# 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

**S** *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-$\zeta$ 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.
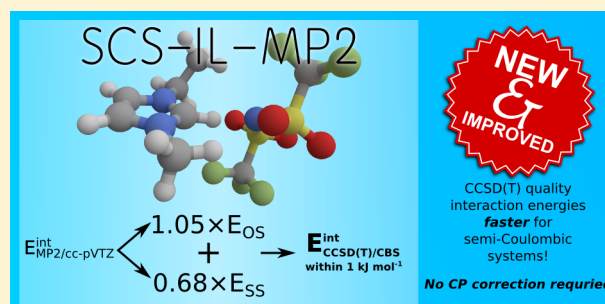
## 1. INTRODUCTION

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—speculated to be in excess of $10^{18}$, considering binary and ternary ILs[2]—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., $\pi$–$\pi$ 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 $\alpha$ and $\beta$ for the opposite-spin and same-spin, respectively.

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

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 $\alpha = 6/5$. 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-$\zeta$ basis set, ultimately yielding $\beta = 1/3$.

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 ($\alpha = 1.3$, $\beta = 0$). 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 ($\alpha = 0.40$, $\beta = 1.29$); 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 ($\alpha = 0$, $\beta = 1.76$). Their study also presented data showing impact of basis set in the choice of coefficients with $\alpha$ ranging from 1.29 to 1.75, and $\beta$ ranging from 0.17 to 0.40 involving double-, triple-, quadruple-$\zeta$, 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 $\alpha$ and $\beta$ without fitting to benchmark levels of theory. In Szabados' study, the average values were $\alpha = 1.12$ and $\beta = 0.84$ for the systems examined. Subsequently, Fink[20] has presented scaling factors where $\alpha = 1.15$ and $\beta = 0.75$, 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,

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

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

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

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;[12] 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, $E_{\text{complex}}$.[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-$\zeta$ quality basis set, BSSE is minimized through a cancellation of errors, with several other
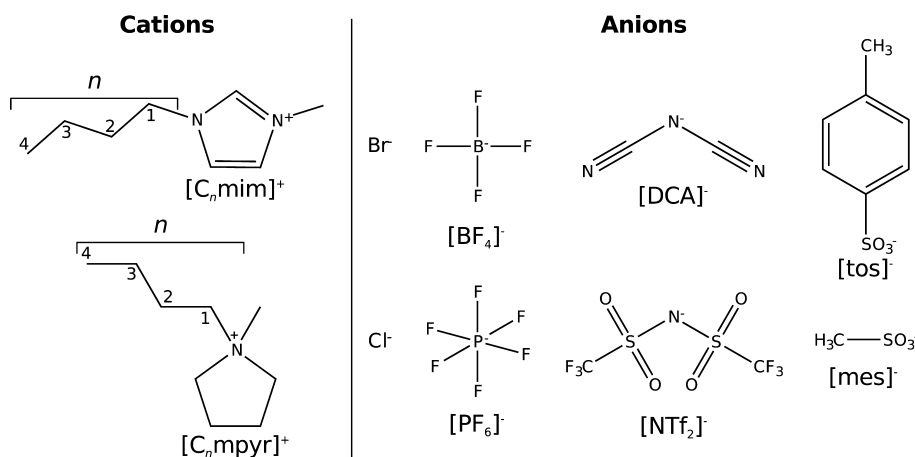
**Figure 1.** Summary of cations (left) and anions (right) used in this study.



**Figure 2.** [C$_2$mim]Br/Cl and [C$_2$mpyr]Br/Cl structures, anion positions labeled as P1−P4.

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$_2$mim]Br/Cl, [C$_2$mpyr]Br/Cl, [C$_2$mim][NTf$_2$], and [C$_2$mpyr][NTf$_2$], which all possess an ethyl group. In these examples, a number of key interactions are highlighted; [C$_2$mim]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
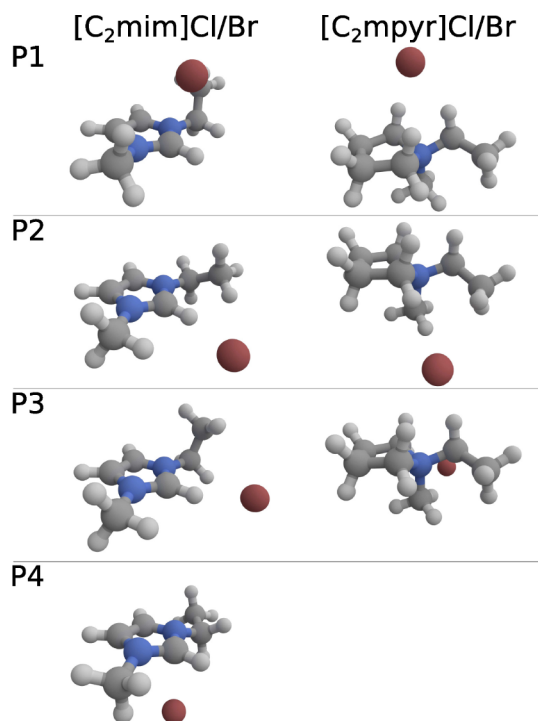
described by the methyl group, nitrogen atom, and the alkyl group (P2 and P3). More complex interactions can be seen in Figure 3 for the [NTf$_2$]$^-$ anion, as in these cases the [NTf$_2$]$^-$ anion can interact either via the imide or sulfonyl groups. Similar to the structures in Figure 2, [C$_2$mim][NTf$_2$] 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$_2$mpyr][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
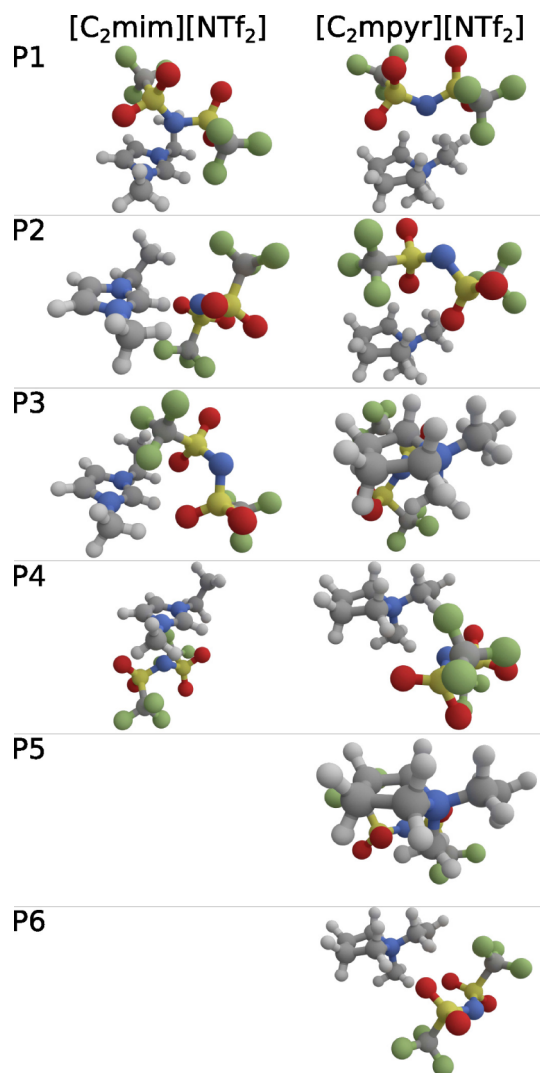
**Figure 3.** $[C_2mim][NTf_2]$ and $[C_2mpyr][NTf_2]$ structures, anion positions labeled as P1–P6.

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 Jurečka 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 $i$th benchmark ion pair interaction energy (e.g., CCSD(T)/CBS energy), $\hat{x}_i$ is the $i$th 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-$\zeta$ → triple-$\zeta$ (D → T) versus a triple-$\zeta$ → quadruple-$\zeta$ (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 ($-0.66 \pm 0.35$ 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 $\pm$ 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
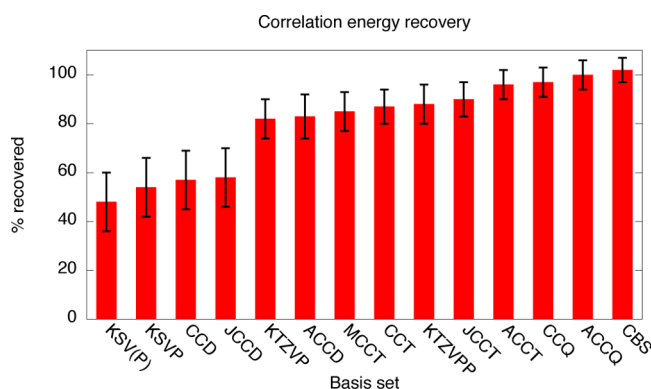


**Figure 4.** Average CP corrected MP2 correlation energy recovery with respect to CCSD(T)/CBS; error bars shown as one standard deviation.

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-$\zeta$ 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
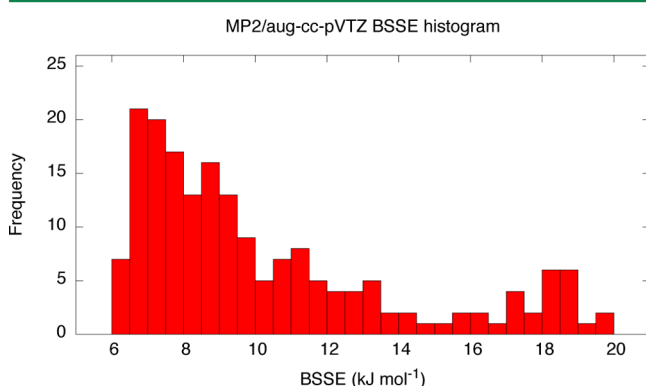


**Figure 5.** Histogram of BSSE for MP2/aug-cc-pVTZ interaction energies.

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 $1.01 \pm 4.52$ kJ mol$^{-1}$ whereas the much more complete aug-cc-pVQZ basis set gives errors of $5.94 \pm 4.36$ kJ mol$^{-1}$. These errors become systematic and follow expected trends after CP correction with errors of $-17.00 \pm 4.47$ kJ mol$^{-1}$ and $0.19 \pm 2.50$ 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
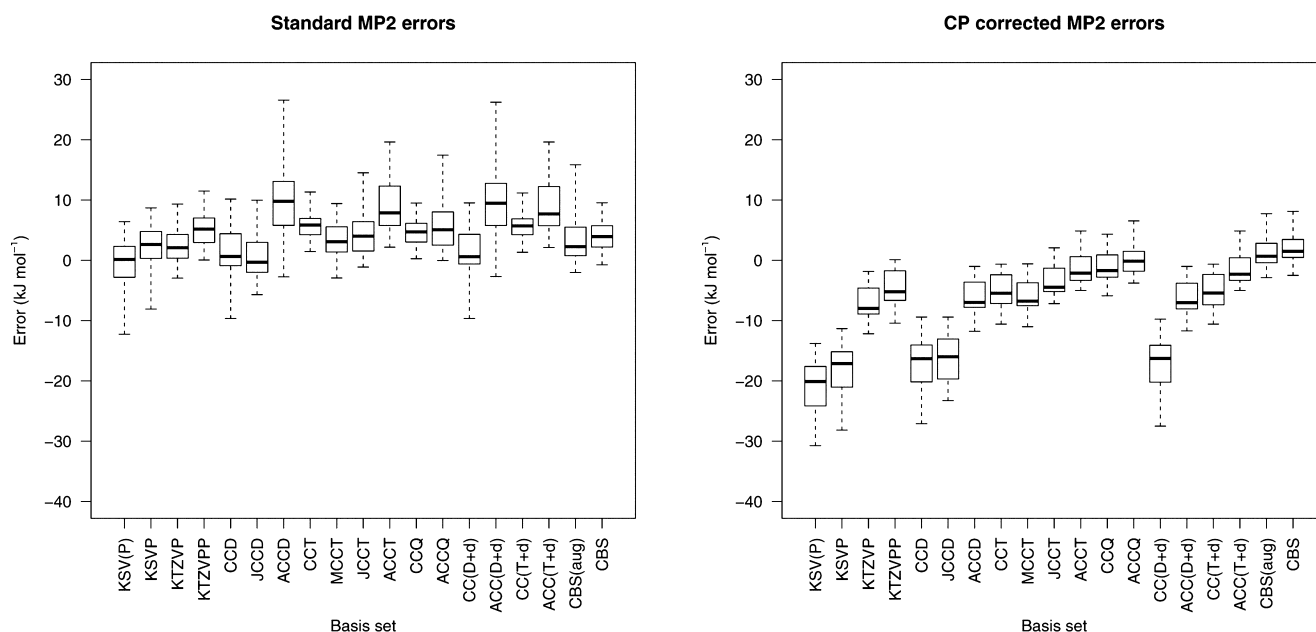


**Figure 6.** Comparison of errors between standard (i.e., no CP correction) (left) and counterpoise corrected (right) MP2 energies.
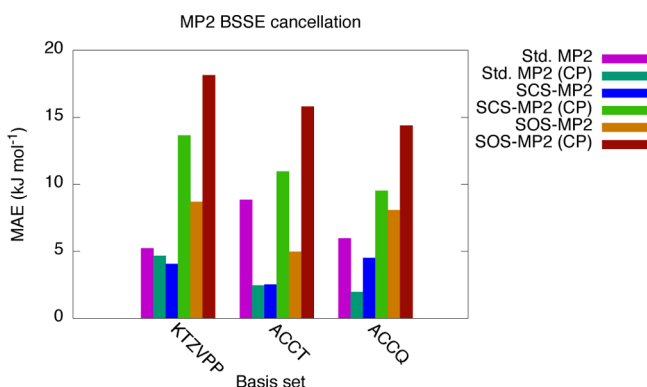
**Figure 7.** Comparison of errors between SCS-MP2 and standard MP2 with and without CP correction.

corrected energies using the aug-cc-pVQZ basis set do not perform as well as the triple-$\zeta$ 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-$\zeta$ basis set. The convenient observation that SCS-MP2 performs best with a triple-$\zeta$ 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 $\Delta E_{\text{OS}}$ and $\Delta E_{\text{SS}}$ 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-$\zeta$ 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 $\alpha$ 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-$\zeta$ 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
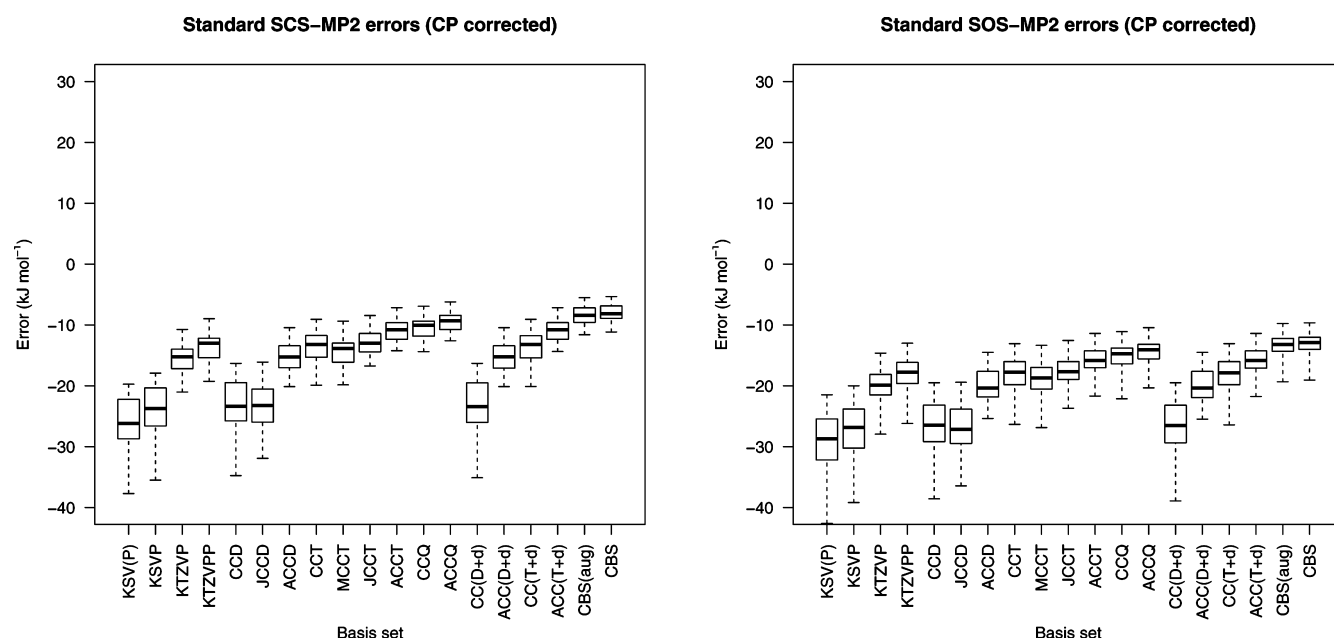
(a) Counterpoise corrected energies
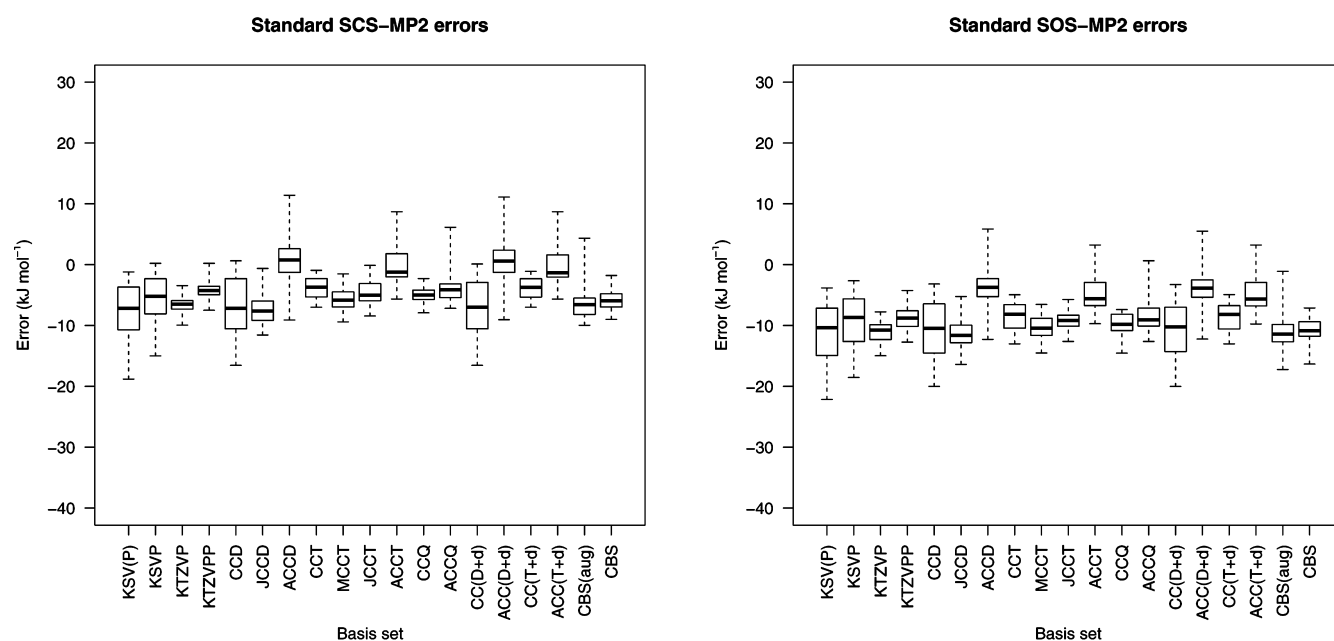


(b) Non-counterpoise corrected energies

**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 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, $\alpha$ and $\beta$ coefficients—are estimated for each random sample population, and overall errors are derived. Bootstrapping is advanta-

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

| | SCS-MP2 | | | | | | SOS-MP2 | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| basis set | $\alpha$ | 95% CI | $\beta$ | 95% CI | mean error (kJ mol$^{-1}$) | std. dev. (kJ mol$^{-1}$) | $\alpha$ | 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 |
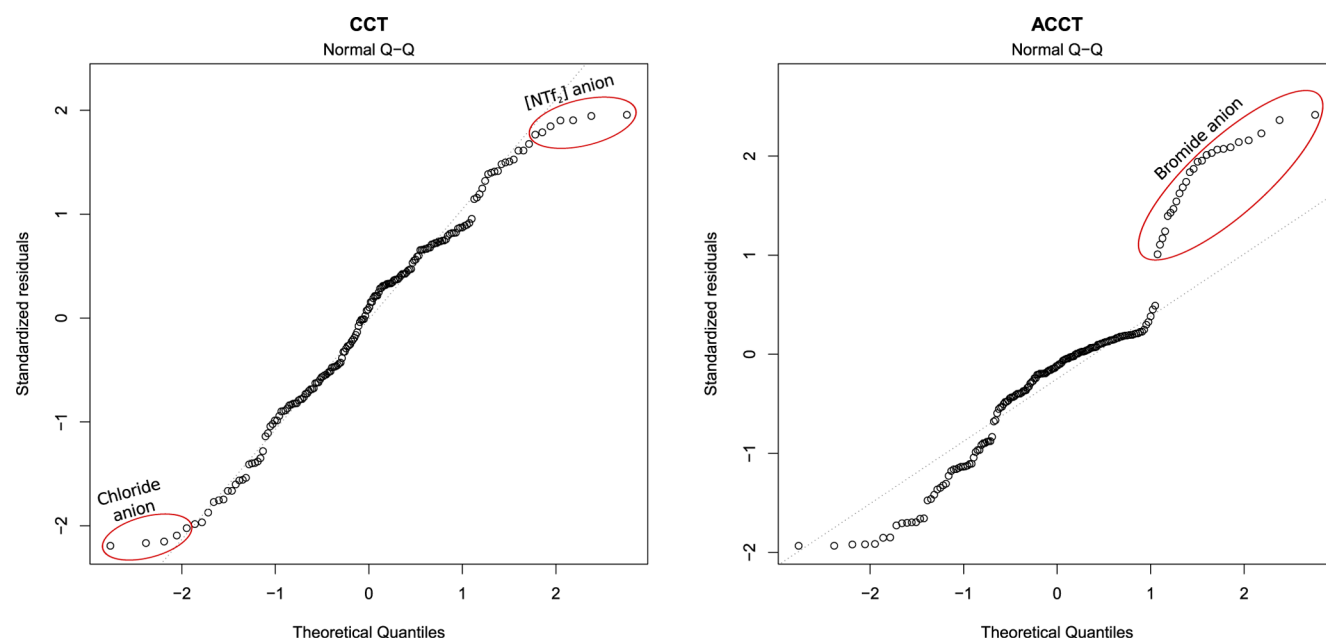


**Figure 9.** Q−Q plots of SCS-IL-MP2 multiple linear regression for cc-pVTZ (left) and aug-cc-pVTZ (right) basis sets.

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
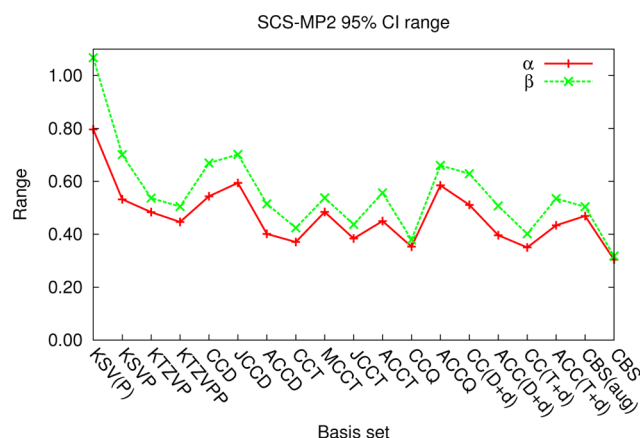
3119

dx.doi.org/10.1021/ct500309x | J. Chem. Theory Comput. 2014, 10, 3111−3122

**Figure 10.** 95% confidence interval range for the opposite- ($\alpha$) and same-spin ($\beta$) component coefficients in the SCS-IL-MP2 scheme.
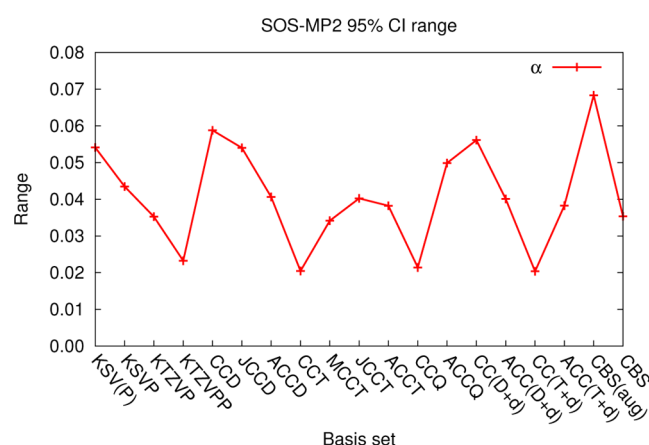


**Figure 11.** 95% confidence interval range for the opposite-spin ($\alpha$) component coefficient in the SOS-IL-MP2 scheme.



**Figure 12.** IL benchmark errors at the MP2/cc-pVTZ level of theory.

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.

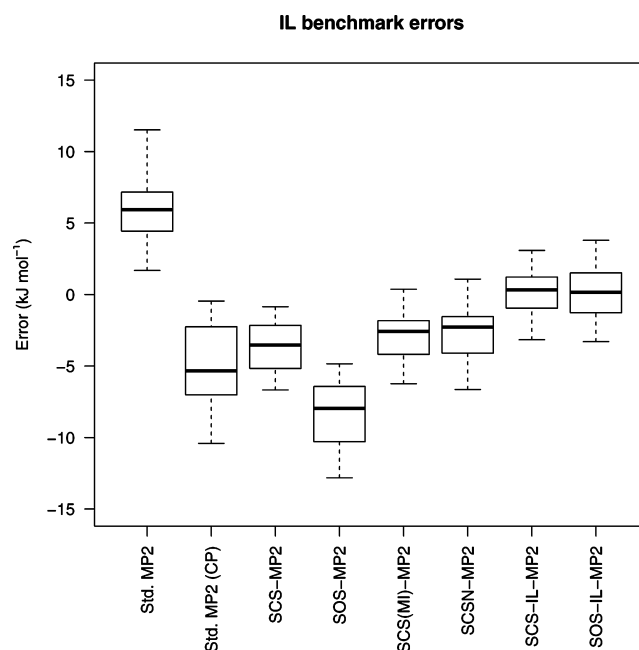(2) For energy calculations where high accuracy is desired ($<1.0$ kJ mol$^{-1}$ on average), MP2/cc-pVTZ along with

the appropriate scaling factors (SCS-IL-MP2: $\alpha = 1.05$, $\beta = 0.68$; SOS-IL-MP2: $\alpha = 1.64$, $\beta = 0.00$) 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
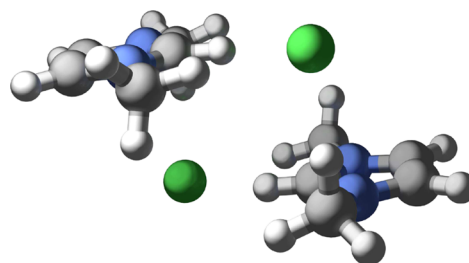


**Figure 13.** Example of a cluster of two [C$_1$mim]Cl ion pairs.

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

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) Jurečka, 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) Jurečka, 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.; Jurečka, 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.