Phase Equilibria in Ionic Liquid Facilitated Liquid–Liquid Extractions

Anand Bharti, Debashis Kundu, Dharamashi Rabari, Tamal Banerjee

COSMO-SAC: A Predictive Model for Calculating Thermodynamic Properties on a-priori Basis

Publication details
Anand Bharti, Debashis Kundu, Dharamashi Rabari, Tamal Banerjee
Published online on: 17 Mar 2017

How to cite:
Accessed on: 28 Nov 2019
3 COSMO-SAC: A Predictive Model for Calculating Thermodynamic Properties on a-priori Basis

3.1 Liquid-Phase Thermodynamics

The prediction of properties of liquid-phase thermodynamics is convention-
ally performed using the Gibbs energy of solvation ($\Delta G^{\text{solv}}$), which is merely
the energy required to transfer a solute molecule from vacuum to a solvent
phase. Two of the important properties computed from this route are the
activity coefficient and the vapor pressures.

$$\ln(x_{i/s}Y_{i/s}) = \frac{\Delta G^{\text{solv}}_{i/S} - \Delta G^{\text{solv}}_{i/i}}{kT} = \frac{\Delta G^{*\text{solv}}_{i/S} - \Delta G^{*\text{solv}}_{i/i}}{kT} + \ln\left(\frac{V_{i/i}}{V_{i/S}}\right)$$ (3.1)

In the equation, $\Delta G^{*\text{solv}}_{i/S}$ is the Gibb’s free energy of solvation of molecule
$i$ in solvent $j$, while $\Delta G^{*\text{solv}}_{i/i}$ refers to the term for molecule $i$ in itself. Here,
the term $\Delta G^{*\text{solv}}_{i/i}$ refers to the non-translational contribution of the Gibb’s
free energy of solvation. This is related by Equations 3.2 and 3.3, as given
below:

$$\Delta G^{*\text{solv}}_{i/S} = \Delta G^{\text{solv}}_{i/S} - kT \ln\left(\frac{N\Lambda^3_i}{V_{i/S}}\right)$$ (3.2)

$$\Lambda_i = \frac{h}{\sqrt{2\pi m_i kT}}$$ (3.3)

Here, Equation 3.3 refers to the De Broglie’s wavelength originating from the
dual particle theory of material, where $m_i$ is the mass of the particle, $h$ is
the Planck’s constant (J.s) and $k$ is the Boltzmann constant (J/K). In a simi-
lar manner, the vapor pressures can be predicted by Equation 3.4, where an
ideal phase has been assumed.
It should be noted that the thermodynamic properties, namely partition coefficient, heat of vaporization and phase equilibria, are computed from $\gamma_{i/S}$ and $P_{i}^{\text{vap}}$. Hence, the objective of COSMO-SAC (Lin and Sandler, 2002) is to determine these properties from quantum chemical and statistical mechanical framework. The model is based on the original pioneering work of Klamt (Klamt, 1995; Klamt, Jonas, Bürger, & Lohrenz, 1998; Klamt & Schuurmann, 1993). The primary contribution from COSMO-SAC is that $\Delta G_{ij}^{\text{solv}}$. $\Delta G_{ij}^{\text{solv}}$ is mainly bifurcated from two contributors, namely electrostatic (ES) and van der Waals (vdW) force (Lin, Chang, Wang, Goddard, & Sandler, 2004; Wang, Lin, Chang, Goddard, & Sandler, 2006; Wang, Sandler, & Chen, 2007; Wang, Stubbs, Siepmann, & Sandler, 2005). Contributions such as those from vibrational, rotational and dipole–dipole moment are ignored. Let us write out the contribution for each part in Equation 3.5.

$$\Delta G_{i/S}^{\text{solv}} = \Delta G_{i/S}^{\text{ES}} + \Delta G_{i/S}^{\text{vdW}} = \Delta G_{i/S}^{\text{ES}} + \left[P \left(V_{i/S}^{\text{liq}} - V_{i}^{\text{ld, gas}}\right)\right]$$

$$= \Delta G_{i/S}^{\text{ES}} + \Delta A_{i/S}^{\text{disp}} + \Delta A_{i/S}^{\text{cav}} - kT$$

$\Delta A_{i/S}^{\text{disp}}$ is the Helmholtz energy of molecule $i$ in the solvent. Assuming the fact that $V_{i}^{\text{ld, gas}} > V_{i/S}^{\text{liq}}$, we further write out the Helmholtz energy as a sum of two parts, namely $\Delta A_{i/S}^{\text{disp}}$ and $\Delta A_{i/S}^{\text{cav}}$. The former refers to the London Dispersion term. This is usually referred to as temporary attractive force, which occurs when electrons in two nearby atoms occupy positions such that they form temporary dipoles. This force is also called ‘induced dipole–dipole’ interaction. $\Delta A_{i/S}^{\text{cav}}$ refers to the amount of energy required to construct a cavity of component $i$ within a solvent $S$. This primarily depends on the shape and size of the molecule.

The predictions of $\Delta G_{ij}^{\text{solv}}$ take a different pathway each for activity coefficients and vapor pressure. The activity coefficients are usually predicted from the difference of their properties in solvent and individual phase, as given below (Burnett, 2012):

$$\ln \left(\frac{P_{i}^{\text{vap}}V_{ij}}{kT} \right) = \frac{\Delta G_{ij}^{\text{solv}}}{kT}$$
By rearranging the terms, we obtain:

\[
\frac{\Delta G^*_{i/S} - \Delta G^*_{i/i}}{kT} + \left( \frac{\Delta A_{i/S}^{vdW} - \Delta A_{i/i}^{vdW}}{kT} \right) + \left( \frac{\Delta A_{i/S}^{cav} - \Delta A_{i/i}^{cav}}{kT} \right) + \ln \left( \frac{V_{i/S}}{V_{i/i}} \right) = \left( \frac{\Delta G^*_{i/S} - \Delta G^*_{i/i}}{kT} \right) + \ln \left( \gamma_{i/S}^{\text{comb}} \right)
\]  

(3.7)

Both the phases considered are liquid; hence, the dispersion energies are almost similar and equal. In other words, their difference is fractionally small as compared with other terms. So, we neglect this contribution, that is, the second term of Equation 3.7. On similar lines, the last two terms reflect the differences in cavity formation and density. These are used through a combinatorial contribution, which is similar to that used in nonrandom two-liquid (NRTL) and UNIQUAC models. Hence, the final expression takes the form (Equation 3.8):

\[
\ln \left( \frac{P_{i}^{\text{vap}}V_{i/S}}{kT} \right) = \left( \frac{\Delta G^*_{i/S} + \Delta A_{i/S}^{vdW} + \Delta A_{i/S}^{cav}}{kT} - kT \right)
\]  

(3.8)

For the estimation of vapor pressures, we require the Gibb’s free energy of solvation for a single phase. Hence, the dispersion terms do not cancel out; this invariably implies that the cavity formation term in Equation 5.8 (second term) is not straightforward. The Equation 3.8 takes the form:

\[
\ln \left( \frac{P_{i}^{\text{vap}}V_{i/i}}{kT} \right) = \left( \frac{\Delta G^*_{i/i} + \Delta A_{i/i}^{vdW} + \Delta A_{i/i}^{cav}}{kT} - kT \right)
\]  

(3.9)

### 3.2 Estimation of Electrostatic Contribution

Within the COSMO-SAC model, the largest contribution, as evident from Equation 3.8, is from the electrostatic contribution, which is due the inherent charges between molecules. The description of \(\Delta G^*_{ES}\) is a tedious job and is usually computed with a continuum solvation model, as defined in quantum mechanics. Here, the solute molecule is first modeled to be surrounded by a van der Waal-like segmented surface. Thereafter, it is immersed in a conductor with infinite dielectric constant, such that it totally screens the molecule’s inherent electronic density. Density Functional Theory (DFT) is then applied on it to solve for the Hamiltonian, so as to determine the properties of each segment, including the charge and area associated with each segment. Thereafter, an ensemble-based theory along with a statistical mechanical framework is used, which relates the segment interactions to useful thermodynamic

---

properties such as chemical potential. In a nutshell, this incorporates steps, namely geometry optimization of the molecule, construction of the vdW segmented surface, immersion of the molecule with the surface into a conductor and, finally, the computation of the segment properties via statistical mechanical framework. We shall now discuss all the steps in detail.

### 3.2.1 Geometry Optimization of the Molecule

The geometry for any given molecule is obtained by minimizing the total energy with respect to the atomic coordinates. This is usually performed by solving the Schrödinger equation, so as to predict the energies. Initially, the structures are drawn in some suitable freeware, such as Avogadro, Molden or GaussView, to name a view. Thereafter, the optimization is done in vacuum by giving a suitable level of theory (HF/DFT/Semi-Empirical) and a basis set in commercial packages such as Gaussian and TURBOMOLE. This is done in vacuum by using where after the end of optimization, an absence of negative frequencies indicates a global minimum. This step also gives us the optimized geometry, its partial charges and the energies of the molecule in vacuum. This optimized geometry is then used for all subsequent calculations. In so doing, the effect of any changes in geometry between vacuum and solvent is assumed negligible.

An interesting question may be: Should the geometry optimization be carried out in a known solvent? However, the choice of solvent may be then biased either towards high dielectric constants such as water or towards low dielectric constants such as hexane. Since neither of them is applicable, universal phenomena such as vacuum are chosen. This also reduces the computational time, as optimization in vacuum is faster. Moreover, reported experimental data (Pye, Ziegler, van Lenthe, & Louwen, 2009) suggest that optimization in vacuum provides better reproducibility.

### 3.2.2 Segmented Surface

The segment surface is created by a vdW-like surface, which is determined by the union of spheres drawn around each of the atoms in the geometry-optimized molecule. Here, each element corresponding to the particular atom is given a specific radius, with which the spheres are created. These radii, which are determined empirically, are called the ‘COSMO radii’ and these are roughly 20% more than their vdW radii. The COSMO radii for 11 elements, namely H, C, N, O, F, Si, P, S, Cl and Br, have been optimized by Klamt et al. (1998). For unknown elements, Klamt and Schuurmann (1993) have recommended a value 20% larger than the vdW radius. The optimized radii of the elements were obtained from 15,000 DFT calculations (Klamt et al., 1998). Given this extensive optimization, the commercial packages directly use these data in their database (Banerjee, Singh, & Khanna, 2006; Grensemann & Gmehling, 2005).

Once the molecular surface or the molecule is defined, it is then divided into tiny polyhedra (segments), each with its own position and area. Figure 3.1
shows a sample segmented sphere. Each segment can be either a pentagon or a hexagon, but for computational affordability, each is assumed circular, with a radius, \( r_a = \sqrt{a_a/\pi} \).

In the initial DMOL3 implementation of COSMO, which was Klamt’s original DFT software, the average segment size is 0.24 Å².

### 3.2.3 Determination of Segment Properties

After the segmented surface is obtained by the union of all the vDW spheres outlined around each atom, where each atom possesses a ‘COSMO radii’, the molecule is then immersed in a homogeneous conductor having infinite dielectric constant (Klamt, 2005). In such a scenario, the screening charges from the conductor migrate to the segmented surface to negate the underlying molecular charge distribution. This is physically modeled by placing a screening charge on each segment. The total energy of solvation \( E_{solv} \) for the entire molecule is then determined from the interaction of the screening charges with the solute nuclei, the solute electronic density and within themselves. The segment properties are now calculated by defining the energy of solvation \( E_{solv} \), which solves the interaction of the segments with the solute nuclei \( Z \) (first term of Equation 3.1), solute electron density (second term of Equation 3.1) and interaction within the segments (third term of Equation 3.1), as given below. \( q_\alpha \) represents the screening charge of \( \alpha \)th segment.

\[
4\pi \varepsilon_0 E_{solv} = \sum_A \sum_\alpha Z_A q_\alpha B_{\alpha A} + \sum_\alpha q_\alpha C_\alpha + \frac{1}{2} \sum_\alpha \sum_\beta q_\alpha q_\beta D_{\alpha \beta} \tag{3.10}
\]

where:

\[
B_{\alpha A} = \frac{1}{|r_A - r_\alpha|}; \quad C_\alpha = \int_V V_\alpha \left( r \right) \rho \left( r \right) dr
\]

\[
D_{\alpha \beta} = \frac{1}{|r_\alpha - r_\beta|} \quad (\alpha \neq \beta); \quad D_{\alpha \alpha} = 1.07 \sqrt{\frac{4\pi}{a_\alpha}}
\]
Here, $B$, $C$, $D$ and $r$ represent vector quantities and represent the contributions due to Coulomb’s law. $Z_\alpha$ and $r_\alpha$ denote the molecular charge and position of nucleus $A$, respectively. $V_\alpha$, possessing a unit of inverse of length, signifies the distance between the segment and the inherent charge density $\rho(r)$ of the molecule. $D_{\alpha\alpha}$ represents the energy required to provide a charge on each segment.

The expression for the self-energy of a segment was made by assuming the molecule as a sphere of radius $R$. In such a situation, the total charge $Q$ is distributed homogenously across its surface. Hence, the electrostatic energy of the sphere is given by

$$E = \frac{Q^2}{8\pi\varepsilon_0 R}$$  \hspace{1cm} (3.11)

As per the COSMO protocol, the surface has to divide into $n$ homogeneously similar segments. This implies that the energy then becomes the total of the sum of segment interaction energies and the segment self-energies; this is given by

$$E = \sum_{\alpha=1}^{n} \sum_{\beta=1}^{n} E_{\alpha\beta} + \sum_{\alpha=1}^{n} E_{\alpha}^{self} = \sum_{\alpha=1}^{n} \sum_{\beta=1}^{n} \frac{q_\alpha q_\beta}{4\pi\varepsilon_0 r_{\alpha\beta}} + \sum_{\alpha=1}^{n} E_{\alpha}^{self}$$  \hspace{1cm} (3.12)

Here, $E_{\alpha\beta}$ is the interaction energy between segments, while $E_{\alpha}^{self}$ is the self-energy of segment. The latter reflects the energy required to maintain a charge of a given density on a particular segment. $q_\alpha$ denotes the charge of segment $\alpha$, while $r_{\alpha\beta}$ provides the distance between segments $\alpha$ and $\beta$. It is assumed that the area and charge of each segment are similar and the summations as above are exact. However, we still do not have an exact expression for $E_{\alpha}^{self}$, since the energy of the sphere must be the same and will not depend on the distribution of charge on the sphere. An explicit way to determine the same is to an expression for can be obtained by equating Equations 3.11 and 3.12. Therefore, by rearranging the above equations, we obtain (Burnett, 2012):

$$\frac{Q^2}{8\pi\varepsilon_0 R} = \frac{n}{8\pi\varepsilon_0} \sum_{\beta=2}^{n} \frac{Q^2}{n^2 r_{\beta}} + nE_{\alpha}^{self}$$  \hspace{1cm} (3.13)
By reorganizing the terms, we obtain:

\[
E_{\alpha}^{\text{self}} = \frac{Q^2}{8\pi\varepsilon_0 nR} \left[ n - \sum_{\beta=2}^n \frac{1}{\eta_{\beta}} \right] \quad (3.14)
\]

Here, \( R = \sqrt{A/4\pi} \) and is equal to \( \sqrt{a_n/4\pi} \), \( A \) is the surface area of the sphere and \( a_n = \sqrt{A/n} \) is the area of each segment. Thus, in a nutshell, Equation 3.14 is exact only when the charges are homogeneously placed on top of the sphere. However, as the charge and area are known from a COSMO calculation, this dependence is limited to only the last part of the equation, that is, \( \sqrt{1/n[n - \sum_{\beta=2}^n \eta_{\beta}]} \). The expression for this term is the highest (i.e. 1.250) at \( n = 4 \), while it steadily drops down to 1.078 at \( n = 20 \) and finally to 1.069 at \( n = 60 \). Hence, it has been recommended to keep this constant at \( n = 20 \) (i.e. 1.070) such that the final expression takes the form:

\[
E_{\alpha}^{\text{self}} = \frac{0.170}{8\pi\varepsilon_0} q_{\alpha} \sqrt{\frac{4\pi}{a_\alpha}} = \frac{0.15}{\varepsilon_0} \left( \frac{q_{\alpha}^2}{a_\alpha} \right) \quad (3.15)
\]

The energy of solvation \( (E_{\text{solv}}) \) is then minimized with respect to the segments’ charges \( (q_{\alpha}) \), as given below:

\[
\left( \frac{dE_{\text{solv}}}{dq_{\alpha}} \right) = \sum_A Z_A B_{A\alpha} + C_{\alpha} + \sum_\beta q_\beta D_{\alpha\beta} = 0 \quad (3.16)
\]

As the molecule is placed on the continuum with infinite dielectric constant, the charges in the nearby region tend to cancel the electron density \( \rho(r) \) of the solute molecule. This causes the molecule to polarize. Hence, this calculation needs to be performed in an iterative manner, so as to achieve consistency. At each step, the screening charge \( q \) is found for the respective \( \rho(r) \), after which \( \rho(r) \) is updated with new values of \( q \). This procedure, which utilizes
an iterative method, is called ‘self-consistent field’ approach. This process of generating and identifying the set of \( \rho(r) \) and \( q \) that are self-consistent is called a ‘conductor-like screening model’ (COSMO).

In addition to \( q \) and area of the segments \( a \) and \( r \), it also provides the potential for each segment \( \phi \). The calculation also provides us with the total energy of the molecule in a conductor, which we will denote as \( (E_{\text{COSMO}}) \). Sample screening charges for different molecules are depicted in Figure 3.2. It should be noted that the positive screening charges depict the negative regions of the molecule. Likewise, the negative screening charges represent the positive regions of the molecule. The screening charge distribution of the molecules is drawn by using the COSMO builder within TURBOMOLE. The package can also evaluate the segment potential for the segments, which we shall discuss in the ensuing section.

For the acetone COSMO file, ‘nps’ represents the number of segments, which is 739. The values of area \( (\text{Å}^2) = 371.79 \) and volume \( (\text{Å}^3) = 561.92 \) represent the total cavity area and volume, respectively, within the conductor, where acetone molecule is placed. The next section denotes the position of all the 10 atoms, along with the total screening charge on each atom ‘COSMOCharge’ and the last column represents the division of this ‘COSMOCharge’ with the area corresponding to each atom, called ‘area’. The last column is the most desirable quantity, as it reflects the screening charge density. It is interesting to note that except for oxygen atom ‘O3’ (COSMOCharge = 0.24046), all other atoms possess a negative charge. This follows the obvious trend, as negatively charged atoms possess positive screening charges, and vice versa. Further, the positive charge on the oxygen atom is then divided into a number of segments for which information is available in the concluding part of the COSMO file. For example, information for the segments of ‘C1’ and ‘H9’ atoms of acetone is depicted in Figure 3.3. The screening charges \( q \) of the solvent medium are usually calculated by a scaling \( q^* \), so that \( q = f(\epsilon) q^* \), where \( f(\epsilon) \) is the scaling factor given by

\[
f(\epsilon) = \frac{\epsilon - 1}{\epsilon + 0.5}
\]  

(3.17)
Gaussian COSMO output

$cosmo_data

fepsi = 1.00
nps = 739
area = 371.79
volume = 561.92

$coord_rad

#atom x y z element radius [A] area COSMOcharge sigma
1 2.42938323473619 -1.15096746670694 0.00000200604284 6 2.00000
18.67965 -0.01508 -0.00081
2 0.00000790190236 0.32678868411764 -0.000001189287756 6 2.00000
8.51712 -0.01501 -0.00176
3 0.00011174943123 2.62986595127750 0.00000320803667 8 1.72000
20.25649 0.24046 0.01187
4 -2.42950234678886 -1.15080747183769 -0.00000101959499 6 2.00000
18.67981 -0.01519 -0.00081
5 -2.53466744940604 -2.3615022845060 -1.65106704759713 1 1.30000
6.35161 -0.03540 -0.00557
6 -4.01918033927670 0.12745236116114 -0.00078429598930 1 1.30000
6.28553 -0.02677 -0.00426
7 -2.53529356852293 -2.3602533097624 1.65128747790871 1 1.30000
6.35202 -0.03540 -0.00557
8 2.53546863518548 2.36133231522858 1.65128747790871 1 1.30000
6.35152 -0.03540 -0.00557
9 4.01819033927670 0.12718985822214 -0.00049579434371 1 1.30000
6.28566 -0.02682 -0.00427
10 2.53505783077930 -2.36055542838593 -1.65182187102383 1 1.30000
6.35194 -0.03540 -0.00557

$screening_charge

<table>
<thead>
<tr>
<th>cosmo</th>
<th>correction</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td>-0.017099</td>
<td>0.017099</td>
<td>0.000000</td>
</tr>
</tbody>
</table>

$cosmo_energy

Total energy corrected [a.u.] = -193.0450861463
Dielectric energy corr. [a.u.] = -0.0105147256

FIGURE 3.3

An excerpt of a COSMO file, as generated from Gaussian03 version C.02 for acetone molecule.

(Continued)
Thus, a simpler boundary condition of a conductor appears in the above equation and also in Figure 3.3 ($\varepsilon_{fepsi} = 1.00$). It should be noted that a conductor has an infinite supply of charge, so as to screen the entire solute molecule. Hence, there is a point of thought that the selection of solvent as a conductor can alter the charges, as evident in Figure 3.2. Therefore, a decision has to be arrived for the selection of a solvent in such a manner so that it is able to screen the entire solute component. This brings out to the point that it should possess sufficient charge, and thus, for obvious reason, it has to be a perfect conductor. Here, the COSMO approximation is exact in the limit of $\varepsilon = \infty$ and is within 0.5% accuracy for strong dielectrics such as water with a permeability of ($\varepsilon = 80$). Even for a lower dielectric limit of solvents such as $\varepsilon = 2$, COSMO coincides with the exact dielectric model within 10%. Hence, the COSMO approach is becoming a standard Continuum Solvation Model (CSM) in quantum chemical codes.

It should be noted that the self-consistent field (SCF) calculation or the COSMO scheme has to be done once for each component. Once these charges are obtained, then a framework has to be created so as to restore these charges when placed in an actual solvent of a known dielectric constant. This is where the derivation of chemical potentials and activity coefficients is performed. Our aim will then be to create a repository of COSMO files, which will include ions, solutes, solvents and inorganic compounds, subject to their dissociation constants. A sample description of an input for generating the COSMO file is given in Figure 3.4.

The different terms and their meaning are as follows:

‘% mem = 540 MW’ refers to the total amount of internal memory required to perform this calculation. The route section with ‘# P BVP86/SVP/DGA1’
indicates the following form ‘DFT Theory/Basis Set/optional’. So, the form of functional used for DFT is the P BVP86, while the basis set used is SVP or Split Valence Polarized. The orbital coefficients for SVP can be obtained by the GFprint and GFInput commands in Gaussian03. The density gradient approximation or ‘optional’ used is DGA1. It expands the electron density of the atoms in the form of atom-centered functions for saving computational time. It only integrates the coulomb interaction in place of all the electron integrals. The ‘SCF = tight’ indicates a full convergence of energy.

3.2.4 Computation of the Restoring Free Energy \( \Delta G_{i/s}^{res} \)

Here, the contribution to the restoring free energy can be divided into two parts, namely \( \Delta G_{i/s}^{res} \) and \( \Delta G_{i}^{iS} \), where the former implies the electrostatic energy due to the component in the solution and the latter explains the change in energy when the molecule is taken from a conductor to a solvent. These are mathematically defined as below:

\[
\Delta G_{i/s}^{res} = \Delta G_{i/s}^{res} + \Delta G_{i}^{iS}
\]

Both the energies in COSMO conductor and vacuum optimization can be obtained as per the procedures mentioned in Sections 3.2.2 and 3.2.1, respectively. For example, the total energy in the conductor, as per Figure 3.3, is
'Total energy corrected [a.u.] = −193.0450861463'. \( \Delta G_{\text{res}}^{\text{IS}} \) represents the restoring free energy for the movement of the solute from the conductor to real solvent. For a real conductor, \( \Delta G_{\text{res}}^{\text{IS}} = 0 \), which implies that \( \Delta G_{\text{res}}^{\text{IS}} = \Delta G_{\text{IS}}^{\text{IS}} \). Hence, the major portion from now on will be the Segment Activity Coefficient (SAC) approach, which will involve the derivation of segment potential, as discussed in the next section.

### 3.2.5 Segment Chemical Potential

Based on statistical thermodynamics, COSMO-SAC assumes the ensemble of molecules by an ensemble of independently interacting segments (Lin and Sandler, 2002; Burnett, 2012). Hence, for any mixture, segments from each component are mixed in proportion with the desired component mole fractions. The resulting interaction is then captured by the energy required to bring them from an infinite point in space. Let us consider an ensemble of \( k \) different types of molecules, each having \( n_i \) segments. Hence, the total number of segments is

\[
\eta_c = \sum_{i=1}^{k} n_i \tag{3.19}
\]

\[
\eta^{\text{tot}} = \sum_{i=1}^{k} b_i n_i = \sum_{\alpha=1}^{n_c} \eta_\alpha
\]

Here, \( b \) represents the total number of molecules of type \( i \) and \( \eta_\alpha \) represents the total number of segments of type \( \alpha \). For a single component, this reduces to \( \eta_\alpha = \eta_b = b_i \). Assuming a constant geometry, where all the segments are allowed to move freely, implies that the movements of the segments are not dependent on neighboring segments. This ensemble then takes a form of mixture of independently interacting segments. We can thus propose any such formation as below (Burnett, 2012):

\[
\begin{bmatrix}
2\eta_{1,1} + \eta_{1,2} + \eta_{1,3} + \ldots \ldots + \eta_{1,n_c} \\
\eta_{2,1} + 2\eta_{2,2} + \eta_{2,3} + \ldots \ldots + \eta_{2,n_c} \\
\ldots \\
\ldots \\
\eta_{n_i,1} + \eta_{n_i,2} + \eta_{n_i,3} + \ldots \ldots + 2\eta_{n_i,n_c}
\end{bmatrix}
= \begin{bmatrix}
\eta_1 \\
\eta_2 \\
\ldots \\
\ldots \\
\eta_{n_c}
\end{bmatrix}
\]

Each row in the above matrix notation denotes the interaction of the first with the remaining segments (same component or of other component). Hence, the total neighbors will be of type ‘I’ and will be \( \eta_i \) such that \( \eta_{\text{total}} = \sum_{i=1}^{n_c} \eta_i \). So, from the same logic, the number of pairs in the mixture will be
\[
n_{\text{pairs}} = \sum_{\alpha=1}^{n_c} n_{\alpha\alpha} + \frac{1}{2} \sum_{\alpha=1}^{n_c} \sum_{\beta=1}^{n_c} n_{\alpha\beta} = \frac{1}{2} \sum_{\alpha=1}^{n_c} n_{\alpha} = \frac{n_{\text{total}}}{2} \quad (3.21)
\]

Thus, the probability of obtaining a segment pair \(n_{\alpha\beta}\) will be

\[
p_{\alpha\beta} = \frac{\langle n_{\alpha\beta} \rangle}{n_{\text{pairs}}} \quad (3.22)
\]

Here, \(\langle n_{\alpha\beta} \rangle\) is an ensemble averaged value approximated from a canonical ensemble. Hence, the whole mixture can be written in terms of a partition function \(Q(N, V, T)\) with a Boltzmann exponential term (see Appendix I for canonical partition function). This is given as below

\[
p_{\alpha\beta} = \omega_{\alpha\beta} \exp \left( \frac{-E_{\alpha\beta}}{kT} \right) \frac{Q\left(n_{\text{pairs}} - n_{\alpha\beta}, V, T\right)}{Q\left(n_{\text{pairs}}, V, T\right)} \quad (3.23)
\]

Here, \(Q\) and \(Q'\) represent the partition functions of the mixture with and without including the segment pair, respectively. Here, \(E_{\alpha\beta}\) represents the interaction between the segment pair \(\alpha\) and \(\beta\). Thus, when \(\alpha = \beta\), then \(\omega_{\alpha\beta} = 1\), since both the segments become indistinguishable. The value of the degeneracy or \(\omega_{\alpha\beta}\) becomes 2 when \(\alpha \neq \beta\). Equation 3.23 takes the following form when average probability is defined:

\[
p_{\alpha\beta} = \omega_{\alpha\beta} \exp \left( \frac{-E_{\alpha\beta}}{kT} \right) \frac{Q\left(n_{\text{total}} - \chi_{\alpha} - \chi_{\beta}, V, T\right)}{Q\left(n_{\text{total}}, V, T\right)} \quad (3.24)
\]

The difference between Equations 3.23 and 3.24 is that they denote a segment pair and an individual segment, respectively. It is a known fact the Helmholtz energy for the ensemble having the two segments \(\chi_{\alpha}, \chi_{\beta}\) takes the form:

\[
A\left(n_{\text{total}} - \chi_{\alpha} - \chi_{\beta}, V, T\right) = -kT \ln Q\left(n_{\text{total}} - \chi_{\alpha} - \chi_{\beta}, V, T\right) \quad (3.25)
\]

From statistical thermodynamics, we know:

\[
\left( \frac{\partial G}{\partial N_i} \right)_{T,P,N_{\text{pairs}}} = \left( \frac{\partial A}{\partial N_i} \right)_{T,V,N_{\text{pairs}}} = -kT \left( \frac{\partial \ln Q}{\partial N_i} \right)_{T,V,N_{\text{pairs}}} = \mu_i \text{(chemical potential)} \quad (3.26)
\]

Hence, Equation 3.25 takes the form:

\[
A\left(n_{\text{total}} - \chi_{\alpha} - \chi_{\beta}, V, T\right) = A\left(n_{\text{total}}\right) - \mu_{\alpha} - \mu_{\beta} \quad (3.27)
\]
Here, $\mu_\alpha$ and $\mu_\beta$ are the chemical potentials for the two segments $\alpha$ and $\beta$, respectively. Hence, we have:

$$Q_{(n_{\text{total}} - \chi_\alpha - \chi_\beta)} \frac{\nabla}{V, T} = \exp\left(-A(n_{\text{total}} - \chi_\alpha - \chi_\beta, V, T)\right)$$

$$\Rightarrow \exp\left(-A(n_{\text{total}}, V, T) - A(n_{\text{total}} - \chi_\alpha - \chi_\beta, V, T)\right)$$

$$\Rightarrow \exp\left(\frac{\mu_\alpha + \mu_\beta}{kT}\right)$$

Equation 3.24 then takes the form:

$$p_{\alpha\beta} = \omega_{\alpha\beta} \exp\left(-\frac{E_{\alpha\beta}}{kT}\right) \exp\left(\frac{\mu_\alpha + \mu_\beta}{kT}\right)$$

The non-translational part of the chemical potential is given as

$$\mu^* = \mu - kT \ln \left(\rho^3\right)$$

This implies that the second part is the translational contribution for chemical potential. So, for ideal gases, the non-translational part is usually zero, or

$$\mu^* = \mu - kT \ln \left(\rho^3\right) \Rightarrow \mu^* = kT \ln \left(\rho^3\right)$$

The final form of Equation 3.29 takes the form:

$$p_{\alpha\beta} = \omega_{\alpha\beta} \exp\left(-\frac{E_{\alpha\beta}}{kT}\right) \exp\left(\frac{\mu_\alpha^* + \mu_\beta^*}{kT}\right)$$

For obtaining the average number of ensemble pairs, we thus use the relation as below:

$$\chi_{\alpha\beta} = n_{\text{pairs}} p_{\alpha\beta} = n_{\text{pairs}} \omega_{\alpha\beta} \exp\left(-\frac{E_{\alpha\beta}}{kT}\right) \exp\left(\frac{\mu_\alpha^* + \mu_\beta^*}{kT}\right)$$

The total number of degrees of freedom for such a system can be shown to be equal to $n^c/2(n^c + 3)$; hence, we require the same number of equations so as
to solve them and get a unique solution. Rewriting Equation 3.22 by diving it with $\omega_{\alpha\beta}$, we obtain:

$$\sum_{\beta=1}^{n_c} \frac{p_{\alpha\beta}}{\omega_{\alpha\beta}} = \left[ \frac{\eta_{\alpha 1}}{2} + \frac{\eta_{\alpha 2}}{1} + \ldots \frac{\eta_{\alpha c}}{1} + \ldots + \frac{\eta_{\alpha \eta c}}{2} \right] \frac{1}{n_{\text{pairs}}}$$

(3.34)

Here, the total probability is divided with the degeneracy and finally with total number of segment pairs. In a way, the above expression represents all possible probabilities of segment $\alpha$ with remaining segments across all the components $c$.

$$= \left[ \eta_{\alpha 1} + \eta_{\alpha 2} + \ldots + \eta_{\alpha c} \right] \frac{1}{n_{\text{total}}}$$

(3.35)

Equation 3.35 points out to the fact that the total interaction with the remaining segments merely points out to the total number of pairs with segment $\alpha$. Hence, we obtain $n$:

$$\Rightarrow \frac{\eta_{\alpha}}{n_{\text{total}}} = x_{\alpha}$$

(3.36)

Further, Equation 3.36 points out to the fact that a probability of locating a particular segment pair in the ensemble is merely locating a single pair or the segment mole fraction $x_{\alpha}$. When applied to Equation 3.29, it takes the form:

$$\sum_{\beta=1}^{n_c} \frac{p_{\alpha\beta}}{\omega_{\alpha\beta}} = \sum_{\beta=1}^{n_c} \frac{\omega_{\alpha\beta}}{\omega_{\alpha\beta}} \exp \left( \frac{-E_{\alpha\beta}}{kT} \right) \exp \left( \frac{\mu_{\alpha}^* + \mu_{\beta}^*}{kT} \right)$$

(3.37)

$$\Rightarrow \exp \left( -\frac{\mu_{\alpha}^*}{kT} \right) \sum_{\beta=1}^{n_c} \exp \left( \frac{-E_{\alpha\beta}}{kT} \right) \exp \left( \frac{\mu_{\beta}^*}{kT} \right)$$

(3.38)

By combining the above two expressions, we obtain:

$$x_{\alpha} \Rightarrow \exp \left( -\frac{\mu_{\alpha}^*}{kT} \right) \sum_{\beta=1}^{n_c} \exp \left( \frac{-E_{\alpha\beta}}{kT} \right) \exp \left( \frac{\mu_{\beta}^*}{kT} \right)$$

(3.39)

Rearranging the equation provides the following form:

$$\exp \left( -\frac{\mu_{\alpha}^*}{kT} \right) \Rightarrow x_{\alpha} \left( \sum_{\beta=1}^{n_c} \exp \left( \frac{-E_{\alpha\beta}}{kT} \right) \exp \left( \frac{\mu_{\beta}^*}{kT} \right) \right)^{-1}$$

(3.40)
This can be rewritten as

$$\Gamma_\alpha = x_\alpha \left( \sum_{\beta=1}^{n_c} \tau_{\alpha\beta} \right)^{-1}$$

(3.41)

where:

$$\Gamma_\beta^* = \exp \left( \frac{-\mu_\alpha^*}{kT} \right) \quad \text{and} \quad \tau_{\alpha\beta} = \exp \left( \frac{-E_{\alpha\beta}}{kT} \right)$$

(3.42)

Thus, the equation is implicit in nature, as $\mu_\alpha^*$ appears on both sides of the equation. Hence, this needs to be solved iteratively. The $E_{\alpha\beta}$ possesses the form as given below:

$$E_{\alpha\beta} = f_{\text{pol}} \frac{0.15a_{\text{eff}}^{3/2}}{e_0} (\sigma_\alpha + \sigma_\beta)^2$$

(3.43)

Here, $a_{\text{eff}}$ is the area of the effective segment and is an adjustable parameter. $\sigma_\alpha = q_\alpha/a_a$ is charge density of segments. Here, $c_{es}$ is a constant and is equal to $f_{\text{pol}} 0.15a_{\text{eff}}^{3/2}/e_0$. The equation above reflects the self-consistency approximation, as discussed earlier. Here, all the segments are assumed to be of the same size as per ‘area of the effective segment; hence, the segment pair itself can be assumed to be an individual identity, with a total charge as the sum of the two charge segments. As the two molecules interact with each other, their electronic densities and subsequently, the wave function get disturbed. To capture this effect, the factor $f_{\text{pol}}$ is defined.

Now, we shall move to the most important aspect of thermodynamics, that is, the derivation of the chemical potential from the information available from the segment charge densities. In such a scenario, we shall assume the entire mixture of $N$ molecules to be divided into number of segments, $n_1, n_2 \ldots n_n$. Assuming that we have large number of molecules implies that removal of a single molecule or segment will define another system. An analogy can be devised from an NPT ensemble, where a deduction of a particular molecule (mol$_i$) or a segment (seg$_\alpha$) gives a pathway for calculating the chemical potential of the particular molecule or a segment, respectively.

$$G(N-\text{mol}_i,P,T) = G(N,P,T) - \mu_{\text{mol}_i}$$

Here, $\mu_{\text{mol}_i}$ refers to the chemical potential of the $i$th molecule from the mixture. Now, if we consider the same ensemble but in terms of segments (seg$_\alpha$), we obtain (Burnett, 2012):
\[ G(N - \text{seg}_a, P, T) = G(N, P, T) - \mu_{\text{seg},a} \]  \hspace{1cm} (3.44)

So, by summing up the contributions from all the segments, we obtain:

\[ \mu_i = \sum_{\alpha=1}^{N} \mu_{\alpha} \]  \hspace{1cm} (3.45)

So, the total energy or the restoring free energy required will be equal to the sum of all segments in a mixture \( S \). This is given as below:

\[ \Delta G_{i/S}^{\text{res}} = \sum_{\alpha=1}^{n_i} \mu_{a/S}^* \]  \hspace{1cm} (3.46)

However, Equation 3.44 shall also be valid when the solvent itself is a perfect conductor. This will be applicable for the following condition:

\[ \Delta G_{nx/kT}^{i/S} = \sum_{\alpha=1}^{n_i} n_i^{\text{eff}} x_{\alpha/i} \left( \mu_{a/S}^* - kT \ln x_{a/S} \right) \]  \hspace{1cm} (3.47)

Here, \( n_i^{\text{eff}} \) denotes the effective number of segments of component \( I \), while \( n_i^{\text{eff}} x_{\alpha/i} \) gives us the effective number of segments of molecule \( I \) of type \( \alpha \). So, when it becomes a perfect conductor, then Equation 3.47 takes the form:

\[ \Delta G_{i/S}^{\text{res}} = 0 \hspace{1cm} (3.48) \]

\[ \sum_{\alpha=1}^{n_i} n_i^{\text{eff}} x_{\alpha/i} \left( \mu_{a/S}^* - kT \ln x_{a/S} \right) = 0 \]

Thus, we have to subtract the term \( kT \ln x_{a/S} \) from the chemical potential to obtain the non-translational contribution.

\[ \mu_{a/S}^* - kT \ln x_{a/S} = 0 \hspace{1cm} (3.49) \]

\[ \Rightarrow \mu_{a/S}^* = kT \ln x_{a/S} \]

This merely reflects the fact that for a perfect conductor, the restoring free energy is zero and the chemical potential is due to its translational contribution. Equation 3.47 is based on the premises that all segment interactions are possible. In order to assume such information, we need to neglect the three-dimensional segment information. Further, no attempt has been made to consider the steric hindrance or the long-range liquid structure. This may be termed the most limiting assumption in the model. On the contrary, this assumption only makes the model simplistic in nature, as it eliminates the need to explicitly sample the full configuration space of the entire ensemble.
In the stated COSMO calculation, the surface of the molecule is dissected in small segments and screening charges are determined for each segment, such that the net potential everywhere at the surface is zero (perfect screening). These charges are averaged in order to make the predictions uniform and also to capture the different shapes and sizes of the segments uniformly. Thus, to obtain apparent screening charges to be used in COSMO-SAC model, the averaging algorithm is proposed as below:

\[
\sigma_m = \frac{\sum_n \sigma_n \frac{r_n^2 r_{\text{eff}}^2}{r_n^2 + r_{\text{eff}}^2} \exp \left( -\frac{d_{mn}^2}{r_n^2 + r_{\text{eff}}^2} \right)}{\sum_n \frac{r_n^2 r_{\text{eff}}^2}{r_n^2 + r_{\text{eff}}^2} \exp \left( -\frac{d_{mn}^2}{r_n^2 + r_{\text{eff}}^2} \right)}
\]  

(3.50)

where:

- \( r_{\text{eff}} = \sqrt{d_{\text{eff}}/\pi} \) is the radius of the standard surface segment
- \( r_n = \sqrt{a_n/\pi} \) is the radius of the segment \( n \)
- \( d_{mn} \) is the distance between segments \( m \) and \( n \)

The averaging of the screening charge densities, therefore, creates a new contribution to the overall Gibb’s free energy. This is provided as below:

\[
\Delta G_{i/0}^{\text{res}} = \Delta G_{i/0}^{\text{res}} + \Delta G_{i}^{\text{Ic}} + \Delta G_{i}^{\text{cc}}
\]  

(3.51)

Here, the third term on the right-hand side is the Gibb’s free energy change occurring due to the averaging process.

\[
\Delta G_{cc}^{*} = f_{\text{pol}}^{1/2} \left[ E_i^{\text{die}}(q) - E_i^{\text{die}}(q^*) \right]
\]  

(3.52)

Thus, the first term, that is, the restoring free energy, is a function of the solvent, while the remaining terms are dependent on the molecule configuration. Here, \( E_i^{\text{die}} \) is given by

\[
E_i^{\text{die}}(q) = \frac{1}{2} \sum_{\mu} \phi_{i\mu} q_{i\mu}
\]  

(3.53)

\( E_i^{\text{die}}(q^*) \) is directly obtained from COSMO output file and is the energy needed for dielectric polarization, that is, the screening charge distribution. The factor has a constant value of \( f_{\text{pol}} = 0.6917 \). These two free energies, namely the second and third terms of Equation 3.51, are pure species properties. So, they have the same values in ideal gas phase (i.e. in vacuum) and in the solvent. Mixture properties are thus obtained by subtracting free energy values of mixture and vacuum phase respectively. This finally leads to the cancellation for mixture property calculation. This is the reason that they are not included in the final expression in solvation free energy.
3.2.6 Sigma Profiles of Simple Compounds

The three-dimensional screening charge density distributions, which are obtained as per the procedure given in Figure 3.3, are first averaged using Equation 3.50. This gives us charges ranging from \(-0.03\) e/Å\(^2\) to \(+0.03\) e/Å\(^2\). These segment charge densities are grouped using a histogram, which is commonly known as \(\sigma\)-profile (\(p(\sigma)\)). This essentially is the probability of locating a surface segment with screening charge density \(\sigma\) by the expression:

\[
p_i(\sigma) = \frac{A_i(\sigma)}{A}
\]

(3.54)

where:
- \(A_i(\sigma)\) is the surface area with a charge density of value \(\sigma\)
- \(A_i\) is the total surface area of species \(i\)

For a mixture, the \(\sigma\)-profile is determined from the area-weighted average of contributions from all the components present in the mixture. This is given as

\[
p_\text{s}(\sigma) = \frac{\sum_i x_i A_i p_i(\sigma)}{\sum_i x_i A_i}
\]

(3.55)

Since the current theme of this book is on ionic liquids (ILs), we shall introduce sigma profiles for simple components and then move towards ILs. It should be noted that computation of segment potential is quite a cumbersome process, concerning the number of segments a molecule possesses. A different and easier approach would be to use the histogram known as \(\sigma\)-profile or \(p(\sigma)\). In such a scenario, Equation 3.50 takes the form:

\[
\exp\left(-\frac{\mu^i(\sigma)}{kT}\right) = p_\text{s}(\sigma) \sum_{\beta=1}^{N_\text{bins}} \exp\left(-\frac{E_{\alpha\beta}}{kT}\right) \exp\left(-\frac{\mu^\ast_s(\sigma)}{kT}\right)
\]

(3.56)

\[
\Delta G_{i/S}^{\text{res}} = n_i \left\{ \sum_{\beta=1}^{N_\text{bins}} p_i(\sigma) \left[ \mu^\ast_s(\sigma) - kT \ln p_\text{s}(\sigma) \right] \right\}
\]

(3.57)

Here, \(\sigma_a\) refers to the charge density within the histogram. Now, it is known that the histograms lie between \(-0.03\) e/Å\(^2\) and \(+0.03\) e/Å\(^2\). So, if we allocate each histogram with a value of 0.001 e/Å\(^2\), we shall then reduce the whole sigma profile to 61 histograms. Hence, \(N_\text{bins}\) takes the number 61. In the above altered equations, \(x_{a/j}\) and \(x_{a/S}\) are invariably replaced by \(p_i(\sigma_a)\) and \(p_5(\sigma_a)\), respectively. The latter two determine the sigma profiles of the component in itself (Equation 3.54) and in the mixture (Equation 3.55), respectively.
However, $x_\alpha$ is segment mole fraction, while $p_S(\sigma_\alpha)$ is based on the segment area fraction. Notwithstanding as Equation 3.47 is not derived, the difference is usually very small when computing $\Delta G_{\alpha/S}^{\text{res}}$ from $\mu_{\alpha/S}$.

It is interesting to note that the exchange energy ($E_{\text{pair}}$) contains contributions from electrostatic interactions. One of them is referred as misfit energy $E_{mf}$ which is the energy computed due to the difference in shape and size between a pair of segments. Another contribution, the hydrogen bonding energy $E_{hb}$ occurs when an electropositive atom (hydrogen) forms additional bonds with electronegative atoms such as oxygen, nitrogen and fluorine. The expressions of the two terms are given by Equations 3.58 and 3.59, respectively.

$$E_{mf} = (\alpha'/2)(\sigma_\alpha + \sigma_\beta)^2$$

$$E_{hb} = c_{hb} \min \left[ 0, \max (0, \sigma_{\text{acc}} - \sigma_{hb}) \min (0, \sigma_{\text{don}} + \sigma_{hb}) \right]$$

$$E_{\text{pair}}(\sigma_\alpha, \sigma_\beta) = E_{mf} + E_{hb}$$

Where, $\alpha'$ is a constant for the misfit energy and is calculated from the surface area of a standard segment.

$$\alpha' = \frac{0.64 \times 0.3 \times a_{\text{eff}}^{3/2}}{e_0}$$

$$e_0 = 2.395 \times 10^{-4} (e^2 \text{ mol})/(\text{kcal} \ \text{Å})$$

Here, $c_{hb}$ is a constant for the hydrogen bonding interactions, while $\sigma_{hb}$ is cutoff value for hydrogen bonding interactions. When Equation 3.59 is used, all possible pairs are first grouped. Now, each pair is dealt with separately, having screening charge densities as $\sigma_\alpha$ and $\sigma_\beta$. Then, $\sigma_{\text{acc}}$ and $\sigma_{\text{don}}$ are allocated the larger and smaller values of $\sigma_m$ and $\sigma_n$. Here max and min indicate the larger and smaller values of their arguments, respectively. Under this definition, the hydrogen bonding contribution is non-zero only if one segment has a negative charge density less than $|\sigma_{hb}|$ and the other has a positive charge density greater than $\sigma_{hb}$. In such a case, the contribution is limited to segment pairs of opposite charge possessing larger $|\sigma_{hb}|$ magnitudes. This contribution will also increase higher difference in charges densities between the pair of segments. The SAC then takes up the form:

$$\ln \Gamma_s(\sigma_\alpha) = -\ln \left\{ \sum_{\sigma_\alpha} p_S(\sigma_\alpha) \Gamma_s(\sigma_\alpha) \exp \left[ -\frac{E_{\text{pair}}(\sigma_\alpha, \sigma_\beta)}{kT} \right] \right\}$$

(3.61)

This equation is then solved iteratively. In the next section, we will focus our attention on the sigma profile and potential of the commonly used cations and anions.
3.2.7 Sigma Profiles and Potentials of Ionic Liquids

The sigma profile and potential of cations and anions are obtained using Equations 3.55 and 3.56, respectively (Figures 3.5 and 3.6). For such an exercise, we have chosen a common cation, that is, 1-alkyl-3-methylimidazolium

![Figure 3.5: Sigma profiles of individual cations](image1)

**FIGURE 3.5**
Sigma profiles of individual cations (EMIM: 1-ethyl-3-methylimidazolium; BMIM: 1-butyl-3-methylimidazolium; OMIM: 1-Octyl-3-methylimidazolium).

![Figure 3.6: Sigma potential of individual cations](image2)

**FIGURE 3.6**
Sigma potential of individual cations.
(\[RMIM\]; R=O, H, B etc.; O:Octyl; H:Hexyl; B:Butyl) (Figure 3.5) and then paired it up with three common anions, as given in Figure 3.7.

The sigma profiles and potentials for cations are shown in Figures 3.5 and 3.6, respectively. In all the normalized profiles, a small part of the sigma profile lies to the left of the hydrogen bonding cutoff radius (i.e. \(\sigma < -0.0082 \text{ e/Å}^2\)), indicating a hydrogen bonding donor capacity. The sigma profiles for \([BMIM]\) and \([OMIM]\) are of the same nature. Most of the prominent peaks lie in the negative side of the sigma profile. The negative screening charges are due to the positive charge residing inside the imidazolium ring. \([EMIM]\), \([BMIM]\), \([OMIM]\), all show peaks at the outmost position in the negative direction. They also show up in the sigma potential (Figure 3.6), that is, negative values of chemical potentials are encountered on the positive side of the screening charge densities (SCDs). The negative values are encountered since extra free energy is gained by forming hydrogen bonds. It should be noted that extra energy is gained with the formation of hydrogen bonds; this energy is negative in value when compared with the free energy required (which is positive in nature) for removing the SCDs.

The sigma profiles and potentials for anions are shown in Figures 3.7 and 3.8, respectively. For the anions, it is evident that a negative contribution to the sigma potential is due to positive screening charges, as obtained from their sigma profiles. It should be noted that the contribution of chemical or segment potential is a sum of two values, that is, the misfit term and the hydrogen bond term. For the anions, a negative value on the profile implies that it is the hydrogen-acceptor bonding that dominates the left-hand side of

**FIGURE 3.7**
Sigma profile of individual anions.
Figure 3.7. However, this is the opposite in the donor region, where the peak rises sharply owing to the difference in the shape and size of the segments or the misfit contribution is known to dominate the total interaction.

For all the anions, peaks are lying on the right of the cutoff zone for hydrogen bonding, that is, $\sigma_{h||} = +0.0082 \, e/\text{Å}^2$. All the prominent peaks for the anions, that is, 0.0085 e/Å$^2$ for [PF$_6$], 0.014 e/Å$^2$ for [CF$_3$SO$_3$], 0.015 e/Å$^2$ for [EtSO$_4$] and 0.011 e/Å$^2$ for [BF$_4$], lie to the right of the cutoff SCD, that is, $\sigma_{h||} = +0.0082$. This is due to the inherent negative charges of the anions. The sigma potential is a similar one as observed for cations, except that negative values of chemical potentials are encountered on the negative side of the SCDs.

### 3.2.8 Restoring Free Energy

Finally, we obtain the restoring free energy of molecule in solvent or mixture as

$$\frac{\Delta G_{\text{res}}^{\Gamma}}{RT} = n_i \sum_{\alpha} p_i (\sigma_{\alpha}) \ln \Gamma_s (\sigma_{\alpha})$$  \hspace{1cm} \text{(3.62)}$$

This original definition was given by Klamt et al. (1998). Although this gave significant improvement for in hydrogen bonded systems, it was not fully consistent with the experimental observations. It should be noted that such policy made the hydrogen bonding contribution biased towards
acceptors, which are limited to small electronegative atoms such as nitrogen, oxygen and fluorine. On the other hand, the donors originates from the hydrogen atoms which are in turn bonded to these electronegative atom(s). Hence, the process needs some course correction, as it leads to occasional meaningful errors. These weaknesses leave the process prone to occasional errors.

To incorporate a more meaningful hydrogen bond model, the definitions of acceptor and donor were modified. We assumed positively charged segments from the surfaces of all nitrogen, oxygen and fluorine atoms to be considered as acceptors. Similarly, the negatively charged segments from the surfaces of hydrogen atoms bound to any one of the acceptor atoms were chosen as HB donors. The modification will be taken up with test cases in Chapter 5 of this book. Now let us revert to Equation 3.7, which is given below:

\[
\frac{\Delta G_{i/S}^{ES} - \Delta G_{i/i}^{ES}}{kT} + \left( \frac{\Delta A_{i/S}^{vdW} - \Delta A_{i/i}^{vdW}}{kT} \right) + \left( \frac{\Delta A_{i/S}^{cav} - \Delta A_{i/i}^{cav}}{kT} \right) + \ln \left( \frac{V_{i/S}}{V_{i/i}} \right)
\]

Here, we have obtained relation for the electrostatic contribution (first part of Equation 3.7, as above).

Now, we move our attention to the second term, that is, the dispersion or the van der Waals expression. This is the contribution due to the Helmholtz free function, which is modeled using a first order mean-free approach, the perturbation approach. For pure component, the expression takes the form:

\[
\frac{\Delta A_{i/i}^{disp}}{NkT} = \frac{1}{kTV_{mix}} \sum_{i} \sum_{k} \varepsilon_{jk} m_{j}^{i} m_{k}^{i}
\]

Here, \( \varepsilon_{jk} = \sqrt{\varepsilon_{ij} \varepsilon_{ik}} \) (K Å\(^3\)) is the pair interaction energy between atom type \( j \) and atom type \( k \), while \( m_{j}^{i} \) is the effective number of atoms of type \( j \) within a species \( i \) molecule, which is obtained from:

\[
m_{j}^{i} = \sum_{a \in i} \left( \frac{S_{a}}{S_{a0}} \right)^{q_{s}}
\]

Where, \( S_{a} \) is the solvent accessible area of atom \( a \) and \( S_{a0} \) is the bare surface area calculated using the set of atomic radii in the quantum chemical calculations. \( q_{s} \) is the scaling factor in the range from 0 to unity. \( k \) is the Boltzmann constant in J/(molecule K), while \( V \) is the volume in Å\(^3\)/molecule. Here, \( r \) denotes the
atomic type: C, H, N, O, F, Cl, S, P, Br or I. For mixtures, it takes the following form (the interaction here is considered using all possible element pairs):

$$\frac{\Delta A_{\text{mix}}^{\text{disp}}}{NkT} = -\frac{1}{TV_{\text{mix}}} \sum_{\alpha=\beta=a} \sum_{i=\alpha}^{r} \sum_{j=\beta}^{r} x_{\alpha} x_{\beta} m_{i} m_{j} \sqrt{\epsilon_{i} \epsilon_{j}}$$

(3.66)

Here, $\Delta A_{\text{mix}}^{\text{disp}}$ is the dispersion free energy for the mixture in joules. $N$ is the number of molecules, where $N = \sum N_{\alpha}$ and $V_{\text{mix}}$ the volume of the ideal mixture given by

$$V_{\text{mix}} = \sum_{\alpha} x_{\alpha} V_{\alpha}$$

(3.67)

$x_{\alpha}$ is the mole fraction given by: $x_{\alpha} = N_{\alpha}/N$. Here, $i$ and $j$ sum over all the components bond types, namely $\alpha$ and $\beta$. $a$ refers to the number of components in the mixture. This contribution is used to account for all possible interactions between the atoms in different ensembles. For a binary mixture (Wang et al., 2007):

$$\frac{\Delta A_{\text{mix}}^{\text{disp}}}{NkT} = -\frac{1}{TV_{\text{mix}}} \left[ x_{a}^{2} \sum_{i=\alpha}^{r} \sum_{j=\alpha}^{r} m_{i} m_{j} \sqrt{\epsilon_{i} \epsilon_{j}} + 2x_{a} x_{b} \sum_{i=\alpha}^{r} \sum_{j=\beta}^{r} m_{i} m_{j} \sqrt{\epsilon_{i} \epsilon_{j}} + x_{b}^{2} \sum_{i=\beta}^{r} \sum_{j=\beta}^{r} m_{i} m_{j} \sqrt{\epsilon_{i} \epsilon_{j}} \right]$$

(3.68)

This can be reduced as

$$\frac{\Delta A_{\text{mix}}^{\text{disp}}}{NkT} = -\frac{1}{TV_{\text{mix}}} \left[ x_{a}^{2} E_{aa} + 2x_{a} x_{b} E_{ab} + x_{b}^{2} E_{bb} \right]$$

(3.69)

Here

$$E_{aa} = \sum_{i=\alpha}^{r} \sum_{j=\alpha}^{r} m_{i} m_{j} \sqrt{\epsilon_{i} \epsilon_{j}} \quad E_{ab} = \sum_{i=\alpha}^{r} \sum_{j=\beta}^{r} m_{i} m_{j} \sqrt{\epsilon_{i} \epsilon_{j}}$$

(3.70)

The dispersion term of the residual chemical potential for species $a$ and $b$ is given as

$$\ln \gamma_{i}(\text{disp}) = \frac{\mu_{i}^{\text{disp}}}{kT}$$

(3.71)
Hence, we obtain:

\[
\frac{\mu_{a}^{\text{disp}}}{kT} = \frac{1}{kT} \left( \frac{\partial \Delta A_{\text{mix}}^{\text{disp}}}{\partial N_{a}} \right)_{T,V,N_{\text{arb}}} \\
\Rightarrow -\frac{1}{T} \left[ \frac{V_{a}^{*}}{V_{\text{mix}}^{*}} \left( x_{a}^{2}E_{a}^{i} + 2x_{a}x_{b}E_{ab}^{i} + x_{b}^{2}E_{bb}^{i} \right) - 2 \left( \frac{x_{a}E_{a}^{i} + x_{b}E_{ab}^{i}}{V_{\text{mix}}^{*}} \right) \right]
\]

(3.72)

Similarly, for bond type \( b \), we have:

\[
\frac{\mu_{b}^{\text{disp}}}{kT} = \frac{1}{kT} \left( \frac{\partial \Delta A_{\text{mix}}^{\text{disp}}}{\partial N_{b}} \right)_{T,V,N_{\text{arb}}} \\
\Rightarrow -\frac{1}{T} \left[ \frac{V_{b}^{*}}{V_{\text{mix}}^{*}} \left( x_{b}^{2}E_{b}^{i} + 2x_{a}x_{b}E_{ab}^{i} + x_{b}^{2}E_{bb}^{i} \right) - 2 \left( \frac{x_{b}E_{b}^{i} + x_{b}E_{ab}^{i}}{V_{\text{mix}}^{*}} \right) \right]
\]

(3.73)

For large complex molecules, additional corrections are required, which depend on the molecular size. It should be noted that the dispersion term is primarily responsible for the pure component calculations. However, it is the reverse for a mixture, where the contribution to the activity coefficient is negligible because of its cancellation between the solvent and pure solute phases. Hence, we shall neglect this contribution when we compute the activity coefficient.

Now again referring back to Equation 3.7, we will now concentrate on the remaining two terms, that is, \((\Delta A_{i}^{\text{cav}} - \Delta A_{i}^{\text{cav}}/kT)\) and \(\ln(V_{i}^{\text{mix}}/V_{i}^{\text{cav}})\). These are calculated using the Staverman–Guggenheim contribution (Smith, Van Ness, & Abbott, 2001), where the liquid volumes can be accessed from the DIPPR database (Design Institute for Physical Properties, Sponsored by AIChE DIPPR Project 801). The contribution was derived based on a lattice model where the activity coefficient is caused by the difference in the shape and size of the molecule. It is given as

\[
\ln\left( \gamma_{i}^{\text{comb}} \right) = 1 - \phi_{i} + \ln\left( \phi_{i} \right) - 5 \frac{A_{i}^{\text{total}}}{A_{i}^{\text{norm}}} \left[ 1 - \frac{\phi_{i}}{\theta_{i}} + \ln\left( \frac{\phi_{i}}{\theta_{i}} \right) \right]
\]

(3.74)

where:

\[
\phi_{i} = \frac{V_{i}^{\text{total}}}{\sum_{j} x_{j}V_{j}^{\text{total}}} \quad \theta_{i} = \frac{A_{i}^{\text{total}}}{\sum_{j} x_{j}A_{j}^{\text{total}}}
\]

(3.75)

Here, \(V_{i}^{\text{total}}\) is the volume of the molecule obtained from the COSMO file output (Figure 3.3). \(A_{i}^{\text{norm}}\) refers to a normalization constant, as \(A_{i}^{\text{total}}\) is not
dimensionless. For molecules having similar shape and sizes, $\ln(\gamma_{i}^{\text{comb}}) \sim 0$, and hence, this term can be neglected. It should also be noted that $\ln(\gamma_{i}^{\text{comb}}) = 0$ for pure component, so it is used to model the cavity formation between a solvent and a solute only. We now are in a position to write the total expression for the activity coefficient of a component in the mixture. Thus, from Equations 3.62 through 3.65, we have the following equation:

$$
\ln(\gamma_{i}^{\text{mix}}) = \left(\frac{\Delta G_{i/S}^{\text{ES}} - \Delta G_{i/i}^{\text{ES}}}{kT}\right) + \left(\frac{\Delta A_{i/S}^{\text{vdW}} - \Delta A_{i/i}^{\text{vdW}}}{kT}\right) + \left(\frac{\Delta A_{i/S}^{\text{cav}} - \Delta A_{i/i}^{\text{cav}}}{kT}\right) + \ln\left(\frac{V_{i/i}}{V_{i/S}}\right)
$$

By writing out each contribution, we have

$$
\frac{\Delta G_{i/S}^{\text{res}}}{RT} = n_{i} \sum_{\alpha} p_{i}(\sigma_{\alpha}) \ln \Gamma_{s}(\sigma_{\alpha})
$$

$$
\left(\frac{\Delta A_{i/S}^{\text{vdW}} - \Delta A_{i/i}^{\text{vdW}}}{kT}\right) = 0 \quad \text{(for mixtures)}
$$

$$
\left(\frac{\Delta A_{i/S}^{\text{cav}} - \Delta A_{i/i}^{\text{cav}}}{kT}\right) + \ln\left(\frac{V_{i/i}}{V_{i/S}}\right) = \ln(\gamma_{i}^{\text{comb}}) = 1 - \varphi_{i} + \ln(\varphi_{i})
$$

$$
-5 \frac{A_{i}^{\text{total}}}{A_{\text{norm}}} \left[ 1 - \frac{\varphi_{i}}{\theta_{i}} + \ln\left(\frac{\varphi_{i}}{\theta_{i}}\right) \right]
$$

Thus, the activity coefficient of a segment that is $\Gamma_{s}(\sigma)$ and sigma potential are connected through the following relation:

$$
p_{s}(\sigma_{\alpha})(\Gamma_{s}(\sigma_{\alpha})) = \exp\left(\frac{\mu_{s}(\sigma_{\alpha})}{kT}\right) \quad (3.76)
$$

The activity coefficient of segment in the mixture and in the pure liquid, $\Gamma_{s}(\sigma)$ and $\Gamma_{i}(\sigma)$, can then be rewritten as

$$
\ln \Gamma_{s}(\sigma_{\alpha}) = -\ln\left\{ \sum_{\alpha_{\Phi}} p_{i}(\sigma_{\alpha}) \Gamma_{s}(\sigma_{\alpha}) \exp\left[ -\frac{E_{\text{pair}}(\sigma_{\alpha}, \sigma_{\Phi})}{kT}\right] \right\} \quad (3.77)
$$

The final expression then takes the form:

$$
\ln \gamma_{i/S} = n_{i} \sum_{\sigma_{\alpha}} p_{i}(\sigma_{\alpha}) \left[ \ln \Gamma_{s}(\sigma_{\alpha}) - \ln \Gamma_{i}(\sigma_{\alpha}) \right] + \ln \gamma_{i/S}^{\text{comb}} \quad (3.78)
$$
This equation is referred to as the COSMO Segment Activity Coefficient (COSMO-SAC) model. Now, we shall move ahead to compute the activity coefficient using COSMO-SAC model (Equation 3.78), as these values will be required to predict the tie line composition in case of ternary or binary system. The entire model is characterized by our earlier COSMO-SAC parameters, $a_{\text{eff}} = 6.32 \, \text{Å}^2$ (surface area of a standard segment), $\alpha' = 8419 \, \text{kcal} \, \text{Å}^4 \, \text{mol}^{-1} \, \text{e}^{-2}$ (misfit energy constant) for misfit energy interaction, $c_{\text{hb}} = 75006 \, \text{kcal} \, \text{Å}^4 \, \text{mol}^{-1} \, \text{e}^{-2}$ (hydrogen bonding energy constant) and $\sigma_{\text{hb}} = 0.0084 \, \text{e}^{-2}$ (hydrogen bonding cutoff).

### 3.3 Predictions of Tie Lines

The equilibrium for any ternary liquid–liquid system is defined by the equation:

$$\gamma_i^I x_i^I = \gamma_i^{II} x_i^{II} \quad \text{where } i = 1,2,3$$

(3.79)

Here, our aim is to compute the $\gamma_i$, the activity coefficient of component $I$, as per Equation 3.78 from COSMO-SAC model. These values will be different for different phases ($I$ and $II$). The corresponding mole fractions in phases $I$ and $II$ will be denoted by $x_i^I$ and $x_i^{II}$, respectively. Assuming an adiabatic operation and that both feed $F$ and solvent $S$ enter a theoretical stage at same temperature, we are then left to account for the heat of mixing. This value for most thermodynamics calculations involving liquid–liquid equilibria (LLE) is sufficiently small, as we observe that a small temperature change occurs. For calculating the LLE split, we adopt the modified Rachford–Rice algorithm (Figure 3.9). Initially, a random feed composition to any known LLE data is used. In any case, we can use the reported data to get an initial feed point. Thus, the feed concentration ($z_i$) is then calculated using the following equation:

$$z_i = \frac{x_i^I + x_i^{II}}{2}$$

(3.80)

Assuming a unity feed rate, that is, $F = 1$, we visualize the process of one feed flow, namely $F$, entering a flash column and then flashing into one liquid phase ($L$) and one vapor phase ($V$), such that $F = L + V$. In the next step, the values of distribution coefficient ($K_i, i = 1,2,3$) are then assigned as follows:

$$K_i = \frac{x_i^{II}}{x_i^I} = \frac{\gamma_i^I}{\gamma_i^{II}}$$

(3.81)
COSMO-SAC: A Predictive Model for Calculating Thermodynamics

Where, $\gamma_i^I$ and $\gamma_i^II$ are predicted using the COSMO-SAC model. With the values of $K_i$, the isothermal flash equation is solved using the conventional flash expression:

\[
 f(\psi) = \sum \frac{z_i(1 - K_i)}{1 + \psi(K_i - 1)} = 0
\]  

subject to,

\[
 Fz_i = L_1x_i^I + L_2x_i^II \quad \text{and} \quad \psi = L_1/F
\]  

FIGURE 3.9
Modified Rachford–Rice algorithm for COSMO-SAC model.
Phase Equilibria in Ionic Liquid Facilitated Liquid–Liquid Extractions

Here, we have assumed \( L = L_1 \) and \( V = L_2 \), as it contains both liquid phases. Both the liquid phases represent the flow rate of the extract \( L_1 \) and the raffinate phase \( L_2 \). Equation 3.83 is nonlinear in nature. This is first iteratively solved for \( \Psi \), which is a fraction and varies between zero and unity. Thereafter, the mole fractions in both phases are calculated via Equations 3.84 and 3.85, respectively.

\[
x_i^I = \frac{z_i}{1 + \Psi(K_i - 1)}
\]

(3.84)

\[
x_i^{II} = K_i x_i^I
\]

(3.85)

For the ionic liquid systems described in the book, the in-house experimental data were used to assume the feed compositions (Equation 3.80). Thus, in a summary, the COSMO-SAC model can predict the tie lines based on the initial input of \( F, z, P, T \). Here, the effect of pressure on LLE computations has been assumed to be negligible. Thus, not only ionic liquids but also any unknown system can be predicted once its COSMO files (Figure 3.3) are generated. This also makes use of the fact that a successful computation of \( \Psi \) by Equation 3.82 indicates a heterogeneous region. A feed point outside the bimodal cure, as observed in a ternary LLE diagram, will usually fail to converge Equation 3.82. Thus, based on any random feed composition \( z_i \), one can generate a tie line data. These predictions can be further used for other higher-order, namely quaternary and quinary, systems. It has been described in detail for the computation of VLLE (vapor–liquid–liquid equilibria) diagrams by making necessary adjustments. This has been discussed in Chapter 4.

### 3.4 Predictions of LLE of Ionic Liquid Systems

Recent experimental work suggests that ILs can be considered as fully dissociated cations and anions. Further, the mechanism of hydrogen bonding (Scheiner, 1997) in the ILs is also not fully available. This fact was used to study the coordination chemistry extraction mechanisms of metal ions, particularly actinide (Cocalia, Gutowski, & Rogers, 2006; Jensen, Neuefeind, Beitz, Skanthakumar, & Soderholm, 2003; MacFarlane & Seddon, 2007). Experimentally, these prove that ILs are completely dissociated into cations and anions when in solution. Also, the definition of IL states it to be a liquid with melting point below 100°C and containing ions exhibiting ionic conductivity. The fact that the ionic conductivity can be measured experimentally proves that ILs consist of ions in solutions. Ionic conductivity studies
have been carried out by Carda-Broch et al. (2003), Berthod and Armstrong (2003) on fluorinated anions such as $[\text{PF}_6]$.

Further, in molecular dynamic simulations, a cubical box is taken, which contains equal number of cations and anions, instead the whole IL molecule. For example, in the work of Morrow and Maginn (2002), the authors studied the thermo-physical properties of ionic liquids by using 300 cations and 300 anions. Similarly, in the work of Lee, Jung and Han (2005), the ionic conductivity of ILs were found using a cubical box of 100 cations and 100 anions. On the contrary, Earle et al. (2006) recently distilled IL without decomposition, at very low pressures. In their pioneering work, they were successful in distilling the $[(\text{CF}_3\text{SO}_2)_2\text{N}]$-based ILs at 100 Pa and 573 K, without decomposition. However, in LLE experiments, ILs are assumed to be in isobaric condition of 1 atm, whereas the distillation is observed with pressures $\sim 0.000009$ atm. Since the operating pressures of LLE are 1 atm, our approach holds true.

Dissociation was also observed during vapor–liquid equilibria (VLE) predictions on reported systems (Kato & Gmehling, 2005; Kato, Krummen, & Gmehling, 2004) in ILs, where we obtained better results than obtained using a single composite molecule (Diedenhofen, Eckert, & Klamt, 2003; Table 3.1). This prompted us to propose that the IL be considered as a dissociated pair of cation and anion for both the VLE and LLE predictions. This novel approach can be applied to ILs containing a limitless combination of

### Table 3.1

<table>
<thead>
<tr>
<th>No.</th>
<th>Reference</th>
<th>System</th>
<th>COSMO-RS</th>
<th>Single Molecule</th>
<th>Cation and Anion</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>(Döker &amp; Gmehling, 2005)</td>
<td>[Emim][($\text{CF}_3\text{SO}_2$)$_2\text{N}$] + Acetone</td>
<td>2.70</td>
<td>1.81</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>(Döker &amp; Gmehling, 2005)</td>
<td>[Emim][($\text{CF}_3\text{SO}_2$)$_2\text{N}$] + 2-Propanol</td>
<td>3.80</td>
<td>2.60</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>(Döker &amp; Gmehling, 2005)</td>
<td>[Emim][($\text{CF}_3\text{SO}_2$)$_2\text{N}$] + Water</td>
<td>9.80</td>
<td>5.26</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>(Döker &amp; Gmehling, 2005)</td>
<td>[Bmim][($\text{CF}_3\text{SO}_2$)$_2\text{N}$] + Acetone</td>
<td>3.50</td>
<td>3.08</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>(Döker &amp; Gmehling, 2005)</td>
<td>[Bmim][($\text{CF}_3\text{SO}_2$)$_2\text{N}$] + 2-Propanol</td>
<td>4.00</td>
<td>2.89</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>(Döker &amp; Gmehling, 2005)</td>
<td>[Bmim][($\text{CF}_3\text{SO}_2$)$_2\text{N}$] + Water</td>
<td>9.80</td>
<td>4.56</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>(Kato &amp; Gmehling, 2005)</td>
<td>[Mmim][($\text{CH}_3$)$_2\text{PO}_4$] + Acetone</td>
<td>5.10</td>
<td>2.50</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>(Kato &amp; Gmehling, 2005)</td>
<td>[Mmim][($\text{CH}_3$)$_2\text{PO}_4$] + Tetrahydrofuran</td>
<td>4.60</td>
<td>2.28</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>(Kato &amp; Gmehling, 2005)</td>
<td>[Mmim][($\text{CH}_3$)$_2\text{PO}_4$] + Water</td>
<td>10.60</td>
<td>6.84</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>(Kato &amp; Gmehling, 2005)</td>
<td>[Mmim][($\text{CF}_3\text{SO}_2$)$_2\text{N}$] + Benzene</td>
<td>8.92</td>
<td>3.66</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>(Kato &amp; Gmehling, 2005)</td>
<td>[Mmim][($\text{CF}_3\text{SO}_2$)$_2\text{N}$] + Cyclohexane</td>
<td>6.26</td>
<td>4.23</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>(Kato et al., 2004)</td>
<td>[Emim][EtSO$_4$] + Benzene</td>
<td>9.75</td>
<td>5.62</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>(Kato et al., 2004)</td>
<td>[Emim][EtSO$_4$] + Cyclohexane</td>
<td>6.94</td>
<td>2.36</td>
<td></td>
</tr>
</tbody>
</table>

Average 6.59 3.66
cation and anion ($10^{18}$). Therefore, one has to do separate quantum mechanical calculations for cations and anions. Thereafter, one can test an IL for a particular application by addition of sigma profiles and sigma potential of cation and anion only.

Moreover, till date, an application concerned with the assumption of complete dissociation of ILs into cations and anions with equimolar concentrations has been carried out earlier by Klamt and Schuurmann (1993) has assumed a complete dissociation of cations and anions for their prediction of infinite dilution activity coefficients (IDAC) of solutes in ILs by COSMO-RS implementation. In their work, the ILs have been described by an equimolar mixture of two distinct ions, the cation and the anion, which finally contribute to the sigma profile of the mixture as two different compounds.

Hence, for the prediction of LLE-based system involving ILs, two modes of approach were assumed: (1) IL as a pure solvent and (2) IL consisting of a pair of cations and anions (Table 3.2). The latter approach assumes a solvent of a mixture of two components: cation and anion. The latter approach has proved to give better phase split with lower root-mean-square deviation. Hence, we proposed a complete dissociation of ILs into cations and anions with equimolar quantities, which is in line with the linear addition of the sigma profiles of the cation and anion.

$$p_{\text{ionic liquid}}(\sigma) = p_{\text{cation}}(\sigma) + p_{\text{anion}}(\sigma)$$ (3.86)

Where, $p_{\text{cation}}(\sigma)$ and $p_{\text{anion}}(\sigma)$ are the sigma profiles for cation and anion, respectively. This is based on the assumption that the IL in a solution is fully dissociated into its respective cations and anions with equimolar concentrations. It should be noted that the entropy of mixing will not exist, since we are dealing with a sigma profile of a single compound/mixture. Thereafter, the linear additions of COSMO area and volume are done, so as to obtain the profile of single mixture/compound.

The results of all the reported IL ternary systems are given in Table 3.2. Using the sigma profile for the single composite molecule, it is observed that for nearly half of the systems, the values of the activity coefficients were not able to predict the split between the extract (IL phase) and raffinate phase for the given tie line. Predictions giving an average RMSD of 36.5% and a maximum root-mean-square deviation (RMSD) of 90% ([OMim][Cl]-Benzene-Heptane) were observed. **NO SPLIT (NS)** condition for 15 systems was encountered. For 13 systems, an RMSD much greater than 10% was observed.

The data sets contain 12 different ILs comprising seven different cations and eight anions. Apart from imidazolium-based ILs, pyridinium-based IL (1-butyl-3-methylpyridinium tetrafluoroborate [Bmpy][BF$_4$]) has also been studied. Switching from composite to additive sigma profile, we notice that a ‘NO SPLIT’ situation has been converted to a ‘SPLIT’ between the two phases for all the 15 cases. It is clear that the improvement is drastic for the IL containing system namely [Bmim][CF$_3$SO$_3$], [OMim][MDeg] and [Bmpy][BF$_4$].
TABLE 3.2
LLE Comparison of Single and Cation–Anion Pair by Using COSMO-RS

<table>
<thead>
<tr>
<th>System No.</th>
<th>T/K</th>
<th>System</th>
<th>Root-Mean-Square Deviation</th>
<th>Experiment</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Single Molecule</td>
<td>Cation + Anion</td>
</tr>
<tr>
<td>1</td>
<td>298.15</td>
<td>[Bmim][CF_{3}SO_{3}] – Ethanol-ethyl-tert-butyl ether</td>
<td>NS</td>
<td>0.095</td>
</tr>
<tr>
<td>2</td>
<td>298.15</td>
<td>[Bmim][CF_{3}SO_{3}] – Ethanol-tert-amyl ethyl ether</td>
<td>NS</td>
<td>0.032</td>
</tr>
<tr>
<td>3</td>
<td>298.15</td>
<td>[Hmim][BF_{4}] – Benzene-heptane</td>
<td>0.711</td>
<td>0.043</td>
</tr>
<tr>
<td>4</td>
<td>298.15</td>
<td>[Hmim][BF_{4}] – Benzene-dodecane</td>
<td>0.087</td>
<td>0.098</td>
</tr>
<tr>
<td>5</td>
<td>298.15</td>
<td>[Hmim][BF_{4}] – Benzene-hexadecane</td>
<td>0.481</td>
<td>0.072</td>
</tr>
<tr>
<td>6</td>
<td>298.15</td>
<td>[Hmim][BF_{4}] – Ethanol-hexene</td>
<td>NS</td>
<td>0.147</td>
</tr>
<tr>
<td>7</td>
<td>298.15</td>
<td>[Hmim][BF_{4}] – Ethanol-heptene</td>
<td>NS</td>
<td>0.125</td>
</tr>
<tr>
<td>8</td>
<td>298.15</td>
<td>[Hmim][PF_{6}] – Benzene-heptane</td>
<td>0.856</td>
<td>0.042</td>
</tr>
<tr>
<td>9</td>
<td>298.15</td>
<td>[Hmim][PF_{6}] – Benzene-dodecane</td>
<td>0.385</td>
<td>0.047</td>
</tr>
<tr>
<td>10</td>
<td>298.15</td>
<td>[Hmim][PF_{6}] – Benzene-hexadecane</td>
<td>0.457</td>
<td>0.059</td>
</tr>
<tr>
<td>11</td>
<td>298.15</td>
<td>[Hmim][PF_{6}] – Ethanol-hexene</td>
<td>NS</td>
<td>0.315</td>
</tr>
<tr>
<td>12</td>
<td>298.15</td>
<td>[Hmim][PF_{6}] – Ethanol-heptene</td>
<td>NS</td>
<td>0.306</td>
</tr>
<tr>
<td>13</td>
<td>298.15</td>
<td>[Omim][Cl] – Methanol-hexadecane</td>
<td>0.029</td>
<td>0.005</td>
</tr>
<tr>
<td>14</td>
<td>298.15</td>
<td>[Omim][Cl] – Ethanol-hexadecane</td>
<td>0.03</td>
<td>0.008</td>
</tr>
<tr>
<td>15</td>
<td>298.15</td>
<td>[Omim][Cl] – Ethanol-tert-amyl ethyl ether</td>
<td>0.382</td>
<td>0.067</td>
</tr>
<tr>
<td>16</td>
<td>298.15</td>
<td>[Omim][Cl] – Benzene-heptane</td>
<td>0.896</td>
<td>0.096</td>
</tr>
<tr>
<td>17</td>
<td>298.15</td>
<td>[Omim][Cl] – Benzene-dodecane</td>
<td>0.058</td>
<td>0.096</td>
</tr>
<tr>
<td>18</td>
<td>298.15</td>
<td>[Omim][Cl] – Benzene-hexadecane</td>
<td>0.146</td>
<td>0.07</td>
</tr>
<tr>
<td>19</td>
<td>298.15</td>
<td>[Emim][C_{6}H_{12}SO_{3}] – Benzene-heptane</td>
<td>0.362</td>
<td>0.07</td>
</tr>
<tr>
<td>20</td>
<td>298.15</td>
<td>[Emim][C_{6}H_{12}SO_{3}] – Benzene-hexadecane</td>
<td>0.456</td>
<td>0.14</td>
</tr>
</tbody>
</table>

(Continued)
systems and also for the alkene-based systems [System No. 6, 7, 11, 12], for which no prediction had been possible when considering the IL as a single molecule. For the [Bmim][CF<sub>3</sub>SO<sub>3</sub>], [Oimim][MDEg] and [Bmim][BF<sub>4</sub>] systems (System No. 1, 2, 21, 22 and 23–29), the RMSD is less than 10%. Out of the 36 systems studied, most of the predictions are excellent, considering the fact that only six systems gave RMSD greater than 10%.

<table>
<thead>
<tr>
<th>System No.</th>
<th>T/K</th>
<th>System</th>
<th>Root-Mean-Square Deviation</th>
<th>Experiment</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Single Molecule</td>
<td>Cation + Anion</td>
</tr>
<tr>
<td>21</td>
<td>298.15</td>
<td>[Oimim][MDEg]&lt;sup&gt;a&lt;/sup&gt; – Benzene-heptane</td>
<td>NS 0.103</td>
<td>0.006</td>
</tr>
<tr>
<td>22</td>
<td>298.15</td>
<td>[Oimim][MDEg]&lt;sup&gt;a&lt;/sup&gt; – Benzene-hexadecane</td>
<td>NS 0.043</td>
<td>0.006</td>
</tr>
<tr>
<td>23</td>
<td>313.15</td>
<td>[Bmpy][BF&lt;sub&gt;4&lt;/sub&gt;] – Xylene-octane</td>
<td>NS 0.03</td>
<td>0.0025</td>
</tr>
<tr>
<td>24</td>
<td>348.15</td>
<td>[Bmpy][BF&lt;sub&gt;4&lt;/sub&gt;] – Xylene-Octane</td>
<td>NS 0.024</td>
<td>0.0025</td>
</tr>
<tr>
<td>25</td>
<td>313.15</td>
<td>[Bmpy][BF&lt;sub&gt;4&lt;/sub&gt;] – Ethylbenzene-octane</td>
<td>NS 0.041</td>
<td>0.0025</td>
</tr>
<tr>
<td>26</td>
<td>348.15</td>
<td>[Bmpy][BF&lt;sub&gt;4&lt;/sub&gt;] – Ethylbenzene-octane</td>
<td>NS 0.029</td>
<td>0.0025</td>
</tr>
<tr>
<td>27</td>
<td>313.15</td>
<td>[Bmpy][BF&lt;sub&gt;4&lt;/sub&gt;] – Benzene-hexane</td>
<td>NS 0.048</td>
<td>0.0025</td>
</tr>
<tr>
<td>28</td>
<td>333.15</td>
<td>[Bmpy][BF&lt;sub&gt;4&lt;/sub&gt;] – Benzene-hexane</td>
<td>NS 0.036</td>
<td>0.0025</td>
</tr>
<tr>
<td>29</td>
<td>298.15</td>
<td>[Bmpy][BF&lt;sub&gt;4&lt;/sub&gt;] – Toluene-heptane</td>
<td>NS 0.075</td>
<td>0.0025</td>
</tr>
<tr>
<td>30</td>
<td>298.15</td>
<td>[Mmim][CH&lt;sub&gt;3&lt;/sub&gt;SO&lt;sub&gt;4&lt;/sub&gt;] – Toluene-heptane</td>
<td>0.455</td>
<td>0.008</td>
</tr>
<tr>
<td>31</td>
<td>298.15</td>
<td>[Emim][C&lt;sub&gt;2&lt;/sub&gt;H&lt;sub&gt;5&lt;/sub&gt;SO&lt;sub&gt;4&lt;/sub&gt;] – Toluene-heptane</td>
<td>0.116</td>
<td>0.031</td>
</tr>
<tr>
<td>32</td>
<td>298.15</td>
<td>[Bmim][CH&lt;sub&gt;3&lt;/sub&gt;SO&lt;sub&gt;4&lt;/sub&gt;] – Toluene-heptane</td>
<td>0.235</td>
<td>0.032</td>
</tr>
<tr>
<td>33</td>
<td>313.15</td>
<td>[Emim][EtSO&lt;sub&gt;4&lt;/sub&gt;] – Ethanol-hexene</td>
<td>NS 0.124</td>
<td>0.005</td>
</tr>
<tr>
<td>34</td>
<td>313.15</td>
<td>[Emim][EtSO&lt;sub&gt;4&lt;/sub&gt;] – Ethanol-heptene</td>
<td>NS 0.023</td>
<td>0.005</td>
</tr>
<tr>
<td>35</td>
<td>313.15</td>
<td>[E-2,3-dmim][EtSO&lt;sub&gt;4&lt;/sub&gt;] – Ethanol-hexene</td>
<td>NS 0.032</td>
<td>0.005</td>
</tr>
<tr>
<td>36</td>
<td>313.15</td>
<td>[E-2,3-dmim][EtSO&lt;sub&gt;4&lt;/sub&gt;] – Ethanol-heptene</td>
<td>NS 0.065</td>
<td>0.005</td>
</tr>
</tbody>
</table>

Note: NS: No splitting of the phase; NA: Not available

<sup>a</sup> MDEG: monomethyldiethyleneglycol
COSMO-SAC: A Predictive Model for Calculating Thermodynamics

Acknowledgments
A part of this chapter has been adapted from my earlier National Program for Technology Enhanced Learning (NPTEL) web course entitled ‘Molecular Simulation in Chemical Engineering’. I am deeply indebted to the Ministry of Human Resource and Development, Government of India, for providing me financial support for the same. Further, I would like to thank Prof. Stanley I. Sandler, University of Delaware, for helping me in learning the basics of COSMO-based modeling. I am also indebted to his student Dr. Russel I. Burnett for helping me out in the COSMO-SAC equations. This visit would not have been possible without the funding from IUSSTF (Indo–U.S. Science and Technology Forum).

Appendix: Canonical Partition Functions
The canonical partition function that is \( Q(N, V, \beta) \) is given as

\[
Q(N, V, \beta) = \sum_{\text{states}} e^{-\beta E_i(N, V)}
\]  

(A.1)

The probability of occurrence of a particular microstate \( i \) with Energy \( E_\alpha \) is given as

\[
p_i(E_\alpha) = \frac{e^{-\beta E_\alpha}}{\sum_{\text{All states}} e^{-\beta E_i}} = \frac{e^{-\beta E_\alpha}}{Q(N, V, \beta)}
\]  

(A.2)

The significance of \( \beta \) can be made from the fact that as \( \beta > 0 \), the state of higher energy is less likely as the state of lower energy. We now shall see the process of evaluating the thermodynamic properties based on canonical partition functions. It should be noted that the probability of finding a system in any microstate of energy \( E_\alpha \) would be the product of the probability that the system is in a particular microstate with energy \( E_\alpha \) and the number of states having that energy (also known as degeneracy). Thus, we have:

\[
p(E_\alpha) = \sum_{\text{States } i \text{ having Energy } E_\alpha} p_i(E_\alpha) = p_i(E_\alpha) \times \text{Degeneracy} \left( \frac{\text{number of states with energy } E_\varepsilon}{\text{number of states}} \right)
\]  

(A.3)
We now define the probability of occurrence of a particular microstate \( i \) with energy \( E_\alpha \) as

\[
p_i(E_\alpha) = \frac{e^{-\beta E_\alpha}}{\sum_{\text{All } i \text{ states}} e^{-\beta E_i}} = \frac{e^{-\beta E_\alpha}}{Q(N, V, \beta)} \tag{A.4}
\]

In a similar manner, the probability of occurrence of the energy level \( E_\alpha \), \( p(E_\alpha) \) is

\[
p(E_\alpha) = \text{Degeneracy} \left( \frac{\text{number of states with energy } E_\alpha}{\text{all states}} \right) \times p_i(E_\alpha) = \frac{\omega(E_\alpha) e^{-\beta E_\alpha}}{Q(N, V, \beta)} \tag{A.5}
\]

Here Degeneracy = \( \omega(E_\alpha) \)

Thus, the partition function can be written in terms of either levels or states, as below:

\[
Q(N, V, \beta) = \sum_{\text{All states}} e^{-\beta E_i} = \sum_{\text{All levels}} \omega(E_i) e^{-\beta E_i} p_i(E_\alpha) \tag{A.6}
\]

Now, once the partition function has been defined, we can derive thermodynamics functions such as internal energy \( U \):

\[
U = \sum_{\text{all states of } k} E_k p(E_k) = \frac{\sum_{\text{all states of } k} E_k p(E_k)}{\sum_{\text{all states of } k} e^{-\beta E_k}} = \frac{\sum_{\text{all states of } k} E_k p(E_k)}{Q} \tag{A.7}
\]

However, from the definition, we have:

\[
\left( \frac{\partial Q}{\partial \beta} \right)_{N, V} = -\sum_{k \text{ states}} E_k e^{-\beta E_k} \tag{A.8}
\]

Thus, the internal energy \( U \) takes the form:

\[
U = \left( \frac{\partial \ln Q}{\partial \beta} \right)_{N, V} = \frac{\sum_{k \text{ states}} E_k e^{-\beta E_k}}{Q} \tag{A.9}
\]
References


Design Institute for Physical Properties, Sponsored by AIChE DIPPR Project 801 – Full Version. Design Institute for Physical Property Data/AIChE.


