Self-consistent continuum solvation (SCCS): The case of charged systems

C. Dupont,1,2,a) O. Andreussi,2 and N. Marzari2

1Institut Carnot de Bourgogne, UMR 6303 CNRS - Université de Bourgogne, BP 47870, F-21078 Dijon, France
2Theory and Simulation of Materials (THEOS), École Polytechnique Fédérale de Lausanne, Station 12, 1015 Lausanne, Switzerland

(Received 26 July 2013; accepted 7 November 2013; published online 4 December 2013)

The recently developed self-consistent continuum solvation model (SCCS) [O. Andreussi, I. Dabo, and N. Marzari, J. Chem. Phys. 136, 064102 (2012)] is applied here to charged species in aqueous solutions. Describing ions in solution represents a great challenge because of the large electrostatic interactions between the solute and the solvent. The SCCS model is tested over 106 monocharged species, both cations and anions, and we demonstrate its flexibility, notwithstanding its much reduced set of parameters, to describe charged species in solution. Remarkably low mean absolute errors are obtained with values of 2.27 and 5.54 kcal/mol for cations and anions, respectively. These results are comparable or better than the state of the art to describe solvation of charged species in water. Finally, differences of behavior between cations and anions are discussed. © 2013 AIP Publishing LLC. [http://dx.doi.org/10.1063/1.4832475]

I. INTRODUCTION

Many chemical or biochemical reactions occur in solution, and in particular in water; it is thus of outmost importance to properly describe solvent effects in quantum simulations. Most often the solvent plays such a fundamental role that chemical processes cannot be properly described when omitting it. Nevertheless, a full explicit description of all solvent molecules is both extremely expensive and often not necessary; and many continuum models have been developed. Among the most popular, one can mention the family of SMx models developed by Truhlar and co-workers,1–6 the widely used PCM model of Tomasi and co-workers,7–9 and the COSMO model of Klamt and Schüürmann.10 While all these developments have been implemented in quantum chemistry codes, recently, starting from the dielectric model proposed by Fattebert et al.,11–13 a new continuum model has been developed by Andreussi et al.,14 building on the previous developments of Scherlis et al.13 This self-consistent continuum model (SCCS) has been shown to offer high accuracy (mean absolute errors around 0.3–0.4 kcal/mol on large sets of neutral molecules and 1.2 kcal/mol on a complete set of 248 molecules), with an elegant and compact formulation, which relies on a much reduced number of parameters, and solving the ill-conditioning issues that affected the original dielectric model (see also Refs. 15 and 16 for slightly different, but not variational, implementations). In addition, the SCCS approach has been designed for first-principles molecular dynamics (MD) simulations, plane wave basis sets, and periodic boundary conditions, thus allowing the extension of continuum solvation to a wider range of applications. We note that recent developments in classical density functional theory (DFT) for liquids have been applied to solvation properties, and joint DFT17,18 also appears as a promising approach to extend the capabilities of solvation models in first-principles simulations.

Despite their high accuracy, continuum models present limitations when dealing with ions and numerous efforts have been dedicated to study solvation of charged species. Nevertheless, no model has allowed to reach the same absolute accuracy that obtained for neutral compounds. For example, applications of SMx4–6 or related models19–21 that lead to mean absolute errors below 1 kcal/mol (often around 0.5 kcal/mol) for neutral molecules give errors for charged species that are at least 4 kcal/mol. Other investigations based either on PCM,9 on the Miertus-Scrocco-Tomasi model,22,23 or on hybrid approaches24 also report errors of several kcal/mol for ions. Hence, reaching chemical accuracy (lower than 1 kcal/mol), as is the case for neutral molecules represents a great challenge.

Of course, charged species are associated to strong electrostatic effects and interactions with the solvent that complicate the description of the solute-solvent interactions. In fact solvation energies for ions are typically one order of magnitude larger than for neutral compounds. Given these strong interactions, the role and definition of the cavity surrounding the solutes play an even more important role than for neutral species. Several efforts9,25,26 have been dedicated to improve the definition of cavities for charged species, but it remains a most critical step in trying to reach high accuracy.

A second source of uncertainty comes from the lack of experimental data for ions. In particular, no direct experimental measurements are possible for the solvation free energies of single ions. Several papers3–27–29 have addressed this point and provided databases for solvation energies, but the field still lacks in its comparison with the case of neutrals.

Despite these limitations, the study of charged species in solution cannot be avoided: ions act as reaction intermediates.
in numerous chemical reactions and charged species in solution are also fundamental in electrochemistry, inorganic chemistry, and biochemistry. In this paper, the SCCS model is applied to charged species — both cations and anions. The definition of the cavity is found to be of crucial importance, and we exploit the great flexibility of the model to improve its accuracy for ions solvation. The article is organized as follows: after a brief description of the main features of the model, new parameterizations are reported for both cations and anions, and a careful analysis of the mean absolute errors (MAE) obtained is performed. The overall qualitative behavior for anions and cations is discussed, together with the main limitations and the possible improvements of the proposed approach.

II. METHODOLOGY

Since a detailed description of SCCS model is reported in the original paper, only its main features are recalled here for the sake of clarity. Both electrostatic and non-electrostatic short-range interactions are defined as functionals of the electronic density of the system, and are optimized self-consistently with the solute in a first-principles calculation. The solvation free energy is defined as

$$\Delta G_{\text{sol}} = \Delta G_{\text{el}}(\epsilon_0, \rho_{\text{min}}, \rho_{\text{max}}) + (\alpha + \gamma) S + \beta V,$$

where all the empirical parameters of the model have been listed explicitly.

In particular, the first term in the above equation $\Delta G_{\text{el}}(\epsilon_0, \rho_{\text{min}}, \rho_{\text{max}})$ refers to the electrostatic contributions, which are computed from the electrostatic interaction of the solute charge density with a dielectric medium surrounding a molecular shaped cavity built around the solute, following Refs. 11 and 12. Such a cavity is defined as an optimally smooth function of the electronic density of the solute in terms of two thresholds, $\rho_{\text{min}}$ and $\rho_{\text{max}}$: regions of space in which the electronic density is greater than $\rho_{\text{max}}$ are assigned a vacuum-like permittivity ($\epsilon = \epsilon_0$), while, as the electronic density falls below $\rho_{\text{min}}$, the static permittivity of bulk dielectric ($\epsilon = \epsilon_0$) is considered. References 11–14 iterated on the optimal form for the dielectric function. The second part of the solvation energy in Eq. (1) corresponds to the non-electrostatic contributions, which in the SCCS model are assumed to be proportional to the quantum surface $S$ and the quantum volume $V$ of the solute. Both $S$ and $V$ are defined in SCCS in terms of isosurfaces of the electronic density and thus enter self-consistently into the problem. Among the non-electrostatic terms, $\gamma S$ refers to the cavitation energy, with $\gamma$ the experimental surface tension, while the sum of $\alpha S + \beta V$ is related to repulsion and dispersion terms, with $\alpha$ and $\beta$ corresponding to two tunable parameters. In summary, of the six parameters which enter the SCCS definition of solvation free energy, two correspond to well-characterized bulk properties of the solvent ($\epsilon_0$ and $\gamma$), while the other four represent intrinsic properties of the solvent-solute interaction. We note that $\gamma$ and $\alpha$ appear together, and only their sum is relevant. We have kept these separate to investigate future non-aqueous solvents and to possibly avoid refitting $\alpha$. While first-principles simulations could provide some estimates of these parameters, the very compact representation of the solvation free energy in terms of such a limited number of parameters allow to fit them on experimental results.

In this study, the SCCS model is applied to a large set of 106 monocharged species — both cations (51) and anions (55) — representative of different chemical families. This set is based on the one previously defined by Cramer et al. for the SM6 model and is detailed in the supplementary material. Compared to the database defined in Ref. 4, one cation and five anions are missing, due to convergence issues. In Truhlar’s studies, these six species all correspond to clustered ions (i.e., inclusive of explicit solvent molecules). In fact, besides these six ions, in the remaining set 7 cations and 18 anions have been clustered in the study of the SM6 model, while, we stress, no explicit water molecules are used in our study. This database was selected for two reasons. First, it allows an accurate comparison with results of the SM6 model. Then, as already mentioned, experimental solvation energies for ions are scarce, and this set is certainly one of the most extended databases available in the literature. To fit the tunable parameters of SCCS, two reduced sets composed of 15 cations and 15 anions are considered and described in Table I, together with their experimental solvation energies. These subseries of anions and cations are chosen to be representative of both the diversity of chemical families available in the whole set and the mean absolute error of the complete database.

All calculations have thus been performed modelling the aqueous solution via the SCCS approach, as implemented in the open sources Quantum ESPRESSO package. This is based on density-functional theory with periodic-boundary conditions using the Martyna-Tuckerman correction to avoid spurious periodic charged images, recently extended to this case of continuum solvation. The generalized gradient approximation (GGA) has been applied in the Perdew-Burke-Ernzerhof (PBE) form for the

<table>
<thead>
<tr>
<th>Cations</th>
<th>Anions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Formula</td>
<td>Experimental solution energy (kcal/mol)</td>
</tr>
<tr>
<td>H$_2$O$^+$</td>
<td>−108.7</td>
</tr>
<tr>
<td>CH$_3$OH$^+$</td>
<td>−91.4</td>
</tr>
<tr>
<td>(C$_2$H$_5$OH)$^+$</td>
<td>−78.1</td>
</tr>
<tr>
<td>CH$_3$CO$_2$H$^+$</td>
<td>−75.5</td>
</tr>
<tr>
<td>CH$_3$NH$_2^+$</td>
<td>−74.8</td>
</tr>
<tr>
<td>(CH$_3$)$_2$NH$^+$</td>
<td>−59.5</td>
</tr>
<tr>
<td>CH$_3$NH$_2^+$</td>
<td>−70.8</td>
</tr>
<tr>
<td>CH$_3$NH$_2^+$</td>
<td>−69.3</td>
</tr>
<tr>
<td>pyridine$^+$</td>
<td>−59.5</td>
</tr>
<tr>
<td>CH$_3$NH$_2$H$_2^+$</td>
<td>−64.4</td>
</tr>
<tr>
<td>CH$_3$CN$^+$</td>
<td>−73.7</td>
</tr>
<tr>
<td>CH$_3$ON$_2$H$_2^+$</td>
<td>−68.0</td>
</tr>
<tr>
<td>CH$_3$CO$_2$H$_2^+$</td>
<td>−72.3</td>
</tr>
<tr>
<td>(CH$_2$)$_2$SH$^+$</td>
<td>−62.9</td>
</tr>
<tr>
<td>m-CIC$_6$H$_4$NH$_3^+$</td>
<td>−73.1</td>
</tr>
</tbody>
</table>


exchange-correlation functional and Vanderbilt ultrasoft pseudo no space potentials (USPP) were used from the PSLibrary of Dal Corso.\textsuperscript{35,36} Plane wave cut offs of 40 and 400 Ry have been used for wave functions and charge densities, respectively. To provide good convergence, simulation-cell sizes have been defined with the following criteria: the cell dimensions need to be 25 Å larger than the largest dimension of the ion, or to be at least 30 Å for small ions.

Calculations have also been performed with GAUSSIAN 03\textsuperscript{37} and GAUSSIAN 09,\textsuperscript{38} using the PBE exchange correlation functional. This allows a comparison of the SCCS model with two flavors of IEF-PCM.\textsuperscript{39,40} While both versions of IEF-PCM are based on the same physical model, they differ, among other things, in the definition of the molecular cavity and in the way polarization charges are represented numerically. For geometry optimizations, the double zeta 6-31G(d) basis set has been considered, while energies have been computed with the triple zeta 6-31+G(d,p) basis set.

III. RESULTS

A. General behavior

In the application of the SCCS model to neutral molecules,\textsuperscript{14} two optimal sets of parameters were proposed, namely, \textit{fitg03} + \(\beta\) and \textit{fitg03'} + \(\beta\), referring to different fits of the electrostatic and non-electrostatic parameters. As the first step in the present study, the same sets of parameters are applied to the case of charged species in solution and the values obtained are reported in Table II.

At first glance, MAEs appear to be very high compared to those of the neutral set: in fact MAE around 10 kcal/mol are obtained for the complete set of ions, as opposed to 1.2 kcal/mol for the case of neutrals. Nevertheless, due to the strong electrostatic interactions existing between the charged solute and the solvent, solvation energies for ions are, on average, ten times larger than those for neutral molecules. Thus it could be more consistent to discuss the percentage error than the MAE: in this case, the value of 14.3 % obtained for ions is clearly of the same order of magnitude than the one of neutral molecules. Besides, these results also compare well with IEF-PCM models. In particular, MAE obtained with both GAUSSIAN 03 and GAUSSIAN 09 are in the same range, with values of 8.09 and 12.7 kcal/mol, respectively. If the SM6 model of Truhlar and co-workers\textsuperscript{3} is considered: a MAE between 4.19 and 14.95 kcal/mol was found for a similar set of ions, and depending on the functional and on the basis set used.

While the above considerations are relevant for an all purpose model, it is worthwhile to analyze the accuracy of the SCCS approach separately for the specific cases of cations or anions. In fact, despite having solvation free energies of the same order, MAE for anions are more than six times larger than the ones for cations. This difference of behavior has already been evidenced in previous studies\textsuperscript{3,4} and can be related to the more localized charge of cations compared to anions.

B. New parameterization of the SCCS model for charged species

1. Cations

A new optimization of the 4 parameters in the SCCS models has been performed iteratively, by alternate minimizations of the MAE for a the reduced set of cations (see Table II) in the two-dimensional subspace of electrostatic (\(\rho_{\text{min}}\) and \(\rho_{\text{max}}\)) parameters and in that of the non-electrostatic (\(\alpha\) and \(\beta\)) parameters. The final iteration on these electrostatic parameters is reported in Figure 2(a). According to the results, an excellent MAE of 2.07 kcal/mol is obtained when using

<table>
<thead>
<tr>
<th>Name</th>
<th>(\rho_{\text{min}}) (a.u.)</th>
<th>(\rho_{\text{max}}) (a.u.)</th>
<th>(\alpha + \gamma) (dyn cm(^{-1}))</th>
<th>(\beta) (GPa)</th>
<th>Neutrals MAE (kcal/mol)</th>
<th>All ions MAE (kcal/mol)</th>
<th>Cations MAE (kcal/mol)</th>
<th>Anions MAE (kcal/mol)</th>
</tr>
</thead>
<tbody>
<tr>
<td>fitg03 + (\beta)</td>
<td>0.00010</td>
<td>0.0050</td>
<td>50</td>
<td>-0.35</td>
<td>1.20</td>
<td>8.43</td>
<td>2.55</td>
<td>17.4</td>
</tr>
<tr>
<td>fitg03' + (\beta)</td>
<td>0.00030</td>
<td>0.0030</td>
<td>20</td>
<td>-0.08</td>
<td>1.28</td>
<td>8.05</td>
<td>2.66</td>
<td>16.9</td>
</tr>
<tr>
<td>GAUSSIAN 03</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>8.07</td>
<td>4.00</td>
<td>10.2</td>
<td></td>
</tr>
<tr>
<td>GAUSSIAN 09</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>12.7</td>
<td>11.9</td>
<td>15.0</td>
<td></td>
</tr>
<tr>
<td>Fit cations</td>
<td>0.00020</td>
<td>0.0035</td>
<td>5</td>
<td>0.125</td>
<td>0.450</td>
<td></td>
<td>2.26</td>
<td></td>
</tr>
<tr>
<td>Fit anions</td>
<td>0.00240</td>
<td>0.0155</td>
<td>0</td>
<td>4.50</td>
<td></td>
<td></td>
<td>5.54</td>
<td></td>
</tr>
</tbody>
</table>
FIG. 1. Calculated solvation energies $\Delta G_{\text{sol,SCC}}$ (in kcal/mol) as a function of experimental data $\Delta G_{\text{sol,exp}}$ (in kcal/mol) for all cations ($\times$) and anions ($\bullet$) considered. The solid line represents the ideal fit (slope = 1), while the dotted one represents the fit for anions.

$\rho_{\text{min}} = 0.0002$ a.u. and $\rho_{\text{max}} = 0.0035$ a.u., for instance. It is important to stress that, as reported for neutral molecules, a continuum of $(\rho_{\text{min}}, \rho_{\text{max}})$ pairs leads to this low MAE, thus pointing to a correlation between the two.

We note in passing that, for the same range of $\rho_{\text{min}}$ and $\rho_{\text{max}}$ plotted in the figure, neutral molecules\cite{14} show a smoother evolution with MAE between 0.4 and 4.6 kcal/mol, while for cations, the MAE is comprised between 2.07 and 13.8 kcal/mol. This result confirms the larger contribution of electrostatic phenomena for charged species compared to neutral.

Starting from the best $(\rho_{\text{min}}, \rho_{\text{max}})$ pair, the last step of optimization of the non-electrostatic parameters is reported in Figure 2(b). Qualitatively, the behavior of cations towards non-electrostatic parameters is similar to the one already mentioned for neutrals,\cite{14} showing a linear relationship between $\alpha + \gamma$ and $\beta$. As suggested for neutral solutes, such a behavior can be rationalized in terms of the small sizes of the systems considered, for which a strong coupling between quantum surface and quantum volume is expected. Quantitatively, no further improvement occurs after this last optimization of non-electrostatic parameters, with only a negligible modification of the MAE is observed with respect to the final value of 2.07 kcal/mol for the selected group. Finally, as the best set of parameters is applied on the whole ensemble of cations, an overall MAE of 2.27 kcal/mol is obtained. This value is remarkable: it corresponds to a percentage error of 3.34%, much better than that obtained for neutral molecules.

2. Anions

The same iterative procedure described above has been applied to model the solvation energies for anions. Due to the comparatively poor results for the initial sets of parameters, a wider range of values for $\rho_{\text{min}}$ and $\rho_{\text{max}}$ has been considered. The final step of the optimization of the electrostatic parameters is reported in Figure 3(a), with $\rho_{\text{min}}$ in the range between 0.00175 and 0.00225 a.u. and $\rho_{\text{max}}$ between 0.0125 and 0.0185 a.u., both one order of magnitude larger than the case of neutral species and cations.

Qualitatively the evolution of the MAE with $\rho_{\text{min}}$ and $\rho_{\text{max}}$ is similar to the one already discussed for cations. In particular, anions present a sensitivity towards $\rho_{\text{min}}$ and $\rho_{\text{max}}$ similar to the one already mentioned for cations. This confirms the stronger influence of electrostatic parameters for charged species. The non-electrostatic parameterization is reported in the supplementary material,\cite{44} and leads to an identical behavior as the one discussed for cations: a linear relationship between $\alpha + \gamma$ and $\beta$. Only the range of values differs: larger $\beta$ leads to lower MAE, a symptom of the larger repulsion for the solvent water molecules associated to the anions. Following these successive optimizations, a very good MAE

FIG. 2. MAE (in kcal/mol) calculated with respect to the experimental solvation energies for the 15 cations of the reduced database: optimization of (a) electrostatic parameters with $\beta = 0.125$ GPa and $\alpha + \gamma = 5$ dyn/cm, and (b) non-electrostatic parameters with $\rho_{\text{min}} = 0.00020$ a.u. and $\rho_{\text{max}} = 0.0035$ a.u.
of 4.71 kcal/mol is obtained for the reduced set. As for cations, this value is somewhat increased when all anions are considered, reaching a global MAE of 5.54 kcal/mol. Qualitatively, the overall behavior is clearly improved compared to the experimental values. In fact, at variance with the results of Figure 1, a more homogeneous distribution is obtained (see Figure 3(b)), with solvation free energies that are both over and underestimated. Nevertheless, some discrepancies still remain. In fact, if most of solvation energies are well described with relative errors well below average, some anions still present high discrepancies (up to 30 kcal/mol for H₂C=CHCH₂O⁻, for instance). Given the relatively small set we have used (only 55 anions), it is difficult to perform an accurate analysis in terms of the different chemical functionalizations of the solutes. However, according to the detailed values reported in the supplementary material, it appears that electronic delocalization (for instance anions with aromatic rings) leads to lower relative errors since a more distributed charge density decreases the interaction between the solvent and the solute, this could lead to higher accuracy. We conclude mentioning also some of the fundamental issues in describing anions in density-functional theory — a notoriously challenging case, where often local or semi-local exchange correlation functionals fail to bind the additional electrons. Hybrid functionals ameliorate this failure, but the correct asymptotic behaviour of the charge density cannot be achieved with functionals that do not have the correct long-range behaviour — hence further studies in this direction would certainly be warranted.

C. Discussion

The performance of the SCCS model can now be discussed in comparison with other continuum solvation models. In particular the IEF-PCM model, implemented in GAUSSIAN 03 and GAUSSIAN 09, the SMx family, the SS(V)PE model, as well as its combination with the FESR correction⁴¹,⁴² are considered here. Results from SMx,⁵ ⁶ SS(V)PE,⁴² and SS(V)PE+FESR⁴² represent a very good point of comparison, since they are based on a closely similar database. Nonetheless, it is important to keep in mind that the simulations reported in Refs. 5 and 6 contain calculations on clustered ions, namely ions surrounded by water molecules, which are not included in the present work and which deliver significant improvements with respect to a purely continuum picture.⁳ All SMx, SS(V)PE, and SS(V)PE+FESR results from the literature are reported in Table III, together with those obtained with SCCS, g03, and g09 in this study.

The SCCS model presents one of the best MAE for cations, with a value of 2.27 kcal/mol. Only the addition of the FESR correction over the SS(V)PE model allows to attain a better MAE. Nevertheless, one has to mention that FESR represents an empirical correction not derived from a variational principle, contrary to the SCCS model. Hence a rigorous comparison should be between the current version of SCCS and the original SS(V)PE model. For anions, the SCCS model clearly compares very well with the IEF-PCM models for which a MAE larger than 10 kcal/mol are obtained for the same set of anions. This MAE is also very comparable with the SMx models, where improved performance is obtained also thanks to the presence of selectively clustered anions in the database. As mentioned previously, anions with a large electronic delocalization present low MAEs and addition of explicit water molecules will either increase delocalization or simply shift the boundary for the continuum from a charged ion to a neutral water molecule. As for the case of cations, the very good performance of the FESR correction should be noted, as it allows to reach MAE below 3 kcal/mol for anions.

It is worth mentioning that while the SCCS results are obtained with only four fitting parameters, these four are
TABLE III. Mean Absolute Error (MAE) for cations and anions. Results for SM5, SM6, SM7, and SM8 are extracted from Ref. 5, while values for SM8-AD are from Ref. 6. In all these cases, 52 cations and 60 anions are studied. These include 8 clustered cations and 23 clustered anions. Values for SS(V)PE and SS(V)PE+FESR are from Ref. 42 and correspond to 51 cations and 60 anions from the Minnesota Solvation Database. Values for PCM / g03, PCM / g09, and SCCS come from this study and correspond to a reduced data set of 51 cations and 55 anions, but without any explicit water molecules.

<table>
<thead>
<tr>
<th>Method</th>
<th>Cations (MAE) kcal/mol</th>
<th>Anions (MAE) kcal/mol</th>
</tr>
</thead>
<tbody>
<tr>
<td>SM5</td>
<td>5.2</td>
<td>6.7</td>
</tr>
<tr>
<td>SM6</td>
<td>2.72</td>
<td>3.56</td>
</tr>
<tr>
<td>SM7</td>
<td>2.8</td>
<td>3.8</td>
</tr>
<tr>
<td>SM8</td>
<td>2.7</td>
<td>3.7</td>
</tr>
<tr>
<td>SM8-AD 6</td>
<td>3.0</td>
<td>5.5</td>
</tr>
<tr>
<td>SS(V)PE</td>
<td>2.80</td>
<td>11.83</td>
</tr>
<tr>
<td>B3LYP/G3large 42</td>
<td>3.87</td>
<td>14.69</td>
</tr>
<tr>
<td>SS(V)PE + FESR</td>
<td>2.01</td>
<td>2.57</td>
</tr>
<tr>
<td>B3LYP/G3large 42</td>
<td>1.97</td>
<td>2.95</td>
</tr>
<tr>
<td>PCM</td>
<td>4.00</td>
<td>10.2</td>
</tr>
<tr>
<td>g03</td>
<td>11.9</td>
<td>15.0</td>
</tr>
<tr>
<td>SCCS</td>
<td>2.27</td>
<td>5.54</td>
</tr>
</tbody>
</table>

strongly correlated, thus confirming the great flexibility of the model to reproduce the chemistry at play with a very minimal set of parameter. For comparison, parameterization of SMx models is based on the intrinsic chemical nature of each atom, obtained first through the optimization of a charge model and then of the SM model itself; as a consequence, several dozen of parameters are required. In addition, the model is fully variational (e.g., as opposed to SS(V)PE-FESR), and so it can be straightforwardly used in first-principles molecular dynamics or in response/perturbation theory (e.g., optical spectra from time-dependent DFT).

In order to understand the different behavior of charged species in SCCS, it is apparent from Table II that anions and cations require different shapes for their cavity, and the optimal thresholds obtained confirm this observation. In fact, values of $\rho_{\text{min}}$ and $\rho_{\text{max}}$ leading to the minimum MAE for anions are around one order of magnitude higher than the ones considered for either cations or neutral molecules. The larger density thresholds for anions lead to smaller cavities, in agreement with previous results, and this is highlighted by the example of acetamide, reported in Figure 4.

By comparing the cavities optimized for the different protonation states of the acetamide molecule, it is possible to rationalize the behavior of the model for the dielectric function and to address its possible limitations. Both anions and cations show dielectric cavities with a sharper interface with respect to the neutral solute, as can be seen from the smaller separation between the two isosurfaces of the electronic density corresponding to the onset of the dielectric (grey surface) and to the dielectric reaching its bulk value (blue surface). Moreover, the two interfaces are shown to be closer to the solute for both the protonated (positive) and the deprotonated (negative) species.

The latter finding is consistent with the expected behavior of a polar solvent on charged systems, which should be dominated by Coulomb monopole-dipole interactions, leading nearby solvent molecules to get closer to the solute. Such a stronger coordination of solvent molecules happens regardless of the sign of the charge of the solute, even though some asymmetries can be expected between solvated cations and solvated anions in water, due to the different coordination mechanisms mediated by oxygen lone-pairs or hydrogen-bonds, respectively. On the other hand, the sharpness of the optimized dielectric interface cannot be explained solely with electrostatic considerations, as the stronger coordination of solvent molecules due to charged solutes would suggest a more slowly varying dielectric function. In particular, the dielectric constant of the environment immediately surrounding a charged species is expected to be much lower than its bulk value, since the first solvation shell of water molecule presents a much reduced rotational freedom due to the coordination with the solute. Non-electrostatic interactions, in particular hydrogen-bonds, are probably responsible for the behaviors of the dielectric functions in charged species.

It is instructive to analyze what would happen to the different charged species using a common set of threshold parameters (as in Figure 5(b)). For this purpose, it is important to keep in mind that, for the dielectric function chosen, the real-space separation of the two isosurfaces of the density is directly proportional to the (DFT) ionization potential of the solute. Since cations show a lower electronic density than the corresponding neutrals, the isosurface corresponding to the onset of the dielectric would get closer to the solute; at the same time, for cations showing higher ionization potentials, the transition between vacuum and bulk dielectric becomes sharper than for neutrals. Exactly the opposite happens for anions, where the first isosurface would end up further away from the solute charge density, and the interface would get smoother. Both these trends go in the wrong direction, as confirmed by the poor results obtained from the initial set of simulations and reported in Sec. III A. It is important to stress that this limitation is implicit of all models which rely on a local definition of the dielectric function in terms of the electronic density alone, and would be particularly challenging when trying, e.g., to describe a zwitterion. While a common
set of parameters cannot be found which provides the correct description of the solute cavity as a function of the solute total charge, the flexibility of the model allows us not only to accurately describe the separated cases, but points to future expansions that could be made aware of the local charge environment.

D. Conclusions

The revised self-consistent continuum solvation model SCCS is applied here to charged species. Its flexibility and accuracy, already demonstrated for neutral molecules, are confirmed in the study of both cations and anions. As expected, due to the stronger interactions into play, higher mean absolute errors, but lower relative errors, are obtained when compared to neutrals. In particular, after optimization of both the electrostatic and non-electrostatic parameters, MAEs of 2.27 and 5.54 kcal/mol are found for standard test sets of cations and anions, respectively. These values place the SCCS model among the best solvation continuum models available. In fact, for anions, only the SM6 model, when tested over a set that includes selectively clustered anions (i.e., including explicit water molecules), and the SS(V)PE model in combination with its non-variational correction (FESR) offer better results, while for cations even SM6 displays a higher MAE (with the SS(V)PE + FESR combination comparable but marginally better than SCCS). We thus conclude that the SCCS model is able to reach accurate solvation energies for charged species, while preserving the variational nature of its formulation and its straightforward implementation in extended or periodic systems and associated plane-wave basis sets. The different parametrization of the model for the case of anions, as opposed to cations and neutrals, warrants further studies on the description of anions with more accurate functionals, able to correctly describe the highest-occupied orbital and the exponential charge-density decay (a very challenging task), or more detailed physical models of the solute-continuum interactions that can preserve the variational nature of the functional formulation.

33. O. Andreussi and N. Marzari, “Electrostatics of continuum solvation models in periodic boundary conditions” (to be published).
35. A. Dal Corso, PSlibrary version 0.2.3, see http://qe-forge.org/projects/pslibrary.
37 M. J. Frisch, G. W. Trucks, H. B. Schlegel et al., GAUSSIAN 03, Revision C.02, Gaussian, Inc., Wallingford, CT, 2004.
38 M. J. Frisch, G. W. Trucks, H. B. Schlegel et al., GAUSSIAN 09, Revision A.02, Gaussian Inc., Wallingford, CT, 2009.
44 See supplementary material at http://dx.doi.org/10.1063/1.4832475 for the whole database of ions and additional results for anions.