS-IL-MP2 scaling factors. The more narrow the interval, the more certainty there is for the coefficients to be generally applicable. Specifically, if the training data set for any particular basis set contains disproportionately influential points, random variation of their quantity in the bootstrapping iterations will broaden the confidence interval. Conversely, should no influential data points be present, the coefficients will be invariant with respect to the randomly resampled data.

As indicated by the range of the confidence intervals, SCS- and SOS-MP2 coefficients are best fitted to MP2 energies calculated with cc-pVTZ, cc-pV(T+d)Z, and cc-pVQZ basis sets, as well as at the CBS limit when extrapolated without augmentation. For SCS-MP2 (SOS-MP2), the overall errors produced for each of these basis sets are qualitatively indistinguishable at −0.03 ± 1.55 kJ mol^{-1} (−0.05 ± 1.71 kJ mol^{-1}), −0.03 ± 1.51 kJ mol^{-1} (−0.06 ± 1.70 kJ mol^{-1}), −0.15 ± 1.37 kJ mol^{-1} (−0.23 ± 1.72 kJ mol^{-1}), and −0.21 ± 1.78 kJ mol^{-1} (−0.48 ± 2.70 kJ mol^{-1}), respectively. The cc-pV(T+d)Z basis set differs from cc-pVTZ only in the addition of tight d-functions on Al−Ar atoms[37] and is shown here to provide essentially no change to the fitting errors for both SCS- and SOS-MP2. The use of cc-pV(T+d)Z is difficult to justify
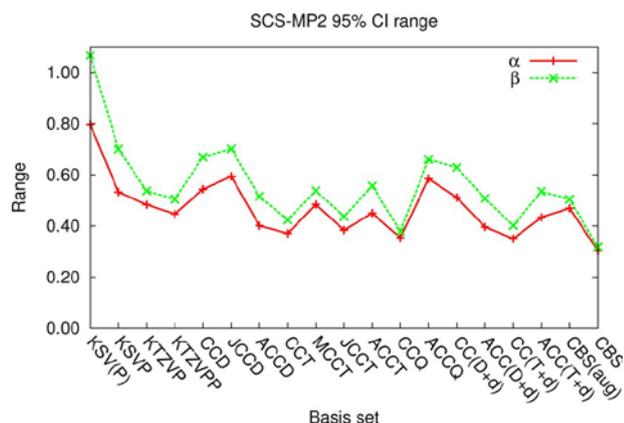
Figure 10. 95% confidence interval range for the opposite- (α) and same-spin (β) component coefficients in the SCS-IL-MP2 scheme.
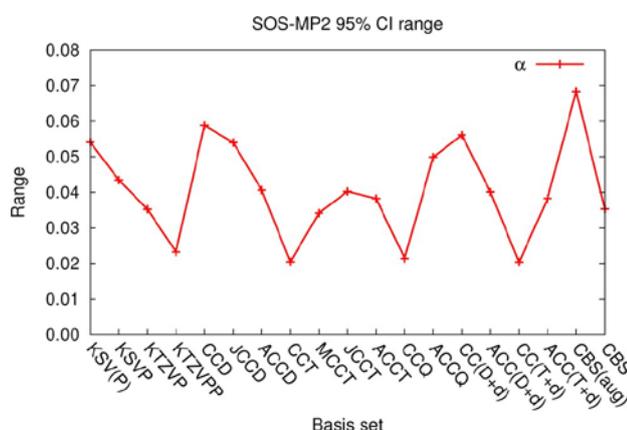
Figure 11. 95% confidence interval range for the opposite-spin (α) component coefficient in the SOS-IL-MP2 scheme.

given the small change in comparison to cc-pVTZ, which provides remarkable accuracy after scaling and is well within the practical limitations of current computer hardware.

The performance of scaled MP2/cc-pVTZ energies are shown in Figure 12, termed "SCS-IL-MP2" and "SOS-IL-MP2", in comparison to standard MP2 (with and without CP correction), standard SCS- and SOS-MP2, as well as the SCSN and SCS(MI) variants. While the range of errors are comparable, it is important to note that the error distribution produced by SCS-IL-MP2 and SOS-IL-MP2 are much more evenly distributed both about the median value and about zero. Further, it is worth noting that the errors for both SCS-IL-MP2 and SOS-IL-MP2 differ by subkilojoule amounts and therefore its use would be advantageous where software permits the omission of same-spin contributions and hence accelerated calculation.

We provide the following guidelines for calculating MP2 energies of semi-Coulombic IL systems using the newly fitted coefficients:

(1) Basis sets with augmentation should be avoided; they typically result in negative coefficients (or confidence intervals that extend into negative values) and the errors tend to be nonuniform.

(2) For energy calculations where high accuracy is desired (<1.0 kJ mol^{-1} on average), MP2/cc-pVTZ along with
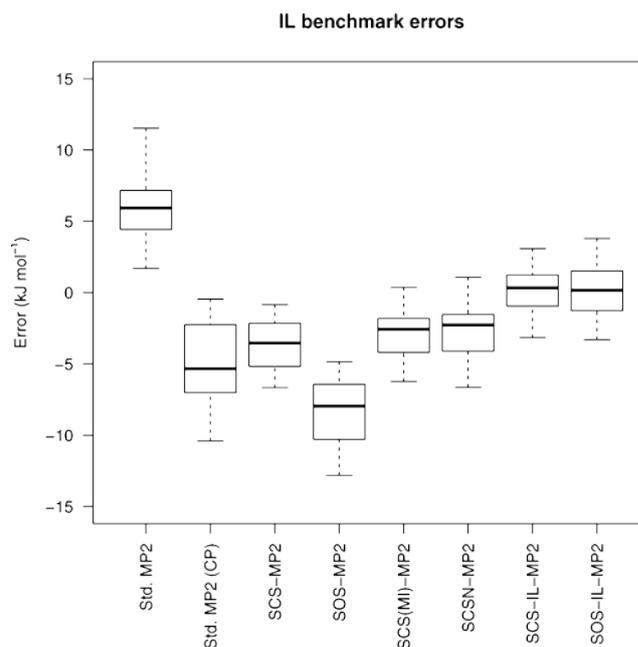
Figure 12. IL benchmark errors at the MP2/cc-pVTZ level of theory.

the appropriate scaling factors (SCS-IL-MP2: α = 1.05, β = 0.68; SOS-IL-MP2: α = 1.64, β = 0.00) should be used.

(3) SOS-IL-MP2 should be used in instances where the calculation is accelerated by omitting the same-spin component.[14]

4. FUTURE WORK

We have presented scaling factors for MP2 that are based not only on a wide variety of IL anion/cation combinations but also a wide variety of energetically favorable conformations. Considering the simple example of a two ion pair cluster of [C$_1$mim]Cl as shown in Figure 13, the two chlorides are
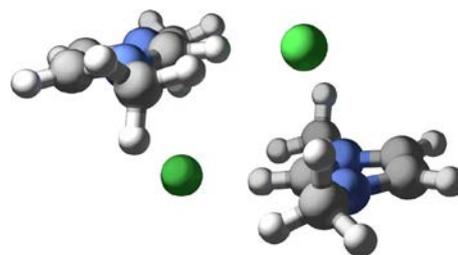
Figure 13. Example of a cluster of two [C$_1$mim]Cl ion pairs.

interacting both in the plane of the imidazolium ring and above/below the ring. These two distinct modes of interaction are included in the fitting data set used in this study, as are many others. Furthermore, both dispersion and BSSE originate from interactions that do not extend significantly beyond neighboring ions; indeed, dispersion interactions are already accounted for sufficiently as a sum of pairwise (two-body) interactions at the MP2 level of theory.[12] It is therefore hypothesized that by including a variety of conformations of ion pairs, the coefficients will be better suited to large scale calculations in which a variety of interaction modes may exist and traditional CP correction is infeasible. In particular, we believe that methods such as the FMO approach that rely on

Article

the explicit inclusion of pairwise interactions to describe energetics of large-scale ionic systems would greatly benefit from increased accuracy of the MP2 method at lower computational cost. The applicability of the recommended SCS-IL and SOS-IL-MP2 coefficients previously discussed to large IL clusters will be addressed in a follow-up paper.

Further, this paper has dealt exclusively with IL conformations in the equilibrium geometry and the question of how this method performs for nonequilibrium structures remains unanswered. Thus, the validity of our scaling factors for ILs over a portion of their potential energy surfaces will be assessed in a follow-up paper reflecting a methodology similar to the analysis of the S66 × 8 data set by Řezáč et al.[51]

5. CONCLUSIONS

Spin-component scaled MP2 methods have historically shown to be advantageous in reproducing CCSD(T)/CBS energies, with a number of different derivations presented in the literature. In this paper, we have presented scaling factors fitted to a series of 174 semi-Coulombic ionic liquid systems in a variety of conformations for 17 popular basis sets, with MP2/cc-pVTZ having been shown to produce the most reliable energies, deviating on average less than 1.0 kJ mol^{-1} from the CCSD(T)/CBS benchmark. A strong basis set dependence was observed with the best-fitting basis set identified as cc-pVTZ. Other basis sets introduced increased error and in some cases produced unrealistic negative coefficients. Coefficients of $\alpha = 1.05$, $\beta = 0.68$ and $\alpha = 1.64$, $\beta = 0$ in combination with MP2/cc-pVTZ spin-components form our "SCS-IL-MP2" and "SOS-IL-MP2" methods, respectively. Our new coefficients have been fitted to CP corrected CCSD(T)/CBS benchmark energies and consequently result in minimal basis set superposition errors. Additionally, the SOS-IL-MP2 method has errors almost identical to the SCS-IL-MP2 and therefore can be used in preference to the SCS-IL-MP2 method owing to the improved efficiency achieved by neglecting the same-spin component as in the original SOS-MP2 implementation. These coefficients will be applied to investigate energetics of larger clusters of ionic materials in view of circumventing the need for time-consuming CP correction. This new SCS-IL-MP2 method can be considered as an improved second order of Møller−Plesset perturbation theory for reliable studies of energetics of any noncovalently bound complex dominated by the interplay of electrostatics, induction, and dispersion forces.

■ ASSOCIATED CONTENT

*Supporting Information

All molecular structures and associated energies used in this study, as well as additional statistical diagnostic plots. This material is available free of charge via the Internet at http://pubs.acs.org/.

■ AUTHOR INFORMATION

Corresponding Author

*Email: Katya.Izgorodina@monash.edu.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

■ REFERENCES

(1) Plechkova, N. V.; Seddon, K. R. Chem. Soc. Rev. 2008, 37, 123−150.

(2) Rogers, R. D.; Seddon, K. R. Science 2003, 302, 792−793.

(3) Izgorodina, E. I.; Golze, D.; Maganti, R.; Armel, V.; Taige, M.; Schubert, T. J. S.; MacFarlane, D. R. Phys. Chem. Chem. Phys. 2014, 16, 7209−7221.

(4) Burke, K. J. Chem. Phys. 2012, 136, 150901.

(5) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. J. Chem. Phys. 2010, 132, 154104.

(6) Steinmann, S. N.; Piemontesi, C.; Delachat, A.; Corminboeuf, C. J. Chem. Theory Comput. 2012, 8, 1629−1640.

(7) Rigby, J.; Izgorodina, E. I. Phys. Chem. Chem. Phys. 2013, 15, 1632−1646.

(8) Grimme, S.; Hujo, W.; Kirchner, B. Phys. Chem. Chem. Phys. 2012, 14, 4875−4883.

(9) Schütz, M.; Hetzer, G.; Werner, H.-J. J. Chem. Phys. 1999, 111, 5691−5705.

(10) Asadchev, A.; Allada, V.; Felder, J.; Bode, B. M.; Gordon, M. S.; Windus, T. L. J. Chem. Theory Comput. 2010, 6, 696−704.

(11) Kitaura, K.; Ikeo, E.; Asada, T.; Nakano, T.; Uebayasi, M. Chem. Phys. Lett. 1999, 313, 701−706.

(12) Izgorodina, E. I.; Rigby, J.; MacFarlane, D. R. Chem. Commun. 2012, 48, 1493−1495.

(13) Grimme, S. J. Chem. Phys. 2003, 118, 9095−9102.

(14) Jung, Y.; Lochan, R. C.; Dutoi, A. D.; Head-Gordon, M. J. Chem. Phys. 2004, 121, 9793−9802.

(15) Distasio, R. A., Jr.; Head-Gordon, M. Mol. Phys. 2007, 105, 1073−1083.

(16) Jurecka, P.; Sponer, J.; Cerny, J.; Hobza, P. Phys. Chem. Chem. Phys. 2006, 8, 1985−1993.

(17) Hill, J. G.; Platts, J. A. J. Chem. Theory Comput. 2007, 3, 80−85.

(18) Szabados, A. J. Chem. Phys. 2006, 125.

(19) Feenberg, E. Phys. Rev. 1956, 103, 1116−1119.

(20) Fink, R. F. J. Chem. Phys. 2010, 133, 174113.

(21) Grimme, S.; Goerigk, L.; Fink, R. F. Wiley Interdiscip. Rev. Comput. Mol. Sci. 2012, 2, 886−906.

(22) Zahn, S.; MacFarlane, D. R.; Izgorodina, E. I. Phys. Chem. Chem. Phys. 2013, 15, 13664−13675.

(23) Boys, S.; Bernardi, F. Mol. Phys. 1970, 19, 553−566.

(24) Fedorov, D.; Kitaura, K. The Fragment Molecular Orbital Method: Practical Applications to Large Molecular Systems; Taylor & Francis: New York, 2009; Chapter 2, pp 5−36.

(25) Antony, J.; Grimme, S. J. Phys. Chem. A 2007, 111, 4862−4868 PMID: 17506533..

(26) King, R. A. Mol. Phys. 2009, 107, 789−795.

(27) Tongying, P.; Tantirungrotechai, Y. J. Mol. Struct.: THEOCHEM 2010, 945, 85−88.

(28) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.;

Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. Gaussian 09 Revision A.1; Gaussian Inc.: Wallingford, CT, 2009.

(29) Welton, T. Chem. Rev. 1999, 99, 2071−2084.

(30) Forsyth, S. A.; Pringle, J. M.; MacFarlane, D. R. Aus. J. Chem. 2004, 57, 113−119.

(31) Turney, J. M.; Simmonett, A. C.; Parrish, R. M.; Hohenstein, E. G.; Evangelista, F. A.; Fermann, J. T.; Mintz, B. J.; Burns, L. A.; Wilke, J. J.; Abrams, M. L.; Russ, N. J.; Leininger, M. L.; Janssen, C. L.; Seidl, E. T.; Allen, W. D.; Schaefer, H. F.; King, R. A.; Valeev, E. F.; Sherrill, C. D.; Crawford, T. D. Wiley Interdiscip. Rev. Comput. Mol. Sci. 2012, 2, 556−565.

(32) DePrince, A. E.; Sherrill, C. D. J. Chem. Theory Comput. 2013, 9, 293−299.

(33) Weigend, F.; Ahlrichs, R. Phys. Chem. Chem. Phys. 2005, 7, 3297−3305.

(34) Dunning, T. H. J. Chem. Phys. 1989, 90, 1007−1023.

(35) Woon, D. E.; Dunning, T. H. J. Chem. Phys. 1993, 98, 1358−1371.

(36) Kendall, R. A.; Dunning, T. H.; Harrison, R. J. J. Chem. Phys. 1992, 96, 6796−6806.

(37) Dunning, T. H.; Peterson, K. A.; Wilson, A. K. J. Chem. Phys. 2001, 114, 9244−9253.

(38) Papajak, E.; Zheng, J.; Xu, X.; Leverentz, H. R.; Truhlar, D. G. J. Chem. Theory Comput. 2011, 7, 3027−3034.

(39) Helgaker, T.; Olsen, J.; Jørgensen, P. Molecular Electronic-Structure Theory; Wiley: New York, 2000; pp 322−324.

(40) Jurecka, P.; Hobza, P. Chem. Phys. Lett. 2002, 365, 89−94.

(41) Efron, B. J. Am. Statist. Assoc. 1987, 82, 171−185.

(42) Shimizu, K.; Tariq, M.; Gomes, M. F. C.; Rebelo, L. P. N.; Lopes, J. N. C. J. Phys. Chem. B 2010, 114, 5831−5834.

(43) Riley, K. E.; Pitoňák, M.; Jurecka, P.; Hobza, P. Chem. Rev. 2010, 110, 5023−5063.

(44) Tsuzuki, S.; Uchimaru, T.; Matsumura, K.; Mikami, M.; Tanabe, K. Chem. Phys. Lett. 2000, 319, 547−554.

(45) Wendler, K.; Dommert, F.; Zhao, Y. Y.; Berger, R.; Holm, C.; Delle Site, L. Faraday Discuss. 2012, 154, 111−132.

(46) Izgorodina, E. I.; Bernard, U. L.; MacFarlane, D. R. J. Phys. Chem. A 2009, 113, 7064−7072 PMID: 19462960..

(47) Izgorodina, E. I.; Forsyth, M.; MacFarlane, D. R. Phys. Chem. Chem. Phys. 2009, 11, 2452−2458.

(48) Johansson, K. M.; Izgorodina, E. I.; Forsyth, M.; MacFarlane, D. R.; Seddon, K. R. Phys. Chem. Chem. Phys. 2008, 10, 2972−2978.

(49) Izgorodina, E. I.; Forsyth, M.; MacFarlane, D. R. Aus. J. Chem. 2007, 60, 15−20.

(50) R Development Core Team. R: A Language and Environment for Statistical Computing; R Foundation for Statistical Computing: Vienna, Austria, 2011; ISBN 3-900051-07-0.

(51) Řezáč, J.; Riley, K. E.; Hobza, P. J. Chem. Theory Comput. 2011, 7, 2427−2438.

SCS-IL-MP2 produces accurate interaction energies for ionic liquid clusters

Jason Rigby,[a] Santiago Barrera Acevedo[a] and Ekaterina I Izgorodina*[a]

Abstract: Accurate energetics of intermolecular interactions are difficult to predict using quantum chemical methods due to their great computational expense. Here the application of our recently developed SCS-IL-MP2 method is reported for ionic liquid (IL) clusters of two and four ion pairs. Interaction energies were on average within 1.5 kJ mol^{-1} per ion pair of the CCSD(T)/CBS benchmark for clusters of both two and four ion pairs. This is a marked improvement by a factor of four to conventional MP2/CBS, which was on average between 5 and 8 kJ mol^{-1} per ion pair in these clusters. SCS-IL-MP2 substantially improves interaction energies for semi-Coulomb systems such as ILs, where electrostatic, dispersion and induction forces play are equally important, with CCSD(T) quality at a greatly reduced expense. SCS-IL-MP2 may be suitable for *ab initio* molecular dynamics or static quantum chemical calculations, for which large-scale ensembles of ionic species are of interest.

Ionic liquids (ILs) are becoming increasingly relevant in a diverse set of applications, including as electrolytes in energy storage devices,[1] synthetic catalysis,[2] mechanical lubricants[3] and pharmaceutically active ingredients,[4] to name just a few. Liquids qualifying as ILs (e.g. with melting points below 100°C) are said to number in the trillions, given common anion and cation combinations that give rise to low melting points, and possible mixtures thereof.[5] This incredibly large number of potential IL formulations means that the existence of suitable candidates for any given application would not be an unreasonable assumption, however identification of such a candidate is a formidable task. As yet, accurate and widely applicable models short of wavefunction-based quantum chemical methods are lacking.[6] Consequently, efforts to reduce the notorious computational cost associated with quantum chemical approaches are advantageous for the development of *a priori* design methodologies for ILs.[7]

Previously, we have reported an extension[8] to the widely used spin-component scaled MP2, SCS-MP2,[9] and scaling opposite-spin MP2, SOS-MP2,[10] approaches in electronic structure theory where the spin components of second-order Møller-Plesset perturbation theory are scaled to reproduce higher levels of theory that are considered benchmark (e.g. coupled-cluster with single, double and perturbative triple excitations, CCSD(T)), shown in eq 1.

$$E_{\text{CCSD(T)/CBS}}^{\text{corr}} \approx \alpha \Delta E_{(\text{OS})} + \beta (\Delta E)_{\text{SS}} \quad (1)$$

This new method, termed SCS-IL-MP2, was parameterised via multiple linear regression techniques to identify the ideal basis set and spin component coefficients for an extensive set of 174 IL ion pairs, which reproduced CCSD(T) energies within 1.0 kJ mol^{-1} on average and largely eliminated the need for basis set superposition error (BSSE) correction as commonly performed via the counterpoise (CP) method.[11] SCS-IL-MP2

involves scaling the opposite- and same-spin components by 1.05 and 0.68, respectively, for MP2 energies calculated with the relatively small Dunning's correlation-consistent triple-ζ quality basis set, cc-pVTZ. As SCS-IL-MP2 can be coupled with efficient algorithms for integral evaluation (*e.g.* the Resolution-of-Identity approximation[12]) and does not require CP correction, exceptional accuracy can be achieved at just *a fraction of the computational cost* that would be required for CP corrected MP2 calculations, let alone the benchmark method, CCSD(T) that scales as N with respect to molecular size.

In this study, we show that in addition to achieving high accuracy for clusters consisting of ion pairs of routinely used ionic liquids, the SCS-IL-MP2 method supersedes conventional CP corrected MP2 for ionic liquid clusters of two and four ion pairs. We assert that SCS-IL-MP2 is consequently an ideal method by which the Fragment Molecular Orbital approach[13] (FMO) may be applied to determine accurate electronic energies with near-linear scalability that would otherwise be hampered by CP correction and be almost certainly inaccessible with current coupled cluster methods. This approach is expected to produce favourable results owing to the fact that the dispersion energy is primarily accounted for as a two-body effect.[7] The newly proposed SCS-IL-MP2 method[8] was originally constructed for single ion pairs that included all possible energetically favourable configurations between the imidazolium or pyrrolidinium cations and a number of routinely used anions. Since BSSE is expected to be a somewhat localised source of error,[8] the SCS-IL-MP2 method seems suited to account for its effect through the scaled opposite- and same-spin components.

Two and four ion pair IL clusters consisting of 1,3-dimethylimidazolium (C$_1$mim), N-methyl-N-ethylpyrrolidinium (C$_2$mpyr) and N-ethylpyridinium (C$_2$py) cations, and tetrafluoroborate (BF$_4$), dicyanamide (N(CN)$_2$), bromide (Br) and chloride (Cl) anions (see Figure 1) were constructed in an recursive way. Energetically favourable ion pair configurations were assembled into ion pair dimers and these dimer units were duplicated and then assembled as ion pair tetramers. At each iteration the ion cluster was optimised. Two- and four-ion pair structures were optimised with the two-body FMO approach at the MP2/cc-pVDZ level of theory using the GAMESS software package.[14]

Cations: C$_1$mim (1,3-dimethylimidazolium), C$_2$mpyr (N-methyl-N-ethylpyrrolidinium), C$_2$py (N-ethylpyridinium)

Anions: BF$_4$ (tetrafluoroborate), N(CN)$_2$ (dicyanamide)
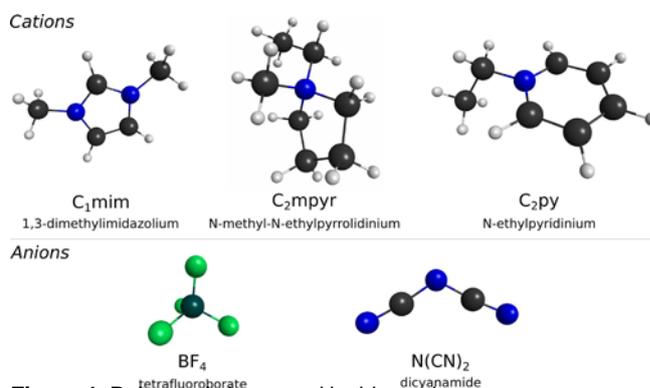
Figure 1. Polyatomic ions used in this study.

It has to be noted that the specific geometries used in this study are not of particular importance as fully optimized minima (and not necessarily global minima) on the potential energy surface are already sufficient to generate a robust set of benchmark energies to assess the performance of SCS-IL-MP2. In doing so,

[a]   Mr Jason Rigby, Dr Santiago Barrera Acevedo and Dr Ekaterina I Izgorodina
      Monash Computational Chemistry Group
      School of Chemistry, Monash University
      Wellington Road, Clayton VIC 3800
      E-mail: katya.izgorodina@monash.edu

the versatility and robustness of SCS-IL-MP2 may be verified in a more diverse set of situations, such as those that may be found in condensed phase molecular dynamics simulations, for example.

Benchmark CCSD(T)/CBS-quality correlation interaction energies were computed by adding the CP corrected interaction energy difference between MP2/aug-cc-pVDZ and CCSD(T)/aug-cc-pVDZ to two-point extrapolated MP2/CBS energies calculated using cc-pVTZ (abbreviated as CCT) and cc-pVQZ (abbreviated as CCQ) basis sets, as shown in eqs. 2 and 3.[15] We have previously shown that the use of augmented basis functions (denoted by the "aug-" prefix) is unnecessary for CBS extrapolation.[8]

$$\Delta E_{\text{MP2/CBS}}^{\text{corr}} = \frac{\Delta E_{\text{MP2}/X}^{\text{corr}} - \Delta E_{\text{MP2}/Y}^{\text{corr}}}{(\quad)} \tag{2}$$

$$\Delta E_{\text{CCSD(T)/CBS}}^{\text{corr}} = \Delta E_{\text{MP2/CBS}}^{\text{corr}} + \Delta E_{\text{CCSD(T)/ACCD}}^{\text{corr}} - \Delta E_{\text{MP2/ACCD}}^{\text{corr}} \tag{3}$$

The MP2/CBS extrapolation scheme shown in eq 1 follows the standard Helgaker approach,[16] in which MP2/X and MP2/Y represent MP2 correlation energies at two correlation-consistent basis sets, cc-pVTZ and cc-pVQZ, and *X* and *Y* are corresponding cardinal numbers ($X = 3$ and $Y = 4$). All energies were calculated with density-fitting, and CCSD(T) calculations were calculated either using the frozen natural orbital approximation[17] or where unfeasible, the domain based local pair natural orbital coupled cluster approximation (DLPNO-CCSD(T)).[18] All single-point energies were performed with the Psi4 software package[19] except for the DLPNO-CCSD(T) energies, which were calculated using ORCA.[20]
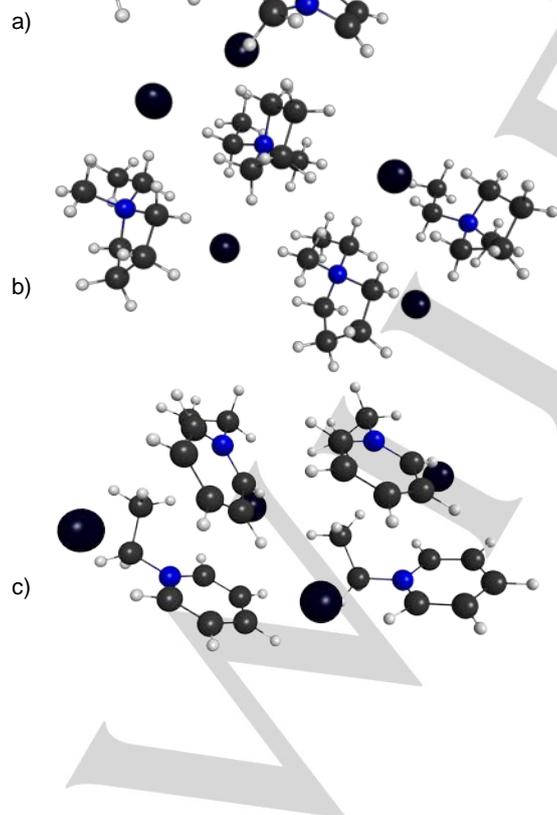
a)

b)

c)

interactions between these non-polar regions of the ions interact. In clusters (a) and (c), π-π stacking interactions between cations is present, further contributing to the non-negligible dispersion contribution found typically in ionic liquids.[7, 21] All clusters show multiple ion binding sites; the imidazolium clusters exhibit interactions inline with, and either side of the plane of the imidazolium ring; pyrrolidinium clusters have several exposed faces about the nitrogen atom, in which the anion may interact; and the pyridinium clusters have interaction sites on either face of the ring and in the plane near the nitrogen atom. The SCS-IL-MP2 method was fitted in view of accounting for the variety of interaction modes with multiple ion pair configurations used in the fitting procedure, so it is expected to perform consistently in clusters, in which multiple interaction sites are unavoidable. It has to be emphasised that the N-ethyl-pyridinium cation was not included in the original study for scaling the opposite- and same-spin components in the SCS-IL-MP2 method. Therefore, clusters with the C_2py cation represent the ultimate test for the performance of SCS-IL-MP2 for ionic liquid clusters.

The overall performance of SCS-IL-MP2 (abbreviated as SCS-IL) can be seen in Table 1 producing lowest mean absolute errors (MAE) compared to the other methods. The standard deviation of just 0.9 kJ mol⁻¹ indicates a narrow error distribution regardless of ionic cluster size and chemical nature of constituent cations and anions.

	Two ion pairs		Four ion pairs	
	MAE	Std. Dev.	MAE	Std. Dev.
SCS-IL	1.5	0.9	1.5	0.9
SOS-IL	1.7	0.9	1.9	0.8
SCS/CCT	7.2	1.8	7.0	2.5
SCS/CCT(CP)	21.2	1.5	23.7	2.2
SCS/CCQ	7.5	1.6	7.2	2.2
MP2/CCT(CP)	6.3	3.5	5.6	3.0
MP2/CCQ(CP)	2.9	2.3	3.5	2.7
MP2/CBS(CP)	5.4	3.3	8.2	3.8

Table 1. Mean absolute error (MAE) and the standard deviation for correlation interaction energies of the two- and four-ion pair given in kJ mol⁻¹ per ion pair. The non-counterpoised corrected SCS-MP2 method (abbreviated here as SCS) is included here for comparison. CP indicates that the energy was counterpoise-corrected using the Boys and Benardi method.

the positions of all other anion types in this series. A number of interaction modes are observed in these clusters; in clusters (b) and (c), alkyl chains are positioned such that dispersion

Figure 2. Four-ion pair cluster examples; a) [C₁mim]Br, b) [C₂mpyr]Br and c) [C₂py]Br.

Figure 2 shows examples of the IL clusters used. The bromide anion is shown here, as it is generally representative of

When relative errors are expressed in terms of the percentage of the total correlation interaction energy as in Table 2, SCS-IL-MP2 gives the lowest average and maximum errors among all methods studied in both two and four ion pair systems. SCS-IL-MP2 produces errors as low as 2.3% for the two ion pair systems, and 1.9% for the four ion pair systems. This indicates that the error relative to the interaction energy of the cluster is largely invariant with cluster size. Indeed, a *decrease* in percentage error is observed due to the increased interaction energy resulting from the many-body effects present in IL clusters.[7]

	Two ion pairs		Four ion pairs	
	Mean error	Max error	Mean error	Max error
SCS-IL	2.3 (1.7)	6.4	1.9 (1.2)	3.6
SOS-IL	2.6 (1.6)	5.4	2.4 (1.1)	4.3
SCS/CCT	10.6 (2.4)	14.4	8.5 (2.1)	12.2
SCS/CCT(CP)	32.0 (5.7)	45.1	29.8 (4.6)	38.8
SCS/CCQ	11.3 (3.0)	19.6	9.0 (2.5)	12.3
MP2/CCT(CP)	10.2 (6.8)	26.4	7.6 (5.0)	18.5
MP2/CCQ(CP)	4.2 (3.6)	14.6	4.1 (2.7)	9.7
MP2/CBS(CP)	7.5 (3.5)	14.7	9.6 (3.5)	15.0

Table 2. Mean, standard deviation in parentheses, and maximum error as a percentage of the total correlation interaction correlation energy for the two- and four-ion pair clusters. The non-counterpoised corrected SCS-MP2 method (abbreviated here as SCS) is included here for comparison. CP indicates that the energy was counterpoise-corrected using the Boys and Benardi method.
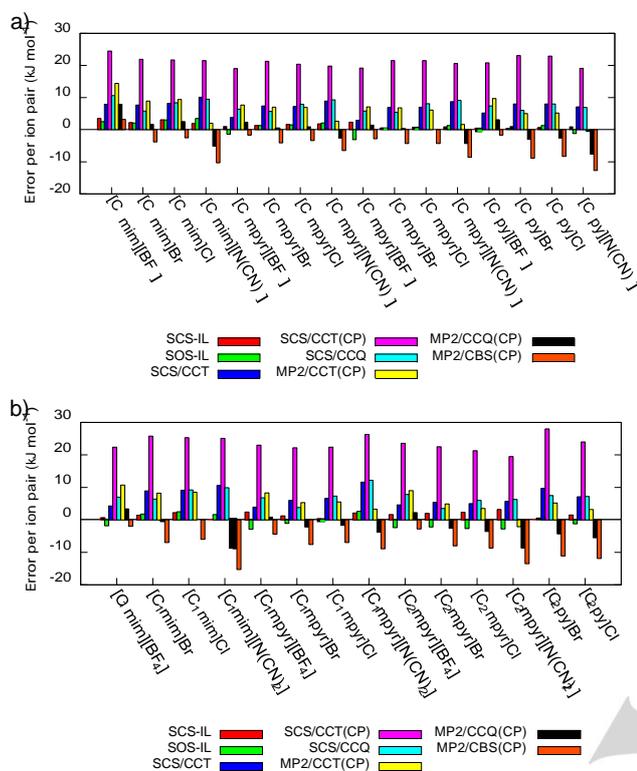
Figure 3. Two-ion pair (a) and four-ion pair (b) cluster interaction energy error given per ion pair. The SCS-MP2 method (abbreviated here as SCS) is included here for comparison. CP indicates that the energy was counterpoise-corrected using the Boys and Benardi method.

Figure 3 shows the errors for all the systems studied, given on a per ion pair basis. It can be seen that SCS-IL-MP2 consistently outperforms the original SCS formulation for both the triple- and quadruple-ζ quality basis sets for the clusters studied. Of particular note are the results from the C$_2$py series that was not part of the original SCS-IL-MP2 fitting set. Here, SCS-IL-MP2 gives errors that are sub-kJ mol^{-1} per ion pair for the two-ion pair set, and less than 2 kJ mol^{-1} per ion pair for the four ion pair set. These results show that the fitted SCS-IL-MP2 coefficients are likely to be suitable to broad range of semi-Coulomb systems, in which electrostatic, dispersion and induction forces are non-negligible and play an equally important role.

The vast majority of systems are described with increased accuracy and consistency by the SCS-IL-MP2 methods, with the performance being generally comparable to CP corrected MP2 energies calculated at the cc-pVQZ basis set, and exceeding MP2/CBS extrapolated energies. Notable outliers for the standard MP2 methods in general are the dicyanamide-based ILs, which often have a substantially overestimated correlation energy contribution that increases with the size of the basis set as was previously shown for ion pairs with the N(CN)$_2$ anion[8] For these systems, SCS-IL-MP2 is able to provide a consistently high level of accuracy that is superior to the standard MP2 method as well as the original SCS formulation. To this end, three conclusions can be drawn: 1) while CP corrected MP2/cc-

pVQZ may reproduce CCSD(T) energies in the majority of cases, outliers such as ILs with the N(CN)$_2$ anion cannot be corrected and *a priori* knowledge of presence "potentially difficult" anions/cations is usually not possible; 2) SCS-IL-MP2 can provide comparable accuracy whilst minimising the likelihood of outliers at only a fraction of the computational expense, as counterpoise correction is not required and the smaller cc-pVTZ basis set is required; 3) SCS-IL-MP2 was independently tested for a set of ionic liquids that were not included in the original dataset.

A key property of SCS-IL-MP2 involves its avoidance of CP correction. Indeed, CP correction for the original SCS-MP2 has been shown to have a detrimental effect on the quality of the interaction energies produced, whereas CP correction for standard MP2 energies are essential.[8] The precise origin of this effect is not well known, although it is hypothesised that there is a favourable balance between the basis set incompleteness errors and rate of correlation energy recovery that results in a uniquely well-fitting regression when MP2 spin components calculated using the cc-pVTZ basis set are fitted to a CP corrected set of CCSD(T)/CBS benchmark energies. It has been established that this is the most well-fitting basis set combination having performed rigorous statistical analysis of 17 popular basis sets in the original SCS-IL-MP2 paper.[8]

In a practical sense, the absence of CP correction means that the analysis of IL clusters becomes ever more computationally accessible; whereas a CP corrected interaction energy would require each monomer (ion) to be calculated in the basis set of the entire clusters, SCS-IL-MP2 requires only the actual cluster to be calculated with each ions being calculated in isolation. Therefore, not only is the overall computational complexity of the calculation reduced, but also the problem becomes very well suited to highly parallel fragmentation approaches, such as the FMO approach. Using standard MP2, the application of the FMO approach, FMO-MP2, can be made of near-linear scalability only for large-scale clusters of ionic liquids,[7] whereas scalability of energy calculations for individual ions in the basis set of the cluster are still hindered by scalability of the underlying MP2 method, *i.e.* N^5 where N is the number of basis functions. Thus, the bottleneck of any interaction energy calculation of large-scale ionic clusters is the CP correction itself. It is obvious that SCS-IL-MP2 does not change how a combination of the FMO approach with standard MP2 is applied. The new method brings a significant short cut through elimination of the time consuming CP step.

The new SCS-IL-MP2 method for calculating the interaction energies of ionic liquid clusters represents a remarkably accurate and efficient route by which the bulk properties of ILs can be characterised. By leveraging an already well-established framework, SCS-MP2, it has been shown that by fine-tuning the scaling factors applied to the spin components of the MP2 correlation interaction energy with only the relatively small cc-pVTZ basis set, accuracy approaching CCSD(T)/CBS quality can be achieved for both ion pairs as well as large-scale clusters of ionic liquids. The latter represents a tremendous achievement as it allows for not only accurate calculations of energetics of novel ionic liquids through highly parallel techniques such as the FMO approach but also development of more accurate forces fields that are extremely needed for molecular dynamics simulations with predictive power.[22]

Acknowledgements

Keywords: ionic liquids • MP2 • coupled cluster • electronic structure • ion clusters

[1] Y. Lu, K. Korf, Y. Kambe, Z. Tu, L. A. Archer, *Angew. Chem.* **2014**, *126*, 498-502.
[2] K. Matuszek, A. Chrobok, F. Coleman, K. R. Seddon, M. Swadzba-Kwasny, *Green Chem.* **2014**, *16*, 3463-3471.
[3] H. Li, R. J. Wood, M. W. Rutland, R. Atkin, *Chem. Commun.* **2014**, *50*, 4368-4370.
[4] I. M. Marrucho, L. C. Branco, L. P. N. Rebelo, *Annu. Rev. Chem. Biomol. Eng.* **2014**, *5*, 527-546.
[5] R. D. Rogers, K. R. Seddon, *Science* **2003**, *302*, 792-793.
[6] E. I. Izgorodina, *Phys. Chem. Chem. Phys.* **2011**, *13*, 4189-4207.
[7] E. I. Izgorodina, J. Rigby, D. R. MacFarlane, *Chem. Commun.* **2012**, *48*, 1493-1495.
[8] J. Rigby, E. I. Izgorodina, *J. Chem. Theory Comput.* **2014**.
[9] S. Grimme, *J. Chem. Phys.* **2003**, *118*, 9095-9102.
[10] R. A. Distasio Jr, M. Head-Gordon, *Mol. Phys.* **2007**, *105*, 1073-1083.
[11] S. F. Boys, F. Bernardi, *Mol. Phys.* **1970**, *19*, 553-566.
[12] M. Feyereisen, G. Fitzgerald, A. Komornicki, *Chem. Phys. Lett.* **1993**, *208*, 359-363.
[13] a) K. Kitaura, E. Ikeo, T. Asada, T. Nakano, M. Uebayasi, *Chem. Phys. Lett.* **1999**, *313*, 701-706; b) D. G. Fedorov, K. Kitaura, *J. Chem. Phys.* **2004**, *120*, 6832-6840; c) D. G. Fedorov, K. Kitaura, *J. Chem. Phys.* **2004**, *121*, 2483-2490.
[14] M. W. Schmidt, K. K. Baldridge, J. A. Boatz, S. T. Elbert, M. S. Gordon, J. H. Jensen, S. Koseki, N. Matsunaga, K. A. Nguyen, S. Su, T. L. Windus, M. Dupuis, J. A. Montgomery, *J. Comput. Chem.* **1993**, *14*, 1347-1363.
[15] a) P. Jurecka, J. Sponer, J. Cerny, P. Hobza, *Phys. Chem. Chem. Phys.* **2006**, *8*, 1985-1993; b) P. Jurečka, P. Hobza, *Chem. Phys. Lett.* **2002**, *365*, 89-94.
[16] T. Helgaker, P. Jorgensen, J. Olsen, *Molecular Electronic-Structure Theory*, Wiley, **2000**.
[17] A. E. DePrince, C. D. Sherrill, *J. Chem. Theory Comput.* **2012**, *9*, 293-299.
[18] C. Riplinger, B. Sandhoefer, A. Hansen, F. Neese, *J. Chem. Phys.* **2013**, *139*, 134101.
[19] J. M. Turney, A. C. Simmonett, R. M. Parrish, E. G. Hohenstein, F. A. Evangelista, J. T. Fermann, B. J. Mintz, L. A. Burns, J. J. Wilke, M. L. Abrams, N. J. Russ, M. L. Leininger, C. L. Janssen, E. T. Seidl, W. D. Allen, H. F. Schaefer, R. A. King, E. F. Valeev, C. D. Sherrill, T. D. Crawford, *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2012**, *2*, 556-565.
[20] F. Neese, *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2012**, *2*, 73-78.
[21] K. Shimizu, M. Tariq, M. F. C. Gomes, L. s. P. N. Rebelo, J. N. C. Lopes, *J. Phys. Chem. B* **2010**, *114*, 5831-5834.
[22] E. J. Maginn, J. R. Elliott, *Ind. Eng. Chem. Res.* **2010**, *49*, 3059-3078.

Entry for the Table of Contents (Please choose one layout)

Layout 2:

COMMUNICATION
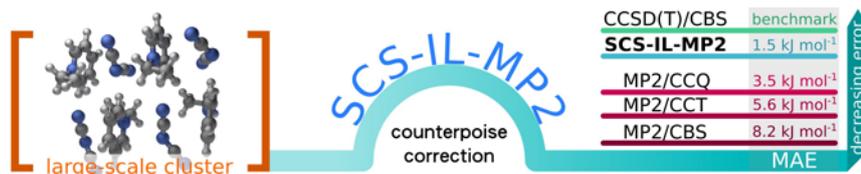
Author(s), Corresponding Author(s)*

Page No. – Page No.

Title

SCS-IL-MP2 provides accurate interaction energies for semi-Coulombic systems, such as ionic liquids, without the need for computationally expensive counterpoise correction.

Supporting Information

Cluster interaction energies

Table S1. CCSD(T)/CBS benchmark correlation energies. Superscript asterisk indicates that the 2---IP cluster was calculated using the DLPNO---CCSD(T) approximation. DLPNO---CCSD(T) was used for the 4---IP clusters. All energies CP corrected.

IL	2---IP		4---IP	
	cluster total	per ion pair	cluster total	per ion pair
[C_1mim][BF$_4$]	---108.9	---54.4	---230.6	---57.7
[C_1mim]Br	---123.0	---61.5	---316.8	---79.2
[C_1mim]Cl	---115.2	---57.6	---299.5	---74.9
[C_1mim][N(CN)$_2$]	---164.3	---82.1	---421.5	---105.4
[C_1mpyr][BF$_4$]*	---105.1	---52.6	---268.3	---67.1
[C_1mpyr]Br	---132.6	---66.3	---299.0	---74.7
[C_1mpyr]Cl	---124.0	---62.0	---298.3	---74.6
[C_1mpyr][N(CN)$_2$]	---152.5	---76.3	---403.9	---101.0
[C_2mpyr][BF$_4$]*	---112.8	---56.4	---271.6	---67.9
[C_2mpyr]Br	---135.9	---68.0	---311.4	---77.8
[C_2mpyr]Cl	---138.1	---69.1	---308.3	---77.1
[C_2mpyr][N(CN)$_2$]*	---165.5	---82.8	---360.4	---90.1
[C_2py][BF$_4$]*	---112.5	---56.3	—	—
[C_2py]Br	---156.9	---78.5	---399.5	---99.9
[C_2py]Cl	---152.5	---76.2	---352.9	---88.2
[C_2py][N(CN)$_2$]	---172.9	---86.5	—	—

Note: Omitted 4---IP energies were unable to be completed due to software limitations

Table S2. MP2/CBS interaction energies

IL	2---IP				4---IP			
	ΔE	ΔE^{CP}	BSSE	ΔE^{CP} per IP	ΔE	ΔE^{CP}	BSSE	ΔE^{CP} per IP
[C$_1$mim][BF$_4$]	---107.3	---102.4	4.9	---51.2	---249.1	---238.5	10.7	---59.6
[C$_1$mim]Br	---144.7	---130.6	14.1	---65.3	---378.7	---344.9	33.8	---86.2
[C$_1$mim]Cl	---126.0	---120.2	5.7	---60.1	---337.3	---323.3	14.0	---80.8
[C$_1$mim][N(CN)$_2$]	---188.2	---184.8	3.4	---92.4	---492.0	---482.6	9.4	---120.7
[C$_1$mpyr][BF$_4$]	---113.8	---108.6	5.3	---54.3	---298.6	---285.9	12.7	---71.5
[C$_1$mpyr]Br	---155.5	---140.8	14.7	---70.4	---365.1	---329.5	35.6	---82.4
[C$_1$mpyr]Cl	---136.7	---130.8	5.9	---65.4	---340.2	---326.4	13.8	---81.6
[C$_1$mpyr][N(CN)$_2$]	---169.6	---165.5	4.2	---82.8	---450.1	---439.6	10.5	---109.9
[C$_2$mpyr][BF$_4$]	---123.9	---118.3	5.6	---59.2	---295.5	---282.8	12.7	---70.7
[C$_2$mpyr]Br	---160.1	---144.5	15.5	---72.3	---380.6	---343.5	37.1	---85.9
[C$_2$mpyr]Cl	---153.0	---146.7	6.4	---73.4	---357.6	---343.1	14.5	---85.8
[C$_2$mpyr][N(CN)$_2$]	---186.7	---182.6	4.1	---91.3	---424.4	---414.6	9.8	---103.7
[C$_2$py][BF$_4$]	---120.6	---116.0	4.6	---58.0	---360.6	---348.3	12.3	---87.1
[C$_2$py]Br	---187.7	---174.8	12.9	---87.4	---476.0	---444.0	32.0	---111.0
[C$_2$py]Cl	---175.0	---169.1	5.9	---84.6	---413.8	---400.4	13.4	---100.1
[C$_2$py][N(CN)$_2$]	---202.0	---198.4	3.6	---99.2	---513.5	---503.6	9.9	---125.9

Table S3. MP2/cc---pVTZ interaction energies

IL	2---IP				4---IP			
	ΔE	ΔE^{CP}	BSSE	ΔE^{CP} per IP	ΔE	ΔE^{CP}	BSSE	ΔE^{CP} per IP
[C$_1$mim][BF$_4$]	---115.4	---80.1	35.3	---40.1	---264.4	---188.0	76.4	---47.0
[C$_1$mim]Br	---135.1	---105.2	29.9	---52.6	---354.9	---283.9	71.0	---71.0
[C$_1$mim]Cl	---124.0	---96.4	27.7	---48.2	---331.5	---265.4	66.2	---66.4
[C$_1$mim][N(CN)$_2$]	---183.7	---160.2	23.5	---80.1	---481.7	---422.1	59.6	---105.5
[C$_1$mpyr][BF$_4$]	---121.7	---89.6	32.1	---44.8	---315.9	---235.2	80.7	---58.8
[C$_1$mpyr]Br	---148.0	---118.7	29.3	---59.4	---346.0	---277.6	68.4	---69.4
[C$_1$mpyr]Cl	---137.4	---110.3	27.1	---55.2	---341.2	---276.5	64.7	---69.1
[C$_1$mpyr][N(CN)$_2$]	---169.9	---147.4	22.5	---73.7	---451.3	---390.8	60.5	---97.7
[C$_2$mpyr][BF$_4$]	---133.2	---98.7	34.6	---49.4	---315.5	---235.6	79.9	---58.9
[C$_2$mpyr]Br	---153.2	---122.4	30.8	---61.2	---363.5	---291.8	71.8	---73.0
[C$_2$mpyr]Cl	---155.3	---125.9	29.5	---63.0	---361.0	---294.3	66.8	---73.6
[C$_2$mpyr][N(CN)$_2$]	---186.8	---162.4	24.4	---81.2	---425.6	---369.0	56.5	---92.3
[C$_2$py][BF$_4$]	---126.1	---93.1	33.0	---46.6	---376.1	---286.3	89.8	---71.6
[C$_2$py]Br	---178.3	---146.9	31.5	---73.5	---454.8	---378.8	75.9	---94.7
[C$_2$py]Cl	---172.4	---142.1	30.2	---71.1	---408.9	---340.1	68.8	---85.0
[C$_2$py][N(CN)$_2$]	---198.9	---174.2	24.7	---87.1	---507.4	---445.5	62.0	---111.4

Table S4. MP2/cc---pVQZ interaction energies

IL	2---IP				4---IP			
	ΔE	ΔE^{CP}	BSSE	ΔE^{CP} per IP	ΔE	ΔE^{CP}	BSSE	ΔE^{CP} per IP
[C$_1$mim][BF$_4$]	---110.7	---93.0	17.7	---46.5	---255.6	---217.2	38.4	---54.3
[C$_1$mim]Br	---140.6	---119.9	20.8	---60.0	---368.7	---319.1	49.5	---79.8
[C$_1$mim]Cl	---125.2	---110.2	15.0	---55.1	---334.8	---298.9	36.0	---74.7
[C$_1$mim][N(CN)$_2$]	---186.3	---174.4	11.9	---87.2	---487.7	---457.1	30.6	---114.3
[C$_1$mpyr][BF$_4$]	---117.2	---100.6	16.6	---50.3	---305.9	---264.5	41.3	---66.1
[C$_1$mpyr]Br	---152.4	---131.5	20.9	---65.8	---357.1	---307.6	49.4	---76.9
[C$_1$mpyr]Cl	---137.0	---122.1	14.8	---61.1	---340.6	---305.3	35.3	---76.3
[C$_1$mpyr][N(CN)$_2$]	---169.7	---157.8	11.9	---78.9	---450.6	---419.0	31.6	---104.8
[C$_2$mpyr][BF$_4$]	---127.9	---110.0	17.8	---55.0	---303.9	---262.8	41.1	---65.7
[C$_2$mpyr]Br	---157.2	---135.2	22.0	---67.6	---373.4	---321.7	51.7	---80.4
[C$_2$mpyr]Cl	---154.0	---137.9	16.1	---69.0	---359.1	---322.5	36.5	---80.6
[C$_2$mpyr][N(CN)$_2$]	---186.8	---174.1	12.7	---87.1	---424.9	---395.4	29.5	---98.9
[C$_2$py][BF$_4$]	---122.9	---106.3	16.6	---53.2	---367.1	---322.1	45.0	---80.5
[C$_2$py]Br	---183.7	---163.0	20.7	---81.5	---467.1	---416.5	50.5	---104.1
[C$_2$py]Cl	---173.9	---157.7	16.2	---78.9	---411.7	---374.9	36.8	---93.7
[C$_2$py][N(CN)$_2$]	---200.7	---188.2	12.5	---94.1	---511.0	---479.1	31.9	---119.8

Table S5. SCS---MP2/cc---pVTZ interaction energies

IL	2---IP				4---IP			
	ΔE	ΔE^{CP}	BSSE	ΔE^{CP} per IP	ΔE	ΔE^{CP}	BSSE	ΔE^{CP} per IP
[C$_1$mim][BF$_4$]	---93.2	---59.8	33.3	---29.9	---213.3	---141.1	72.2	---35.3
[C$_1$mim]Br	---107.6	---79.1	28.5	---39.6	---281.3	---213.4	67.9	---53.4
[C$_1$mim]Cl	---98.9	---71.8	27.0	---35.9	---262.9	---198.2	64.8	---49.6
[C$_1$mim][N(CN)$_2$]	---144.0	---121.2	22.8	---60.6	---379.2	---321.0	58.2	---80.3
[C$_1$mpyr][BF$_4$]	---97.4	---67.2	30.3	---33.6	---252.6	---176.3	76.3	---44.1
[C$_1$mpyr]Br	---117.9	---90.0	27.9	---45.0	---275.1	---210.1	65.0	---52.5
[C$_1$mpyr]Cl	---109.6	---83.2	26.4	---41.6	---272.1	---208.8	63.2	---52.2
[C$_1$mpyr][N(CN)$_2$]	---134.7	---113.0	21.7	---56.5	---357.6	---298.5	59.0	---74.6
[C$_2$mpyr][BF$_4$]	---107.1	---74.3	32.7	---37.2	---253.2	---177.3	75.9	---44.3
[C$_2$mpyr]Br	---122.2	---92.9	29.3	---46.5	---289.7	---221.4	68.3	---55.4
[C$_2$mpyr]Cl	---124.1	---95.2	28.9	---47.6	---288.4	---222.9	65.4	---55.7
[C$_2$mpyr][N(CN)$_2$]	---148.1	---124.3	23.8	---62.2	---337.5	---282.5	55.0	---70.6
[C$_2$py][BF$_4$]	---102.1	---70.9	31.2	---35.5	---301.8	---216.3	85.6	---54.1
[C$_2$py]Br	---141.0	---110.8	30.2	---55.4	---360.6	---287.4	73.2	---71.9
[C$_2$py]Cl	---136.4	---106.8	29.7	---53.4	---324.6	---257.1	67.5	---64.3
[C$_2$py][N(CN)$_2$]	---158.7	---134.8	23.9	---67.4	---406.7	---346.1	60.6	---86.5

Table S6. SCS-MP2/cc-pVQZ interaction energies

IL	2-IP				4-IP			
	ΔE	ΔE^{CP}	BSSE	ΔE^{CP} per IP	ΔE	ΔE^{CP}	BSSE	ΔE^{CP} per IP
[C$_1$mim][BF$_4$]	-87.6	-70.2	17.4	-35.1	-202.5	-164.8	37.7	-41.2
[C$_1$mim]Br	-111.5	-91.5	19.9	-45.8	-291.2	-243.5	47.8	-60.9
[C$_1$mim]Cl	-98.4	-83.4	15.0	-41.7	-262.7	-226.5	36.1	-56.6
[C$_1$mim][N(CN)$_2$]	-145.3	-133.3	12.0	-66.7	-381.8	-350.7	31.0	-87.7
[C$_1$mpyr][BF$_4$]	-92.4	-76.2	16.2	-38.1	-241.3	-200.6	40.7	-50.2
[C$_1$mpyr]Br	-121.2	-101.1	20.1	-50.6	-283.7	-236.2	47.5	-59.1
[C$_1$mpyr]Cl	-108.3	-93.4	14.8	-46.7	-269.2	-233.8	35.4	-58.5
[C$_1$mpyr][N(CN)$_2$]	-134.0	-122.0	12.0	-61.0	-355.0	-322.9	32.1	-80.7
[C$_2$mpyr][BF$_4$]	-101.2	-83.7	17.5	-41.9	-240.5	-199.9	40.5	-50.0
[C$_2$mpyr]Br	-125.2	-104.1	21.1	-52.1	-297.3	-247.5	49.8	-61.9
[C$_2$mpyr]Cl	-121.8	-105.6	16.2	-52.8	-284.1	-247.4	36.8	-61.9
[C$_2$mpyr][N(CN)$_2$]	-147.3	-134.5	12.9	-67.3	-335.1	-305.2	29.9	-76.3
[C$_2$py][BF$_4$]	-97.8	-81.5	16.2	-40.8	-290.2	-245.7	44.5	-61.4
[C$_2$py]Br	-144.9	-124.8	20.1	-62.4	-369.3	-320.1	49.2	-80.0
[C$_2$py]Cl	-136.5	-120.2	16.3	-60.1	-324.1	-287.1	37.1	-71.8
[C$_2$py][N(CN)$_2$]	-159.2	-146.6	12.6	-73.3	-406.9	-374.5	32.4	-93.6

Table S7. SCS-IL-MP2 interaction energies

IL	2-IP				4-IP			
	ΔE	ΔE^{CP}	BSSE	$\Delta E^{non\text{-}CP}$ per IP	ΔE	ΔE^{CP}	BSSE	$\Delta E^{non\text{-}CP}$ per IP
[C$_1$mim][BF$_4$]	-101.9	-68.7	33.2	-51.0	-233.4	-161.5	72.0	-58.4
[C$_1$mim]Br	-118.7	-90.4	28.2	-59.4	-311.2	-244.0	67.2	-77.8
[C$_1$mim]Cl	-109.0	-82.6	26.4	-54.5	-290.7	-227.5	63.2	-72.7
[C$_1$mim][N(CN)$_2$]	-160.4	-138.0	22.4	-80.2	-421.2	-364.3	56.9	-105.3
[C$_1$mpyr][BF$_4$]	-107.1	-76.9	30.2	-53.6	-277.9	-201.9	76.0	-69.5
[C$_1$mpyr]Br	-130.0	-102.3	27.7	-65.0	-303.7	-239.2	64.6	-75.9
[C$_1$mpyr]Cl	-120.8	-94.9	25.9	-60.4	-299.9	-238.0	61.8	-75.0
[C$_1$mpyr][N(CN)$_2$]	-149.0	-127.6	21.4	-74.5	-395.6	-337.8	57.8	-98.9
[C$_2$mpyr][BF$_4$]	-117.4	-84.9	32.6	-58.7	-278.0	-202.5	75.4	-69.5
[C$_2$mpyr]Br	-134.6	-105.6	29.1	-67.3	-319.4	-251.6	67.8	-79.9
[C$_2$mpyr]Cl	-136.6	-108.4	28.2	-68.3	-317.5	-253.6	63.9	-79.4
[C$_2$mpyr][N(CN)$_2$]	-163.8	-140.5	23.3	-81.9	-373.2	-319.3	53.9	-93.3
[C$_2$py][BF$_4$]	-111.5	-80.4	31.1	-55.8	-331.3	-246.5	84.8	-82.8
[C$_2$py]Br	-156.2	-126.4	29.8	-78.1	-398.8	-326.7	72.1	-99.7
[C$_2$py]Cl	-151.0	-122.1	28.9	-75.5	-358.7	-292.9	65.9	-89.7
[C$_2$py][N(CN)$_2$]	-174.8	-151.3	23.5	-87.4	-446.8	-387.6	59.2	-111.7

Table S8. SOS---IL---MP2 interaction energies

IL	2---IP				4---IP			
	ΔE	ΔE^{CP}	BSSE	$\Delta E^{non\text{---}CP}$ per IP	ΔE	ΔE^{CP}	BSSE	$\Delta E^{non\text{---}CP}$ per IP
[C$_1$mim][BF$_4$]	---103.8	---62.9	40.9	---51.9	---237.7	---149.1	88.5	---59.4
[C$_1$mim]Br	---118.8	---83.7	35.1	---59.4	---309.5	---225.7	83.8	---77.4
[C$_1$mim]Cl	---109.2	---75.4	33.8	---54.6	---289.4	---208.5	80.9	---72.4
[C$_1$mim][N(CN)$_2$]	---157.2	---128.8	28.4	---78.6	---415.1	---342.5	72.6	---103.8
[C$_1$mpyr][BF$_4$]	---107.9	---70.8	37.1	---54.0	---279.7	---186.1	93.7	---69.9
[C$_1$mpyr]Br	---130.1	---95.7	34.3	---65.1	---303.3	---223.3	80.0	---75.8
[C$_1$mpyr]Cl	---121.1	---88.2	32.9	---60.6	---300.6	---221.7	78.9	---75.2
[C$_1$mpyr][N(CN)$_2$]	---148.3	---121.4	26.9	---74.2	---393.4	---319.7	73.7	---98.4
[C$_2$mpyr][BF$_4$]	---118.9	---78.7	40.2	---59.5	---281.1	---187.7	93.4	---70.3
[C$_2$mpyr]Br	---135.1	---99.1	36.0	---67.6	---320.0	---235.9	84.2	---80.0
[C$_2$mpyr]Cl	---137.3	---101.2	36.1	---68.7	---319.0	---237.2	81.8	---79.8
[C$_2$mpyr][N(CN)$_2$]	---163.0	---133.3	29.6	---81.5	---371.5	---302.9	68.6	---92.9
[C$_2$py][BF$_4$]	---114.0	---75.7	38.3	---57.0	---335.0	---229.6	105.4	---83.8
[C$_2$py]Br	---154.9	---117.5	37.4	---77.5	---396.9	---306.2	90.7	---99.2
[C$_2$py]Cl	---149.9	---112.8	37.1	---75.0	---357.6	---273.1	84.5	---89.4
[C$_2$py][N(CN)$_2$]	---175.4	---145.7	29.7	---87.7	---451.1	---375.4	75.7	---112.8

Section 4.5. *Supporting information: SCS-IL-MP2 produces accurate interaction*
Chapter 4. *Implicit BSSE correction with a modified SCS-MP2 approach*
energies for ionic liquid clusters (abridged)
76
76

References

[1] Stefan Grimme. Improved second-order Møller–Plesset perturbation theory by separate scaling of parallel- and antiparallel-spin pair correlation energies. *The Journal of Chemical Physics*, 118(20):9095–9102, 2003.

[2] S.F. Boys and F. Bernardi. The calculation of small molecular interactions by the differences of separate total energies. some procedures with reduced errors. *Molecular Physics*, 19(4):553–566, 1970.

[3] Stefan Zahn, Douglas R. MacFarlane, and Ekaterina I. Izgorodina. Assessment of Kohn-Sham density functional theory and Møller-Plesset perturbation theory for ionic liquids. *Physical Chemistry Chemical Physics*, 15:13664–13675, 2013.

[4] Jens Antony and Stefan Grimme. Is spin-component scaled second-order MøllerPlesset perturbation theory an appropriate method for the study of noncovalent interactions in molecules? *The Journal of Physical Chemistry A*, 111(22):4862–4868, 2007. PMID: 17506533.

[5] Ekaterina I. Izgorodina, Jason Rigby, and Douglas R. MacFarlane. Large-scale ab initio calculations of archetypical ionic liquids. *Chemical Communications*, 48:1493–1495, 2012.

Chapter 5

Revisiting density functional theory:
The improvement and application of empirical dispersion corrections

5.1 Introduction

The use of Kohn-Sham density functional theory (DFT)[1] in quantum chemistry has become routine and is considered by many to be the most popular approach to molecular energetics. DFT, however, is known for being notoriously poor in reproducing electron correlation effects as the exact relationship between electron density and the corresponding energy is unknown, although formally proven to exist.[2] As such, significant effort has been made in producing functionals that approximate this relationship. The forms taken by these functionals vary significantly amongst the myriad of those that are currently available, however many rely on the assumption that the properties of accepted models, such as the well-characterised uniform electron gas[2] or the two-electron helium atom,[3] can be transferred to more complex chemical systems. While such assumptions have made significant inroads to accurate molecular descriptions, it appears these simplified approximations are unlikely to produce consistently accurate results amongst the wide variety of chemical systems of interest, including charge-transfer complexes[4] and ionic liquids (ILs).[5] Indeed, the classification of DFT as an "*ab initio*" method may be considered somewhat of a misnomer given that functionals are often tailored to a particular class of chemical system or specific test sets,[6, 7] which limits

their broad applicability. Indeed, Becke introduced the concept of the "density functional approximation" in order to make the distinction between DFT, which is exact, and its application, which is approximate and may contain empirical or non-empirical parameters.[8]

Given that the inherent limitation of DFT is in its description of the van der Waals interaction,[9] it has been proposed by Grimme that pairwise atom empirical dispersion coefficients be developed such that the dispersion energy can be described as the power series expansion such as in eq. 5.1. At the time of writing, there have been three versions of the empirical dispersion correction methodology proposed by Grimme;[10–12] this chapter is concerned exclusively with the latest version, DFT-D version 3 (DFT-D3).[12]

The original two-body DFT-D3 formulation defined the dispersion correction as:

$$E_{\text{disp}}^{\text{DFT-D3(zero)}} = -\frac{1}{2} \sum_{i \neq j} \sum_{n \geq 6, n \text{ even}} s_n \frac{C_n^{ij}}{r_{ij}^n} f_{\text{damp}}(n, r_{ij}) \tag{5.1}$$

This summation is usually truncated at $n = 8$ and is summed over all atom pairs, i and j. In this formulation, f_{damp} is a distance dependent damping function that attempts to correct for divergence of the dispersion energy as $r_{ij} \rightarrow 0$, given in eq. 5.2, known as "zero damping." The other variable in this expression is the s_n term, which has been labelled a density functional (DF) dependent term and is a global scaling factor for the nth order dispersion energy. The s_n terms are required because the DF will already contribute some approximation of the true dispersion energy and therefore only some proportion of the empirical correction is used.

$$f_{\text{damp}}(n, r_{ij}) = \frac{1}{1 + 6(r_{ij}/(s_{r,n}R_{ij}^0))^{-\alpha_n}} \tag{5.2}$$

The damping function given above depends on an atom pair cut-off radius, R_{ij}^0, which is defined as the sum of the van der Waals radii of the atom pair. This approximation has an additional correction via a scaling factor, $s_{r,n}$, which is specific to the approximation order ($n = 6$ or 8). This was originally fitted via a least-squares regression method for $s_{r,6}$ and set to unity for $s_{r,8}$. The remaining parameter, α_n, has been set to 14 for $n = 6$ and for $n > 6$, $\alpha_{n+2} = \alpha_n + 2$.

An alternative Becke-Johnson (BJ) damping function has been proposed,[13] which takes the following form:

$$E_{\text{disp}}^{\text{DFT-D3(BJ)}} = -\frac{1}{2} \sum_{i \neq j} \sum_{n \geq 6, n \text{ even}} s_n \frac{C_n^{ij}}{r_{ij}^n + [f(R_{ij}^0)]^n} \tag{5.3}$$

$$f(R^0_{ij}) = a_1 R^0_{ij} + a_2 \tag{5.4}$$

In this case, the BJ damping functions results in a finite dispersion correction at $r_{ij} = 0$. Here, the atom cut-off term proposed by Johnson and Becke[14] was adopted and is defined as:

$$R^0_{ij} = \sqrt{\frac{C^{ij}_8}{C^{ij}_6}} \tag{5.5}$$

In both the zero damping and BJ damping forms of DFT-D3, the DF dependent terms, s_6 and s_8 are tunable. For the zero damping function, tunable terms include $s_{r,6}$ and $s_{r,8}$, and for the BJ damping function, a_1 and a_2. In the official DFT-D3 parameter set, s_6 was held at unity, and only the s_8 term was fitted.   For the zero damping function, only $s_{r,6}$ was proposed to be scaled with $s_{r,8}$ held at unity. When fitting the BJ damping function, both a_1 and a_2 were allowed to vary. Parameters for the original DFT-D3 formulation were optimised against the S66 data set[15] using CCSD(T)/CBS benchmark energies.

In this study, the tunable coefficients are refitted in order to reproduce CCSD(T)/CBS energies for an extensive set of ionic liquid ion pairs (IPs) by providing empirical corrections for pure Hartree-Fock (HF) energies, and energies calculated using the GGA-type PBE[16, 17] and BLYP[3, 18, 19] DFs, which are used routinely in *ab initio* molecular dynamics (AIMD).[20–24] Both zero and BJ damping functions are assessed in cases where all coefficients are permitted to vary, and where some restrictions are applied, including as described in Grimme's original methodology.[12, 13] The refitted DFT-D3 parameters are then validated against a subset of ILs that are scanned along a coordinate of the potential energy surface (PES) in order to evaluate the performance at non-equilibrium geometries, with benchmark interaction PESs calculated at CCSD(T)/CBS. Ionic liquid specific SCS-MP2 (SCS-IL-MP2)[25] PESs have also been produced for comparison. This study embodies the concepts raised in Chapter 4 where it is asserted that the application of any empirical corrections are necessarily fitted to only a subset of chemical systems; that is, the only way to consistently reproduce high-quality energies is simply to employ higher levels of theory. As a consequence, the refined DFT-D3 parameters are designed specifically for semi-Coulombic systems in exactly the same sense as the SCS-IL-MP2 method.

5.2   Theoretical procedures

5.2.1   Fitting data

The IPs analysed consisted of a set of 174 ILs involving several energetically favourable configurations for each. The cations and anions consisted of $[C_{(1-4)}mim]^+$ and $[C_{(1-4)}mpyr]^+$ combined with Br^-, Cl^-, tetrafluoroborate ($[BF_4]^-$), hexafluorophosphate ($[PF_6]^-$), di-cyanamide ($[DCA]^-$), tosylate ($[tos]^-$), mesylate ($[mes]^-$) and bis(trifluoromethylsulfonyl)amide ($[NTf_2]^-$). Structures and associated interaction energies are available as part of the Electronic Supporting Information in reference [25] and described in detail in Chapter 4, Section 4.3. Examples of the conformations included in the fitting set are shown in Figures 5.1 and 5.2.
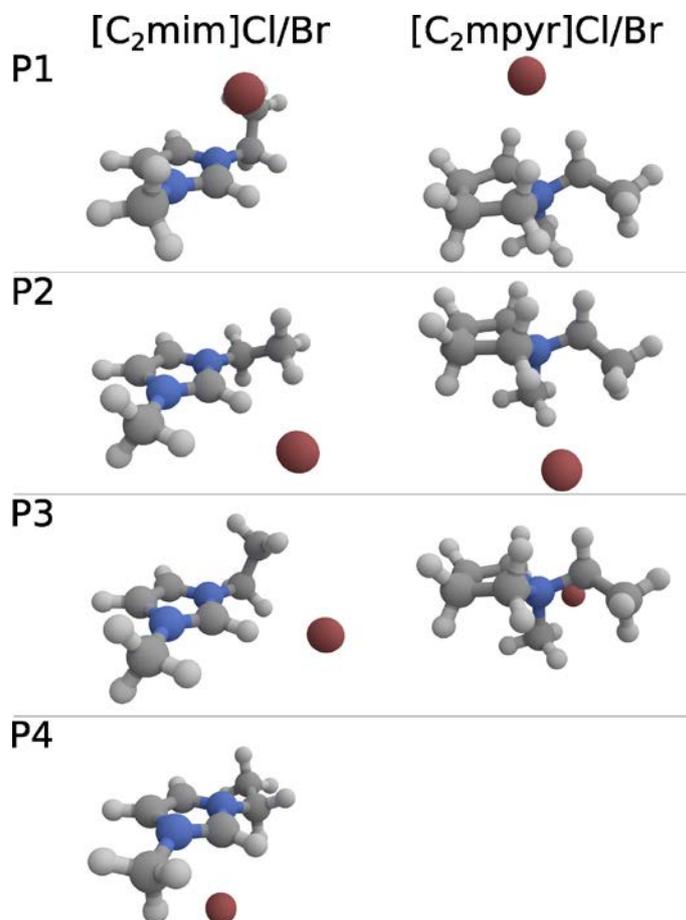

Figure 5.1: Examples of IL IP conformations for Br^- and Cl^- anions. Reprinted with permission from Rigby *et al.*[25] Copyright 2014 American Chemical Society.

The DFT-D3 program (V3.1 Rev 0)[12] produced by the Grimme group was interfaced with the R statistical package[26] such that appropriate parameters were passed from R to the DFT-D3 program dynamically (s_6, s_8, $s_{r,6}$ and $s_{r,8}$ for zero damping, and s_6, s_8, a_1 and a_2 for BJ damping). A Nedler and Mead optimisation procedure[27] was then
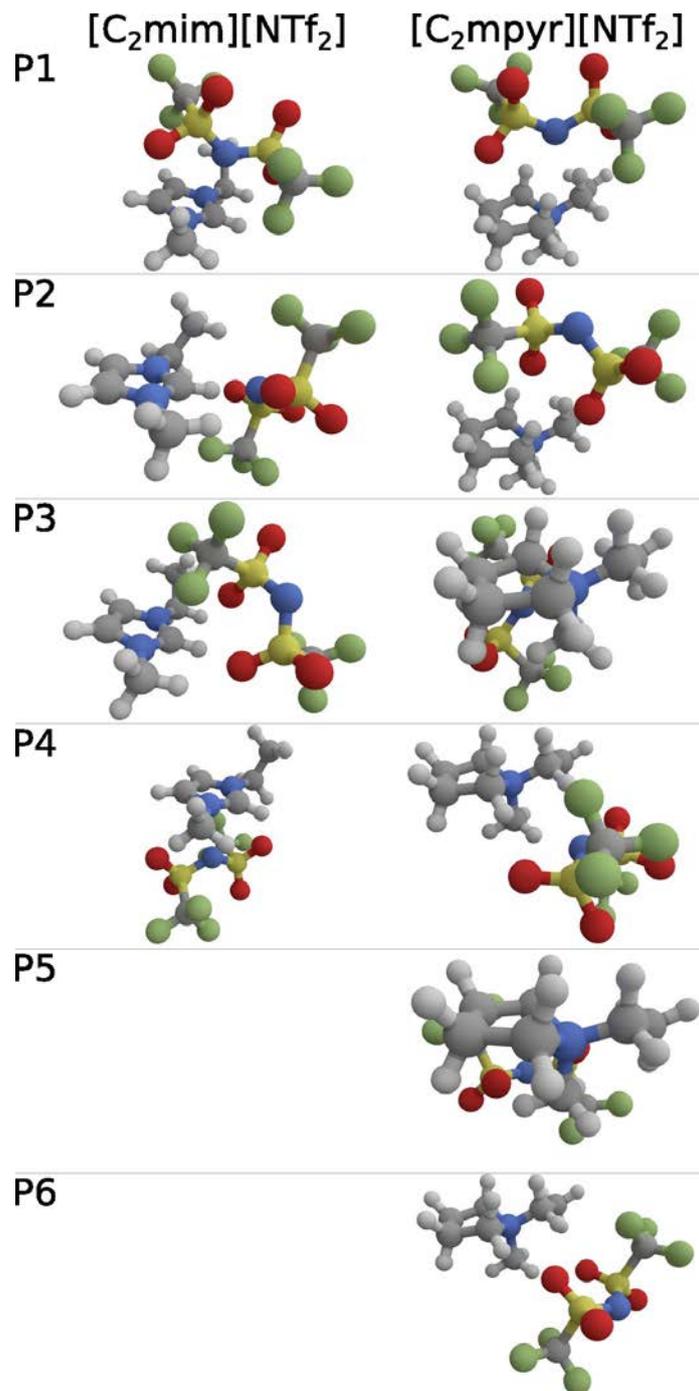
Figure 5.2: Examples of IL IP conformations for the [NTf$_2$]$^-$ anion. Reprinted with permission from Rigby *et al.*[25] Copyright 2014 American Chemical Society.

performed to minimise the sum of the squared residuals. Fitting was performed using the structures and CCSD(T)/CBS quality benchmark interaction energies previously published as part of the SCS-IL-MP2 study (see Chapter 4 and reference [25]). Parameters fitted for HF-D3 were fitted against the CCSD(T) correlation energy component only, whereas the BLYP and PBE functionals (both used in conjunction with the aug-cc-pVTZ basis set) were fitted against the CCSD(T)/CBS correlation energy added to

the HF/aug-cc-pVQZ interaction energy, which is considered to be at the CBS limit for the HF wavefunction.

For the zero damping function, parameters were fitted as:

(A) all unrestricted

(B) s_6 fixed at unity

(C) s_6 and $s_{r,8}$ fixed at unity (as in reference [12])

And for the BJ damping function:

(A) all unrestricted

(B) s_6 fixed at unity (as in reference [13])

The above letter codes are used throughout the remainder of this chapter to refer to the fitting conditions used for each set of coefficients proposed. Where the damping function used is not evident in the context of the discussion, it may be referred to in parentheses alongside these letter codes; e.g. BLYP-D3 (A,BJ).

In order to characterise the performance of the DFT-D3 corrections, zeroth-order symmetry-adapted perturbation theory (SAPT0)[28, 29] interaction energy decompositions were performed for the 174 IL IPs using the Psi4 software package[30] and the jun-cc-pVDZ basis set,[31] which has been recommended previously for SAPT0 calculations.[32]

5.2.2   Potential energy surfaces

PESs were generated for a selection of IPs consisting of the $[C_{(1,2)}mim]^+$ and $[C_{(1,2)}mpyr]^+$ cations, and all the previously mentioned anions in Section 5.2.1. Between one and five conformations of each IP, totalling 49 IL IPs, were scanned at displacements from −0.1 Å to +0.7 Å at 0.1 Å increments. Displacements were made along the vector defined by the atoms closest to the geometric midpoints of each ion, and were relaxed at each step using a counterpoise (CP) corrected gradient. Optimisations were performed with MP2/def2-TZVPP using the Turbomole 6.5 software package.[33] Single-point energies at CCSD(T)/CBS, spin-component scaled second-order Møller Plesset perturbation theory (SCS-MP2),[34] SCS-IL-MP2,[25] BLYP/aug-cc-pVTZ and PBE/aug-cc-pVTZ were calculated using the Psi4 software package.[30] CCSD(T)/CBS energies were calculated according to the method by Jurečka *et al.*,[35, 36] shown in eq. 5.7 where the MP2/CBS interaction energy[37] (see eq. 5.6) is added to the CCSD(T) correction as calculated

with a smaller basis set (i.e. aug-cc-pVDZ). For the two-point second-order Møller Plesset perturbation theory (MP2) extrapolation scheme given in eq. 5.6, X (3) and Y (4) correspond to the cardinal numbers of the basis sets X (aug-cc-pVTZ) and Y (aug-cc-pVQZ). The frozen natural orbital approximation was used in order to accelerate the CCSD(T) calculation with minimal loss in accuracy.[38]

$$\Delta E^{corr}_{MP2/CBS} = \frac{X^3 \Delta E^{corr}_{MP2/X} - Y^3 \Delta E^{corr}_{MP2/Y}}{X^3 - Y^3} \tag{5.6}$$

$$\Delta E^{corr}_{CCSD(T)/CBS} = \Delta E^{corr}_{MP2/CBS} + \underbrace{(\Delta E^{corr}_{CCSD(T)/ACCD} - \Delta E^{corr}_{MP2/ACCD})}_{\text{CCSD(T) correction}} \tag{5.7}$$

All interaction energies were corrected for basis set superposition error (BSSE) according to the Boys and Bernardi approach[39] as shown below in eqs. 5.8 – 5.10, where "All BFs" denotes calculations performed in the full basis set of the cluster, and index i denotes each monomer (ion) in the cluster.

$$\Delta E_{int} = E_{complex} - \sum E_i \tag{5.8}$$

$$\Delta E^{CP}_{int} = E_{complex} - \sum_i \dot{E}_i \bigg|_{\text{All BFs}} \tag{5.9}$$

$$\Delta E_{BSSE} = \Delta E^{CP}_{int} - \Delta E_{int} \tag{5.10}$$

5.2.3   Ionic liquid clusters

DFT-D3 corrections to the interaction energy for the HF wavefunction, and BLYP and PBE functionals were calculated for a series of two- and four-IP clusters created from IPs of the 1,3-dimethylimidazolium ([C$_1$mim]$^+$), *N,N*-dimethylpyrrolidinium ([C$_1$mpyr]$^+$), *N*-methyl-*N*-ethylpyrrolidinium ([C$_2$mpyr]$^+$) and *N*-ethylpyridinium ([C$_2$py]$^+$) cations in combination with the Br$^-$, Cl$^-$, [BF$_4$]$^-$ and [DCA]$^-$ anions. These ions represent some of the more commonly used cations, and a set of anions that encompass a number of distinctive characteristics including: (a) monoatomic (Cl$^-$ and Br$^-$), (b) larger spherical ([BF$_4$]$^-$), and (c) delocalised charge ([DCA]$^-$) anions. The cations included: (a) delocalised ([C$_1$mim]$^+$, [C$_2$py]$^+$) and (b) localised ([C$_1$mpyr]$^+$, [C$_2$mpyr]$^+$) charge. In particular, the [C$_2$py]$^+$ cation was not included in the initial fitting set and is therefore an indicator of the broader applicability of the fitted coefficients.

These IL clusters were previously used for the submitted paper concerning the performance of SCS-IL-MP2 for IL clusters provided in Chapter 4, Section 4.4 and their conformations are described in detail therein. In brief, the clusters were optimised using the fragment molecular orbital (FMO) approach truncated at the two-body term, in conjunction with the MP2 wavefunction and the cc-pVDZ basis set. In order to evaluate the performance of the DFT-D3 interaction energy correction, CP corrected DFT interaction energies were determined for the PBE and BLYP functional in conjunction with the aug-cc-pVTZ basis set. These DFT interaction energies were then added to their corresponding DFT-D3 correction, except for the HF-D3 correction, which was compared directly with the benchmark correlation interaction energies. Benchmark correlation interaction energies at an approximate CCSD(T)/CBS level of theory, previously determined in the submitted publication, were used again here.[*] Owing to the increased system size, benchmark CCSD(T)/CBS quality total interaction energies were calculated by adding CP corrected HF/cc-pVQZ interaction energies to the CCSD(T)/CBS correlation interaction energies. Here it is assumed that the small mean absolute error (MAE) between HF/cc-pVQZ and HF/aug-cc-pVQZ, found to be only 1.14 ± 0.90 kJ mol^{-1} for the IP fitting set, is reproduced in the IL clusters.

5.3   Results and discussion

5.3.1   Fitting set

Summarised in Tables 5.1 and 5.2 are the DFT-D3 coefficients considered in this study for HF wavefunction, PBE and BLYP DFs for the zero and BJ damping functions, respectively. The letters in parentheses indicate the fitting conditions, and these correspond with those listed in Section 5.2.1. In the majority of cases, the coefficients obtained appeared within a reasonable range, with the exception of the freely fitted HF-D3 in conjunction with the zero damping function. In this case, the s_8 coefficient was determined to be -119.7570, which appears anomalous when compared to all other cases. To this end, the "HF-D3 (A)" fitting conditions have been omitted from the analysis that follows.

5.3.1.1   Choice of functional and damping function

Figures 5.3a and 5.3b show the error distribution for both the zero and BJ damping functions. The box boundaries indicate the first and third quartiles, the line bisecting

---

[*]Benchmark energies are given in the supporting information in Section 4.5

DFT-D3 type	s_6	$s_{r,6}$	s_8	$s_{r,8}$
HF-D3 orig.	1.0000^F	1.1500	1.7460	1.0000^F
HF-D3 (A)	1.4698	0.6682	−119.7570	2.2507
HF-D3 (B)	1.0000^F	0.6862	0.4443	0.7062
HF-D3 (C)	1.0000^F	0.6067	0.7129	1.0000^F
PBE-D3 orig.	1.0000^F	1.2170	0.7220	1.0000^F
PBE-D3 (A)	3.6840	1.0844	−9.4507	1.2329
PBE-D3 (B)	1.0000^F	0.9858	−1.1022	1.4316
PBE-D3 (C)	1.0000^F	0.9987	−0.0555	1.0000^F
BLYP-D3 orig.	1.0000^F	1.0940	1.6820	1.0000^F
BLYP-D3 (A)	0.2540	0.4621	2.1983	0.9929
BLYP-D3 (B)	1.0000^F	0.9412	0.6393	0.8544
BLYP-D3 (C)	1.0000^F	0.8513	0.6092	1.0000^F

Table   Summary of DFT-D3 coefficients for the zero damping function. ^{F} indicates that this coefficient was fixed during the optimisation.

DFT-D3 type	s_6	s_8	a_1	a_2
HF-D3 orig.	1.0000^F	0.9171	0.3385	2.8830
HF-D3 (A)	0.6488	1.3599	−0.1459	4.8801
HF-D3 (B)	1.0000^F	0.8877	−0.5293	6.6430
PBE-D3 orig.	1.0000^F	0.7375	0.4289	4.4407
PBE-D3 (A)	−0.5092	4.1538	1.2479	0.4698
PBE-D3 (B)	1.0000^F	0.6778	1.0148	1.3323
BLYP-D3 orig.	1.0000^F	2.6996	0.4298	4.2359
BLYP-D3 (A)	0.1749	2.9878	0.6583	2.3956
BLYP-D3 (B)	1.0000^F	1.3139	0.4333	3.2632

Table 5.2: Summary of DFT-D3 coefficients for the BJ damping function. ^{F} indicates that this coefficient was fixed during the optimisation.

the box indicates the median value, and the whiskers indicate the extrema. These error distributions are expressed using standard statistical measures in Table 5.3 in terms of the MAE, maximum absolute error and standard deviation. Based on these data, it can be seen that for the original DFT-D3 parameters, the performance of the BJ damping function reflects the assertions by Grimme *et al.* that there is a marginal improvement in errors obtained using this damping function as compared with the zero damping function.[13] Interestingly, there is no clearly superior damping function when the respective coefficients are re-fitted, although the following observations can be made: (1) HF-D3 errors are generally smaller by all measures for the BJ damping function and does not result in any spurious coefficients with unrealistic values; (2) there is little distinction between the zero and BJ functions for the DFs in terms of the MAE, although the smallest errors are observed with the zero damping function (PBE-D3 (A)). These results reinforce the finding by Grimme *et al.* that the damping function form is not critical in obtaining accurate DFT-D3 interaction energies,[13] however after coefficient optimisation, the appropriate damping function is likely to be dependent on the DF

(a) Zero damping function

BJ damping function

(b)
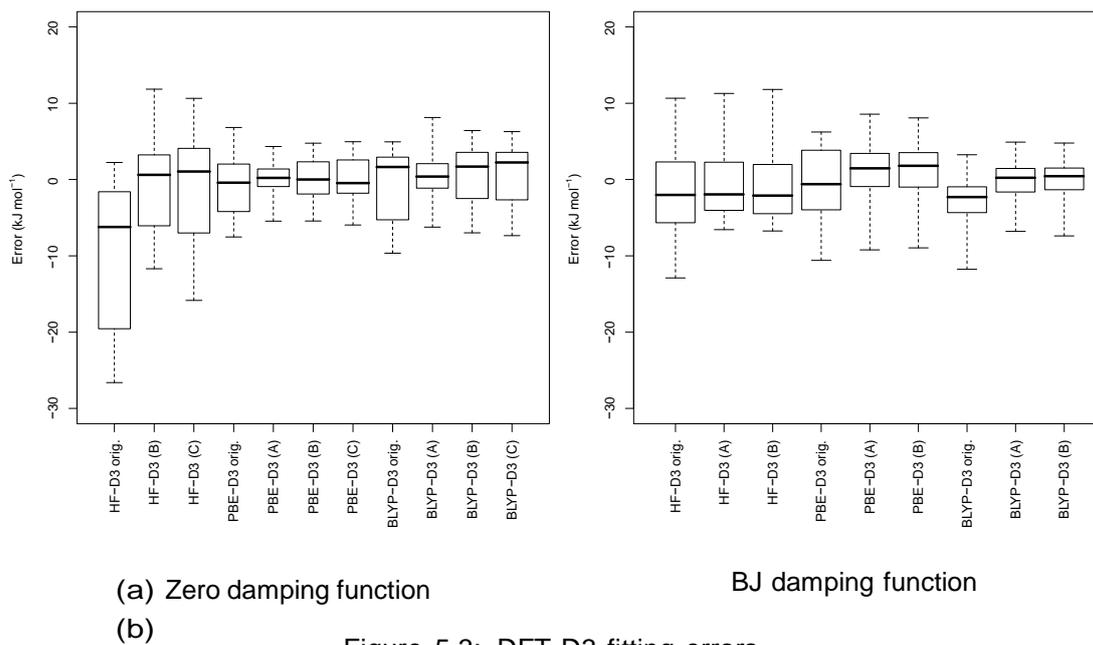
Figure 5.3: DFT-D3 fitting errors

used.

As the HF wavefunction is considered non-correlated, it can be used as a pure measure of the damping function suitability. The behaviour of DFT-D3 for HF interaction energies indicates that the BJ damping function indeed provides a more sound description of the short-ranged correlation energy, which is again in agreement with Grimme *et al*.[13] This is particularly evident given that the fitting procedure results in "well-behaving" coefficients whereas the zero damping function produced physically unrealistic values for s_8 (−119.7570). It is argued here, however, that a dispersion correction model that is more physically realistic should not necessarily be accepted over one that produces improved results, and that improved results do not necessarily follow from a more physically realistic model. Given that the DFT-D3 correction is necessarily correcting a physically unrealistic approximation for correlation effects – the DF itself – the DFT-D3 correction functions not only as an additional dispersion energy contribution, but also as a form of error cancellation that can counter any shortcomings of the DF. Error cancellations must therefore mirror the physical unreality of the flawed approximation and thus should not be restricted to one that adheres to the prescriptive definitions of the chemical interactions in question (e.g. that the dispersion energy should converge to a finite value at $r = 0$ [14]). Therefore, the exact form that the DFT-D3 correction should take is more importantly connected to the ability for it to produce accurate energetics and not how faithful the formulation is to a rigorous description of dispersion. Consequently, it should be reiterated that it would be naïve to suggest any particular damping function as the most superior for *all* DFs and chemical systems. However, it can be suggested here that the BJ damping function is most suitable for HF-D3 and BLYP-D3 corrections, while

damping function	DFT-D3 type	MAE	Max. Error	St. Dev.
	HF-D3 orig.	10.63	26.61	9.26
	HF-D3 (B)	5.31	11.83	3.51
	HF-D3 (C)	6.35	15.84	4.24
	PBE-D3 orig.	3.39	7.52	1.86
	PBE-D3 (A)	1.56	5.46	1.23
Zero	PBE-D3 (B)	2.26	5.43	1.34
	PBE-D3 (C)	2.42	5.96	1.38
	BLYP-D3 orig.	3.70	9.66	2.32
	BLYP-D3 (A)	2.28	8.12	1.95
	BLYP-D3 (B)	3.07	6.98	1.58
	BLYP-D3 (C)	3.27	7.36	1.40
	HF-D3 orig.	5.52	12.90	3.79
	HF-D3 (A)	3.99	11.29	2.56
	HF-D3 (B)	4.03	11.81	2.67
	PBE-D3 orig.	4.08	10.59	2.73
BJ	PBE-D3 (A)	3.31	9.22	2.38
	PBE-D3 (B)	3.52	8.96	2.27
	BLYP-D3 orig.	3.18	11.77	2.50
	BLYP-D3 (A)	1.82	6.79	1.49
	BLYP-D3 (B)	1.91	7.39	1.60

Table 5.3: MAE, maximum error and standard deviation for the original and fitted DFT-D3 parameters given in kJ mol^{-1}

the zero damping function is ideal for the PBE-D3 corrections. In particular, the zero damping function in combination with the PBE functional with coefficients optimised without restriction (PBE-D3 (A)) results in the smallest errors of 1.56 $\pm$ 1.23 kJ mol^{-1}.

5.3.1.2 Predictors of error

Multiple linear regression was performed in order to determine the most significant predictors of error for the DFT-D3 energies with respect to the absolute values of SAPT0 interaction energy components – electrostatic (ES), exchange (EX), induction (IND), induction-exchange (IND-EX), dispersion (DISP), and dispersion-exchange (DISP-EX) – as described by eq. 5.11. Correlation coefficients given in Table 5.4 indicate how the strength of particular components of the interaction energy might affect the quality of the DFT-D3 correction. The *p*-values given are a measure of significance based on the Student's *t* distribution, and may be interpreted as the probability that the relationship described by the coefficient occurred randomly (i.e. is not real); a *p*-value of 0.05 would therefore signify that there is a 95% chance that the given coefficient is a true description of the relationship. In the ideal case, results should indicate that the error is independent from the interaction energy composition (correlation coefficients close to zero), and any correlation that *is* found should be insignificant (*p*-values close to unity). While

SAPT0 is a low-order interaction energy decomposition method neglecting higher-order correlation contributions, it is used here as to generate an approximate "fingerprint" of the overall nature of the interaction. From this it was found that regardless of the damping function, the quality of the DFT-D3 correction was generally dictated by the strength of the exchange contribution to the dispersion energy (dispersion-exchange), which was most evident for the corrected HF wavefunction.

$$\Delta E_{\text{CCSD(T)/CBS}} - \Delta E_{\text{DFT-D3}} \approx \alpha |E_{\text{ES}}| + \beta |E_{\text{EX}}| + \gamma |E_{\text{IND}}| + \delta |E_{\text{IND-EX}}| + c |E_{\text{DISP}}| + \zeta |E_{\text{DISP-EX}}|$$

$$(5.11)$$

The correlation between the HF-D3 error and the magnitude of the dispersion-exchange contribution is indicative of a fundamental incompatibility when equating the power series expansion of the dispersion energy employed by the DFT-D3 approach with the overall correlation energy. By expressing the dispersion energy as in eq. 5.1, the role of the damping function becomes not only one of avoiding a singularity at $r = 0$ but also to implicitly account for exchange effects, which are short-ranged.[40, 41] It is unsurprising that the damping function cannot wholly account for short-ranged effects on dispersion interactions that are purely the result of electron correlation, however it highlights the need to account for these effects by other means, i.e. the DF. This does not imply that HF-D3 energies are without value; the overall error was only 3.99 ± 2.56 kJ mol^{-1} for the least restrictive fitting conditions and the BJ damping function (HF-D3 (B)). In this case errors originated primarily from Cl$^-$ and Br$^-$ anions, which are strongly coordinating anions[42] with significant orbital overlap resulting in higher exchange-dispersion contributions. HF-D3 may provide a means by which correlation energies are rapidly estimated, and could produce reliable geometries, although this remains to be seen.

The regression coefficients indicate no strong relationships exist between the SAPT0 energy decompositions and the DFT-D3 errors as applied to DFs (as opposed to the HF wavefunction). In particular, the BLYP-D3 energies fitted with the zero damping function show exceptionally weak (coefficient near zero) and insignificant (higher p-values) correlation between exchange-dispersion energies, and the PBE functional shows a significant yet weak negative correlation. This is consistent with the low errors observed for these functionals with the DFT-D3 correction.

damping function	DFT-D3 type	$\lvert E_{ES}\rvert$ coeff.	p	$\lvert E_{EX}\rvert$ coeff.	p	$\lvert E_{IND}\rvert$ coeff.	p	$\lvert E_{IND\text{-}EX}\rvert$ coeff.	p	$\lvert E_{DISP}\rvert$ coeff.	p	$\lvert E_{DISP\text{-}EX}\rvert$ coeff.	p
Zero	HF-D3 orig.	−0.0246	0.0095	0.4723	0.0000	−0.3337	0.0001	0.1815	0.0658	−0.3440	0.0001	1.2256	0.1289
	HF-D3 (B)	0.0003	0.9519	−0.1514	0.0002	0.1376	0.0007	−0.2328	0.0000	−0.0183	0.6505	2.1829	0.0000
	HF-D3 (C)	0.0038	0.4253	−0.1567	0.0002	0.1277	0.0024	−0.1934	0.0001	−0.0591	0.1588	2.3021	0.0000
	PBE-D3 orig.	−0.0062	0.0265	0.0744	0.0023	0.0110	0.6524	0.0283	0.3342	0.1117	0.0000	−1.0501	0.0000
	PBE-D3 (A)	0.0078	0.0000	0.0603	0.0002	−0.0849	0.0000	0.0765	0.0001	−0.0128	0.4290	−0.2637	0.0936
	PBE-D3 (B)	0.0077	0.0001	0.0679	0.0000	−0.0410	0.0141	0.0580	0.0037	0.0145	0.3872	−0.7935	0.0000
	PBE-D3 (C)	0.0081	0.0001	0.0493	0.0041	−0.0131	0.4471	0.0381	0.0661	0.0166	0.3404	−0.7920	0.0000
	BLYP-D3 orig.	0.0008	0.7893	0.1312	0.0000	−0.1632	0.0000	0.1059	0.0005	0.0123	0.6237	0.0523	0.8302
	BLYP-D3 (A)	0.0122	0.0002	−0.0121	0.6597	−0.0273	0.3287	0.0281	0.3995	−0.0006	0.9841	0.0269	0.9214
	BLYP-D3 (B)	0.0077	0.0007	0.0560	0.0039	−0.0775	0.0001	0.0594	0.0115	−0.0409	0.0385	0.1152	0.5459
	BLYP-D3 (C)	0.0092	0.0000	0.0350	0.0599	−0.0565	0.0031	0.0472	0.0370	−0.0067	0.7247	−0.0260	0.8877
BJ	HF-D3 orig.	−0.0097	0.0063	−0.0931	0.0024	0.1522	0.0000	−0.2629	0.0000	−0.0950	0.0025	2.3450	0.0000
	HF-D3 (A)	−0.0012	0.7010	−0.1123	0.0001	0.1135	0.0001	−0.2640	0.0000	−0.0340	0.2300	2.1692	0.0000
	HF-D3 (B)	0.0043	0.2342	−0.1165	0.0003	0.0970	0.0027	−0.2269	0.0000	−0.0681	0.0355	2.1324	0.0000
	PBE-D3 orig.	0.0062	0.1093	0.1141	0.0008	−0.1471	0.0000	0.0927	0.0232	−0.0463	0.1774	0.2478	0.4559
	PBE-D3 (A)	0.0093	0.0130	0.0513	0.1117	−0.0595	0.0698	−0.0085	0.8280	−0.0807	0.0151	0.5007	0.1184
	PBE-D3 (B)	0.0106	0.0035	0.0459	0.1403	−0.0582	0.0665	0.0134	0.7214	−0.0387	0.2253	0.1846	0.5497
	BLYP-D3 orig.	0.0174	0.0000	0.1131	0.0005	−0.1606	0.0000	0.1078	0.0056	−0.1129	0.0006	0.0778	0.8051
	BLYP-D3 (A)	0.0128	0.0000	0.0343	0.0858	−0.0605	0.0032	0.0755	0.0021	−0.0110	0.5903	−0.4261	0.0325
	BLYP-D3 (B)	0.0162	0.0000	0.0275	0.1877	−0.0681	0.0015	0.0887	0.0006	−0.0187	0.3809	−0.4308	0.0386

Table 5.4: Regression coefficients and *p*-values using the absolute values of the SAPT0 energy decomposition as predictors of the DFT-D3 error.

5.3.2 Potential energy surfaces

The PESs for a total of 49 IL IPs were evaluated at nine points of 0.1 Å increments at a displacement of −0.1 Å to +0.7 Å from the equilibrium distance, optimised at each step using a CP corrected gradient. The errors resulting from the original and fitted DFT-D3 parameters, calculated with respect to a CCSD(T)/CBS benchmark, were determined at each point. Appendix B, Section B.2 presents detailed histograms to describe these error distributions. The MAEs for each PES generated are shown in Figure 5.6 as a function of displacement from the equilibrium geometry, given in kJ mol^{-1}.

By far the most accurate method of all presented is SCS-IL-MP2. Here, the MAE observed ranged from 0.8 to 1.3 kJ mol^{-1} with a standard deviation consistently below 1 kJ mol^{-1}. Furthermore, Figure B.23 in Appendix B shows the error distribution histogram as extremely narrow and without outlying values. This strongly substantiates claims made in the submitted publication, Section 4.4, showing the suitability of SCS-IL-MP2 for IL clusters and consequently highlights the versatility of this approach. Amongst the dispersion corrected methods, BLYP-D3 (A) with the zero damping function yields the lowest errors of the dispersion corrected DFs, ranging from 1.2 to 1.6 kJ mol^{-1} and standard deviations of between 0.8 and 1.4 kJ mol^{-1}. For the best-fitting DF identified in Section 5.3.1.1, PBE-D3 (A) with the zero damping function, errors ranged from 2.0 to 2.3 kJ mol^{-1} with standard deviations ranging from 1.6 to 1.7 kJ mol^{-1}. Given the difference between these two dispersion corrected DFs is at or below 1 kJ mol^{-1}, it would be premature to favour one over the other by this measure.
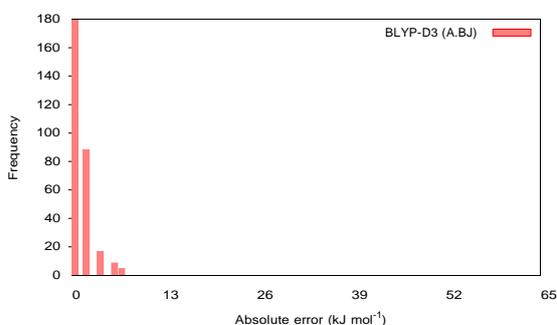
Figure 5.4: BLYP-D3 (A) BJ damping function potential energy surface error histogram
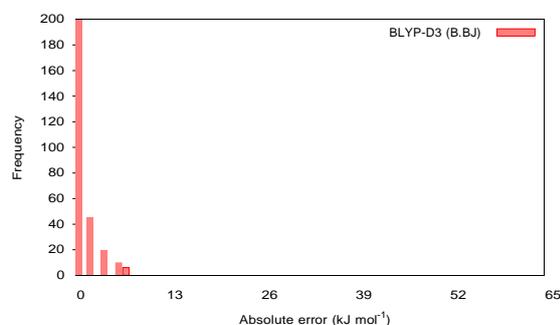
Figure 5.5: BLYP-D3 (B) BJ damping function potential energy surface error histogram

The BJ damping function with the $s_6 = 1$ fitting restriction applied (denoted by the (B) suffix) appears to produce marginally improved PESs compared to the freely fitted case (denoted by the (A) suffix). This effect is most distinct when comparing the error distributions for BLYP-D3 (A) and (B) shown in Figures 5.4 and 5.5, which include the residuals of all 49 ILs at each of the nine points of the PES. The distributions show a notable shift towards lower errors with s_6 fixed at unity, indicating that the error minimisation performed for the fitting set (see Section 5.2.1) may apply only locally

at the equilibrium structures used for fitting and is not suitable for other geometries, resulting in an increased overall error. For the two DFs corrected using the zero damping function, the most accurate values were obtained with no restriction in fitting. Based on this criterion there appears to be no compelling reason to restrict any coefficents for this damping function except for the HF wavefunction where the s_8 coefficient did not optimise to a realistic value.


(a) MP2 methods

(b) Uncorrected methods

(c) HF-D3

(d) PBE-D3

(e) BLYP-D3
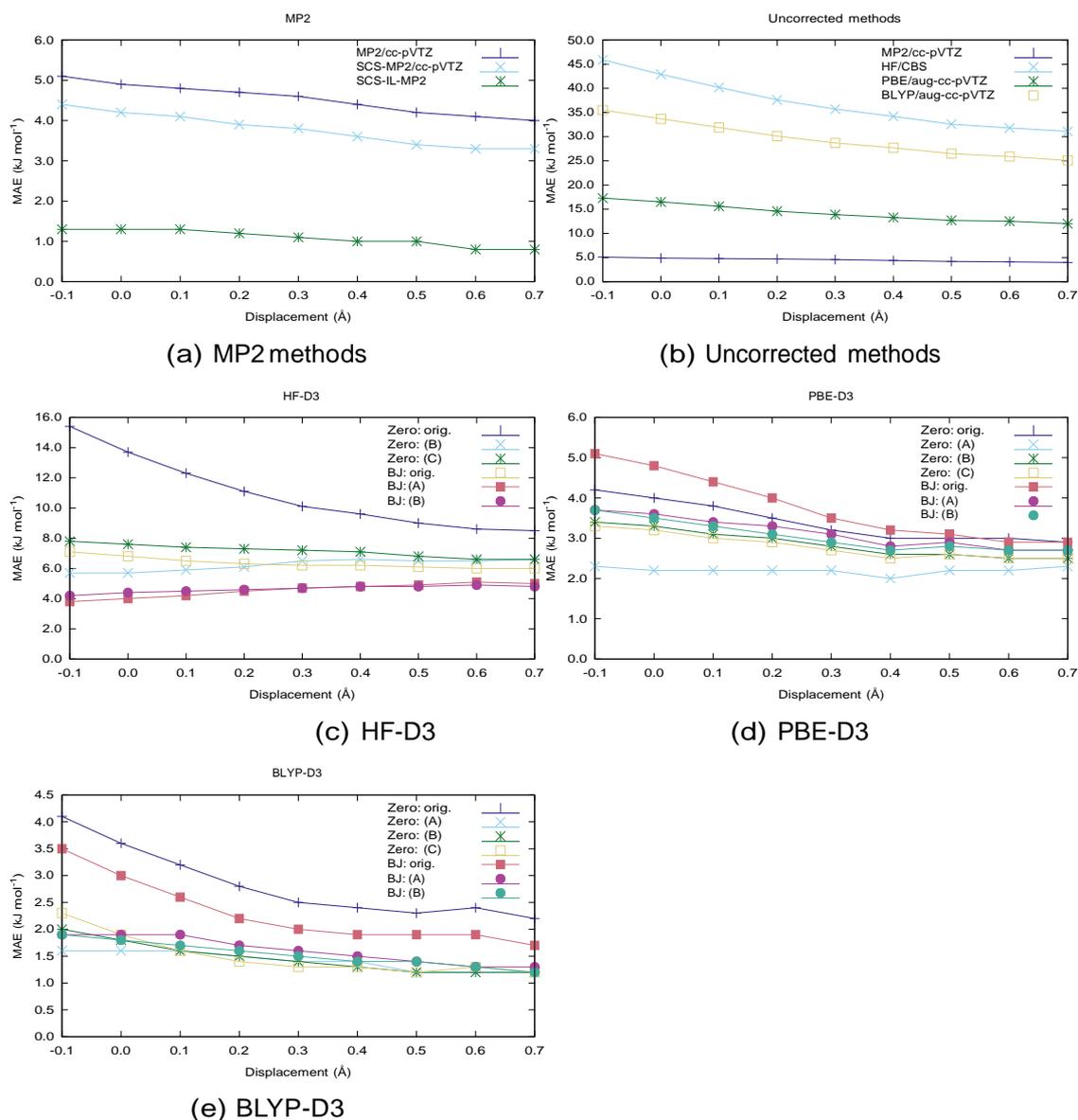
Figure 5.6: Errors over a 0.8 Å PES

It can be seen in Figure 5.6 and Table 5.5 that in the significant majority of cases (16 of 20), a reduction in error is observed as the displacement is increased; only the refitted HF-D3 parameters (zero damping: (C), BJ damping: (A) and (B)) exhibited slight increase in error with distance. It is worth noting that the PBE-D3 (A) PES exhibits an error gradient of zero, indicating that the error is, on average, less sensitive to changes

damping function	DFT-D3 type	error grad. (kJ mol^{-1} Å^{-1})
	HF-D3 orig.	−8.5
	HF-D3 (B)	1.2
	HF-D3 (C)	−1.6
	PBE-D3 orig.	−1.7
	PBE-D3 (A)	0.0
Zero	PBE-D3 (B)	−1.2
	PBE-D3 (C)	−1.1
	BLYP-D3 orig.	−2.2
	BLYP-D3 (A)	−0.7
	BLYP-D3 (B)	−1.0
	BLYP-D3 (C)	−1.1
	HF-D3 orig.	−1.3
	HF-D3 (A)	1.6
	HF-D3 (B)	0.7
	PBE-D3 orig.	−3.0
BJ	PBE-D3 (A)	−1.3
	PBE-D3 (B)	−1.3
	BLYP-D3 orig.	−2.0
	BLYP-D3 (A)	−0.9
	BLYP-D3 (B)	−0.8

Table 5.5: Gradient of the error assuming a linear fit along the PES

in geometry. Consequently it is possible that the PBE functional may be better suited to AIMD as the PES generated includes mainly unbiased systematic error that could potentially produce superior trajectories although this aspect should be the subject of future studies.

5.3.3 Application to clusters

Tables 5.6 and 5.7 (pp. 94 – 95) show the interaction energy errors with respect to CCSD(T)/CBS with all errors mentioned in this section given per ion pair. The DFT-D3 approaches were analysed in terms of their MAE, as well as the proportion of ILs that yield an error less than 1, 5 and 10 kJ mol^{-1}, expressed as a percentage. In so doing, not only the performance of the DFT-D3 correction on average can be evaluated, but the distribution of the errors may too be understood.

5.3.3.1 Performance of the zero damping function

The zero damping function yielded the smallest MAEs for the PBE-D3 (B) fitting conditions for both the two and four IPs of 2.5 and 3.4 kJ mol^{-1}, respectively, with 100% of the clusters under 10 kJ mol^{-1}. Interestingly, while PBE-D3 (C) produced a marginally

elevated MAE (2.6 and 3.8 kJ mol^{-1} for two and four IPs, respectively), the distribution of errors favoured lower errors suggesting that in this case outliers bias the MAE. PBE-D3 (C) produces the same proportion of sub-kilojoule errors as PBE-D3 (B) for two IPs (31%), however an increase is observed for the four IP clusters of 7% (from 14% to 21%). For errors less than 5 kJ mol^{-1}, the proportion of two IP clusters increased by 6% from PBE-D3 (B) (75%) to (C) (81%) and was unchanged for the four IP clusters at 79%. Surprisingly, the original PBE-D3 result produced results with errors below PBE-D3 (A), which was identified as the best-fitting case for the initial fitting set in Section 5.2.1. HF represents the most improvement after fitting with increased proportion of IL clusters with errors under 10 kJ mol^{-1} from just 25% to 69% for two IPs and from 21% to 57% for four IPs for HF-D3 (B).

5.3.3.2   Performance of the Becke-Johnson damping function

The performance of the BJ damping function for the HF wavefunction showed a substantial improvement in all cases over the zero damping function; Grimme's original HF-D3 parameters resulted in a decrease in MAE from 18.7 to 6.9 kJ mol^{-1} and 17.8 to 6.8kJ mol^{-1} for the two- and four-IP clusters, respectively. After fitting, HF-D3 (B) showed remarkably low errors of just 4.5 and 3.3 kJ mol^{-1} for the two- and four-IP clusters, with all errors less than 10 kJ mol^{-1}. The best performing case for the BJ damping functional was the original PBE parameters, with errors of approximately 3 kJ mol^{-1}, although there was one outlier for the two-IP clusters of 11 kJ mol^{-1}. The BLYP-D3 (B) parameters, despite a marginal increase (only 0.5 kJ mol^{-1}) in MAE for two-IP clusters, provide a subtly improved description of the energetics; no clusters show errors greater than 10 kJ mol^{-1}, and the proportion of four-IP clusters with errors less than 1 kJ mol^{-1} increases from 14 to 50%.

	HF orig.		HF(B)		HF(C)		PBE orig.		PBE(A)		PBE(B)		PBE(C)		BLYP orig.		BLYP(A)		BLYP(B)		BLYP(C)	
	2IP	4IP	2IP	4IP	2IP	4IP	2IP	4IP	2IP	4IP	2IP	4IP	2IP	4IP	2IP	4IP	2IP	4IP	2IP	4IP	2IP	4IP
[C$_1$mim][BF$_4$]	4.3	0.2	4.1	7.0	4.2	7.8	6.8	2.9	2.9	2.2	5.3	1.6	4.7	0.7	3.8	0.4	5.5	2.8	4.9	1.7	4.9	1.6
[C$_1$mpyr][BF$_4$]	2.3	4.5	8.7	11.8	8.5	13.6	3.7	8.0	9.8	18.3	5.6	9.3	6.2	10.8	6.1	12.7	5.0	7.3	6.9	10.6	7.3	11.0
[C$_2$mpyr][BF$_4$]	0.8	0.1	10.8	12.0	10.8	12.1	3.8	5.0	10.5	13.0	5.4	5.7	6.3	7.3	6.9	8.2	5.1	4.1	7.5	8.1	8.0	8.6
[C$_2$py][BF$_4$]	5.9	–	0.1	–	1.2	–	5.2	–	0.4	–	3.3	–	2.9	–	1.3	–	2.4	–	2.7	–	2.6	–
[C$_1$mim]Br	22.5	22.5	11.9	12.6	14.3	13.7	4.3	4.4	8.6	5.5	4.7	4.0	4.7	3.7	9.5	7.7	7.2	7.1	9.8	9.2	11.1	10.3
[C$_1$mpyr]Br	24.4	21.8	11.7	10.2	13.7	11.5	0.9	4.1	2.3	1.5	1.1	3.8	1.2	4.3	4.8	1.4	1.4	0.3	2.9	0.7	3.6	1.5
[C$_2$mpyr]Br	24.6	21.6	11.2	8.9	13.4	10.4	0.6	3.7	4.0	0.7	0.0	2.8	0.4	3.7	5.4	1.8	2.6	1.0	3.1	1.0	3.8	1.7
[C$_2$py]Br	26.3	27.6	10.3	13.1	12.8	14.2	1.2	0.8	0.8	0.3	2.6	0.4	2.5	0.9	3.2	3.1	0.6	2.5	1.6	4.1	2.5	4.9
[C$_1$mim]Cl	22.1	21.4	6.5	6.9	9.5	8.5	5.9	5.7	5.8	1.4	5.2	3.9	5.1	3.6	12.5	10.1	10.3	9.9	11.4	9.9	12.0	10.4
[C$_1$mpyr]Cl	25.0	22.9	7.6	5.5	10.5	7.8	1.1	1.8	0.1	6.9	0.0	2.9	0.1	3.4	8.3	5.0	5.0	3.9	4.0	1.9	4.0	1.9
[C$_2$mpyr]Cl	25.2	20.3	5.1	1.0	8.4	3.5	1.8	2.5	0.1	7.6	0.5	3.6	0.2	4.4	8.4	3.5	5.8	3.3	4.5	0.8	4.4	0.6
[C$_2$py]Cl	26.7	25.1	6.0	5.6	9.1	7.6	1.6	0.2	2.1	5.4	1.1	2.3	0.8	2.6	7.2	4.6	3.3	3.2	4.0	2.7	4.2	2.6
[C$_1$mim][DCA]	22.3	23.7	6.8	10.0	7.4	9.2	3.5	5.0	5.1	7.4	0.4	1.6	0.3	0.3	3.5	2.5	6.6	9.7	3.8	5.2	3.0	4.2
[C$_1$mpyr][DCA]	20.0	22.0	1.9	0.8	3.4	1.8	0.6	2.9	6.3	11.5	1.1	0.6	2.2	1.1	2.0	2.4	4.8	8.1	0.1	0.9	1.3	0.8
[C$_2$mpyr][DCA]	21.2	15.7	0.3	3.6	1.6	3.2	2.9	1.8	8.7	16.8	0.7	5.1	1.3	6.3	2.5	2.9	3.7	1.6	0.5	4.4	1.8	5.8
[C$_2$py][DCA]	25.8	–	8.5	–	8.9	–	7.2	–	2.6	–	3.3	–	3.0	–	5.8	–	6.4	–	4.6	–	3.7	–
MAE	18.7	17.8	7.0	7.8	8.6	8.9	3.2	3.5	4.4	7.0	2.5	3.4	2.6	3.8	5.7	4.7	4.7	4.6	4.5	4.4	4.9	4.7
% E < 1 kJ mol^{-1}	6	14	13	7	0	0	19	14	25	14	31	14	31	21	0	7	6	14	13	21	0	14
% E < 5 kJ mol^{-1}	19	21	25	21	25	21	75	86	56	36	75	79	81	79	44	71	50	64	75	64	75	64
% E < 10 kJ mol^{-1}	25	21	69	57	63	57	100	100	94	71	100	100	100	93	94	86	94	100	94	93	88	79

Table 5.6: Absolute errors for DFT-D3 energies applied to clusters of two and four IPs using the zero damping function given per IP in kJ mol^{-1}. The MAE and percentage of ILs with errors less than 1, 5 and 10 kJ mol^{-1} are also given.

	HF orig.		HF(A)		HF(B)		PBE orig.		PBE(A)		PBE(B)		BLYP orig.		BLYP(A)		BLYP(B)	
	2IP	4IP	2IP	4IP	2IP	4IP	2IP	4IP	2IP	4IP	2IP	4IP	2IP	4IP	2IP	4IP	2IP	4IP
$[C_1mim][BF_4]$	0.9	4.4	1.6	1.2	1.6	1.7	11.0	7.6	6.7	3.9	6.3	2.5	14.0	10.7	9.7	7.4	9.3	6.2
$[C_1mpyr][BF_4]$	3.3	7.9	1.3	0.8	2.3	1.0	1.0	1.0	4.6	5.6	5.3	9.3	5.6	3.3	0.9	0.5	0.3	2.2
$[C_2mpyr][BF_4]$	5.5	6.4	0.3	0.1	0.6	0.3	1.2	1.2	4.5	3.6	5.7	6.9	5.1	5.3	0.7	2.0	0.4	0.5
$[C_2py][BF_4]$	1.2	–	3.3	–	2.9	–	8.9	–	5.4	–	4.7	–	11.1	–	7.2	–	6.8	–
$[C_1mim]Br$	11.6	10.6	3.5	4.1	3.5	3.2	3.7	3.9	8.6	8.7	7.0	6.2	8.4	8.0	8.3	8.1	7.0	6.4
$[C_1mpyr]Br$	11.5	9.2	5.5	3.9	5.4	3.2	1.0	3.6	1.2	0.2	0.1	2.9	5.6	3.1	2.9	1.6	2.4	0.2
$[C_2mpyr]Br$	11.2	8.2	5.5	3.6	5.2	2.6	0.7	3.2	2.0	0.5	0.2	2.9	5.5	2.9	3.5	2.2	2.5	0.2
$[C_2py]Br$	9.6	11.1	1.7	5.2	1.1	3.5	2.6	0.1	0.5	2.8	1.4	0.2	2.2	3.3	0.2	2.8	1.2	0.6
$[C_1mim]Cl$	9.1	7.4	4.0	3.9	3.7	2.6	5.1	5.1	6.5	5.9	5.4	4.0	11.6	10.9	9.4	8.8	8.7	7.5
$[C_1mpyr]Cl$	10.0	7.2	7.1	5.3	7.2	4.5	0.8	1.5	0.3	1.4	1.0	4.3	9.1	7.0	4.5	3.7	4.4	2.1
$[C_2mpyr]Cl$	8.5	3.7	6.1	3.0	6.1	2.1	1.5	2.0	0.3	3.3	1.2	6.0	9.3	5.8	4.7	2.4	4.3	0.7
$[C_2py]Cl$	7.8	6.3	2.9	3.7	2.4	2.5	0.1	1.0	1.2	2.0	1.5	3.6	6.4	4.8	1.6	1.2	1.4	0.0
$[C_1mim][DCA]$	6.9	8.0	10.3	13.3	8.8	10.0	1.7	3.3	3.0	0.7	3.5	3.2	2.9	2.2	0.8	2.6	0.3	0.1
$[C_1mpyr][DCA]$	3.5	2.3	7.8	9.6	6.9	7.7	0.0	2.6	5.1	2.5	6.7	6.7	2.2	3.8	0.6	1.9	2.3	1.6
$[C_2mpyr][DCA]$	1.9	2.9	6.3	3.3	5.7	1.4	1.7	2.5	4.7	7.7	5.8	11.1	3.3	1.2	1.3	3.6	2.5	6.6
$[C_2py][DCA]$	8.2	–	10.7	–	9.1	–	4.7	–	0.1	–	0.3	–	4.4	–	1.1	–	0.4	–
MAE	6.9	6.8	4.9	4.4	4.5	3.3	2.9	2.8	3.4	3.5	3.5	5.0	6.7	5.2	3.6	3.5	3.4	2.5
% E < 1 kJ mol^{-1}	6	0	6	14	6	14	31	14	25	21	19	7	0	0	31	7	25	50
% E < 5 kJ mol^{-1}	31	29	50	71	50	86	81	86	69	71	56	57	31	57	75	79	75	71
% E < 10 kJ mol^{-1}	81	86	88	93	100	100	94	100	100	100	100	93	81	86	100	100	100	100

Table 5.7: Absolute errors for DFT-D3 energies applied to clusters of two and four IPs using the BJ damping function given per IP in kJ mol^{-1}. The MAE and percentage of ILs with errors less than 1, 5 and 10 kJ mol^{-1} are also given.

5.4 Conclusions

This chapter has examined the refinement of the DFT-D3 parameters for the HF wavefunction and the PBE and BLYP density functionals in order to reproduce CCSD(T)/CBS interaction energies of ILs from IPs and clusters of up to four IPs. This has involved both freely fitting, and fitting with restrictions in accordance with the original methodology, the s_6 and s_8 coefficients as well as relevant coefficients for the zero[12] ($s_{r,6}$ and $s_{r,8}$) and BJ damping functions (a_1 and a_2).[13] The fitted coefficients were then tested for their ability to reproduce PESs and ion clusters.

In general, the differences between the fitted DFT-D3 parameters were insignificant, varying within only 1 or 2 kJ mol^{-1} for most measures. Of particular note, however, was the significant improvement when fitting the HF-D3 correction. The original parameters for the HF wavefunction gave unacceptably high errors up to 27 kJ mol^{-1} for the zero damping function, and 13 kJ mol^{-1} for the BJ damping function per ion pair. After fitting with the BJ damping function, HF-D3 (B) produced interaction energies with errors less than 5 kJ mol^{-1} for ion pairs, as well as clusters on a per ion pair basis. This result, while not as accurate as the fitted DFT functionals, is quite remarkable and may be a worthwhile approach to rapidly calculate properties such as interaction energies, interaction energy decompositions or geometries.

The margin of error is low for each fitted set of DFT-D3 parameters and there remains no clearly superior DFT-D3 functional/coefficient combination that performs comparatively well. That is, the fitted dispersion correction is able to compensate for the deficiencies present in the DFs. This notwithstanding, there are still noteworthy observations that have been made. While it is logical that allowing the maximum number of coefficients to vary during coefficient optimisation inevitably results in the smallest errors, transferring these fitted parameters does not necessarily correspond to similarly reduced errors when applied to the PES or clusters. It was found that for the zero damping function, freely fitted parameters prevailed with the most accurate PESs, however for the BJ damping function, it was essential to fix s_6 to unity. It should be noted that fixing s_6 to unity does not necessarily mean that unity is the "correct" value; from this the only claim that can be made with confidence is that s_6 should be excluded from the fitting procedure. Interestingly, although the DFT-D3 parameters that produced the lowest errors was BLYP-D3 (B,BJ), the PBE-D3 (A,zero) produced only a very subtle increase in MAE and an error gradient of zero; that is, an error insensitive to changes in geometry. Therefore it would be reasonable to assert that the increase in error for the PBE-D3 (A,zero) approach is simply a fluctuation resulting from a less diverse set of ILs used for the PESs. The small differences in error mean that it cannot be said with confidence which of these two corrected DFs do indeed produce the superior result.

When considering IL clusters, the best performing DFT-D3 approach was PBE-D3 (B) with the zero damping function; MAEs of 2.5 and 3.5 kJ mol^{-1} per ion pair for two and

four IP systems and in no cases did the errors exceed 10 kJ mol^{-1}. For the BJ damping function, the original PBE-D3 parameters gave the lowest MAE for the two IP case, however there was one system in excess of 10 kJ mol^{-1}. BLYP-D3 (B), on the other hand, shows only a slight increase of the two IP MAE and the lowest four IP MAE with all of the errors less than 10 kJ mol^{-1}. On these grounds it is argued the best performing DFT-D3 approach for the BJ damping function when applied to clusters.

The question of which approach is the most reliable based on these data is evidently not straightforward and may indeed be of no consequence given the relatively small differences between them. Nevertheless, it is possible to make some recommendations, and these are detailed as follows. First and foremost, it is clear that the SCS-IL-MP2 approach is predictable and consistent in all aspects outlined in this chapter and therefore is preferred over all DFT-D3 approaches. In terms of the DFT-D3 options presented here, the BLYP-D3 (B) approach in conjunction with the BJ damping function is a reasonable compromise, in view of accurately treating large ionic clusters. This DFT-D3 parameter set provides errors of 1.91 $\pm$ 1.60 kJ mol^{-1} (compared with 1.56 $\pm$ 1.23 kJ mol^{-1} for PBE-D3 (A,zero)), is almost indistinguishable in terms of the MAEs (see Figure 5.6e) and produces cluster energies where all systems investigated were below 10 kJ mol^{-1} of which over 70% of clusters were below 5 kJ mol^{-1}.

A description of the changes to the original DFT-D3 source code is provided as a Unix diff-style patch in Appendix C.

5.5   Acknowledgements

References

[1] W. Kohn and L. J. Sham. Self-consistent equations including exchange and correlation effects. *Physical Review*, 140:A1133–A1138, 1965.

[2] P. Hohenberg and W. Kohn. Inhomogeneous electron gas. *Physical Review*, 136:B864–B871, 1964.

[3] Chengteh Lee, Weitao Yang, and Robert G. Parr. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Physical Review B*, 37:785–789, 1988.

[4] Stephan N. Steinmann, Cyril Piemontesi, Aurore Delachat, and Clemence Corminboeuf. Why are the interaction energies of charge-transfer complexes challenging for DFT? *Journal of Chemical Theory and Computation*, 8(5):1629–1640, 2012.

[5] Ekaterina I. Izgorodina, Uditha L. Bernard, and Douglas R. MacFarlane. Ion-pair binding energies of ionic liquids: Can DFT compete with ab initio-based methods? *The Journal of Physical Chemistry A*, 113(25):7064–7072, 2009. PMID: 19462960.

[6] Axel D. Becke. Density-functional thermochemistry. v. systematic optimization of exchange-correlation functionals. *The Journal of Chemical Physics*, 107(20):8554–8560, 1997.

[7] Fred A. Hamprecht, Aron J. Cohen, David J. Tozer, and Nicholas C. Handy. Development and assessment of new exchange-correlation functionals. *The Journal of Chemical Physics*, 109(15):6264–6271, 1998.

[8] Axel D. Becke. Perspective: Fifty years of density-functional theory in chemical physics. *The Journal of Chemical Physics*, 140(18), 2014.

[9] T. A. Wesolowski, O. Parisel, Y. Ellinger, and J. Weber. Comparative study of benzene$\cdots$X (X = O_2, N_2, CO) complexes using density functional theory: the importance of an accurate exchangecorrelation energy density at high reduced density gradients. *The Journal of Physical Chemistry A*, 101(42):7818–7825, 1997.

[10] Stefan Grimme. Accurate description of van der waals complexes by density functional theory including empirical corrections. *Journal of Computational Chemistry*, 25(12):1463–1473, 2004.

[11] Stefan Grimme. Semiempirical GGA-type density functional constructed with a long-range dispersion correction. *Journal of Computational Chemistry*, 27(15):1787–1799, 2006.

[12] Stefan Grimme, Jens Antony, Stephan Ehrlich, and Helge Krieg. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *The Journal of Chemical Physics*, 132(15), 2010.

[13] Stefan Grimme, Stephan Ehrlich, and Lars Goerigk. Effect of the damping function in dispersion corrected density functional theory. *Journal of Computational Chemistry*, 32(7):1456–1465, 2011.

[14] Erin R. Johnson and Axel D. Becke. A post-Hartree-Fock model of intermolecular interactions: Inclusion of higher-order corrections. *The Journal of Chemical Physics*, 124(17), 2006.

[15] Jan Řezáč, Kevin E. Riley, and Pavel Hobza. S66: A well-balanced database of benchmark interaction energies relevant to biomolecular structures. *Journal of Chemical Theory and Computation*, 7(8):2427–2438, 2011.

[16] John P. Perdew, Kieron Burke, and Matthias Ernzerhof. Generalized gradient approximation made simple. *Physical Review Letters*, 77:3865–3868, 1996.

[17] John P. Perdew, Kieron Burke, and Matthias Ernzerhof. Generalized gradient approximation made simple [physical review lett. 77, 3865 (1996)]. *Physical Review Letters*, 78:1396–1396, 1997.

[18] Burkhard Miehlich, Andreas Savin, Hermann Stoll, and Heinzwerner Preuss. Results obtained with the correlation energy density functionals of Becke and Lee, Yang and Parr. *Chemical Physics Letters*, 157(3):200 – 206, 1989.

[19] A. D. Becke. Density-functional exchange-energy approximation with correct asymptotic behavior. *Physical Review A*, 38:3098–3100, 1988.

[20] Martin Thomas, Martin Brehm, Oldamur Hollóczki, Zsolt Kelemen, László Nyulászi, Tibor Pasinszki, and Barbara Kirchner. Simulating the vibrational spectra of ionic liquid systems: 1-ethyl-3-methylimidazolium acetate and its mixtures. *The Journal of Chemical Physics*, 141(2), 2014.

[21] Dzmitry S. Firaha and Barbara Kirchner. CO_2 absorption in the protic ionic liquid ethylammonium nitrate. *Journal of Chemical & Engineering Data*, Article ASAP, 2014.

[22] Rajdeep Singh Payal and Sundaram Balasubramanian. Dissolution of cellulose in ionic liquids: an ab initio molecular dynamics simulation study. *Physical Chemistry Chemical Physics*, 16:17458–17465, 2014.

[23] P. Ganesh, De-en Jiang, and P. R. C. Kent. Accurate static and dynamic properties of liquid electrolytes for Li-ion batteries from ab initio molecular dynamics. *The Journal of Physical Chemistry B*, 115(12):3085–3090, 2011. PMID: 21384941.

[24] Alfonso S. Pensado, Martin Brehm, Jens Thar, Ari P. Seitsonen, and Barbara Kirchner. Effect of dispersion on the structure and dynamics of the ionic liquid 1-ethyl-3-methylimidazolium thiocyanate. *ChemPhysChem*, 13(7):1845–1853, 2012.

[25] Jason Rigby and Ekaterina I. Izgorodina. New SCS- and SOS-MP2 coefficients fitted to semi-coulombic systems. *Journal of Chemical Theory and Computation*, 10(8):3111–3122, 2014.

[26] R Development Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2011. ISBN 3-900051-07-0.

[27] J. A. Nelder and R. Mead. A simplex method for function minimization. *The Computer Journal*, 7(4):308–313, 1965.

[28] Bogumil Jeziorski, Robert Moszynski, and Krzysztof Szalewicz. Perturbation theory approach to intermolecular potential energy surfaces of van der Waals complexes. *Chemical Reviews*, 94(7):1887–1930, 1994.

[29] Edward G. Hohenstein and C. David Sherrill. Density fitting and cholesky decomposition approximations in symmetry-adapted perturbation theory: Implementation and application to probe the nature of π-π interactions in linear acenes. *The Journal of Chemical Physics*, 132(18), 2010.

[30] Justin M. Turney, Andrew C. Simmonett, Robert M. Parrish, Edward G. Hohenstein, Francesco A. Evangelista, Justin T. Fermann, Benjamin J. Mintz, Lori A. Burns, Jeremiah J. Wilke, Micah L. Abrams, Nicholas J. Russ, Matthew L. Leininger, Curtis L. Janssen, Edward T. Seidl, Wesley D. Allen, Henry F. Schaefer, Rollin A. King, Edward F. Valeev, C. David Sherrill, and T. Daniel Crawford. Psi4: an open-source ab initio electronic structure program. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 2(4):556–565, 2012.

[31] Ewa Papajak, Jingjing Zheng, Xuefei Xu, Hannah R. Leverentz, and Donald G. Truhlar. Perspectives on basis sets beautiful: Seasonal plantings of diffuse basis functions. *Journal of Chemical Theory and Computation*, 7(10):3027–3034, 2011.

[32] Edward G. Hohenstein and C. David Sherrill. Wavefunction methods for non-covalent interactions. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 2(2):304–326, 2012.

[33] TURBOMOLE V6.5 2013, a development of University of Karlsruhe and Forschungszentrum Karlsruhe GmbH, 1989-2007, TURBOMOLE GmbH, since 2007; available from
http://www.turbomole.com.

[34] Stefan Grimme. Improved second-order Møller–Plesset perturbation theory by separate scaling of parallel- and antiparallel-spin pair correlation energies. *The Journal of Chemical Physics*, 118(20):9095–9102, 2003.

[35] Petr Jurečka and Pavel Hobza. On the convergence of the (δECCSD(T)δEMP2) term for complexes with multiple H-bonds. *Chemical Physics Letters*, 365(1–2):89–94, 2002.

[36] Petr Jurečka, Jiri Sponer, Jiri Cerny, and Pavel Hobza. Benchmark database of accurate (MP2 and CCSD(T) complete basis set limit) interaction energies of small model complexes, DNA base pairs, and amino acid pairs. *Physical Chemistry Chemical Physics*, 8:1985–1993, 2006.

[37] T. Helgaker, J. Olsen, and P. Jørgensen. *Molecular Electronic-Structure Theory*. Wiley: New York, 2000.

[38] A. Eugene DePrince and C. David Sherrill. Accurate noncovalent interaction energies using truncated basis sets based on frozen natural orbitals. *Journal of Chemical Theory and Computation*, 9(1):293–299, 2013.

[39] S.F. Boys and F. Bernardi. The calculation of small molecular interactions by the differences of separate total energies. some procedures with reduced errors. *Molecular Physics*, 19(4):553–566, 1970.

[40] A.J. Stone. *The Theory of Intermolecular Forces*. International Series of Monographs on Chemistry. Clarendon Press, 1997.

[41] Matthew P. Hodges and Anthony J. Stone. A new representation of the dispersion interaction. *Molecular Physics*, 98(5):275–286, 2000.

[42] Kevin J. Fraser, Ekaterina I. Izgorodina, Maria Forsyth, Janet L. Scott, and Douglas R. MacFarlane. Liquids intermediate between "molecular" and "ionic" liquids: Liquid ion pairs? *Chemical Communications*, pages 3817–3819, 2007.

Chapter 6

Conclusions and future work

6.1 Conclusions

This thesis has aimed to justify, and develop methodology, for the accurate calculation of intermolecular energetics of ionic liquids (ILs) on a large scale using methods based on *ab initio* quantum chemical techniques. This was achieved in the following ways:

Identifying shortcomings of classical large-scale approaches
Chapter 2 sought to challenge the assumptions made when characterising the electrostatic interaction as applied to molecular dynamics (MD) simulations. A number of atomic partial charge schemes in widespread use were selected and tested against a set of desirable characteristics: that (a) charges converge with increasing basis set size; (b) the charges are invariant with changes to the coordinate system; (c) minor structural changes do not unduly affect the resulting charges; (d) charge transfer effects present in ILs are adequately captured; and (e) the symmetry of charges is preserved in symmetric molecules. It was found that the most reliable atomic partial charge scheme was the "geodesic" scheme belonging to the restrained electrostatic potentials (RESP) family, which are often not adequately explained in the literature. In addition, dipole moments were derived from these partial charge schemes that gave weight to the importance of using polarisable classical models, or indeed *ab initio* quantum mechanical models, since ILs are highly polarisable. This necessitates some flexibility of the atomic partial charges such that they may adapt to the local environment of the individual ions during the course of the simulation, which conventional MD simulations do not allow. While it was found that charges can vary dramatically depending on the scheme used, the careful use of atomic partial charge schemes may still produce reliable forcefields, or at least serve as a rapid diagnostic tool to quantify electrostatic interactions and charge transfer.

Demonstrating the need for large-scale *ab initio* calculations
The linear-scaling fragment molecular orbital (FMO) approach in conjunction with the

second-order Møller Plesset perturbation theory (MP2) level of theory was used in Chapter 3 for a series of ILs arranged in clusters of increasing size. Here it was shown that the many-body effects can be substantial with the overall interaction energy on a per ion pair basis increasing rapidly as the cluster size is increased. It has been shown that at least three-body effects are necessary to reproduce a full-wavefunction MP2 energy using the FMO approach, and only at eight ion pairs (IPs) does the dispersion component of the interaction energy begin to plateau on a per ion pair basis. Of particular interest is the proportion of this energy arising from dispersion forces, which was found to only be around 8% for a single ion pair, but as much as 20% for an eight ion pair cluster. As the dispersion force is an effect originating purely from electron correlation, this chapter serves to portray *ab initio* quantum mechanical approaches as an essential tool in deriving *a priori* intermolecular energetics where the dispersion component of the interaction energy is poorly understood.

Developing a quantum chemistry toolbox for semi-Coulombic chemical systems

Chapters 3, 4 and 5 have developed a set of methodologies to pave the way for large-scale IL calculations using quantum chemical approaches. In addition to establishing the need for large-scale calculations as described above, Chapter 3 quantifies the reliability of the FMO framework for semi-Coulombic IL systems where it was shown that a conventional correlated electronic structure theory, MP2, could be reproduced with sub-kJ mol^{-1} accuracy by truncating at the three-body term of the many-body expansion. As the FMO framework is designed to leverage highly parallel computing architectures, boasting near-linear scalability with respect to the number of central processing units (CPUs), this represents a tool that is central to the application of correlated levels of theory to large chemical systems. These are normally limited by the "exponential wall" that is a consequence of poor scalability with respect to system size, upwards of scaling as the seventh power for CCSD(T). With this, only counterpoise (CP) correction remains non-trivial as this method of basis set superposition error (BSSE) correction is not amenable to fragmented techniques.

Chapter 4 tailors the widely used spin-component scaled second-order Møller Plesset perturbation theory (SCS-MP2) in order to reproduce CCSD(T)/CBS interaction energies of ILs by refitting the spin component scaling coefficients to a representative set of ILs. This has resulted in the new SCS-IL-MP2 approach. The SCS-IL-MP2 approach is unique in that an explicit design goal was to implicitly account for BSSE, thus avoiding the need for CP correction that was identified as a key bottleneck. This was achieved by both fitting to CP corrected benchmark interaction energies and identifying the ideal basis set that produced the best fitting coefficients; the mean absolute error (MAE) for the cc-pVTZ basis set was less than 1 kJ mol^{-1}. With the use of density-fitting, not only do MP2 energies at the cc-pVTZ basis set represent a trivial case for small- to medium-sized systems, but in conjunction with the FMO approach, the bottleneck is

shifted to one of CPU count rather than the "exponential wall," disk or memory requirements. The performance of SCS-IL-MP2 was shown to retain its accuracy beyond the ion pair in Section 4.4 where performance was at least equivalent to MP2/cc-pVQZ if not better.

Chapter 5 investigated an alternative approach to reproducing CCSD(T)/CBS energies by utilising type empirical dispersion corrections. In this chapter, the original formulation was compared against parameters refitted against the same set as in Chapter 4 designed to reproduce benchmark energies under a number of different fitting conditions. These were determined for the PBE and BLYP functionals, as well as for the Hartree-Fock (HF) wavefunction. The two density functionals (DFs) were selected as they are in widespread use when performing *ab initio* molecular dynamics (AIMD) simulations, particularly for ILs. The performance of these dispersion corrections were analysed by a number of measures: firstly, the ability to produce a reduced MAE, maximum error and standard deviation for the fitting set was assessed. Then, a subset of ILs were scanned over a potential energy surface (PES) with the interionic distance ranging from -0.1 Å to $+0.7$ Å relative to the equilibrium geometry in view of identifying the DFT-D version 3 (DFT-D3) parameters that provide the most accurate surface. Finally, these parameters were tested against clusters of two and four IPs in size, with the best performing parameters identified in terms of their MAE and error distribution. It was found that in most cases, the selection of the DF and associated DFT-D3 parameters was often of little consequence with all reaching within 1 to 2 kJ mol^{-1} per IP, however refitted HF-D3 parameters showed a substantial improvement, particularly when used with the Becke-Johnson (BJ) damping function. Refitted HF-D3 and the BJ damping function was able to consistently provide interaction energy errors below 5 kJ mol^{-1} per IP. It should be noted that both HF and DFT levels of theory are easily performed within the FMO framework and easily applied to large chemical systems.

6.2 Future work

By providing a set of rigorously tested approaches and refinements to accepted and widely used formalisms, namely the FMO, spin component scaled, and DFT-D3 approaches, one is well positioned to further investigate the physicochemical properties of ILs. Therefore a major aspect of future work involves implementing periodic boundary conditions for the FMO approach and then performing AIMD simulations in order to (a) determine suitable simulation box sizes, and (b) ultimately reproduce bulk properties of ILs. It is worthwhile investigating how accurately properties may be reproduced by means of SCS-IL-MP2 and DFT-D3 approaches, as well as further investigation into conventional and polarisable classical MD approaches.

Insofar as the dispersion corrected DFT approaches are concerned, the ability for these to produce accurate geometries – particularly in the case of HF-D3 – should be examined as a means to rapidly generate either valid, or reasonable starting guess structures for further   investigation.

Appendix A

Chemistry in the cloud

A.1   Background

The term "cloud computing" has been defined by the National Institute of Standards and Technology, U.S. Department of Commerce as "...a model for enabling ubitquitous, convenient, on-demand network access to a shared pool of configurable computing resources that can be rapidly provisioned and released with minimal management effort or service provider interaction."[1] More specifically, cloud computing services are typically categorised as one of either "Software as a Service," "Platform as a Service," or "Infrastructure as a Service." These are abbreviated as SaaS, PaaS and IaaS, respectively. The differences in these models is essentially one of user control. SaaS providers offer computational resources configured to run only designated software applications, PaaS providers allow user-designed software to be provisioned but restrict access to the underlying operating system, and finally IaaS providers require that the user configure and manage the operating system and installed applications on virtual hardware. This Appendix will address exclusively IaaS services.

A.2   The NeCTAR research cloud

The National eResearch Collaboration Tools and Resources (NeCTAR) research cloud[2] is a government funded IaaS provider, providing federated compute resources established in eight locations, totalling approximately 30,000 cores at the time of writing. Cloud resources are provisioned and managed by OpenStack software.[3] As the NeCTAR resources (and OpenStack service providers in general) operate under the IaaS model, the allocation and usage of compute resources is made in terms of instances and cores. An instance is a self-contained "machine" with its own virtual hardware that is equivalent to an individual computer, and the number of cores is no different to the number of cores

in a physical computer. The term "virtual" here is used because the physical machines used to power the cloud services may emulate several individual machines, but this is invisible to the end user.

Given that the user of an IaaS service is responsible for the configuration of the instances at an operating system level, leveraging these distributed resources is challenging. Whereas traditional compute clusters tend to offer a preconfigured software stack and compute nodes that are tightly integrated, a cloud provider simply makes available hardware platforms upon which user-configured operating systems and software may run – a significant disincentive for those unfamiliar with how to replicate a traditional cluster and manage a queuing system such as Torque[4] or Slurm.[5]

A.3   A cloud-friendly queueing system

In order to simplify and automate the use of cloud resources to assist in undertaking research included in this thesis, a new cloud resource manager was developed known as "openstack-queue." Designed using the Apache jClouds toolkit,[6] openstack-queue is able to automatically instantiate cloud instances, deploy software, initiate computation, collect results and terminate instances on-demand. In this way, cloud resources are made available to other users of the cloud infrastructure while no jobs are running, and the method of operation is very similar to Torque in that commands such as qsub and qstat may be loosely emulated.

Openstack-queue is available as an open source software package, including full source code documentation, on Github here: https://github.com/jasonrig/openstack-queue

References

[1] Peter Mell and Timothy Grance. The NIST definition of cloud computing special publication 800-145. http://csrc.nist.gov/publications/nistpubs/800-145/SP800-145.pdf, 2011.

[2] National eResearch Collaboration Tools and Resources research cloud. http://nectar.org.au/research-cloud.

[3] OpenStack: The open source cloud operating system. http://www.openstack.org/.

[4] Torque resource manager. http://www.adaptivecomputing.com/products/open-source/torque/.

[5] SLURM: A highly scalable resource manager. https://computing.llnl.gov/linux/slurm/.

[6] Apache jclouds. https://jclouds.apache.org/.

Appendix B

DFT-D3 error histograms

B.1 Fitting set histograms

B.1.1 Zero damping function
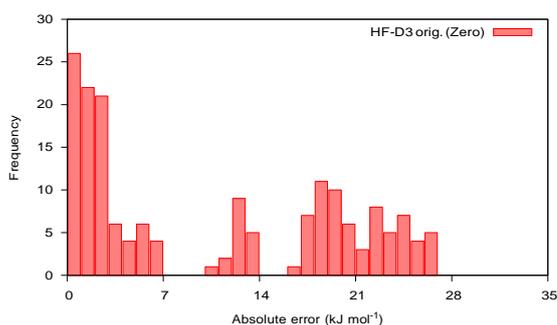

Figure B.1: HF-D3 orig. zero damping function fitting set error histogram
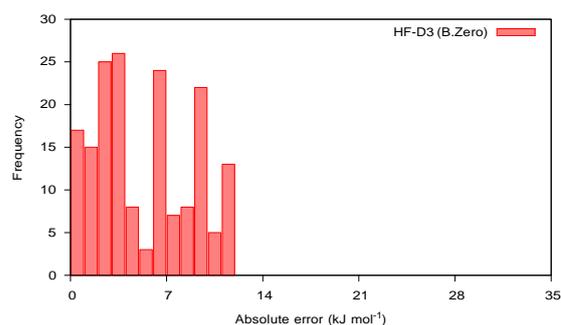
Figure B.2: HF-D3 (B) zero damping function fitting set error histogram
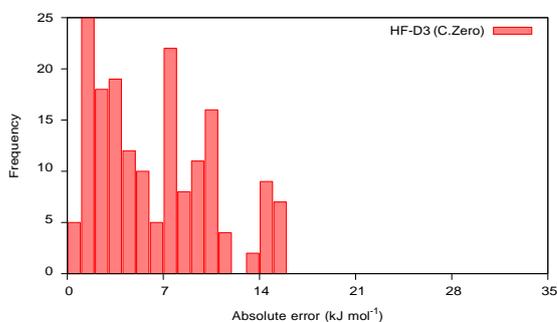
Figure B.3: HF-D3 (C) zero damping function fitting set error histogram
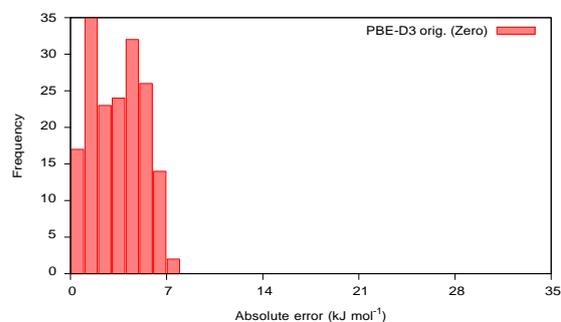
Figure B.4: PBE-D3 orig. zero damping function fitting set error histogram
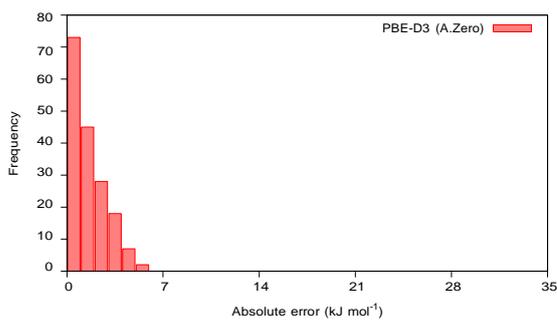
Figure B.5: PBE-D3 (A) zero damping function fitting set error histogram
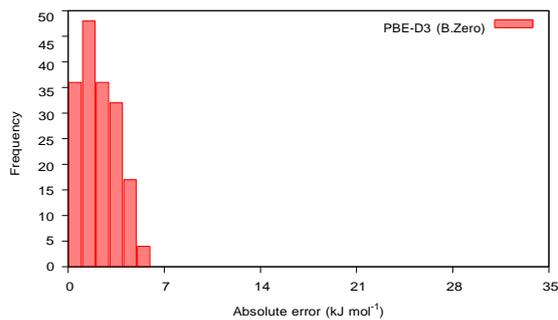
Figure B.6: PBE-D3 (B) zero damping function fitting set error histogram
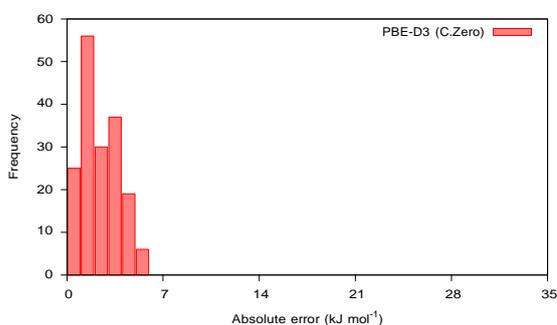
Figure B.7: PBE-D3 (C) zero damping function fitting set error histogram
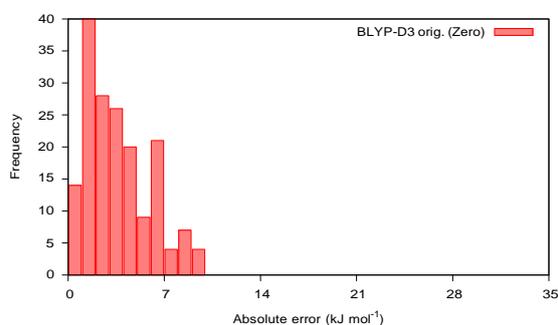
Figure B.8: BLYP-D3 orig. zero damping function fitting set error histogram
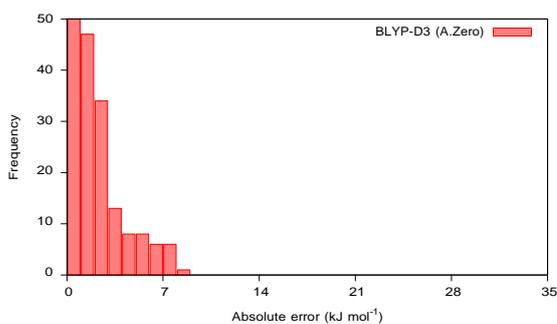
Figure B.9: BLYP-D3 (A) zero damping function fitting set error histogram
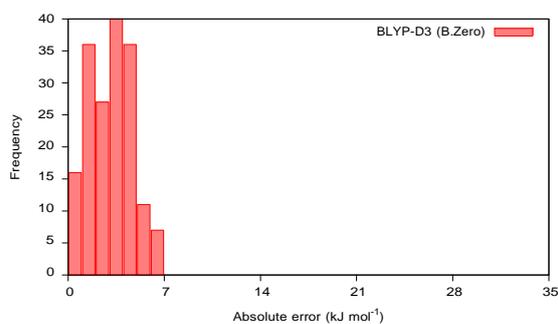
Figure B.10: BLYP-D3 (B) zero damping function fitting set error histogram
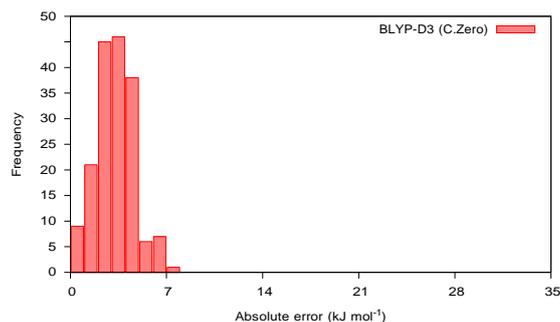
Figure B.11: BLYP-D3 (C) zero damping function fitting set error histogram

B.1.2   Becke-Johnson damping function
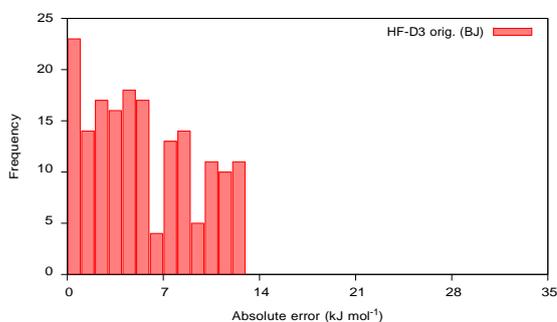
Figure B.12: HF-D3 orig. BJ damping function fitting set error histogram
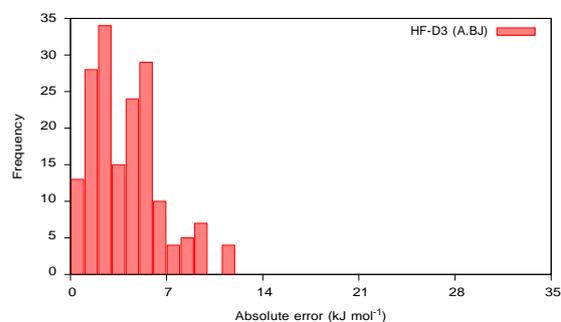
Figure B.13: HF-D3 (A) BJ damping function fitting set error histogram
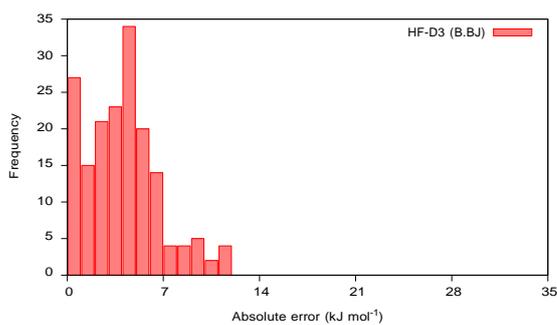
Figure B.14: HF-D3 (B) BJ damping function fitting set error histogram
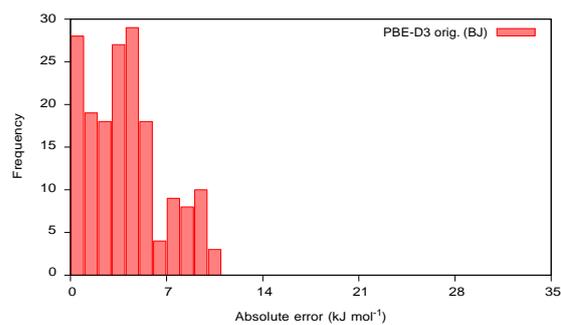
Figure B.15: PBE-D3 orig. BJ damping function fitting set error histogram
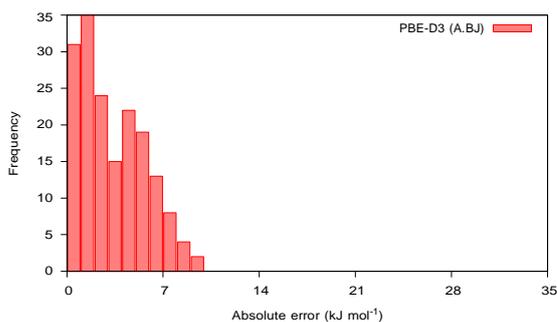
Figure B.16: PBE-D3 (A) BJ damping function fitting set error histogram
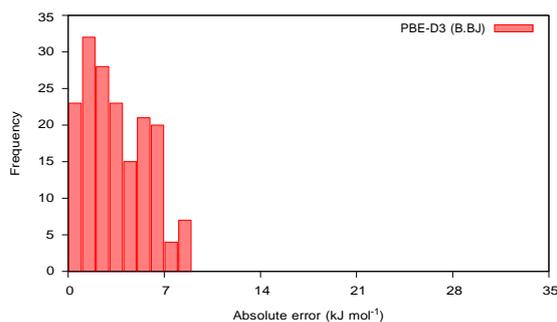
Figure B.17: PBE-D3 (B) BJ damping function fitting set error histogram
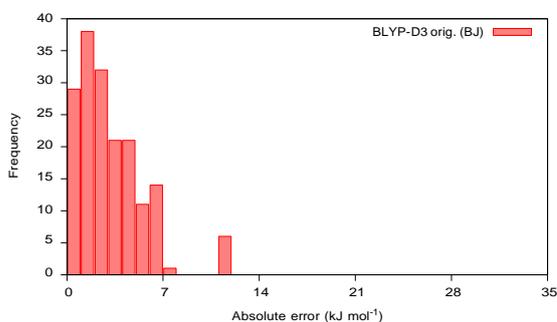
Figure B.18: BLYP-D3 orig. BJ damping function fitting set error histogram
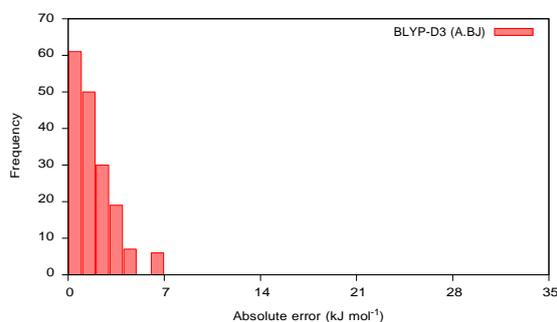
Figure B.19: BLYP-D3 (A) BJ damping function fitting set error histogram
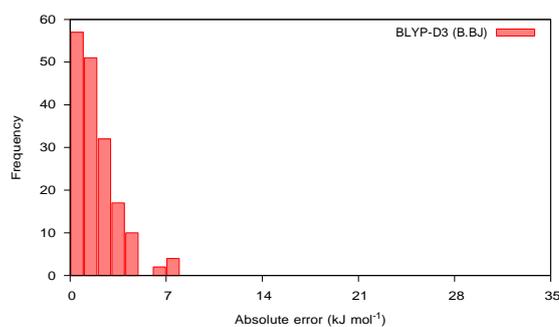
Figure B.20: BLYP-D3 (B) BJ damping function fitting set error histogram

B.2   Potential energy surface histograms

The following error histograms for the potential energy surfaces discussed in Chapter 5 include errors for each point of each ionic liquid (IL), totalling 441 single point energies (49 ILs at nine points along the surface). Therefore these histograms serve to evaluate the goodness of fit overall and not the performance of the DFT-D3 correction as a

function of distance; that is, the distinction between performance at a particular distance versus performance for particular ion pair or subset of ion pairs cannot be made.

B.2.1 MP2 approaches
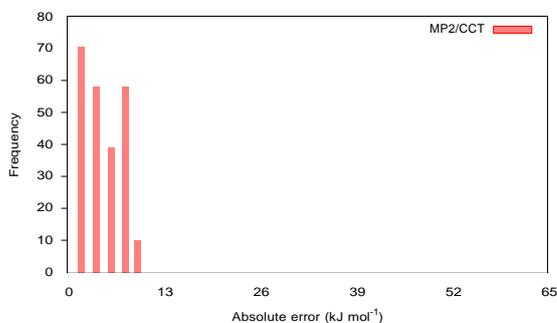
Figure B.21: MP2/cc-pVTZ potential energy surface error histogram
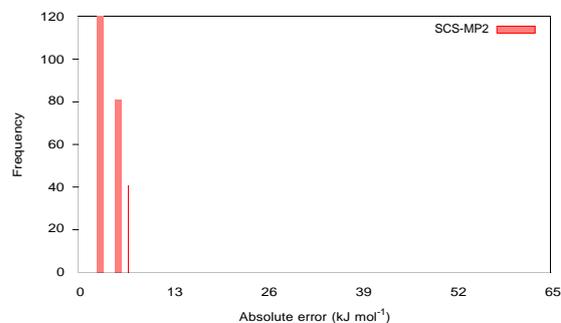
Figure B.22: SCS-MP2/cc-pVTZ potential energy surface error histogram
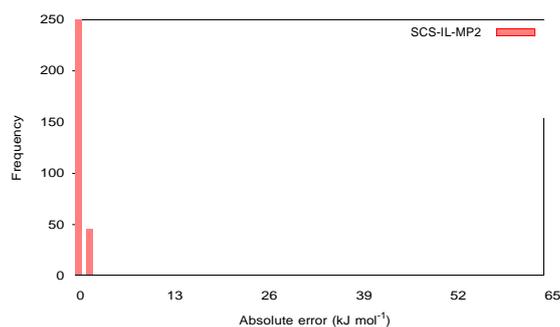
Figure B.23: SCS-IL-MP2 potential energy surface error histogram

B.2.2 Uncorrected DFT functionals
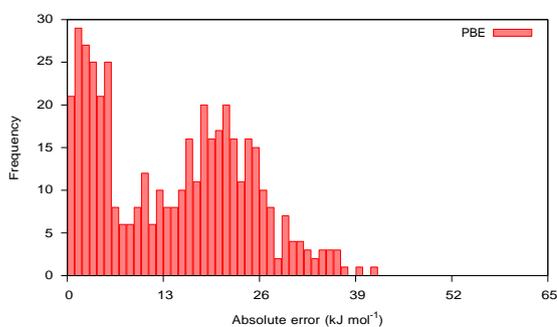
Figure B.24: PBE potential energy surface error histogram
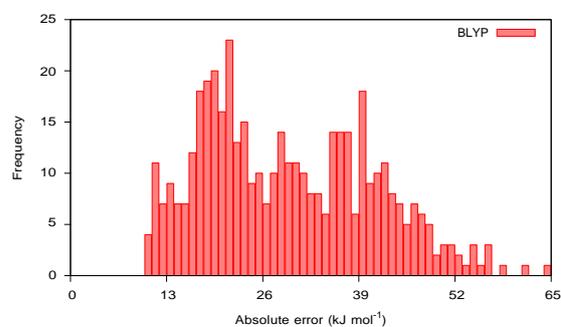
Figure B.25: BLYP potential energy surface error histogram

B.2.3 Corrected DFT functionals and HF wavefunction

B.2.3.1 Zero damping function
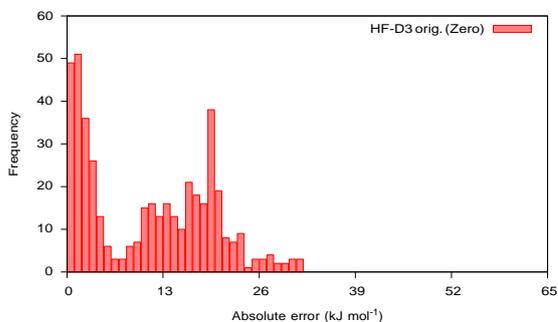
Figure B.26: HF-D3 orig. zero damping function potential energy surface error histogram
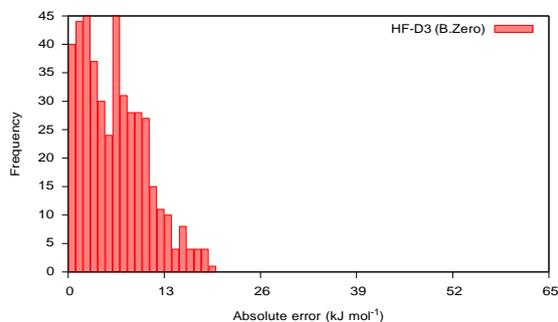
Figure B.27: HF-D3 (B) zero damping function potential energy surface error histogram
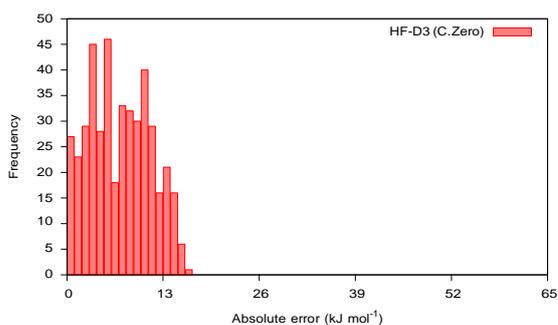
Figure B.28: HF-D3 (C) zero damping function potential energy surface error histogram
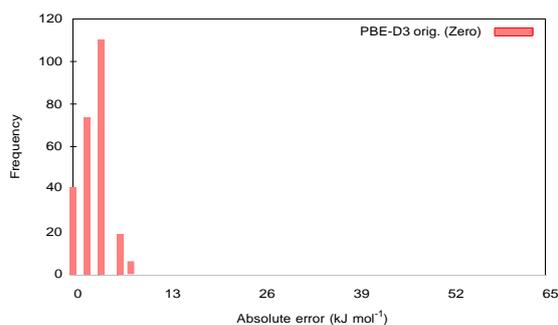
Figure B.29: PBE-D3 orig. zero damping function potential energy surface error histogram
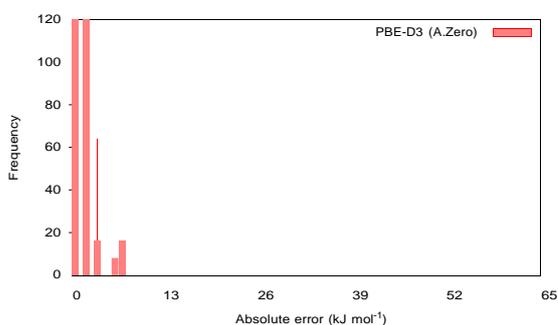
Figure B.30: PBE-D3 (A) zero damping function potential energy surface error histogram
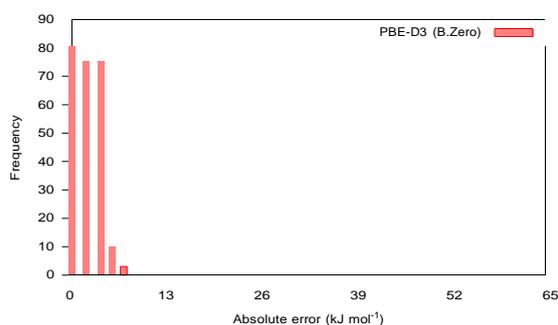
Figure B.31: PBE-D3 (B) zero damping function potential energy surface error histogram
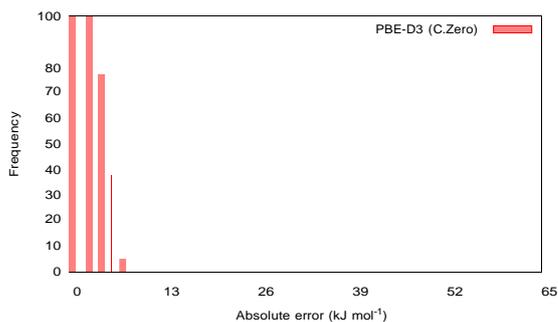
Figure B.32: PBE-D3 (C) zero damping function potential energy surface error histogram
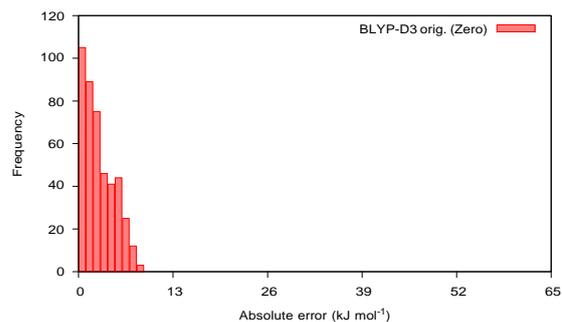
Figure B.33: BLYP-D3 orig. zero damping function potential energy surface error histogram
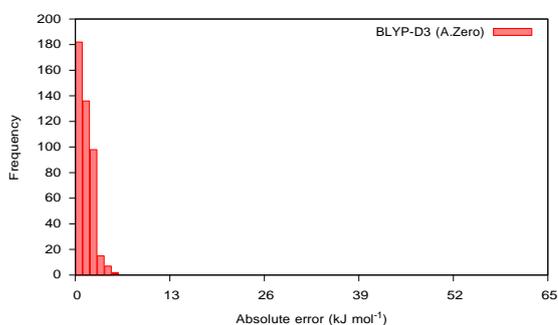
Figure B.34: BLYP-D3 (A) zero damping function potential energy surface error histogram
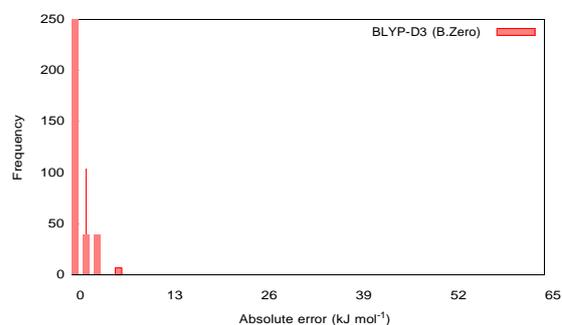

Figure B.35: BLYP-D3 (B) zero damping function potential energy surface error histogram
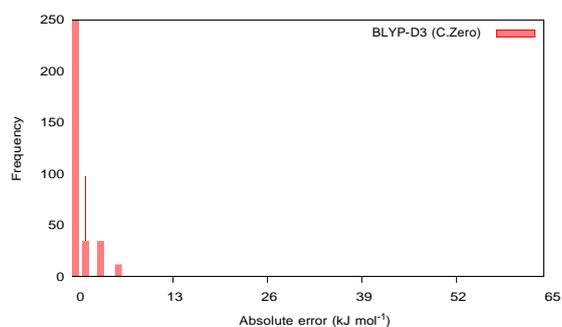
Figure B.36: BLYP-D3 (C) zero damping function potential energy surface error histogram

B.2.3.2 Becke-Johnson damping function
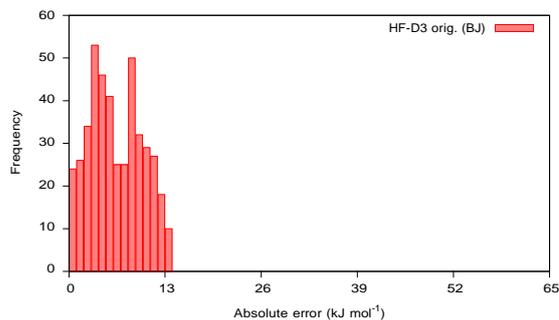
Figure B.37: HF-D3 orig. BJ damping function potential energy surface error histogram
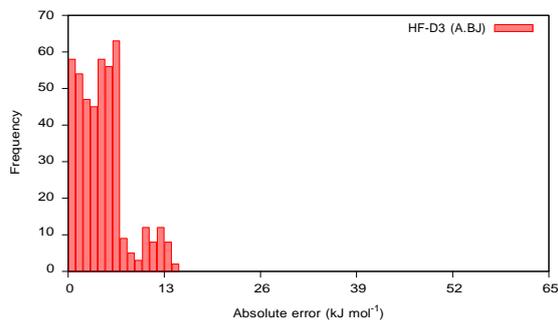
Figure B.38: HF-D3 (A) BJ damping function potential energy surface error histogram
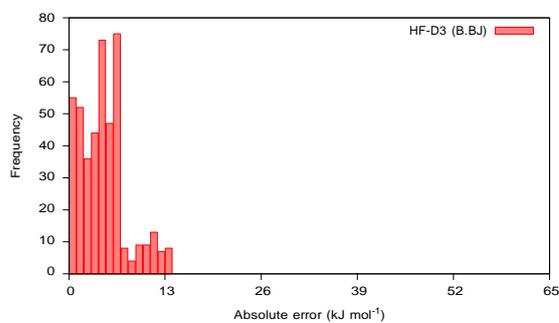
Figure B.39: HF-D3 (B) BJ damping function potential energy surface error histogram
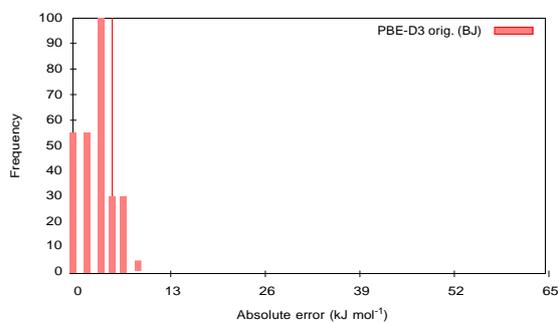
Figure B.40: PBE-D3 orig. BJ damping function potential energy surface error histogram
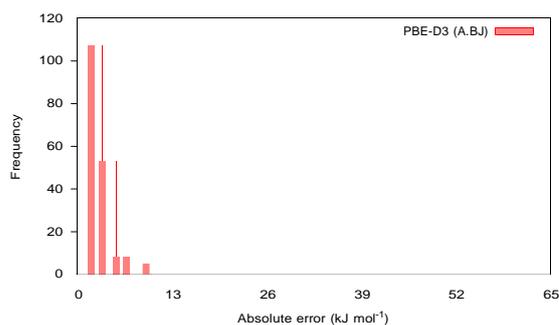
Figure B.41: PBE-D3 (A) BJ damping function potential energy surface error histogram
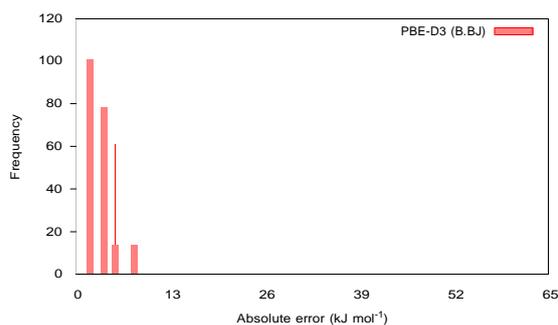

Figure B.42: PBE-D3 (B) BJ damping function potential energy surface error histogram
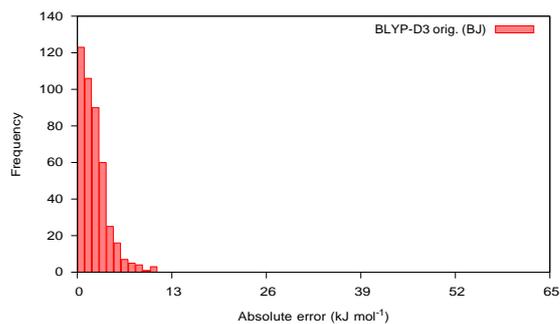
Figure B.43: BLYP-D3 orig. BJ damping function potential energy surface error histogram
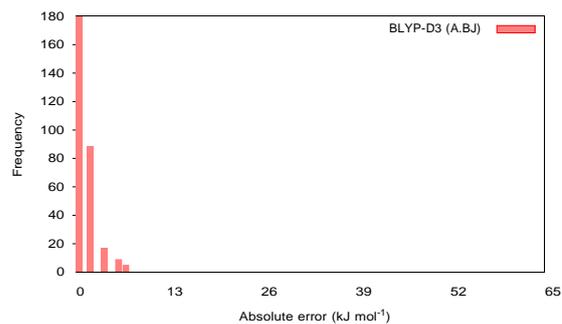

Figure B.44: BLYP-D3 (A) BJ damping function potential energy surface error histogram
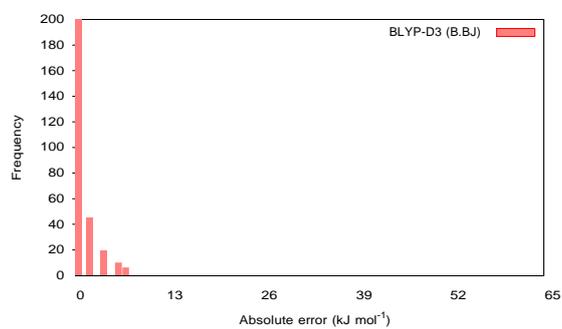

Figure B.45: BLYP-D3 (B) BJ damping function potential energy surface error histogram

Appendix C

DFT-D3 code patch for new coefficients

The following code patch may be applied to the dftd3.f file that is part of the official DFT-D3 package, version 3.3.1.0 that is available from http://www.thch.uni-bonn.de/tc/index.php?section=downloads&subsection=DFT-D3&lang=english at the time of writing. The patch was generated using the Unix diff command.

After compiling the patched code, the corrections outlined in Section 5.3.1, Tables 5.1 and 5.2 may be invoked by executing dftd3 <coord filename> -func <functional> -zero or dftd3 <coord filename> -func <functional> -bj for the zero or BJ damping functions, respectively. The <functional> parameter may be replaced with the text in quotes in the patch detailed below. For example, for HF-D3 (A) with BJ damping one may execute dftd3 <coord filename> -func hf-il-a -bj.

```
--- dftd3-original/dftd3.3.1.0/dftd3.f          2014-06-30 22:40:02.000000000 +1000
+++ dftd3-custom/dftd3.3.1.0/dftd3.f            2014-09-12 12:04:17.758600712 +1000
@@ -909,6 +909,33 @@
          alp =14.0d0

          select case (func)
+          case  ("hf-il-a")
+                 s6   =0.6488
+                 rs6  =-0.1459
+                 s18  =1.3599
+                 rs18=4.8801
+          case   ("pbe-il-a")
```

```
+                s6   =-0.5092
+                rs6  =1.2479
+                s18  =4.1538
+                rs18=0.4698
+         case    ("blyp-il-a")
+                s6   =0.1749
+                rs6  =0.6583
+                s18  =2.9878
+                rs18=2.3956
+         case    ("hf-il-b")
+                rs6  =-0.5293
+                s18  =0.8877
+                rs18=6.6430
+         case    ("pbe-il-b")
+                rs6  =1.0148
+                s18  =0.6778
+                rs18=1.3323
+         case    ("blyp-il-b")
+                rs6  =0.4333
+                s18  =1.3139
+                rs18=3.2632
          case  ("b-p")
                 rs6  =0.3946
                 s18  =3.2822
@@ -1145,6 +1172,37 @@
 c default def2-QZVP (almost basis set limit)
        if(.not.TZ)  then
        select case (func)
+         case    ("hf-il-b")
+                rs6  =0.6862
+                s18  =0.4443
+                rs18=0.7062
+         case    ("hf-il-c")
+                rs6  =0.6067
+                s18  =0.7129
+         case    ("pbe-il-a")
```

```
+                 s6   =3.6840
+                 rs6  =1.0844
+                 s18  =-9.4507
+                 rs18=1.2329
+          case   ("pbe-il-b")
+                 rs6  =0.9858
+                 s18  =-1.1022
+                 rs18=1.4316
+          case   ("pbe-il-c")
+                 rs6  =0.9987
+                 s18  =-0.0555
+          case   ("blyp-il-a")
+                 s6   =0.2540
+                 rs6  =0.4621
+                 s18  =2.1983
+                 rs18=0.9929
+          case   ("blyp-il-b")
+                 rs6  =0.9412
+                 s18  =0.6393
+                 rs18=0.8544
+          case   ("blyp-il-c")
+                 rs6  =0.8513
+                 s18  =0.6092
           case   ("slater-dirac-exchange")
                  rs6 =0.999
                  s18 =-1.957
```