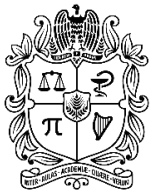


UNIVERSIDAD
NACIONAL
DE COLOMBIA

Parameter Estimation on Molecular Models of Complex Fluids by Stochastic Optimization Techniques

Estimación de parámetros en modelos moleculares de Fluidos Complejos usando técnicas estocásticas de optimización

Carlos Fernando Ospina Trujillo

Universidad Nacional de Colombia
Facultad de Ciencias Exactas y Naturales,
Departamento de Matemáticas y Estadística
Manizales, Colombia
2017

Parameter Estimation on Molecular Models of Complex Fluids by Stochastic Optimization Techniques

Estimación de parámetros en modelos moleculares de Fluidos Complejos usando técnicas estocásticas de optimización

Carlos Fernando Ospina Trujillo

Tesis de maestría presentada como requisito parcial para optar al título de:
Magister en Ciencias - Matemática Aplicada

Director:
Ph.D. Juan Carlos Riaño Rojas
Codirector:
Ph.D Felipe Antonio Perdomo Hurtado

Grupo de Investigación:
PCM Computational Applications

Universidad Nacional de Colombia
Facultad de Ciencias Exactas y Naturales,
Departamento de Matemáticas y Estadística
Manizales, Colombia
2017

Dedico este trabajo a
mis padres, Cielo y Rodrigo,
para demostrarles mi amor y gratitud.

Acknowledgments

I want to thank to my great professor, advisor and friend Juan Carlos Riaño Rojas for his useful comments. I recognize the appropriate help of the professor Felipe Antonio Perdomo Hurtado, who encouraged me to continue working and gave me an effective focal point in this work. Also, I want to recognize the *Universidade Federal do Rio de Janeiro* where the experimental data were measured, and the Chemical Institute from the *Universidad Nacional Autónoma de México* where these were characterized. Finally, but not less important, I want to express my thanks to the *Facultad de Ciencias Exactas y Naturales* and my Institution *Universidad Nacional de Colombia*, for supporting me during all my formation time.

Abstract

With the rising of the study of some ionic liquids, many mathematical problems have appeared, as, for example, complex roots of nonlinear systems and the best parameters for fitting experimental data. This work presents a schematic way to solve these problems, but based on a thermodynamic model wich has all of the mentioned obstacles. The scheme uses stochastic techniques as a central tools. The results, by using these techniques, allow to obtain good parameters that can predict some facts in the interaction of the ionic liquid with low concentrations of water. Another important fact of the present work is that by thense of this scheme, it is possible to save computational time and obtain results in a matter of hours.
Keywords: Stochastic optimization, stochastic search techniques, simulated annealing, simplex optimization, ionic liquids.

Resumen

La importancia de estudiar ciertos líquidos iónicos trae consigo diferentes problemas matemáticos, algunos problemas que pueden surgir son: determinar las raíces de ciertos sistemas de ecuaciónes altamente no lineales y determinar los mejores parámetros que permitan dar un ajuste preciso a las propiedades que experimentalmente se pueden dar. Este trabajo presenta un esquema que ayuda a resolver estos problemas, basado en el estudio de un modelo termodinámico que tiene todos los obstáculos mencionados. El esquema presentado, usa técnicas estocásticas como herramienta central. Los resultados que se obtuvieron al usar estas técnicas proveen buenos parámetros que predicen algunos hechos en la interacción del líquido iónico con bajas concentraciónes de agua. Como resultado adicional de este trabajo, el uso de estas técnicas permiten ahorrar tiempo computacional y obtener resultados en cuestión de horas.
Palabras Clave: Optimización estocástica, técnicas de búsquedas estocástica, simulated annealing, simplex optimization, líquidos iónicos.

Contents

List of Figures

List of Algorithms

Introduction

The Optimization is about choosing the best option among others according to a criterion, what makes it a useful tool in many applications of technology and science; for example, it is widely used in structure designs, business and economy, image processing, optimal control, raw material etc. It is also impressive how optimization is present in the daily life; each one wants to save time, or it is desirable to arrive from one point to another by using the fastest and safest path; optimization is also present when you want to buy the best stuff without wasting too much money and so on.

This work is focused on an application related to determining the best values over thermo-physical properties of certain fluids. The issue to determine these values can be approached by different mathematical techniques, as Perdomo et al. in [43], they approximate these parameters by genetic algorithms. The intention here is to do a revision over stochastic methods, that can carry out extensive examinations, as well as low computational cost, and wich finally provide results close enough to the exact value that satisfies the application needs.

An optimization problem can be stated as:

$$x_{\min} = \arg\min\{f(x) : x \in \Omega\}, \tag{1}$$

where f is a real value function defined over a set $\Omega \subseteq \mathbb{R}^n$, usually, the function f is named the objective function and the set Ω is the feasible set of solutions. When $\Omega = \mathbb{R}^n$, the problem stated by the equation (1) is usually named *Global Optimization Problem* or simply an optimization problem without constraints, in general, when $\Omega \subset \mathbb{R}^n$, we say that (1) is named an optimization problem with constraints. The set Ω can be defined in some cases of the form:

$$\Omega = \{x \in [\alpha_1, \beta_1] \times \cdots \times [\alpha_n, \beta_n] \text{ and } g_r(x) \geq 0 \text{ for each } r \in \{1, 2, \ldots, m\}\}.$$

In this case, g_r are mappings and represent the constraints of the problem and m represents the quantity of constraints, consequently there may not be a relationship between the dimension of the problem n and the number of constraints m. Other problems require to maximize an objective function g instead of minimizing an objective function f, over a particular admissible solution set. Nevertheless, this issue is equivalent to a problem (1) by considering $f = -g$. Due to this, we will only deal with minimization, as the target in every technique described in this document.

If f in (1) is enough smooth, we say $f \in C^2(\Omega)$. Techniques as Newton-Raphson can provide a partial solution to (1), as long as this is locally convergent (the initial step should be really near to a solution) and it does not distinguish between maximum (local) and minimum (local), the Newton-Raphson method is not a useful technique for the intentions in this work. There are other useful techniques as the conjugate gradient

or the descent gradient (that will be explained in the next chapter), but they become inefficient when the dimension of Ω is increased. For example, W. Press in [44] approaches these methods.

In general, the deterministic algorithms can fail in at least one of the following categories:

- The difficulty to converge globally
- The requirement over f to be smooth
- The problem of dimension

The issues above justify the use of stochastic methods because they are suitable techniques to avoid them.

0.1 Brief Note on Optimization Techniques

The types of algorithms we are interested in, are the stochastic type. The 20th century brought many developments, and the most remarkable fact is that with the advances in computers, more problems and types of methods are studied. In regarding deterministic approaches, in linear programming, George Bernard Dantzig established the famous simplex method. It seems that the Russian mathematician Leonid Vitaliyevich Kantorovich found the same method but kept it in secret until Dantzig made it public the algorithm [15]; whereas that in the area of Non-Linear Programming, the famous conditions in Karush-Kuhn-Tucker appears. In addition, Richard Bellman proposed the Bellman equation as a necessary condition in dynamic programming. Another famous problem is The Traveling Salesman Problem that states the following question: "Given a number of cities all of them connected, what is the shortest path that visits each city exactly one time?", a general statement of this problem is attributed to Karl Menger (the same mathematician of the Menger Sponge) in [36], with comments on defining a new curve length. This problem is very important in operation research and theoretical computer science, in fact it is classified as an NP-hard (non-deterministic polynomial-time hard).

There is a relation between the development of Optimization and the electronic Era; new branches in algorithms appeared, they are: stochastic and evolutionary. Both differ from the deterministic algorithms by using random walks to approach the solution, and that they do not usually require differentiability on the objective function.

From one hand, with stochastic algorithms one can find the family of Stochastic Approximation methods. Here, these algorithms try to find the extrema of a function, when the objective function or even the derivatives of this, can not be evaluated directly and only noisy estimations are possible. The initial method was proposed by Herbert Robbins and Sutton Monro in 1951, in the so-called Robbins-Monro algorithm [49]. Successive works inspired in the Robbins-Monro algorithm emerged, as the work of Venter Sacks [52] and [57]. Another noticeable method was proposed by Kiefer and Wolfowitz [28]; as a procedure for estimating the maximum of an objective function, that in contrast to the Robbins-Monro algorithm, it does not need that the function being continuous. Other very important but different algorithm was developed, the Monte-Carlo method appeared in 1953, see ref. [38] with the main goal of addressing physics problems, but applications in optimization appeared, and the Monte-Carlo method can provide a good solution by evaluating the objective function in a sequence of random points and choosing

the extreme value of this sequence. Another important generalization was due to W. K. Hastings by studying the initial proposal of Monte-Carlo (this technique is also named Metropolis-Hastings) for several dimensions and changing the Boltzmann distribution by a general one [22].

In Operation Research, one of the initial adaptations of Metropolis-Hastings method to solve problems in this field was presented by S. Kirkpatrick et al. in [29] to solve combinatorial problems as the Traveler Salesman Problem, with several nodes. The idea is to provide sequences of solutions, saying

$$s_0, s_1, s_2, \ldots, s_n, \ldots$$

where this sequence is supposed to converge to the shortest path. The interesting fact here is that the sequence is a Markov process, which uses the Boltzmann distribution to allow the transition from a partial solution s_i to the next s_{i+1}. To solve continuous problems, authors as Bohachevsky et al. in [8] made proposals but no theoretical results over convergence where given. Nevertheless, this implementation is not far for providing good results in many practical problems. Other works on theoretical results for applying the technique to continuous cases are provided by A. Dekkers and E. Aarts in [16] or M. Locatelli in [32].

On the other hand, there were some people that inspired in nature proposed what is known as the Artificial Intelligence. For example the theory of natural selection and the human social behavior inspired some techniques to approach optimization problems. On one side, inspired by natural selection, the genetic algorithm executes three basic operations to simulate the natural process: mutation, crossover and selection. Initial ideas to this technique were provided by Alan Turin [55] when he was realizing on ideas to simulate the adult or child human mind and he noticed that trying to produce an artificial intelligence, could be made by carrying out an evolutionary process.

The research in genetic algorithms was widely developed, a key reference to these techniques can be found in the book [24]. On the other side, the social behavior of animals, for example a shoal or a herd, have inspired algorithms as the Ant Colony System [17] or the Particle Swarm Optimization [18]. These algorithms belong to the swarm intelligence algorithms, in general, the idea is simulating these groups when they try to find food or being protected from predators, all this oriented to approximate an optimal solution.

Additionally, there exist a kind of very interesting heuristics based on geometrical patterns. One of these heuristics is commonly denominated Pattern search. This technique is convergent locally, but the remarkable fact is that the objective functions do not require differentiability. This heuristic was initially proposed by Robert Hooke and T. A. Jeeves in [25]. There exist some studies of global convergences under certain conditions in [54]. The other technique that will be explained later in one of the next chapters, is the Nelder-Mead technique [39]; here the geometric pattern that explores the solutions is named simplex, it has one more vertex that the dimension of the solutions (a triangle in the plane, tetrahedron in the space, etc) and it moves with four basic movements: reflection, expansion, and two types of contraction. The two principal differences between these algorithms are: first, the geometric pattern used for exploring,and the second difference is that one more movement in the simplex method is added (a contraction to one of the vertices). To clarify, the geometric figure in Pattern Search can be imagined as the middle point of a hypercube, joined with segments by the middle point of their faces, for example in the plane it is a cross.

0.2 Background and Motivation

My research team PCM *computational applications* have lately carried out, different studies that combine mathematics, physics and chemistry, all of them focused on models for physical needs. Among these studies, there are applications of signal processing, coatings, image processing oriented to medical applications and simulation of physical properties by Monte-Carlo, finite elements and molecular dynamic [48, 2, 56, 46].

As we have mentioned before, solving the problem (1) can present some difficulties in terms of non-linearities, many global minima and complex feasible sets. This is present in our main application goal, which is described in later chapters. Our purpose is to find a set of parameters that *better* fit a set of data. These data are obtained from experimental results of measuring some properties to a mixture of chemical potential, pressure, physical work, density, etc. We proceed by relating the characteristics according to the theory SAFT-VRE see ref. [20] and finally by constructing the objective function and the constraints, in order to find the parameters. There exists background on theses kind of problems, as [42, 43, 41]; nevertheless theses works have not taken into account some characteristics of the compounds, some of them are the contribution of the ionization, its work and potential, among others. In addition, the stochastic methods were chosen to develop the fitting because they do not require the objective function to be differentiable or continuous; even more, the construction of the objective function has, as a result a function very difficult to evaluate and a behavior difficult to predict.

The main purpose of this work is to provide an estimation of a set of parameters, that describe a system of ionic liquids. In order to do this, the present work is organized as follow: Chapter 1 presents some of the most important stochastic techniques that will be implemented for carrying out the estimation, Chapter 2 analyzes and develops the SAFT-VRE approach based on literature and making some assumptions we think best describe the ionic liquid, Chapter 3 presents the results of adjusting the parameter to the model and discusses some changes presented in the literature, and finally, Chapter 4 presents the most important facts and conclusions of this work.

Chapter 1

Optimization Methods Based on Stochastic Searches

In this chapter, there is a description of different algorithms used to find the global optimum of a problem by using stochastic searches. These optimization methods make use of random steps, this causes to be less sensitive to errors and additionally to be non-deterministic methods. The reason to have chosen these techniques is because they are suitable to solve problems with many local minima and for highly nonlinear functions. In addition, most of the methods presented are used to find solutions to problems where the objective is not differentiable, or even the feasible set is disconnected.

The most important advantages of these methods are: the simplicity of implementing them in computational languages, and the lack in using the derivative of the objective function. We will start mentioning the Simulated Annealing method, some of its implementations, and finally other variations motivated by this technique.

1.1 Simulated Annealing

As it is mentioned in the historical notes, Metropolis and some other colleagues proposed the Monte-Carlo technique as an strategy to model physical behaviors, by using the new electronic computers and the generator of random numbers, see ref. [38]. As the time passed new approaches emerged, as the theoretical study provided by Hastings in paper [22] and more specifically the idea of using the Boltzmann probability, in order to give positive probabilities to accept a random point, whose function evaluation is greater than the evaluation of a previous iteration when finding the minimal solution for the Traveling Salesman Problem is executed [29].

In agreement with certainly principles of physics, a system with the enough freedom will tend to the state of minimum energy. For example, the temperature of a piece of heated metal tends to be cool. Therefore, it is possible to extract a mathematical model of such a system, in order to find the optimum value of certain function.

In the case of the optimization problem (1), the approach of Simulated Annealing is based on generating samples on the set Ω that can be seen as possible solution due to they belong to the set of feasibility. The technique reaches the optimum value by producing simple iterations $x_0, x_1, \ldots, x_j$ that belong to the set Ω, being j a label used to count the points generated by the method.

One initial solution is proposed to solve the problem, consequently other different

solution is proposed; evidently, one can decide if one of the solutions is better than the other by comparing the images. Nevertheless, the idea is to give the possibility to bad solutions to update the iterations. For our purposes, this method can avoid get trapped in local solutions, accepting an iteration that corresponds to an increase in the function value, which can cause later to make an iteration to arrive to suitable values.

This is done by using the Boltzmann probability transition:

$$P(x_j, x_{j+1}, T) = \min\{1, e^{-\Delta f/kT}\} \tag{1.1}$$

where $\Delta f = f(x_{j+1}) - f(x_j)$ and the constants k and T are the Boltzmann constant and certain positive number named temperature. The expression (1.1) was proposed by Metropolis in the general scheme of Monte-Carlo. The quantity $P(x_j, x_{j+1}, T)$ (the probability to change x_j by x_{j+1} at temperature T) is always positive, but the most important is that, if $f(x_{j+1}) < f(x_j)$, this means that $\Delta f < 0$ hence $P(x_j, x_{j+1}, T) = 1$. The other hand when $f(x_{j+1}) > f(x_j)$, the difference Δf is negative and this implies that the probability of transition from the point x_j to the point x_{j+1} at certain temperature T is positive or that $P(x_j, x_{j+1}, T) > 0$. The above shows why positive probabilities are given to bad movements of the stochastic path. However, the remarkable fact is that a lower but still positive temperatures T the transitions are more difficult to carry out than higher temperatures. The pseudocode of this technique is presented in the algorithm 1, based in one of the proposals of Bohachevsky [8]. This is the simplest proposal of Simulated Annealing, but it has large disadvantages as the most complex. These issues are the difficulty of tunning the parameters: step size δ (for the range of the next search) and temperature T, in order to find the best solution in the less quantity of iterations. In the algorithm 1 step 2, the set $\partial B_\delta(x)$ represents the surface of a sphere with radius δ and center in the x:

$$\partial B_\delta(x) = \{y \in \mathbb{R}^n : ||x - y|| = \delta\},$$

being $||\cdot||$ the euclidean norm.

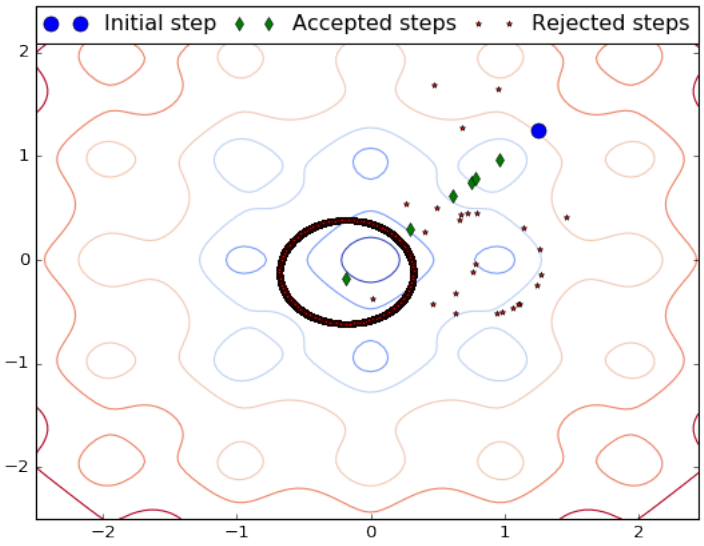
Algorithm 1 Standard Simulated Annealing (SSA)

1: Choose randomly an initial point $x_j \in \Omega$, with $j = 0$.
2: Choose a random point $x_{new} \in \partial B_\delta(x_j) \cap \Omega$.
3: Generate a random number ρ with uniform distribution from the interval $(0, 1)$.
4: Finally,

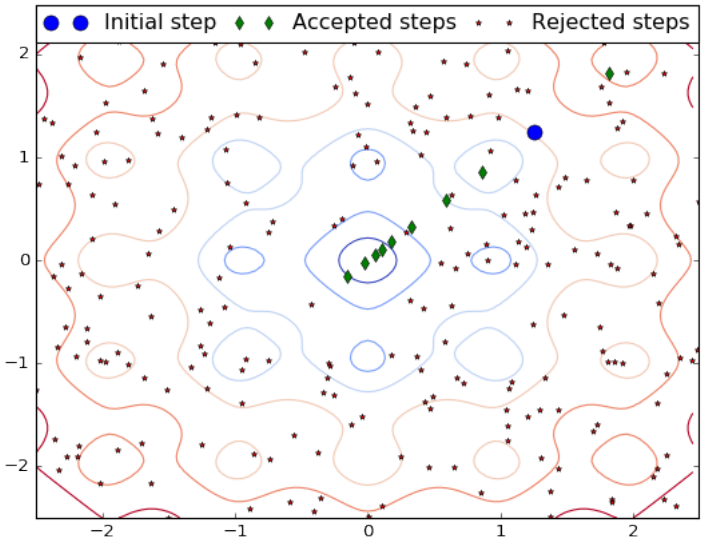
$$x_{j+1} = \begin{cases} x_{new} & \text{if } \rho < P(x_j, x_{new}, T), \\ x_j & \text{otherwise} \end{cases}$$

5: If any stopping criterion is satisfied stop, else return to the step 2.

One of the best alternatives to the this first approach of Simulated Annealing is studied for example by Locatelli in [32] but later than C. Bélisle in [5]. There, the new part is the implementation of cooling schedules. A cooling schedule, since a general view, is a succession of functions $\{\tau_j\}$ defined on $\mathbb{R}^j$ that assigns temperatures T_{j+1}. According to the knowledge of the pass iterations, the schedules do $T_j = \tau_j(x_0, x_1, \ldots, x_{j-1})$, but under the condition that $\lim_{j\to\infty} \tau_j = 0$. To avoid tunning a step size δ, Locatelli and Bélisle also studied the general form of choosing new candidates for updating the succession by certain distribution. This new Simulated Annealing is presented in the algorithm 2.



(a) A iterations carried out by the algorithm 1, using the Ackley as test function. $\delta = 0.4$, $T = 0.9$.



(b) A iterations carried out by the algorithm 2. With a uniform distribution on Ω and a cooling schedule equal to 0.99^j.

Figure 1.1: An example of the iterations carried out by different Simulated Annealing approaches

Algorithm 2 Most common Simulated Annealing

1: Choose randomly an initial point $x_j \in \Omega$, with $j = 0$.
2: Choose a random point $x_{new} \in \Omega$, according to certain distribution.
3: Generate a random number ρ with uniform distribution from the interval $(0, 1)$.
4: According to the cooling schedule, do $T_{j+1} = \tau_{j+1}(x_0, \ldots, x_j)$
5: Finally,

$$x_{j+1} = \begin{cases} x_{new} & \text{if } \rho < P(x_j, x_{new}, T_{j+1}), \\ x_j & \text{otherwise} \end{cases}$$

6: If any stopping criterion is satisfied stop, else return to the step 2.

1.1.1 Cooling Schedules

There exists different cooling schedules in the literature, some of them are presented in the references [29, 1]. In general, down the conditions presented by [5, 32], it is possible to use any scheme of cooling, even those that come from the initial Metropolis-Hastings technique, or those used for solving combinatorial problems by SA.

1.2 Improving Hit-and-Run

Initially, the algorithms type Hit-and-Run were designed as a method for sampling points over bounded subsets of $\mathbb{R}^n$. These are Markov chains type and have the characteristic to be uniformly distributed over the sample space and additionally with good asymptotic properties. The reason to be called in that form is that the algorithm behaves as a driver, crashing points continuously on certain region. Roughly speaking, these algorithms work as random generators according to a general distribution.

Hit-and-Run was presented by R. Smith in ref. [53], he proved the convergence in variation to the uniform distribution. Independently Boneh and Golan in ref. [9] did it, but they did not show results over asymptotic uniformity, it was uniquely conjectured. A important generalization of this algorithm was made by Bélisle et al. in [6]; the idea was to take any arbitrary distribution v and other objective distribution π, under certain conditions of π the sampling converges to it. Other important studies on this were made in [27, 58].

The sampling method proposed by hit-and-run inspired different applications, for example it was proposed for constructing random walks in different regions, as convex sets in ref. [7] or on geometric objects as polytopes in [37]. Zabinsky et al. in ref. [59] proposed a method named Improving Hit-and-Run for finding global optimum of nonlinear functions, it means that the algorithm samples points over the set Ω with a rejection-acceptance step in order to obtain minimum values.

The algorithm 3 describes each step of the Improving Hit-and-Run method. The matrix H can be chosen as the Hessian of the objective function when f is an elliptic function[1]. Normally, the matrix H is not easy to compute. However, it is possible to approximate this matrix with the procedures used in DFP, BFGA in quasi-Newtonian local search as suggested by the authors, or even as SR1 proposed in [31, 11, 40]. Also, this procedure can be simplified using $H = I$ and normalizing the direction vector $\vec{d}$,

[1] In an elliptic problem, the objective function can be written of the form $f(x) = h(r)$ with $r = ||x - x_{\min}||_A$ and A is a non singular square matrix, h is non decreasing for $r > 0$ and the norm is defined as: $||x||_A = ||Ax||$ being $||\cdot||$ the euclidean norm.

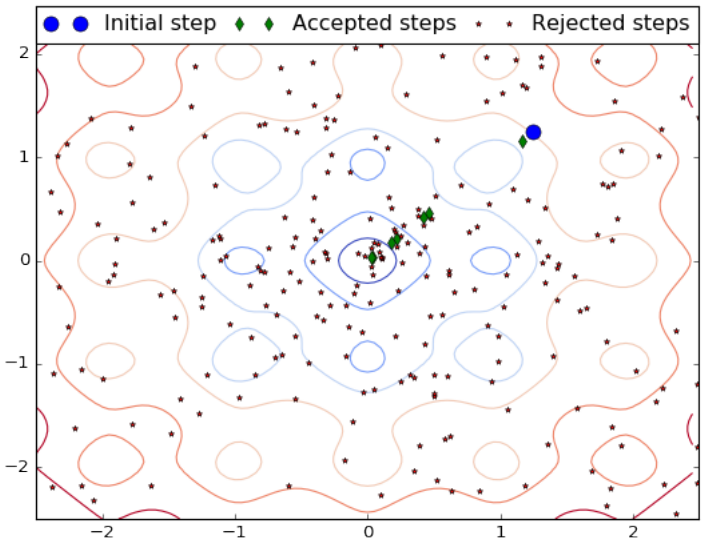


Figure 1.2: An example of the iterations using Improving Hit and Run on the Ackley test function and $H = I$.

which causes that the distribution of the direction vector be uniform over the hypersphere of unitary radius and center in the current iteration [35].

Algorithm 3 Improving Hit and Run for optimization

1: Choose randomly an initial point $x_j \in \Omega$, with $j = 0$.
2: Choose a random direction $\vec{d_j}$ normally distributed on with mean $\mathbf{0}$ and covariance matrix H^{-1} (H is a matrix definite positive).
3: Find the set $L_j = \{\lambda \in \mathbb{R} : x_j + \lambda \vec{d_j} \in \Omega\}$ and choose a random number λ_j uniformly distributed over L_j.
4: Finally,

$$x_{j+1} = \begin{cases} x_j + \lambda_j \vec{d_j} & \text{if } f(x_j + \lambda_j \vec{d_j}) < f(x_j) \\ x_j & \text{otherwise} \end{cases}$$

5: If any stopping criterion is satisfied stop, else return to the step 2.

The figure 1.2 shows an example of the procedure 3 applied to the Ackely test function. Here, one can note the similar results with the example carried out by the Simulated Annealing. This is due to Improving Hit and Run samples uniformly over Ω. The advantage of this method is the lack of using a cooling schedule. But despite of this fact, the high quantity of evaluations turn this method useless over functions difficult to evaluate.

1.3 Hide-and-Seek

This algorithm for optimizing is similar, in essence to the Improving Hit-and-Run procedure. The difference between them lies in the implementation of the Metropolis criterion, as an acceptation-rejection probability of transition to the random walk see ref.

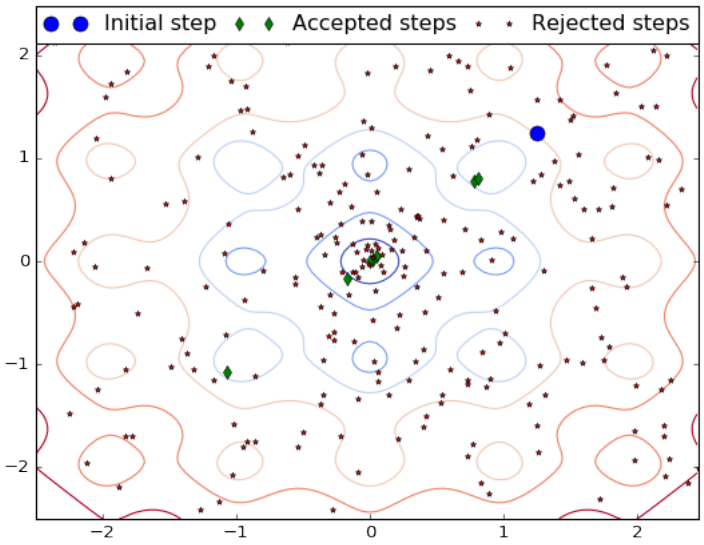


Figure 1.3: An example of Hide and Seek on the Ackely test function. The cooling schedule used is of the form 0.99^j.

[50]. With this transition probability, the algorithm behaves as the game child, in the sense it tries to find a global minimum in a region where it could or could not be placed. The pseudo code is presented in 2; there, the random direction was modified to carry out the results. We use a uniform distribution to choose the direction vector over a n-dimensional hyper-sphere.

Algorithm 4 Hide-and-Seek scheme for optimization

1: Choose randomly an initial point $x_j \in \Omega$, with $j = 0$.
2: Choose a random direction $\vec{d}_j$ with each coordinate distributed uniformly over the feasible region.
3: Find the set $L_j = \{\lambda \in \mathbb{R} : x_j + \lambda \vec{d}_j \in \Omega\}$ and choose a random number λ_j uniformly distributed over L_j.
4: State $T_j = \tau_{j+1}(x_0, \ldots, x_j)$ according to some cooling schedule $\{\tau_j\}$.
5: Finally,

$$x_{j+1} = \begin{cases} x_j + \lambda_j \vec{d}_j & \text{with prob. } \min\{1, e^{\{f(x_j) - f(x_j + \lambda_j \vec{d}_j)\}/T_j}\} \\ x_j & \text{otherwise} \end{cases}$$

6: If any stopping criterion is satisfied stop, else return to the step 2.

In the figure 1.3, the method Hide-and-Seek was applied on the Ackley test function. The cooling scheme is 0.99^j, where j is the label of each iteration. This means that the temperature for the transition probability in each iteration is given by

$$T_j = \tau_j(x_0, x_1, \ldots, x_{j-1}) = 0.99^j$$

and it is in fact a cooling schedule, because $0.99^j \to 0$ as $j \to \infty$. As well as the Improving

Hit-and-Run method, Hide-and-Seek has good precision. Due to the sampling is uniform on the set Ω, the method requires to evaluate many times the objective function to find the desired solution, this is a disadvantage when the function is difficult to evaluate, and it can cause issues for solving hard problems. The efficiency of both methods is compared using different cooling schemes in reference [50].

1.4 Nelder-Mead Simplex Method

The method presented in this section is different to the "simplex method" proposed by Dantzig in [15] for linear programming problems. The name of the method presented here, is motivated by the geometric figure used to explore the region of feasible solutions, a simplex is a generalization of a triangle or a tetrahedron to higher dimensions. Roughly speaking, given $n+1$ points (vertices) $x_1, x_2, \ldots, x_{n+1} \in \mathbb{R}^n$ down the condition that the n vectors of the form:

$$x_2 - x_1, x_3 - x_1, \ldots, x_{n+1} - x_1 \tag{1.2}$$

are linearly independent, then a simplex is the polytope formed by the set:

$$S = \left\{ \lambda_1 x_1 + \lambda_2 x_2 + \cdots + \lambda_{n+1} x_{n+1}, \text{ s.t. } \lambda_i \geq 0 \text{ and } \sum_{i=1}^{n+1} \lambda_i = 1 \right\} \tag{1.3}$$

The requirement for the vectors in (1.2) to be L.I. is always necessary to carry out a non-degenerated search, as it will be shown in the appendix A.

The initial idea of this method was presented by J. A. Nelder and R. Mead in [39] as a proposal to find the minimum value of a function as in (1), this technique is clearly deterministic, and different papers show it to be locally convergent. This last is a disadvantage when the global minimum of a function with many local minimum is required.

Initially, the way for this technique to approach a minimum, is by carrying out some movements of the simplex:

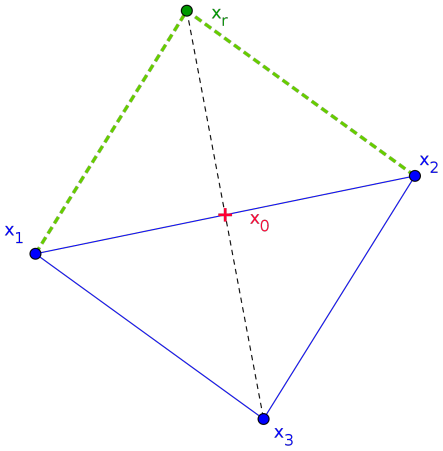
1. Reflection
2. Expansion
3. Contraction
4. Reduction

All this movements are illustrated in figure 1.4. After establishing these movements, one initial simplex is proposed, having in mind that the vertices can not be degenerated, this is assured by using the condition (1.2), where these are labeled with subinidices $(x_1, x_2, \ldots, x_{n+1})$ according to the inequalities:

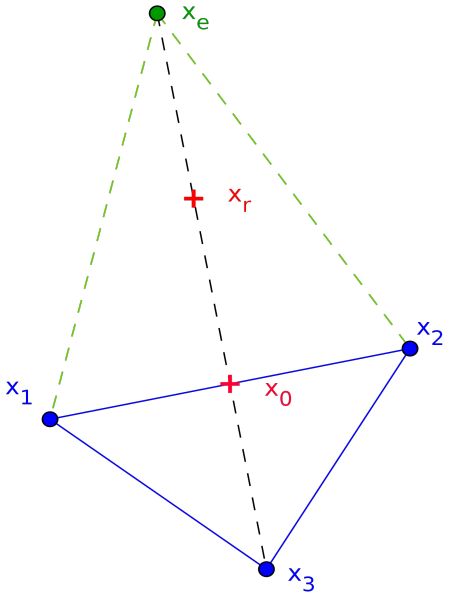
$$f(x_1) \leq f(x_2) \leq \cdots \leq f(x_{n+1}). \tag{1.4}$$

Finally the technique replaces the last vertex labeled with the mentioned movements, these cause the simplexes to approximate an interior local minimum. The algorithm 5 shows in detail the process that this technique executes.

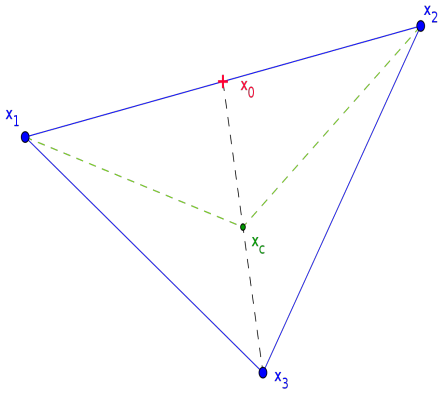
Generally in the algorithm 5 the constants are chosen as: $\alpha = 1, \gamma = 2, \rho = -1/2$ and $\sigma = 1/2$, but for example in [23], these are chosen if holds the inequalities: $0 < \alpha \leq 1$, $2 \leq \gamma$, $-1 < \rho < 0$ and $0 < \sigma < 1$, in order to do a more exhaustive search.



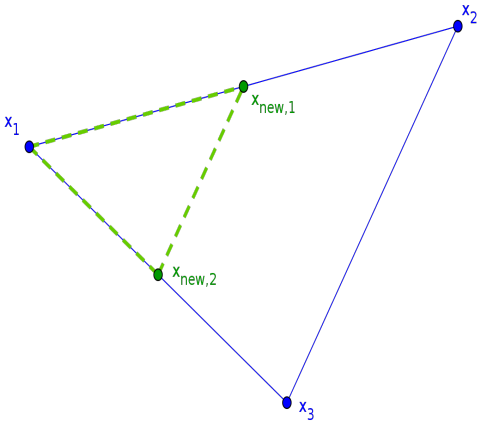
(a) Reflection.



(b) Expansion through a centroid.



(c) Contraction of one vertex.



(d) Reduction of the volume.

Figure 1.4: Possible movements of a simplex with the Simplex Nelderl-Mead Method 5

Algorithm 5 Nelder and Mead Method

1: Sort the vertices as in equation (1.4).
2: Compute the centroid $x_0 = \sum_{i=1}^{n} x_i/n$.
3: Compute the reflected point $x_r = x_0 + \alpha(x_0 - x_{n+1})$.
4: **if** $f(x_1) \leq f(x_r) < f(x_n)$, **then**
5: **Reflection:** $x_{n+1} \leftarrow x_r$ and go to step 1.
6: **else if** $f(x_r) < f(x_1)$, **then**
7: compute $x_e = x_0 + \gamma(x_0 - x_{n+1})$.
8: **if** $f(x_e) < f(x_r)$ **then**
9: **Expansion:** $x_{n+1} \leftarrow x_e$ and go to step 1.
10: **else**
11: $x_{n+1} \leftarrow x_r$ and got to step 1.
12: **end if**
13: **else**
14: Compute the contracted point $x_c = x_0 + \rho(x_0 - x_{n+1})$,
15: **if** $f(x_c) < f(x_{n+1})$ **then**
16: **Contraction:** $x_{n+1} \leftarrow x_c$ and go to step 1.
17: **else**
18: **Reduction:** $x_i = x_1 + \sigma(x_i - x_1)\, \forall i \in \{2, \ldots, n+1\}$ and go to step 1.
19: **end if**
20: **end if**

These technique is very similar to the Pattern Search method proposed by Robert Hooke and T. A. Jeeves in [25]. Their proposal include a geometric pattern as well, but the only difference is that this pattern only follows three movements: reflection, expansion or reduction.

This method is sensible to the initial simplex, for example, the figure 1.5(a) shows a development of the iterations converging to a local optimum, using an initial simplex with the vertices:

$$v_1 = [-1.90999, 1],\ v_2 = [-1.8, 2],\ v_3 = [-1, 2] \tag{1.5}$$

However, the figure 1.5(b) shows a global convergence, with the set of vertices:

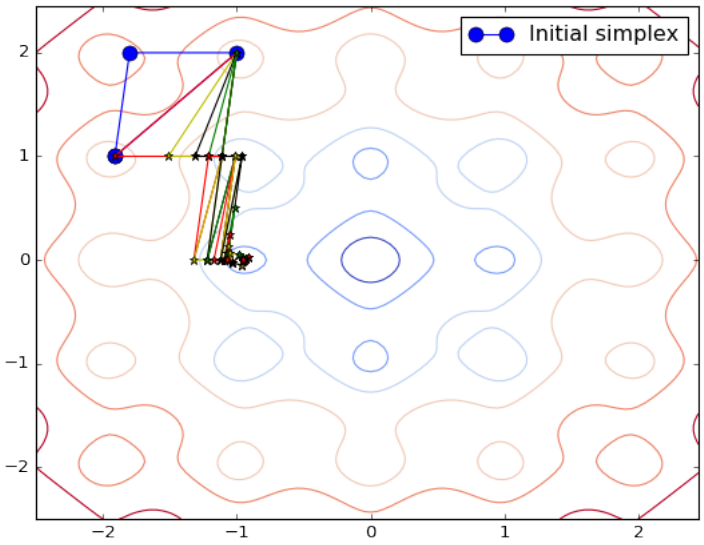
$$v_1 = [-1.8, 1],\ v_2 = [-1.8, 2],\ v_3 = [-1, 1.8] \tag{1.6}$$

the same algorithm reaches the global optimum.

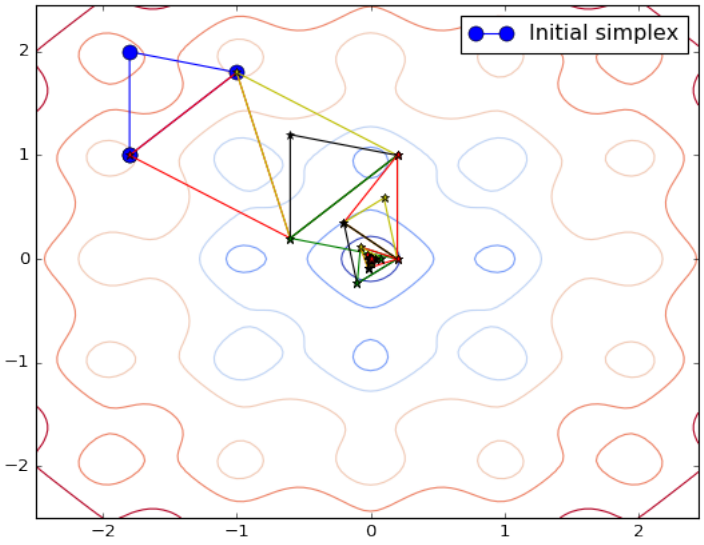
1.4.1 Simplex Simulated Annealing

As a solution to deal with the convergence to local minimum, there exists an approach motivated by Simulated Annealing, that in addition produces a method which explores solutions by a stochastic path. W. Press and S. Teukolsky in [45] proposed a novel method that combined the Simplex method and Simulated Annealing. The idea is to establish certain temperature T and adding the proportional logarithm of a random number normally distributed on $[0, 1]$ to the temperature on each value of vertices as in equation (1.7), and also subtract it from the value of the replacement vertices as in (1.8):

$$\tilde{f}_i = f(x_i) - T \ln(r), \tag{1.7}$$



(a) An example of local convergence for the Nelder and Mead algorithm



(b) An example of global convergence for the Nelder and Mead algorithm

Figure 1.5: Showing the sensibility on the initial simplex. This is an important issue when using the Nelder and Mead method. Here the function used is the Ackely test.

$$\tilde{f}_{rep} = f(x_{rep}) + T\ln(r), \tag{1.8}$$

From one hand, the equation (1.7) has two main consequences in the algorithm, first it perturbs the images to higher values, and second with these perturbed values, the vertices are sorted as in equation (1.4) but randomly. On the other hand the equation (1.8) perturbs new values to others lower, when movements of the simplex are happening. In conclusion, all these happen in order to accept bad movements while the temperature is reducing, which later can cause achieving a global optimum.

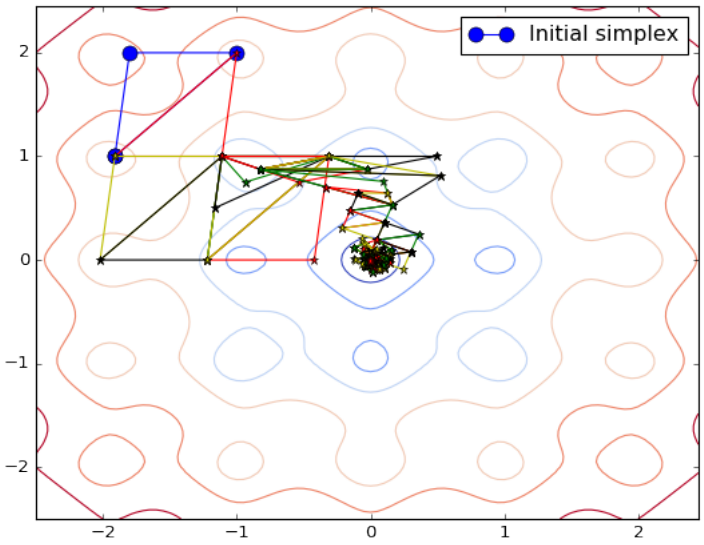
There are not schematic proofs that analyze the behavior of this approach, the only theoretical explanation to justify the convergence to minimum values, is that it is reduced to the Nelder-Mead simplex method as $T \to 0$. For example M. Cardoso et al. in [12] provided some results over test functions; additionally they provided schemes for reducing the temperature, and the most important contribution was a way to deal with constrained problems, where for example the set Ω is formed by nonlinear constrains.

The algorithm 6 shows the most simple proposal for this technique. In this algorithm, the tilde symbol ($\sim$) on the objective function represents, either the perturbation (1.7) (for ordering the vertices in the actual simplex) or either the perturbation (1.8) over the images of proposal vertices in the movement of the simplex.

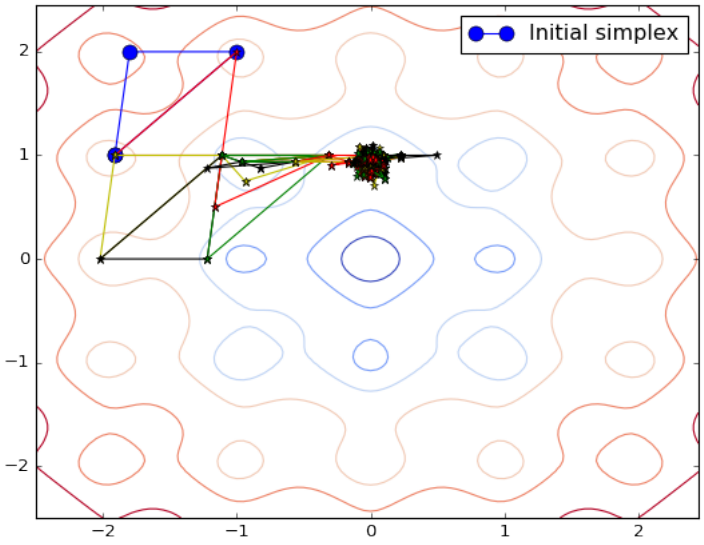
Algorithm 6 Simplex Simulated Annealing

1: Label the vertices such that $\tilde{f}_1 < \tilde{f}_2 < \ldots < \tilde{f}_{n+1}$.
2: Compute the centroid $x_0 = \sum_{i=1}^{n} x_i/n$.
3: Compute the reflected point $x_r = x_0 + \alpha(x_0 - x_{n+1})$ and its perturbed value $\tilde{f}_r$ given by equation (1.8).
4: **if** $\tilde{f}_1 \leq \tilde{f}_r < \tilde{f}_n$ **then**
5: $\quad x_{n+1} \leftarrow x_r$ and got to step 1
6: **else if** $\tilde{f}_r < \tilde{f}_1$ **then**
7: $\quad$ compute $x_e = x_0 + \gamma(x_0 - x_{n+1})$ and $\tilde{f}_e$
8: $\quad$ **if** $\tilde{f}_e < \tilde{f}_r$ **then**
9: $\quad\quad x_{n+1} \leftarrow x_e$ and go to step 1.
10: $\quad$ **else**
11: $\quad\quad x_{n+1} \leftarrow x_r$ and go to step 1.
12: $\quad$ **end if**
13: **else**
14: $\quad$ Compute $x_c = x_0 + \rho(x_0 - x_{n+1})$ and $\tilde{f}_c$
15: $\quad$ **if** $\tilde{f}_c < \tilde{f}_{n+1}$ **then**
16: $\quad\quad x_{n+1} \leftarrow x_c$ and go to step 1.
17: $\quad$ **else**
18: $\quad\quad x_i = x_1 + \sigma(x_i - x_1)\, \forall i \in \{2, \ldots, n+1\}$ and go to step 1.
19: $\quad$ **end if**
20: **end if**

An example of this technique is presented in the figure 1.6. It shows different answers with the same set of vertices in (1.5). The figure 1.6(a) shows a global convergence, while the figure 1.6(b) shows a convergence to a local minimum.



(a) An example of global convergence for Simplex Simulated Annealing



(b) An example of local convergence for Simplex Simulated Annealing

Figure 1.6: Using the Simplex Simulated Annealing on the Ackley test function. The cooling schedule is of the form 0.99^j.

Chapter 2

Developing a Molecular Model for Ionic Liquids Using SAFT-VRE Approach

This chapter will start by describing the main purpose. We will approximate some important parameters in a model of ionic liquids, this is composed by water (H_2O) which we call "specie one" with a number of molecules expressed in moles N_1; the second component are molecules with positive charge (cation) that are called "specie two" with a quantity N_2, and the final component of our main mixture is a compound of molecules charged negatively (anion) with a number of molecules N_3. An important term in the work is the relative density of each compound. Due to the high number of moles in certain experiments, we will associate to each specie a number called molar fraction. If we assume that N is the total number of molecules in the mixture, this fraction is expressed as:

$$x_i = \frac{N_i}{N} = \frac{N_i}{\sum_{i=1}^{3} N_i}. \tag{2.1}$$

Moreover, the additional densities $\rho = \frac{N}{V}$ and $\rho_i = \frac{N_i}{V}$ are remarkable.

2.1 Helmholtz Free Energy

The experiments are carried out in environments where the volume, the temperature and the quantity of the compounds are constant; then, in theory we can model the mixture using a thermodynamical potential that measures the work of the compounds in this environment. The potential is called 'Helmholtz free energy', which was developed by the German physicist Hermann von Helmholtz and usually is represented by the letter A, which is the initial letter of the word *Arbeit* which in German language means 'work'.

According to the SAFT (Statistical Association Fluid Theory), we assume that the dimensionless Helmholtz free energy of a fluid $A/(NkT)$ can be decomposed in terms of the ideal, monomer segment, associative and ionic contributions as

$$\frac{A}{NkT} = \frac{A^{IDEAL}}{NkT} + \frac{A^{MONO}}{NkT} + \frac{A^{ASSOC}}{NkT} + \frac{A^{IONS}}{NkT}. \tag{2.2}$$

There, k is the Boltzmann constant and T is the temperature. Each term on the right side in the equation (2.2) represents a different contribution in the compound.

For our purpose, it is necessary to compute the chemical potential (μ^x), which is defined as the change of the energy with respect to the number of moles of each component in our mixture. Using x as a label for any contribution, according to the relation $\frac{A^x}{kT} = Na^x$ then

$$\begin{aligned} \frac{\mu^x}{kT} &= \left[\frac{\mu_j^x}{kT}\right]_{j=1,2,3} \\ &= \left[\frac{\partial(Na^x)}{\partial N_j}\right]_{j=1,2,3} \end{aligned} \tag{2.3}$$

Also notice that

$$\frac{\mu_j^x}{kT} = \frac{\partial(Na^x)}{\partial N_j} = a^x + N\frac{\partial a^x}{\partial N_j}, \tag{2.4}$$

due to $\partial N/\partial N_j = 1$. For further computations, also notice that

$$\frac{\partial x_i}{\partial N_j} = \frac{N\delta_{ij} - N_i}{N^2} = \frac{1}{N}(\delta_{ij} - x_i) \tag{2.5}$$

and

$$\frac{\partial x_i}{\partial \rho} = \frac{\partial(N_i/(\rho V))}{\partial \rho} = -\frac{x_i}{\rho} \tag{2.6}$$

where δ_{ij} represents the Kronecker delta function.

Finally using the equations (2.2) and (2.3) the formulation for the chemical potential is

$$\frac{\mu}{kT} = \frac{\mu^{IDEAL}}{kT} + \frac{\mu^{MONO}}{kT} + \frac{\mu^{ASSOC}}{kT} + \frac{\mu^{IONS}}{kT} \tag{2.7}$$

2.1.1 Initial Considerations

We will model each atom of specie i as a sphere of certain diameter, the measure of these diameters are labeled as σ_{ii}. For the case of water (species 1), we define: the radius r_d as the distance from the center of the molecule to the association sites, and r_c as diameter of each site; see figure 2.1. Additionally, the attraction between particles of water is carried out in sites of different charges an is considered of short distance; it means that the interaction only occurs when there are already bonded, and additionally, is considered as an intermolecular hydrogen bonding.

On the other hand, the SAFT model in ref. [33] considers that the interaction between a molecule of type i with another of type k is described in the square-well form

$$u_{ik}(r) = \begin{cases} +\infty & \text{if} \quad r < \bar{\sigma}_{avg}, \\ -\epsilon_{ik} & \text{if} \quad \bar{\sigma}_{avg} \leq r \leq \lambda_{ik}\bar{\sigma}_{avg}, \\ 0 & \text{if} \quad r \geq \lambda_{ik}\bar{\sigma}_{avg}, \end{cases} \tag{2.8}$$

where r denotes the distance between the centers of the two molecules, $\bar{\sigma}_{avg} = (\sigma_{ii} + \sigma_{kk})/2$ and $\lambda_{ik} = (\sigma_{ii}\lambda_{ii} + \sigma_{kk}\lambda_{kk})/(\sigma_{ii} + \sigma_{kk})$

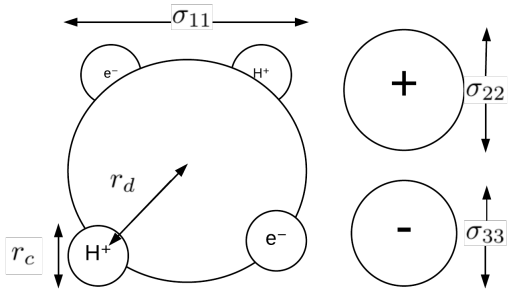


Figure 2.1: General scheme for the diameters σ_{11}, σ_{22}, σ_{33} in each specie; r_d as the distance between the center of the water molecule and the association sites, r_c is the diameter of the sites H^+ and e^-.

2.1.2 Ideal Contribution

The free energy of a ideal mixture is given by the expression [20, 51]

$$\frac{A^{IDEAl}}{NkT} = \sum_{i=1}^{n} x_i \ln\left(\rho_i \Lambda_i^3\right) - 1 \tag{2.9}$$

Where n is the number of species, $\rho_i = N_i/V$ is the molar density of specie i and Λ_i is the Broglie wavelength of the specie i. The relation for ideal chemical potential becomes in to

$$\frac{\mu^{IDEAL}}{kT} = \left[a^{IDEAL} + N\frac{\partial a^{IDEAL}}{\partial N_j}\right]_{j=1,2,3} \tag{2.10}$$

additionally,

$$\begin{aligned}
\frac{\partial a^{IDEAL}}{\partial N_j} &= \sum_{i=1}^{n} \frac{\partial\left\{x_i \ln\left(\rho_i \Lambda_i^3\right)\right\}}{\partial N_j} \\
&= \sum_{i=1}^{n} \left\{ \frac{\partial x_i}{\partial N_j} \ln\left(\rho_i \Lambda_i^3\right) + x_i \frac{\partial\left\{\ln\left(\rho_i \Lambda_i^3\right)\right\}}{\partial N_j} \right\} \\
&= \sum_{i=1}^{n} \left\{ \frac{1}{N}(\delta_{ij} - x_i) \ln\left(\rho_i \Lambda_i^3\right) + x_i \frac{1}{\rho_i \Lambda_i^3} \frac{\Lambda_i^3}{V} \frac{\partial N_i}{\partial N_j} \right\} \\
&= \frac{1}{N} \sum_{i=1}^{n} \delta_{ij} \ln\left(\rho_i \Lambda_i^3\right) - \frac{1}{N} \sum_{i=1}^{n} x_i \ln\left(\rho_i \Lambda_i^3\right) + \frac{1}{V} \sum_{i=1}^{n} x_i \frac{1}{\rho_i} \delta_{ij} \\
&= \frac{1}{N} \ln\left(\rho_j \Lambda_j^3\right) - \frac{1}{N} \sum_{i=1}^{n} x_i \ln\left(\rho_i \Lambda_i^3\right) + \frac{1}{N}
\end{aligned} \tag{2.11}$$

replacing (2.11) in (2.10), the ideal chemical potential of our mixture is

$$\frac{\mu^{IDEAL}}{kT} = \left[\ln\left(\rho_j \Lambda_j^3\right)\right]_{j=1,2,3} \tag{2.12}$$

2.1.3 Monomer Contribution

The monomer free energy is expressed as [20, 21, 19]

$$\frac{A^{MONO}}{NkT} = \left(\sum_{i=1}^{n} x_i m_i\right) \frac{A^M}{N_s kT}, \tag{2.13}$$

m_i is the number of spherical segments in each specie i and N_s is the total number of spherical segments. We assume m_i as 1, then $N_s = N$, usually the fraction $\frac{A^M}{N_s kT} = a^M$ which is monomer free energy per segment is expanded as in ref. [3, 4]:

$$a^M = a^{HS} + \beta a_1 + \beta^2 a_2 + \ldots, \tag{2.14}$$

being a^{HS} the free energy for a mixture of hard spheres, $\beta = 1/kT$ is the thermodynamic perk and a_1 and a_2 are both perturbations related to the attractive energy $-\epsilon_{ij}$, the first describes a mean of attractive energy, while the second describes fluctuations in the energy. In order to compute the monomer free energy, we will only have into account the first three terms in (2.14).

Therefore, the chemical potential in the monomer contribution is

$$\begin{aligned}
\frac{\mu^{MONO}}{kT} &= \left[a^M + N\frac{\partial a^M}{\partial N_j}\right]_{j=1,2,3} \\
&= \left[\left(a^{HS} + N\frac{\partial a^{HS}}{\partial N_j}\right) + \beta\left(a_1 + N\frac{\partial a_1}{\partial N_j}\right)\right. \\
&\qquad \left. + \beta^2\left(a_2 + N\frac{\partial a_2}{\partial N_j}\right)\right]_{j=1,2,3} \\
&= \frac{\mu^{HS}}{kT} + \beta\frac{\mu_1}{kT} + \beta^2\frac{\mu_2}{kT}
\end{aligned} \tag{2.15}$$

The free energy for a mixture of hard spheres is obtained with the expression [10, 34]

$$\begin{aligned}
a^{HS} &= \frac{6}{\pi\rho}\left\{\left(\frac{\zeta_2^3}{\zeta_3^2} - \zeta_0\right)\ln(1-\zeta_3) + \frac{3\zeta_1\zeta_2}{1-\zeta_3} + \frac{\zeta_2^3}{\zeta_3(1-\zeta_3)^2}\right\} \\
&= \frac{6}{\pi\rho}\Gamma(\zeta_0, \zeta_1, \zeta_2, \zeta_3)
\end{aligned} \tag{2.16}$$

where the reduced densities ζ_m are defined as

$$\zeta_m = \frac{\pi\rho}{6}\sum_{i=1}^{n} x_i \sigma_{ii}^m \tag{2.17}$$

Now, we will compute $\partial a^{HS}/\partial N_j$. First, note that:

$$\begin{aligned}
\frac{\partial a^{HS}}{\partial N_j} &= -\frac{6}{\pi\rho N}\Gamma + \frac{6}{\pi\rho}\frac{\partial\Gamma}{\partial N_j} \\
&= -\frac{1}{N}a^{HS} + \frac{6}{\pi\rho}\sum_{m=0}^{3}\frac{\partial\Gamma}{\partial\zeta_m}\frac{\partial\zeta_m}{\partial N_j}
\end{aligned} \tag{2.18}$$

and additionally,

$$\begin{aligned}\frac{\partial \zeta_m}{\partial N_j} &= \frac{\pi}{6}\frac{\partial \rho}{\partial N_j}\sum_{i=1}^{n} x_i \sigma_{ii}^m + \frac{\pi}{6}\rho \sum_{i=1}^{n}\frac{\partial x_i}{\partial N_j}\sigma_{ii}^m \\ &= \frac{\pi}{6}\frac{1}{V}\sum_{i=1}^{n} x_i \sigma_{ii}^m + \frac{\pi}{6}\rho \sum_{i=1}^{n}\frac{\delta_{ij}-x_i}{N}\sigma_{ii}^m \\ &= \frac{\pi}{6}\frac{1}{V}\left\{\cancel{\sum_{i=1}^{n} x_i \sigma_{ii}^m} + \sum_{i=1}^{n}\delta_{ij}\sigma_{ii}^m - \cancel{\sum_{i=1}^{n} x_i \sigma_{ii}^m}\right\} \\ &= \frac{\pi}{6}\frac{1}{V}\sigma_{jj}^m.\end{aligned} \tag{2.19}$$

Second,

$$\begin{aligned}\frac{6}{\pi\rho}\frac{\partial \Gamma}{\partial N_j} &= \frac{1}{N}\sum_{m=0}^{3}\frac{\partial \Gamma}{\partial \zeta_m}\sigma_{jj}^m \\ &= \frac{1}{N}\left\{-\ln(1-\zeta_3) + \frac{3\zeta_2}{1-\zeta_3}\sigma_{jj}\right. \\ &\quad + \left(\frac{3\zeta_1}{1-\zeta_3} + \frac{3\zeta_2^2\ln(1-\zeta_3)}{\zeta_3^2} + \frac{3\zeta_2^2}{(1-\zeta_3)^2\zeta_3}\right)\sigma_{jj}^2 \\ &\quad + \left(\frac{3\zeta_1\zeta_2}{(1-\zeta_3)^2} - \frac{\zeta_2^3 - \zeta_0\zeta_3^2}{\zeta_3^2(1-\zeta_3)}\right. \\ &\quad \left.\left. - \frac{2\zeta_2^3\ln(1-\zeta_3)}{\zeta_3^3} - \frac{\zeta_2^3}{(1-\zeta_3)^2\zeta_3^2} + \frac{2\zeta_2^3}{(1-\zeta_3)^3\zeta_3}\right)\sigma_{jj}^3\right\}.\end{aligned} \tag{2.20}$$

Finally, using the equations (2.15), (2.18), (2.19), (2.20) the chemical potential due to hard spheres is

$$\begin{aligned}\frac{\mu^{HS}}{kT} &= \left[-\ln(1-\zeta_3) + \frac{3\zeta_2}{1-\zeta_3}\sigma_{jj}\right. \\ &\quad + \left(\frac{3\zeta_1}{1-\zeta_3} + \frac{3\zeta_2^2\ln(1-\zeta_3)}{\zeta_3^2} + \frac{3\zeta_2^2}{(1-\zeta_3)^2\zeta_3}\right)\sigma_{jj}^2 \\ &\quad + \left(\frac{3\zeta_1\zeta_2}{(1-\zeta_3)^2} - \frac{\zeta_2^3 - \zeta_0\zeta_3^2}{\zeta_3^2(1-\zeta_3)} - \frac{2\zeta_2^3\ln(1-\zeta_3)}{\zeta_3^3}\right. \\ &\quad \left.\left. - \frac{\zeta_2^3}{(1-\zeta_3)^2\zeta_3^2} + \frac{2\zeta_2^3}{(1-\zeta_3)^3\zeta_3}\right)\sigma_{jj}^3\right]_{j=1,2,3}\end{aligned} \tag{2.21}$$

Now, we turn back to compute the perturbation terms. On one hand, the mean of attractive energy with range λ_{ik} and potential well constant $-\epsilon_{ik}$ is given by

$$a_1 = \sum_{i=1}^{n}\sum_{k=1}^{n} x_i x_k a_1^{ik} \tag{2.22}$$

where

$$a_1^{ik} = -\rho\alpha_{ik}^{VDW} g_{ik}^{HS}[\sigma_{ik}; \zeta_3^{\text{eff}}], \tag{2.23}$$

α_{ik}^{VDW} is the Van Der Waals attractive constant for the ij interactions; these interactions are generalized as

$$\alpha_{ik}^{VDW} = \frac{2\pi}{3}\epsilon_{ik}\sigma_{ik}^3(\lambda_{ik}^3 - 1), \tag{2.24}$$

the mixing rule $g_{ik}^{HS}[\sigma_{ik};\zeta_3^{\text{eff}}]$ used is the MX3b [19], which states a radial distribution function as

$$g_{ik}^{HS}[\sigma_{ik};\zeta_3^{\text{eff}}] = \frac{1}{1-\zeta_3^{\text{eff}}} + 3\frac{D_{ik}\zeta_3^{\text{eff}}}{(1-\zeta_3^{\text{eff}})^2} + 2\frac{(D_{ik}\zeta_3^{\text{eff}})^2}{(1-\zeta_3^{\text{eff}})^3} \tag{2.25}$$

where ζ_3^{eff} is an effective packing given by

$$\zeta_3^{\text{eff}}(\zeta_3,\lambda_{ik}) = c_1\zeta_3 + c_2\zeta_3^2 + c_3\zeta_3^3, \tag{2.26}$$

the coefficients of the polynomial expression above, are given by the matrix product:

$$\begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} = \begin{pmatrix} 2.25855 & -1.50349 & 0.249434 \\ -0.669270 & 1.40049 & -0.827739 \\ 10.1576 & -15.0427 & 5.30827 \end{pmatrix} \begin{pmatrix} 1 \\ \lambda_{ik} \\ \lambda_{ik}^2 \end{pmatrix} \tag{2.27}$$

and D_{ik} is

$$D_{ik} = \frac{\sigma_{ii}\sigma_{kk}}{\sigma_{ii}+\sigma_{kk}} \frac{\sum_{i=1}^n x_i\sigma_{ii}^2}{\sum_{i=1}^n x_i\sigma_{ii}^3} \tag{2.28}$$

In the monomer contribution, we are only considering the interactions between species of the same type; however, we are not neglecting the interaction between water-ion. It is taken into account, by the dielectric constant of a continuum medium and supposing, that both ion types interact inside the solvent (water). Therefore, the equation (2.22) becomes in to

$$a_1 = x_1^2 a_1^{11} + x_2^2 a_1^{22} + x_3^2 a_1^{33}. \tag{2.29}$$

This is because of $\epsilon_{12} = \epsilon_{21} = \epsilon_{13} = \epsilon_{31} = \epsilon_{23} = \epsilon_{32} = 0$. In the equation (2.24), $\sigma_{ik} = (\sigma_{ii}+\sigma_{kk})/2$ and $\lambda_{ik} = (\lambda_{ii}\sigma_{ii} + \lambda_{kk}\sigma_{kk})/(\sigma_{ii}+\sigma_{kk})$.

Now, we should compute the chemical potential provided by this mean perturbation. First of all, observe that:

$$\beta\frac{\mu_1}{kT} = \beta\left[\frac{\partial(Na_1)}{\partial N_j}\right]_{j=1,2,3} = \beta\left[a_1 + N\frac{\partial a_1}{\partial N_j}\right]_{j=1,2,3}. \tag{2.30}$$

Second, the partial derivative in the equation (2.30) above is equal to

$$\begin{aligned}
\frac{\partial a_1}{\partial N_j} &= \sum_{i=1}^n\sum_{k=1}^n \frac{\partial(x_i x_k a_1^{ik})}{\partial N_j} \\
&= \sum_{i=1}^n\sum_{k=1}^n \left\{ \frac{1}{N}(\delta_{ij}-x_i)x_k a_1^{ik} + \frac{1}{N}(\delta_{kj}-x_k)x_i a_1^{ik} + x_i x_k \frac{\partial a_1^{ik}}{\partial N_j} \right\} \\
&= \sum_{i=1}^n\sum_{k=1}^n \frac{1}{N}(\delta_{ij}-x_i)x_k a_1^{ik} + \sum_{i=1}^n\sum_{k=1}^n \frac{1}{N}(\delta_{kj}-x_k)x_i a_1^{ik} \\
&\qquad + \sum_{i=1}^n\sum_{k=1}^n x_i x_k \frac{\partial a_1^{ik}}{\partial N_j} \\
&= \frac{2}{N}(1-x_j)\sum_{i=1}^n x_i a_1^{ji} + \sum_{i=1}^n\sum_{k=1}^n x_i x_k \frac{\partial a_1^{ik}}{\partial N_j}.
\end{aligned} \tag{2.31}$$

The third fact is that the partial derivative $\partial a_1^{ik}/\partial N_j$ in the equation (2.31) can be calculated as

$$\begin{aligned}\frac{\partial a_1^{ik}}{\partial N_j} &= -\frac{\partial \rho}{\partial N_j}\alpha_{ik}^{VDW} g_{ik}^{HS} - \rho \cancelto{0}{\frac{\partial \alpha_{ik}^{VDW}}{\partial N_j}} g_{ik}^{HS} - \rho \alpha_{ik}^{VDW} \frac{\partial g_{ik}^{HS}}{\partial N_j} \\ &= -\frac{1}{V}\alpha_{ik}^{VDW} g_{ik}^{HS} - \rho \alpha_{ik}^{VDW} \left\{ \frac{\partial g_{ik}^{HS}}{\partial \zeta_3^{\text{eff}}} \frac{\partial \zeta_3^{\text{eff}}}{\partial \zeta_3} \frac{\partial \zeta_3}{\partial N_j} + \frac{\partial g_{ik}^{HS}}{\partial D_{ik}} \frac{\partial D_{ik}}{\partial N_j} \right\}\end{aligned} \tag{2.32}$$

In addition, it is necessary to compute some derivatives in the right hand of the expression (2.32). These are:

$$\frac{\partial g_{ik}^{HS}}{\partial \zeta_3^{\text{eff}}} = \frac{1 + 3D_{ik}}{(1-\zeta_3^{\text{eff}})^2} + 6\frac{D_{ik}\zeta_3^{\text{eff}}}{(1-\zeta_3^{\text{eff}})^3} + 4\frac{D_{ik}^2\zeta_3^{\text{eff}}}{(1-\zeta_3^{\text{eff}})^3} + 6\frac{(D_{ik}\zeta_3^{\text{eff}})^2}{(1-\zeta_3^{\text{eff}})^4}, \tag{2.33a}$$

$$\frac{\partial g_{ik}^{HS}}{\partial D_{ik}} = 3\frac{\zeta_3^{\text{eff}}}{(1-\zeta_3^{\text{eff}})^2} + 4\frac{D_{ik}(\zeta_3^{\text{eff}})^2}{(1-\zeta_3^{\text{eff}})^3}, \tag{2.33b}$$

$$\frac{\partial \zeta_3^{\text{eff}}}{\partial \zeta_3} = c_1 + 2c_2\zeta_3 + 3c_3\zeta_3^2, \tag{2.33c}$$

$$\frac{\partial D_{ik}}{\partial N_j} = \frac{1}{N}\frac{\sigma_{ii}\sigma_{kk}}{\sigma_{ii}+\sigma_{kk}} \frac{\sigma_{jj}^2 \sum_{i=1}^n x_i\sigma_{ii}^3 - \sigma_{jj}^3 \sum_{i=1}^n x_i\sigma_{ii}^2}{(\sum_{i=1}^n x_i\sigma_{ii}^3)^2}. \tag{2.33d}$$

Mixing the equations (2.30), (2.31) and (2.32) we finally get the chemical potential provided by the mean perturbation as

$$
\begin{aligned}
\beta\frac{\mu_1}{kT} &= \beta\Bigg[a_1 + 2(1-x_j)\sum_{i=1}^{n} x_i a_1^{ji} \\
&\quad + \sum_{i=1}^{n}\sum_{k=1}^{n} x_i x_k \Big\{-\rho\alpha_{ik}^{VDW} g_{ik}^{HS} - N\rho\alpha_{ik}^{VDW} \\
&\qquad\qquad \times \left[\frac{\partial g_{ik}^{HS}}{\partial \zeta_3^{\text{eff}}}\frac{\partial \zeta_3^{\text{eff}}}{\partial \zeta_3}\frac{\partial \zeta_3}{\partial N_j} + \frac{\partial g_{ik}^{HS}}{\partial D_{ik}}\frac{\partial D_{ik}}{\partial N_j}\right]\Big\}\Bigg]_{j=1,2,3} \\
&= \beta\Bigg[2(1-x_j)\sum_{i=1}^{n} x_i a_1^{ji} \\
&\quad - \sum_{i=1}^{n}\sum_{k=1}^{n} x_i x_k \left\{N\rho\alpha_{ik}^{VDW}\left[\frac{\partial g_{ik}^{HS}}{\partial \zeta_3^{\text{eff}}}\frac{\partial \zeta_3^{\text{eff}}}{\partial \zeta_3}\frac{\partial \zeta_3}{\partial N_j} + \frac{\partial g_{ik}^{HS}}{\partial D_{ik}}\frac{\partial D_{ik}}{\partial N_j}\right]\right\}\Bigg]_{j=1,2,3} \\
&= \beta\Bigg[2(1-x_j)\sum_{i=1}^{n} x_i a_1^{ji} - \sum_{i=1}^{n}\sum_{k=1}^{n} x_i x_k \rho\alpha_{ik}^{VDW}\left\{\frac{\partial g_{ik}^{HS}}{\partial \zeta_3^{\text{eff}}}\frac{\partial \zeta_3^{\text{eff}}}{\partial \zeta_3}\frac{\pi}{6}\rho\sigma_{jj}^3\right. \\
&\quad \left. + \frac{\partial g_{ik}^{HS}}{\partial D_{ik}}\left[\frac{\sigma_{ii}\sigma_{kk}}{\sigma_{ii}+\sigma_{kk}}\frac{\sigma_{jj}^2\sum_{i=1}^{n} x_i\sigma_{ii}^3 - \sigma_{jj}^3\sum_{i=1}^{n} x_i\sigma_{ii}^2}{(\sum_{i=1}^{n} x_i\sigma_{ii}^3)^2}\right]\right\}\Bigg]_{j=1,2,3} \\
&= \beta\Bigg[2(1-x_j)\sum_{i=1}^{n} x_i a_1^{ji} - \frac{\pi}{6}\rho\sigma_{jj}^3\sum_{i=1}^{n}\sum_{k=1}^{n} x_i x_k \rho\alpha_{ik}^{VDW}\frac{\partial g_{ik}^{HS}}{\partial \zeta_3^{\text{eff}}}\frac{\partial \zeta_3^{\text{eff}}}{\partial \zeta_3} \\
&\quad - \left\{\frac{\sigma_{jj}^2\sum_{i=1}^{n} x_i\sigma_{ii}^3 - \sigma_{jj}^3\sum_{i=1}^{n} x_i\sigma_{ii}^2}{(\sum_{i=1}^{n} x_i\sigma_{ii}^3)^2}\right\} \\
&\qquad\qquad \times \sum_{i=1}^{n}\sum_{k=1}^{n}\frac{\sigma_{ii}\sigma_{kk}}{\sigma_{ii}+\sigma_{kk}} x_i x_k \rho\alpha_{ik}^{VDW}\frac{\partial g_{ik}^{HS}}{\partial D_{ik}}\Bigg]_{j=1,2,3}
\end{aligned}
\tag{2.34}
$$

On the other hand, the fluctuation of the energy for a mixture is given by

$$
a_2 = \sum_{i=1}^{n}\sum_{k=1}^{n} x_i x_k a_2^{ik} \tag{2.35}
$$

where by [3]:

$$
a_2^{ik} = \frac{1}{2}K^{HS}\epsilon_{ik}\rho\frac{\partial a_1^{ik}}{\partial \rho} \tag{2.36}
$$

and K^{HS} the isothermal compressibility of a mixture of hard spheres, we use by [47]

$$
K^{HS} = \frac{\zeta_0(1-\zeta_3)^4}{\zeta_0(1-\zeta_3)^2 + 6\zeta_1\zeta_2(1-\zeta_3) + 9\zeta_2^3}. \tag{2.37}
$$

By the equations (2.40a) and (2.40b), we can reduce the general summation term in the equation (2.36) by saying that

$$
\frac{\partial a_1^{ik}}{\partial \rho} = -\alpha_{ik}^{VDW} g_{ik}^{HS} - \rho\alpha_{ik}^{VDW}\cancelto{0}{\frac{\partial g_{ik}^{HS}}{\partial \rho}}, \tag{2.38}
$$

this is followed by the fact that the partial derivative in equation (2.38) can be calculated as:

$$\frac{\partial g_{ik}^{HS}}{\partial \rho} = \frac{\partial g_{ik}^{HS}}{\partial \zeta_3^{\text{eff}}} \frac{\partial \zeta_3^{\text{eff}}}{\partial \zeta_3} \frac{\partial \zeta_3}{\partial \rho} + \frac{\partial g_{ik}^{HS}}{\partial D_{ik}} \frac{\partial D_{ik}}{\partial \rho}. \tag{2.39}$$

This last equation (2.39) is the reason we should compute the expressions:

$$\frac{\partial \zeta_3}{\partial \rho} = \frac{\pi}{6} \sum_{i=1}^{n} x_i \sigma_{ii}^3 + \frac{\rho\pi}{6} \sum_{i=1}^{n} \frac{\partial x_i \cdot \sigma_{ii}^3}{\partial \rho} = \cancelto{0}{\frac{\pi}{6} \sum_{i=1}^{n} x_i \sigma_{ii}^3} - \frac{\rho\pi}{6} \sum_{i=1}^{n} \frac{x_i \sigma_{ii}^3}{\rho} \tag{2.40a}$$

$$\frac{\partial D_{ik}}{\partial \rho} = \frac{\sigma_{ii}\sigma_{kk}}{\sigma_{ii} + \sigma_{kk}} \cancelto{0}{\frac{-\frac{1}{\rho}\left(\sum_{i=1}^{n} x_i \sigma_{ii}^2 \sum_{i=1}^{n} x_i \sigma_{ii}^3 - \sum_{i=1}^{n} x_i \sigma_{ii}^2 \sum_{i=1}^{n} x_i \sigma_{ii}^3\right)}{\left(\sum_{i=1}^{n} x_i \sigma_{ii}^3\right)^2}} \tag{2.40b}$$

Hence, it is clear that the fluctuation energy in the monomer contribution has the form

$$\begin{aligned} a_2 &= \frac{1}{2} \sum_{i=1}^{n} \sum_{k=1}^{n} x_i x_k K^{HS} \epsilon_{ik} \left(-\rho \alpha_{ik}^{VDW} g_{ik}^{HS}\right) \\ &= \frac{1}{2} K^{HS} \sum_{i=1}^{n} \sum_{k=1}^{n} x_i x_k \epsilon_{ik} a_1^{ik}. \end{aligned} \tag{2.41}$$

As the mean energy, the chemical potential contributed by the fluctuation energy is given by

$$\begin{aligned} \beta^2 \frac{\mu_2}{kT} &= \beta^2 \left[a_2 + N \frac{\partial a_2}{\partial N_j} \right]_{j=1,2,3} \\ &= \beta^2 \left[a_2 + \frac{N}{2} \left\{ \frac{2}{N}(1 - x_j) K^{HS} \sum_{i=1}^{n} x_i \epsilon_{ij} a_1^{ij} \right.\right. \\ &\quad \left.\left. + K^{HS} \sum_{i=1}^{n} \sum_{k=1}^{n} x_i x_k \epsilon_{ik} \frac{\partial a_1^{ik}}{\partial N_j} + \frac{\partial K^{HS}}{\partial N_j} \sum_{i=1}^{n} \sum_{k=1}^{n} x_i x_k \epsilon_{ik} a_1^{ik} \right\} \right]_{j=1,2,3} \\ &= \beta^2 \left[(1 - x_j) K^{HS} \sum_{i=1}^{n} x_i \epsilon_{ij} a_1^{ij} \right. \\ &\quad - \frac{1}{2} K^{HS} \sum_{i=1}^{n} \sum_{k=1}^{n} x_i x_k \epsilon_{ik} \rho \alpha_{ik}^{VDW} \left(N \frac{\partial g_{ik}^{HS}}{\partial N_j} \right) \\ &\quad \left. + \frac{1}{2} \left(N \frac{\partial K^{HS}}{\partial N_j} \right) \sum_{i=1}^{n} \sum_{k=1}^{n} x_i x_k \epsilon_{ik} a_1^{ik} \right]_{j=1,2,3} \end{aligned} \tag{2.42}$$

but, having into account that

$$\begin{aligned}\frac{\partial K^{HS}}{\partial N_j} &= \sum_{m=0}^{3} \frac{\partial K^{HS}}{\partial \zeta_m}\frac{\partial \zeta_m}{\partial N_j} = \frac{\pi}{6}\frac{1}{V}\sum_{m=0}^{3}\frac{\partial K^{HS}}{\partial \zeta_m}\sigma_{jj}^m \\ &= \frac{\pi}{6}\frac{1}{V}\left\{\frac{(1-\zeta_3)^4}{\zeta_0(1-\zeta_3)^2+6\zeta_1\zeta_2(1-\zeta_3)+9\zeta_2^3}\right. \\ &\quad -\frac{\zeta_0(1-\zeta_3)^6}{(\zeta_0(1-\zeta_3)^2+6\zeta_1\zeta_2(1-\zeta_3)+9\zeta_2^3)^2} \\ &\quad -\frac{6\zeta_0\zeta_2(1-\zeta_3)^5}{(\zeta_0(1-\zeta_3)^2+6\zeta_1\zeta_2(1-\zeta_3)+9\zeta_2^3)^2}\sigma_{jj} \\ &\quad -\frac{\zeta_0(27\zeta_2^2+6\zeta_1(1-\zeta_3))(1-\zeta_3)^4}{(\zeta_0(1-\zeta_3)^2+6\zeta_1\zeta_2(1-\zeta_3)+9\zeta_2^3)^2}\sigma_{jj}^2 \\ &\quad +\frac{\zeta_0(6\zeta_1\zeta_2+2\zeta_0(1-\zeta_3))(1-\zeta_3)^4}{(\zeta_0(1-\zeta_3)^2+6\zeta_1\zeta_2(1-\zeta_3)+9\zeta_2^3)^2}\sigma_{jj}^3 \\ &\quad \left.-\frac{4\zeta_0(1-\zeta_3)^3}{\zeta_0(1-\zeta_3)^2+6\zeta_1\zeta_2(1-\zeta_3)+9\zeta_2^3}\sigma_{jj}^3\right\}.\end{aligned} \tag{2.43}$$

2.1.4 Association Contribution

We will consider the free energy due to association in the sites of water as the only contribution of this type into the Helmholtz free energy expression. Since the theory of Wertheim [26, 14], the association contribution is given by the equation [20]

$$\frac{A^{ASSOC}}{NkT} = \sum_{i=1}^{n} x_i \left\{\sum_{a\in s_i}\left(\ln X_{a,i} - \frac{X_{a,i}}{2}\right) + \frac{\#[s_i]}{2}\right\}. \tag{2.44}$$

In the expression (2.44), the first sum is made over each specie i and the second over types of sites s_i of each specie i. $X_{a,i}$ represents the fraction of molecules of type i that are not bonded at a site a and are obtained from the mass action equation [13, 14]

$$X_{a,i} = \frac{1}{1+\sum_{k=1}^{n}\sum_{b\in s_i, b\neq a}\rho x_k X_{b,k}\Delta_{a,b,i,k}}. \tag{2.45}$$

The function $\Delta_{a,b,i,k}$ characterize the association between a molecule of specie i with other k in the sites of type a and b respectively,

$$\Delta_{a,b,i,k} = (K_{a,b,i,k})(f_{a,b,i,k})(g^M(\sigma_{ik};\zeta_3)) \tag{2.46}$$

$K_{a,b,i,k}$ is the available volume for bonding in terms of the distance, to the center of the molecule r_d and the diameter of the sites r_c, $f_{a,b,i,k}$ is the Mayer function $f_{a,b,i,k} = e^{-\Phi_{a,b,i,k}/kT} - 1$ of the $a-b$ interaction $\Phi_{a,b,i,k}$, and $g^M(\sigma_{ik};\zeta_3)$ is a contact value of the radial distribution function of the reference unbounded square-well fluid. It is obtained as the expansion [21, 19]

$$g^M(\sigma_{ik};\zeta_3) = g_0^{HS}(\sigma_{ik};\zeta_3) + \beta\epsilon_{ik}g_1(\sigma_{ik};\zeta_3) \tag{2.47}$$

where by [10]

$$g_0^{HS}(\sigma_{ik};\zeta_3) = \frac{1}{1-\zeta_3} + 3\frac{D_{ik}\zeta_3}{(1-\zeta_3)^2} + 2\frac{(D_{ik}\zeta_3)^2}{(1-\zeta_3)^3} \tag{2.48}$$

and

$$g_1(\sigma_{ik};\zeta_3) = g_0^{HS}(\sigma_{ik};\zeta_3) + (\lambda_{ik}^3 - 1)\frac{\partial g_{ik}^{HS}[\sigma_{ik};\zeta_3^{\text{eff}}]}{\partial \zeta_3^{\text{eff}}}\left(\frac{\lambda_{ik}}{3}\frac{\partial \zeta_3^{\text{eff}}}{\partial \lambda_{ik}} - \zeta_3\frac{\partial \zeta_3^{\text{eff}}}{\partial \zeta_3}\right) \tag{2.49}$$

having in mind that

$$\frac{\partial \zeta_3^{\text{eff}}}{\partial \lambda_{ik}} = d_1\zeta_3 + d_2\zeta_3^2 + d_3\zeta_3^3 \tag{2.50}$$

and that the coefficients of the polynomial expression (2.50) above, are obtained by using the matrix equation (2.51):

$$\begin{pmatrix} d_1 \\ d_2 \\ d_3 \end{pmatrix} = \begin{pmatrix} -1.50349 & 0.249434 \\ 1.40049 & -0.827739 \\ -15.0427 & 5.30827 \end{pmatrix}\begin{pmatrix} 1 \\ 2\lambda_{ik} \end{pmatrix} \tag{2.51}$$

As we have mentioned at the beginning of the subsection, we will consider only the association among the molecules of water, but in addition we will assume that the fraction $X_{e^-,1}$ and $X_{H^+,1}$ are identical, it means that $X_{e^-,1} = X_{H^+,1} = X_1$. Therefore, since $s_1 = \{e^-, e^-, H^+, H^+\}$, the association free energy contribution in the equation (2.44) becomes into

$$\begin{aligned} \frac{A^{ASSOC}}{NkT} &= x_1\sum_{i=1}^{n}\left\{\sum_{a\in s_i}\left(\ln X_{a,i} - \frac{X_{a,i}}{2}\right) + \frac{\#[s_i]}{2}\right\}. \\ &= x_1\sum_{a\in s_1}\left(\ln X_{a,1} - \frac{X_{a,1}}{2}\right) + \frac{4}{2} \\ &= x_1\left\{4\left(\ln X_1 - \frac{X_1}{2}\right) + 2\right\}. \end{aligned} \tag{2.52}$$

Also, under the same assumptions, the fraction of molecules of water that are not bonded (no matter in what site) is equal to

$$X_1 = \frac{1}{1 + 2\rho x_1 X_1 \Delta_{11}}. \tag{2.53}$$

As consequence of the equation (2.53), an explicit expression with physical meaning (by obviating the negative root) for the fraction X_1 is

$$X_1 = \frac{-1 + \sqrt{1 + 8\rho x_1 \Delta_{11}}}{4\rho x_1 \Delta_{11}} \tag{2.54}$$

Now we compute the chemical potential due to the free association energy. Using the expressions (2.3) and (2.52) we obtain:

$$\begin{aligned} \frac{\mu^{ASSOC}}{kT} &= \left[a^{ASSOC} + (\delta_{1j} - x_1)\left\{4\left(\ln X_1 - \frac{X_1}{2}\right) + 2\right\}\right. \\ &\quad \left. + x_1\left\{4\left(\frac{1}{X_1} - \frac{1}{2}\right)N\frac{\partial X_1}{\partial N_j}\right\}\right]_{j=1,2,3} \\ &= \left[\delta_{1j}\left\{4\left(\ln X_1 - \frac{X_1}{2}\right) + 2\right\} + 4x_1\left(\frac{1}{X_1} - \frac{1}{2}\right)N\frac{\partial X_1}{\partial N_j}\right]_{j=1,2,3} \end{aligned} \tag{2.55}$$

where, from (2.53) we compute

$$N\frac{\partial X_1}{\partial N_j} = -2\rho X_1^2 \left\{ \frac{x_1\Delta_{11} + (\delta_{1j} - x_1)\Delta_{11} + x_1\left(N\frac{\partial \Delta_{11}}{\partial N_j}\right)}{1 + 4\rho x_1 X_1 \Delta_{11}} \right\} \tag{2.56}$$

and

$$\begin{aligned} N\frac{\partial \Delta_{11}}{\partial N_j} &= N\frac{\partial K_{11}}{\partial N_j}^{0} f_{11} g^M(\sigma_{11};\zeta_3) + NK_{11} f_{11} \frac{\partial f_{11}}{\partial N_j}^{0} g^M(\sigma_{11};\zeta_3) \\ &\quad + NK_{11} f_{11} \frac{\partial g^M(\sigma_{11};\zeta_3)}{\partial N_j} \\ &= K_{11} f_{11} \left\{ N\frac{\partial g_0^{HS}(\sigma_{11};\zeta_3)}{\partial N_j} + \beta\epsilon_{11} N\frac{\partial g_1(\sigma_{11};\zeta_3)}{\partial N_j} \right\} \\ &= K_{11} f_{11} \left\{ (1+\beta\epsilon_{11}) N\frac{\partial g_0^{HS}}{\partial N_j} \right. \\ &\quad + \beta\epsilon_{11}(\lambda_{11}^3 - 1)\left[N\frac{\partial^2 g_{11}^{HS}}{\partial N_j \partial \zeta_3^{\text{eff}}} \left(\frac{\lambda_{11}}{3}\frac{\partial \zeta_3^{\text{eff}}}{\partial \lambda_{11}} - \zeta_3 \frac{\partial \zeta_3^{\text{eff}}}{\partial \zeta_3} \right) \right. \\ &\quad \left. \left. + \frac{\partial g_{11}^{HS}}{\partial \zeta_3^{\text{eff}}} \left(\frac{\lambda_{11}}{3} N\frac{\partial^2 \zeta_3^{\text{eff}}}{\partial N_j \partial \lambda_{11}} - \frac{\pi}{6}\rho\sigma_{jj}^3 \frac{\partial \zeta_3^{\text{eff}}}{\partial \zeta_3} - \zeta_3 N\frac{\partial^2 \zeta_3^{\text{eff}}}{\partial N_j \partial \zeta_3} \right) \right] \right\} \end{aligned} \tag{2.57}$$

Additional derivatives are presented in the equation (2.58)

$$\begin{aligned} N\frac{\partial g_0^{HS}}{\partial N_j} &= \left\{ \frac{1+3D_{11}}{(1-\zeta_3)^2} + 6\frac{D_{11}\zeta_3}{(1-\zeta_3)^3} + 4\frac{D_{11}^2\zeta_3}{(1-\zeta_3)^3} + 6\frac{(D_{11}\zeta_3)^2}{(1-\zeta_3)^4} \right\} \frac{\pi}{6}\rho\sigma_{jj}^3 \\ &\quad + \left\{ 3\frac{\zeta_3}{(1-\zeta_3)^2} + 4\frac{D_{11}\zeta_3^2}{(1-\zeta_3)^3} \right\} \frac{\sigma_{11}}{2} \frac{\sigma_{jj}^2 \sum_{i=1}^n x_i\sigma_{ii}^3 - \sigma_{jj}^3 \sum_{i=1}^n x_i\sigma_{ii}^2}{(\sum_{i=1}^n x_i\sigma_{ii}^3)^2} \end{aligned} \tag{2.58a}$$

$$\begin{aligned} N\frac{\partial^2 g_{11}^{HS}}{\partial N_j \partial \zeta_3^{\text{eff}}} &= \frac{\partial^2 g_{11}^{HS}}{(\partial \zeta_3^{\text{eff}})^2}(c_1 + 2c_2\zeta_3 + 3c_3\zeta_3^2)\frac{\pi}{6}\rho\sigma_{jj}^3 \\ &\quad + \frac{\partial^2 g_{11}^{HS}}{\partial D_{11}\partial \zeta_3^{\text{eff}}} \frac{\sigma_{11}}{2} \frac{\sigma_{jj}^2 \sum_{i=1}^n x_i\sigma_{ii}^3 - \sigma_{jj}^3 \sum_{i=1}^n x_i\sigma_{ii}^2}{(\sum_{i=1}^n x_i\sigma_{ii}^3)^2} \end{aligned} \tag{2.58b}$$

$$N\frac{\partial^2 \zeta_3^{\text{eff}}}{\partial N_j \partial \lambda_{11}} = (d_1 + 2d_2\zeta_3 + 3d_3\zeta_3^2)\frac{\pi}{6}\rho\sigma_{jj}^3 \tag{2.58c}$$

$$N\frac{\partial^2 \zeta_3^{\text{eff}}}{\partial N_j \partial \zeta_3} = (2c_2 + 6c_3\zeta_3)\frac{\pi}{6}\rho\sigma_{jj}^3 \tag{2.58d}$$

$$\frac{\partial^2 g_{11}^{HS}}{(\partial \zeta_3^{\text{eff}})^2} = \frac{4D_{11}^2 + 12D_{11} + 2}{(1-\zeta_3^{\text{eff}})^3} + \frac{6\zeta_3^{\text{eff}}(4D_{11}^2 + 3D_{11})}{(1-\zeta_3^{\text{eff}})^4} + \frac{24(D_{11}\zeta_3^{\text{eff}})^2}{(1-\zeta_3^{\text{eff}})^5} \tag{2.58e}$$

$$\frac{\partial^2 g_{11}^{HS}}{\partial D_{11}\partial \zeta_3^{\text{eff}}} = \frac{3}{(1-\zeta_3^{\text{eff}})^2} + \frac{8\zeta_3^{\text{eff}}(D_{11}+1)}{(1-\zeta_3^{\text{eff}})^3} + \frac{12(\zeta_3^{\text{eff}})^2 D_{11}}{(1-\zeta_3^{\text{eff}})^4} \tag{2.58f}$$

2.1.5 Ionic Contribution

In this contribution we have assume that the anion and cation species are both inside a continuum medium with a dielectric constant (D); as we have mentioned both species are represented by charged spheres with interaction given by the equation (2.8). The common diameter have been defined as the average

$$\tilde{\sigma} = \sum_{i=2}^{n} \tilde{x}_i \sigma_{ii} \tag{2.59}$$

there $\tilde{x}_i$ expresses a corrected molar fraction of a specie i, in other words as

$$\tilde{x}_i = \frac{N_i}{\sum_{i=2}^{n} N_i} = \frac{x_i}{\sum_{i=2}^{n} x_i} \quad i \neq 1. \tag{2.60}$$

Therefore, the excess free energy is given by [20, 30]

$$\frac{A^{IONS}}{NkT} = \ominus \frac{3x^2 + 6x + 2 - 2(1+2x)^{3/2}}{12\pi\rho\tilde{\sigma}} \tag{2.61}$$

where $x = \kappa\tilde{\sigma}$, and

$$\kappa^2 = \frac{4\pi}{DkT} \sum_{i=2}^{n} \rho_i q_i^2 \tag{2.62}$$

To compute the chemical potential caused by the ionic contribution, it is important to state two facts.

First, the partial derivatives of the equations (2.60), (2.59) and (2.62) are followed respectively by expressions as:

$$\begin{aligned} \frac{\partial \tilde{x}_i}{\partial N_j} &= \frac{1}{N} \frac{(\delta_{ij} - \cancel{x_i}) \sum_{i=2}^{n} x_i - x_i \sum_{i=2}^{n} (\delta_{ij} - \cancel{x_i})}{(\sum_{i=2}^{n} x_i)^2} \\ &= \frac{1}{N} \frac{\delta_{ij} - \tilde{x}_i \sum_{i=2}^{n} \delta_{ij}}{\sum_{i=2}^{n} x_i} \end{aligned} \tag{2.63a}$$

$$\frac{\partial \tilde{\sigma}}{\partial N_j} = \sum_{i=2}^{n} \frac{\partial \tilde{x}_i}{\partial N_j} \sigma_{ii} = \frac{1}{N} \frac{\sum_{i=2}^{n} \delta_{ij} (\sigma_{ii} - \tilde{\sigma})}{\sum_{i=2}^{n} x_i} \tag{2.63b}$$

$$2\kappa \frac{\partial \kappa}{\partial N_j} = \frac{4\pi}{DkT} \left\{ \frac{\rho}{N} \sum_{i=2}^{n} (\delta_{ij} - x_i) q_i^2 + \frac{1}{V} \sum_{i=2}^{n} x_i q_i^2 \right\}. \tag{2.63c}$$

Second, using the expressions (2.63b) and (2.63c), the product between the number of moles in the mixture and the partial derivative $\partial x / \partial N_j$ is given in the equation:

$$\begin{aligned} N \frac{\partial x}{\partial N_j} &= N \frac{\partial \kappa}{\partial N_j} \tilde{\sigma} + \kappa N \frac{\partial \tilde{\sigma}}{\partial N_j} \\ &= \frac{2\pi}{DkT} \left\{ \rho \sum_{i=2}^{n} \delta_{ij} q_i^2 \right\} \frac{\tilde{\sigma}}{\kappa} + \kappa \left\{ \frac{\sum_{i=2}^{n} \delta_{ij} (\sigma_{ii} - \tilde{\sigma})}{\sum_{i=2}^{n} x_i} \right\} \end{aligned} \tag{2.64}$$

Therefore, using these details we obtain an expression for the chemical potential in the ionic contribution:

$$\begin{aligned}
\frac{\mu^{IONS}}{kT} &= \left[a^{IONS} + N \left\{ \frac{\partial a^{IONS}}{\partial \rho}\frac{\partial \rho}{\partial N_j} + \frac{\partial a^{IONS}}{\partial \kappa}\frac{\partial \kappa}{\partial N_j} + \frac{\partial a^{IONS}}{\partial \tilde{\sigma}}\frac{\partial \tilde{\sigma}}{\partial N_j} \right\} \right]_{j=1,2,3} \\
&= \left[\cancel{a^{IONS}} - \cancel{a^{IONS}\frac{1}{\rho}N\frac{1}{v}} + \frac{6x+6-6(1+2x)^{1/2}}{12\pi\rho\tilde{\sigma}}\tilde{\sigma}N\frac{\partial \kappa}{\partial N_j} \right. \\
&+ \frac{(6x+6-6(1+2x)^{1/2})\kappa(12\pi\rho\tilde{\sigma})}{(12\pi\rho\tilde{\sigma})^2}N\frac{\partial \tilde{\sigma}}{\partial N_j} \\
&+ \left. -\frac{(3x^2+6x+2-2(1+2x)^{3/2})(12\pi\rho)}{(12\pi\rho\tilde{\sigma})^2}N\frac{\partial \tilde{\sigma}}{\partial N_j} \right]_{j=1,2,3} \\
&= \left[\frac{6x+6-6(1+2x)^{1/2}}{12\pi\rho\tilde{\sigma}} \left\{ \tilde{\sigma}N\frac{\partial \kappa}{\partial N_j} + \kappa N\frac{\partial \tilde{\sigma}}{\partial N_j} \right\} - \frac{a^{IONS}}{\tilde{\sigma}}N\frac{\partial \tilde{\sigma}}{\partial N_j} \right]_{j=1,2,3}
\end{aligned} \tag{2.65}$$

2.2 Objective Function

The main goal is to determine the parameter set $\Theta = \{\sigma_{22}, \sigma_{33}, \lambda_{22}, \lambda_{33}, \epsilon_{22}, \epsilon_{33}\}$ that provides the best fitting to experimental data. The labels 22 and 33 refer to the cation and anion respectively, while σ_{ii}, λ_{ii} and ϵ_{ii} refer to the molecular diameter, the variable range and the depth of the potential of interaction in equation (2.8). There is no need to compute the parameters related to the water, because they have been already determined in early works, for example see ref. [20].

The complete description for the thermodynamical model ends when a relation among the potential A in equation (2.2), the density ρ and the chemical potential μ in equation (2.7) is stated. This relation is given by the total pressure [42]:

$$\frac{P}{\rho kT} = \sum_{i=1}^{n} x_i \frac{\mu_i}{kT} - \frac{A}{NkT}. \tag{2.66}$$

To construct an objective function which allows us to estimate the parameter set, this work had into account the equilibrium equations:

$$\begin{aligned}
\mu^L_{\text{water}}(T, \rho^L, x_w; \Theta) &= \mu^V_{\text{water}}(T, \rho^V, x_w = 1; \Theta) \\
\mu^L_{\text{IL}}(T, \rho^L, x_{IL}; \Theta) &= \mu^V_{\text{IL}}(T, \rho^V, x_{IL} = 0; \Theta) = 0 \\
P^L(T, \rho^L, x_w; \Theta) &= P^V(T, \rho^V, x_w = 1; \Theta)
\end{aligned} \tag{2.67}$$

The first two lines in the equation system (2.67) represent a chemical equilibrium in the liquid and vapor phases, while the last line represents the equilibrium of pressures in the mentioned physical phases, this last equilibrium is obtained from the equation (2.66). Remembering that SAFT-VRE defines $\sum_{i=1}^{3} x_i = 1$, we represent $x_w = x_1$ as the water concentration and $x_2 = x_3 = x_{IL}/2$, in simple words the ions are in the same concentrations and they can be isolated by only knowing the water concentration.

The Ionic Liquid (IL) studied in this work is 1-Allyl-3-methylimidazolium Chloride; the measurements comes from the liquid-vapor equilibrium and the corresponding computation of the osmotic pressure for the IL in solution of water.

Table 2.1: Experimental equilibrium temperatures and water concentrations.

T_i^{exp}[K]	373.23	373.25	373.28	373.76	374.25	374.66	375.05	375.46	375.82	376.77	377.25
x_w	0.995	0.99	0.984	0.976	0.967	0.96	0.953	0.946	0.938	0.929	0.92
T_i^{exp}[K]	378.09	378.9	379.69	380.21	381.57	382.36	385.67	388.81	390.74	394.73	399.88
x_w	0.908	0.894	0.886	0.877	0.865	0.851	0.817	0.794	0.78	0.737	0.729

Table 2.2: Experimental data of the density for the Ionic liquid and corresponding temperature

ρ_k^{exp} [g/cm^3]	1.3000	1.2967	1.2933	1.2900	1.2866	1.2833	1.2799	1.2766	1.2732
T [K]	293.15	298.15	303.15	308.15	313.15	318.15	323.15	328.15	333.15
ρ_k^{exp} [g/cm^3]	1.2699	1.2665	1.2632	1.2598	1.2565	1.2531	1.2498	1.2464	
T [K]	338.15	343.15	348.15	353.15	358.15	363.15	368.15	373.15	

Therefore, according to the equation system (2.67), using the optimized parameters Θ, one can compute using the expressions (2.67) the same temperature of equilibrium ($T = T^{calc}$) by finding the roots T^{calc}, ρ^L, ρ^V. It is assumed that, at the equilibrium temperature, the unique compound in vapor phase is water, then its concentration in vapor phase is equal to one ($x_w = 1$) and the chemical potential of the ionic liquid in vapor phase is zero.

The first part of our main objective function is expressed as:

$$f_1 = \sum_{i=1}^{n_1} \left(\frac{T_i^{exp} - T_i^{calc}}{T_i^{exp}} \right)^2 \tag{2.68}$$

In (2.68), T_i^{exp} is equilibrium temperature measured experimentally but at a certain water concentration x_w, the table 2.1 contains the experimental data. The terms T_i^{calc} represents the equilibrium temperature obtained from the roots $T_i^{calc}, \rho^L, \rho^V$ of the system (2.67) at certain x_w and parameters Θ.

The second part of the main objective function is made by comparing again theoretical with experimental results. To the ionic liquid (x_{IL}), its density (ρ^{exp}) is measured at certain temperature, see table 2.2. The fitting of these data is carried out by applying the second part of the main objective function:

$$f_2 = \sum_{k=1}^{n_2} \left(\frac{\rho_k^{exp} - \rho_k^{calc}}{\rho_k^{exp}} \right)^2 \tag{2.69}$$

being ρ_k^{calc} a root in the equation:

$$1 - \frac{P}{\rho k T} \left(\sum_{i=1}^{n} x_i \frac{\mu_i}{kT} - \frac{A}{NkT} \right)^{-1} = 0, \tag{2.70}$$

leaving constant the preasure P at 1 atmosphere, $x_1 = 0$ and $x_i = x_{IL}/2$ for $i = 2, 3$.

Table 2.3: Adimensionalization for the variables

$\sigma_{ij}^* = \frac{\sigma_{ij}}{\sigma_{11}}$	$\lambda_{ij}^* = \frac{\lambda_{ij}}{\lambda_{11}}$	$\epsilon_{ij}^* = \frac{\epsilon_{ij}}{\epsilon_{11}}$
$T^* = \frac{T}{\epsilon_{11}/k}$	$\rho^* = \rho\sigma_{11}^3$	$P^* = \frac{P\sigma_{11}^3}{\epsilon_{11}}$

The way one can join the equations (2.68) and (2.69) is by certain positive weights w_1 and w_2:

$$f = w_1 \sum_{i=1}^{n_1} \left(\frac{T_i^{exp} - T_i^{calc}}{T_i^{exp}} \right)^2 + w_2 \sum_{k=1}^{n_2} \left(\frac{\rho_k^{exp} - \rho_k^{calc}}{\rho_k^{exp}} \right)^2 . \tag{2.71}$$

There exists different approaches to choose the weights, the most common way is to choose them, such that both satisfy the conditions:

- $w_1, w_2 > 0$,
- $w_1 + w_2 = 1$

For example, assuming that $w_1 = 0.7$ and $w_2 = 0.3$ satisfies the conditions, this implies that the optimization process, will give more importance to fit the set of experimental values in the table 2.1 than those in the table 2.2.

Another approach to choose the weights, takes into account the number of experimental data in each sum, saying that

$$n_1 w_1 = n_2 w_2.$$

This means that as much experimental data contains a set, should be lower the corresponding weight in the equation (2.71). However, if both sets contain the same number of data, the weights should be equal. In our case, by replacing $n_1 = 22$ and $n_2 = 17$, the relation of the weights is given by:

$$w_1 = \frac{17}{22} w_2 \approx 0.772 w_2.$$

2.3 Dimensionless Equations

Every equation that appears until now depends on variables with physical definition, but the problem here is that their units represent a big computational problem. To illustrate this imagine the ratios of the water, the anion and the cation as well as the variable ranges; all of them are given in Ångström units, this measure expressed in the SI system is equivalent to 10^{-10} meters. To solve these kind of problems, we work with the change of variables in the table 2.3.

The reason to provide all the equations in the table 2.3 in terms of water properties is simple: there already exist reports as [43, 20] where the parameters as molecular ratio, association potential and variable ranges are given. For example, the water molecular ratio σ_{11} is reported in 3.036Å, the variable range λ_{11} is 1.8Å and the association energy ϵ_{11}/k is estimated at 253.3 K.

In the case of the ionic contribution, a reduction of the model is made. The normalization, on the equation (2.62) can be done on the next form:

$$\begin{aligned}\kappa^2 &= \frac{4\pi\rho e^2}{DkT}\sum_{i=2}^{n} x_i \\ &= \frac{4\pi\rho\sigma_{11}^3 e^2}{D\sigma_{11}^3\epsilon_{11}\frac{kT}{\epsilon_{11}}}\sum_{i=2}^{n} x_i = \frac{4\pi\rho^* e^2}{D\sigma_{11}^3\epsilon_{11}T^*}\sum_{i=2}^{n} x_i\end{aligned} \tag{2.72}$$

After multiplying the term (2.72) with the normalization of the equation (2.59) to the square, we obtain a new x in the equation (2.61). In this new term, the dielectric constant can be simplified as $D^* = (e^2/(D\sigma_{11}\epsilon_{11}))^{-1}$. As usual, the small trick used to obtain that value is presented in the next procedure:

$$D^* = \left(\frac{e^2}{D\sigma_{11}\epsilon_{11}}\right)^{-1} = \left(\frac{e^2}{D\sigma_{11}\frac{\epsilon_{11}}{k}k}\right)^{-1} \tag{2.73}$$

The constants: e (electron charge), D (dielectric constant of water) and k (Boltzmann's constant) in the equation (2.73) implemented in this work, were $1.60217662 \times 10^{-19}$ in Joules, $\frac{9\times10^9}{74.373*298.15}$ and 1.3806×10^{-23} J/K.

Other important normalization should be done over the equation (2.70), in order to compute reasonable values according to the dimensionless of the parameters. The next equation illustrates the procedure for computing this:

$$\begin{aligned}1 - \frac{P\sigma_{11}^3}{\rho\sigma_{11}^3\epsilon_{11}\frac{kT}{\epsilon_{11}}}\left(\sum_{i=1}^{n} x_i\frac{\mu_i^*}{T^*} - a^*\right)^{-1} &= 1 - \frac{P\sigma_{11}^3}{\rho\sigma_{11}^3\epsilon_{11}T^*}\left(\sum_{i=1}^{n} x_i\frac{\mu_i^*}{T^*} - a^*\right)^{-1} \\ &= 1 - \frac{P\sigma_{11}^3}{\rho^*\frac{\epsilon_{11}}{k}kT^*}\left(\sum_{i=1}^{n} x_i\frac{\mu_i^*}{T^*} - a^*\right)^{-1}.\end{aligned} \tag{2.74}$$

In addition, due to the experimental results of the density of the ionic liquid were obtained at 1 atmosphere of pressure in the equation (2.74), we use $P = 101325$Pa; on the other side, the computations were made with $\frac{\epsilon_{11}}{k} = 253.3$K as usual, the Boltzmann constant $k = 1.38064 \times 10^{-23}$J/K and the molecular diameter of water $\sigma_{11} = 3.036$Å.

Chapter 3

Results

The purpose of this chapter is to provide in a detailed description of the methodology to complete a successful estimation of the objective parameters. The first part explains how to construct some auxiliar functions that accelerate the evaluation process of the objective function, while the final part presents the fitting results.

3.1 Computational Details

The mathematical challenge of the problem is dealing with the roots of the equations (2.67) and (2.70). Hence, we present a strategy to obtain the roots, which uses the algorithms in chapter 1, taking advantage of the derivative-free procedures and the lack of continuity of the functions. We do not work directly with a Newtonian algorithm, because it is necessary to start the process from an initial point near of the root, or at least under certain conditions.

In the case of finding the roots of the equilibrium equations (2.67), this problem can be written in a simpler way for a specific set of parameters Θ as

$$F(\mathbf{X}, \Theta) = \mathbf{0}, \tag{3.1}$$

where $\mathbf{X} = (T^*, \rho_L^*, \rho_V^*)$ and $F : \mathbb{R}^3 \to \mathbb{R}^3$. But, an auxiliary function can be obtained from F; we call this as $f_{aux,1}$ and is defined as

$$f_{aux,1} = ||F(\mathbf{X}, \Theta)||^2 \tag{3.2}$$

being $||\cdot||$ the euclidean norm. According to the stochastic method for solving a global optimization problem where $f_{aux,1}$ is the objective function, the solution $\mathbf{X_0}$ establishes that $||F(\mathbf{X_0}, \Theta)||^2$ is equal or very near to 0, then $F(\mathbf{X_0}, \Theta) \approx \mathbf{0}$. Once the initial point $\mathbf{X_0}$ is stated, the exact roots of the system (2.67) can be computed by the Newton-Raphson method. At first glance, carrying out all this procedure seems unnecessary, but it is justified when for one single set of parameters Θ, we should determine n_1 equilibrium roots of the system (2.67) that allow us to compute the least square fitting function f_1 in equation (2.68). To remember, n_1 represents the number of experimental equilibrium temperatures; in our case, $n_1 = 22$, see table 2.1.

Analogously, the roots of the equation (2.70) are obtained using Newton-Raphson. The initial point is obtained from the auxiliary function:

$$f_{aux,2} = \left| 1 - \frac{P}{\rho kT} \left(\sum_{i=1}^{n} x_i \frac{\mu_i}{kT} - \frac{A}{NkT} \right)^{-1} \right|, \tag{3.3}$$

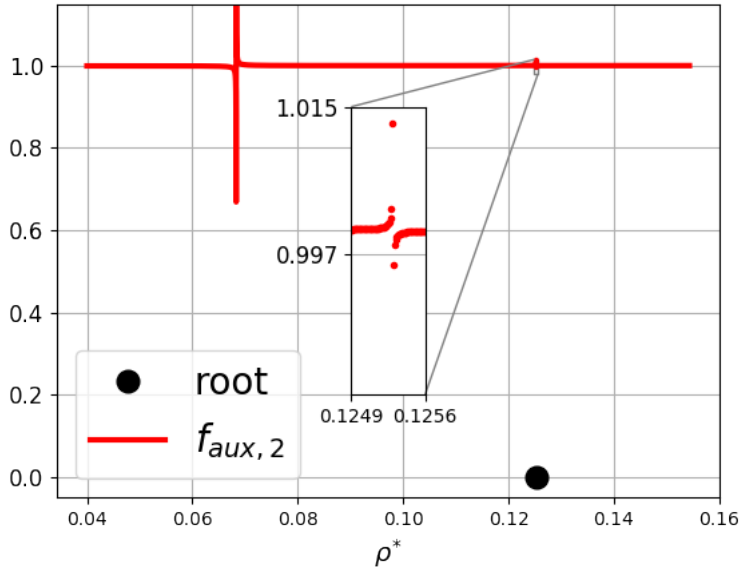


Figure 3.1: Region where root of equation (3.3). This figure was obtained with parameters in table 3.1 and $T = 379.0$K

after applying Simulated Annealing or Improving Hit-and-Run. Due to in this case the density belonging to $\mathbb{R}$, Hide-and-Seek and Simulated Annealing are the same.

In this case, for each set of parameters, the number of densities ρ_k^{calc} is $n_2 = 17$. A graphical justification for using this combination of methods is explicitly presented in the figure 3.1. Just finding an initial point in the augmented region is very difficult, and this region also changes when the parameters Θ are modified. However, the stochastic methods can provide initial points that satisfy convergence conditions for Newton-Raphson.

Some other problems can appear solving the equation (2.70), for example there are certain values for the parameters for which this equation does not have a solution in the interval $[a, b] = [0.120, 0.126]$. This interval is relevant because it contains the experimental values of the dimensionless density $\rho_k^{exp,*}$. To compute the extremes of this interval, we convert the lower and the higher values from the table 2.2 and a and b are chosen such that:

$$a < 1.2464 \times (3.036 \times 10^{-8})^3 \times \frac{6.023 \times 10^{23}}{174} = 0.12073,$$
$$b > 1.3000 \times (3.036 \times 10^{-8})^3 \times \frac{6.023 \times 10^{23}}{174} = 0.12592.$$

An example of a set of parameters is described in the figure 3.2. In this figure, one clearly observe at least three roots, from which the greater is the most interesting because represents the density of the IL, but in liquid state. However, the figure illustrates the problem that this value is not near enough to the experiments.

Another common problem is that, for certain combination of parameters the equation (2.70) does not have solutions for near values of the experimental results, for example in the figure 3.3. This figure shows the behavior of the equation (2.70), where the most relevant fact is that there are no real images for supposed values of the density but

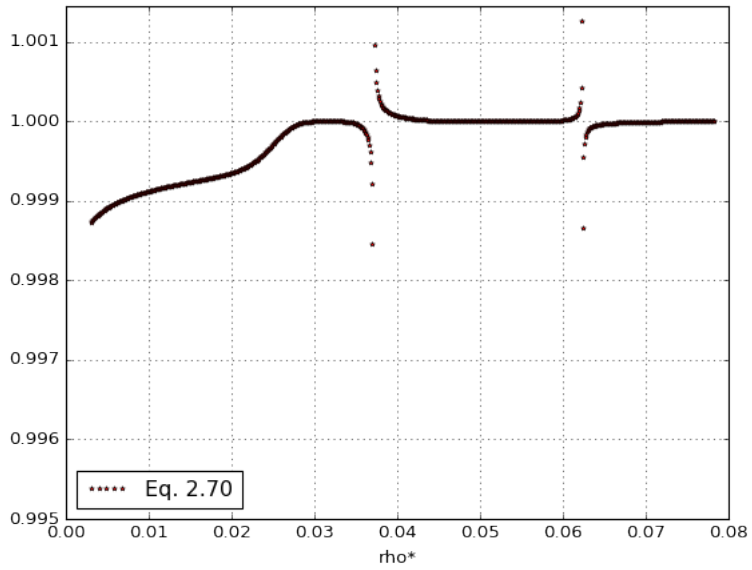


Figure 3.2: Behavior of the right side of the equation (2.70) at 298.15 K and the set of parameters $\sigma^*_{22} = 2.497$, $\sigma^*_{33} = 3.212$, $\lambda^*_{22} = 0.385$, $\lambda^*_{33} = 2.623$, $\epsilon^*_{22} = 2.037$ and $\epsilon^*_{33} = 2.071$.

greater than 0.1.

The expected graphic of the equation (2.70) can be obtained using the parameters in the table 3.2, this is represented in the figure 3.4. In this figure, the expected root falls in the gap between the experimental values.

3.2 Parameter Estimation

This part of the work deals with the problem of the estimation of the parameters such that, one can predict an equilibrium temperature, as well as a density of the ionic liquid. Initially, several computations were carried out using the objective function (2.71), and the schemes of the weights presented below of this function. These initial results show no fitting in both data sets at the same time, this could be caused mainly by the model; different reasons could explain this lack of fitting, among those: an assumption of a spherical model for the ions and an association only between water molecules. In spite of this problem, the fitting for each data set was computed.

On one hand the best values that can predict the equilibrium temperature are presented in table 3.1; these values were found using a minimization scheme for the equation (2.68). These parameters show some interesting facts. First of all, the molecular diameters are larger than the molecular diameter of water, in the case of the cation, it is 243% bigger than the water. It means that in the SI system of units, the molecule studied, has a molecular ratio of about 7.377Å, the other side the molecular ratio for the anion molecule is about 6.628Å. The second fact to notice is the variable range of association, which is also longer than the range of water; in other words, this implies that the association between molecules can happen even when the molecules are very far

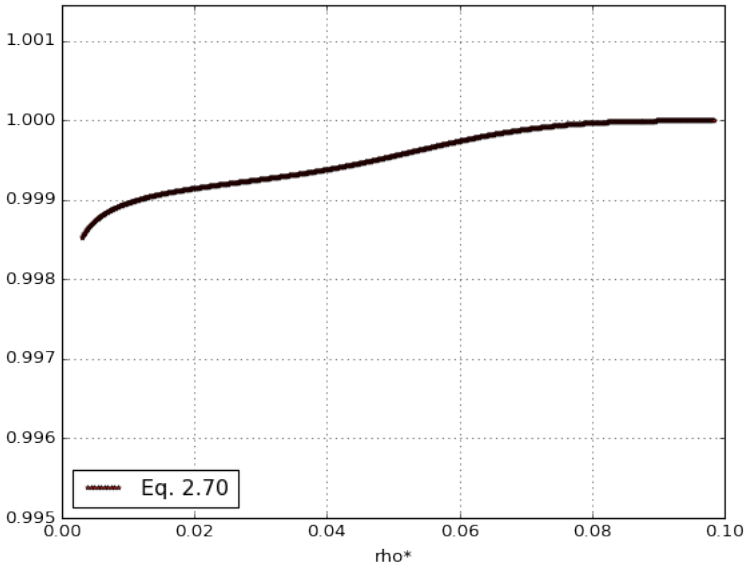


Figure 3.3: Behavior of the right side of the equation (2.70) at 298.15 K and the set of parameters $\sigma_{22}^* = 2.15$, $\sigma_{33}^* = 3.068$, $\lambda_{22}^* = 1.86$, $\lambda_{33}^* = 1.359$, $\epsilon_{22}^* = 2.982$ and $\epsilon_{33}^* = 1.922$.

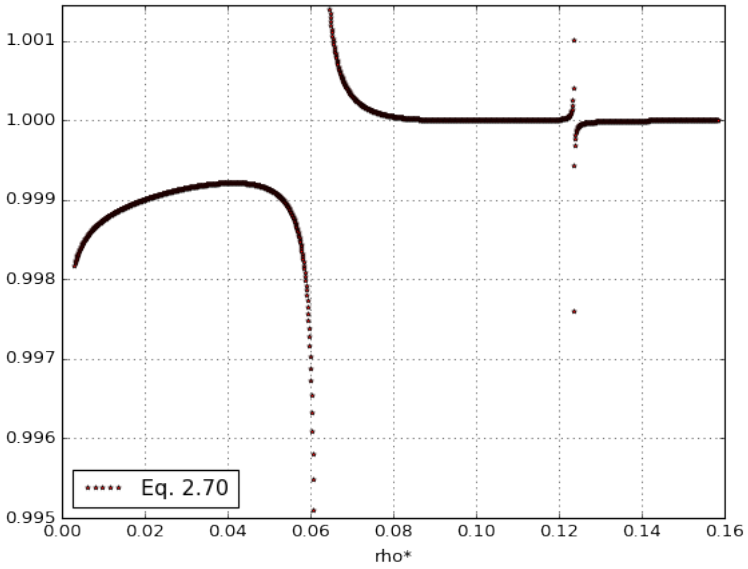


Figure 3.4: Behavior of the right side of the equation (2.70) at 298.15 K and using the parameters in the table 3.2.

Table 3.1: Estimated parameters for predicting the equilibrium temperature

Ion	σ_{ii}^*	λ_{ii}^*	ϵ_{ii}^*
Cation	2.43	1.264	1.643
Anion	2.183	2.523	2.391

from each other. According to the reports already mentioned, the variable association range for water is about 1.8Å, then by comparing the obtained results, this range for the cation was estimated in 2.275Å and for the anion in about 2.7414Å. The final fact that is important to mention is that the depths in interaction between the cation and anion molecules are also greater than the water, which means that the molecules can be bounded stronger and they could even have a dipole moment, much higher than the water one.

Figure 3.5 shows the computations for the equilibrium of the system (2.67). The first chart shows the approximation for the experimental temperature of equilibrium to certain water concentration, while the second chart shows how the density of the liquid and vapor phases behave. Considering the average absolute deviation (% AAD) as

$$\%\ \text{AAD} = \frac{1}{n_1}\sum_{i=1}^{n_1}\left|\frac{T_i^{exp,*} - T_i^{cal,*}}{T_i^{exp,*}}\right| \times 100,$$

the % AAD between the experimental and theoretical temperatures of equilibrium is 2.486. An important fact to notice is that the model presents a good approximation for relative concentrations lower than 0.95. This means that for mixtures of water and the ionic liquid, where the concentration of water is lower than the 95%, the temperature in the chemical equilibrium is close to those experimental results. The lack of approximating good equilibrium temperatures to higher concentrations of water could happen for different reasons, one could be that in the computations have not taken into account well tuning constants as the diameter of water in equation (2.73) or the bonding constant of water K_{11} in equation (2.57). This effect is important because the model that is presented is highly nonlineal and this cause an important mismatch in the results. A clear example of this effect can be noticed by comparing the figures 3.5 and 3.8, both figures were made with the values in tables 3.1 and 3.2 respectively.

In addition, we compute the density of the IL to certain temperatures, the figure 3.6 shows the results using the parameters in table 3.1. In this case the % AAD is 1.8461. The problem is that the parameters do not allow to estimate the density of the IL properly, despite that the obtained values are near to the experimental values and show the same behavior to decrease as the temperature increases. Note that the average absolute deviation in this case is computed by the expression:

$$\%\ \text{AAD} = \frac{1}{n_2}\sum_{k=1}^{n_2}\left|\frac{\rho_k^{exp,*} - \rho_k^{cal,*}}{\rho_k^{exp,*}}\right| \times 100,$$

The other hand the parameters presented in the table 3.2 were computed by using a scheme of minimization on the objective function (2.69). These results are not very different at first glance to those presented in the table 3.1. These values also show that the diameters of the ions are much bigger than the diameter of a water molecule. The variable ranges of association are also greater than the water one, this fact implies that

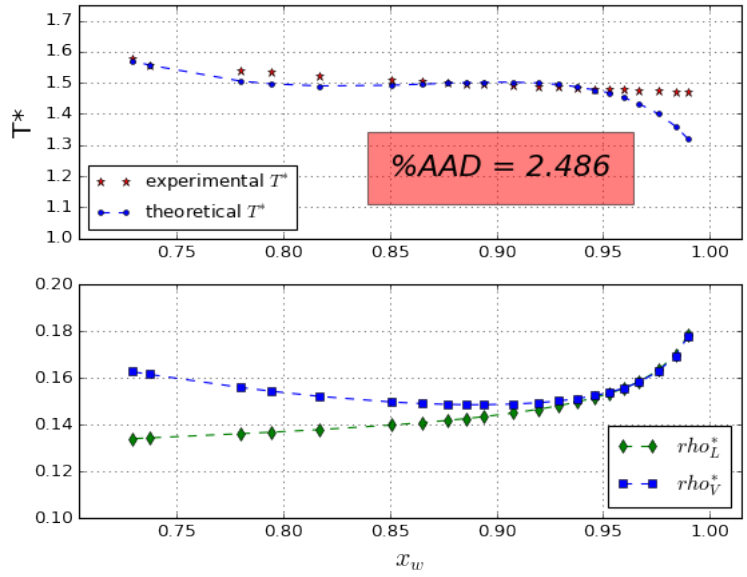


Figure 3.5: Prediction of equilibrium values for the nonlinear system (2.67) using the parameter on the table 3.1.

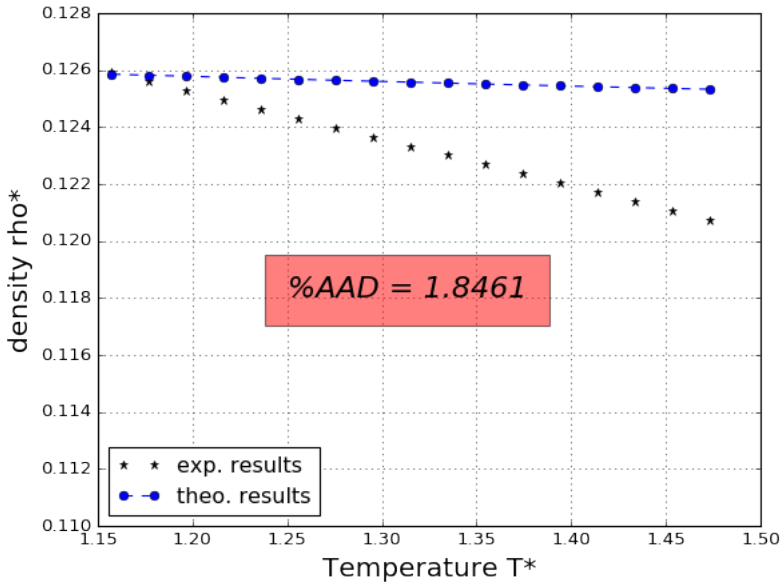


Figure 3.6: Density vs. Temperature for the ionic liquid using the parameters on table 3.1.

Table 3.2: Estimated parameters for predicting the density.

Ion	σ^*_{ii}	λ^*_{ii}	ϵ^*_{ii}
Cation	2.327	1.246	1.647
Anion	2.258	2.631	2.38

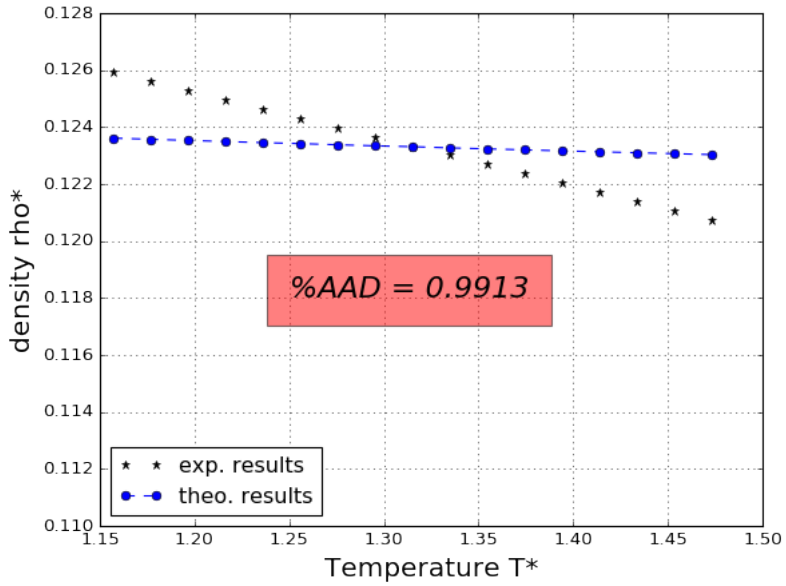


Figure 3.7: Density vs. Temperature for the ionic liquid using the parameters on table 3.2.

the attraction in the interaction cation-cation and anion-anion can happen even where those molecules are really separated. Finally, the depth of the potential interactions are greater than the water, this also confirms what was stated by the first set of parameters that were presented in the table 3.1. Figure 3.7 shows the prediction for the density of the ionic liquid to a certain temperature. Despite the % AAD between the experimental results and the theoretical ones is 0.9913, and the theoretical results are given in the same gap, this set of parameters do not estimate a good density at a certain temperature, as it happened with the parameters that are proposed to estimate the equilibrium temperatures. Additionally, those parameters fail in approximating the experimental equilibrium temperatures as the figure 3.8 shows, the % AAD in this fitting was 29.348.

Finally, motivated by these results, we conclude that the SAFT-VRE model is only admissible until certain concentration of the solvent, and we only estimated the parameters for fitting equilibrium temperatures, that correspond to water concentrations lower than 0.95. The table 3.3 contains the mentioned parameters, while the figure 3.9 contains the respective data fitting to the equilibrium temperatures, and the figure 3.10 presents the fitting to the density data.

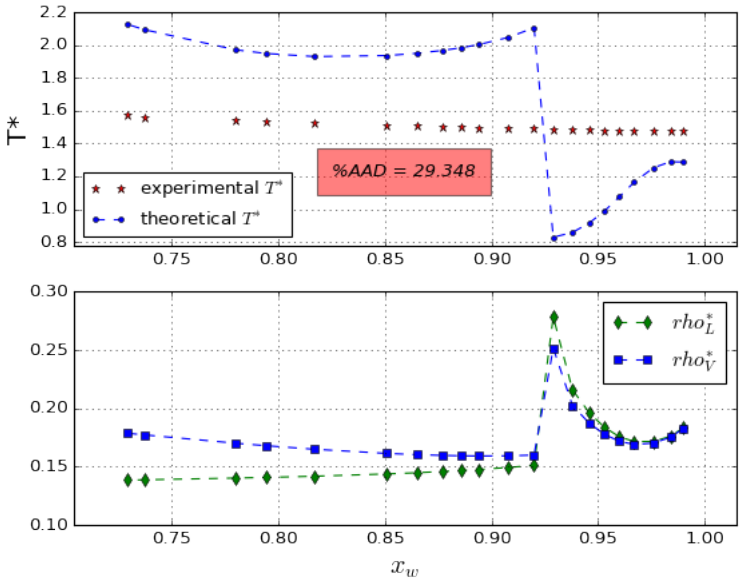


Figure 3.8: Prediction of equilibrium values for the nonlinear system (2.67) using the parameter on the table 3.2.

Table 3.3: The most suitable parameters for estimating the equilibrium temperatures.

Ion	σ_{ii}^*	λ_{ii}^*	ϵ_{ii}^*
Cation	2.443	1.232	1.667
Anion	2.186	2.598	2.279

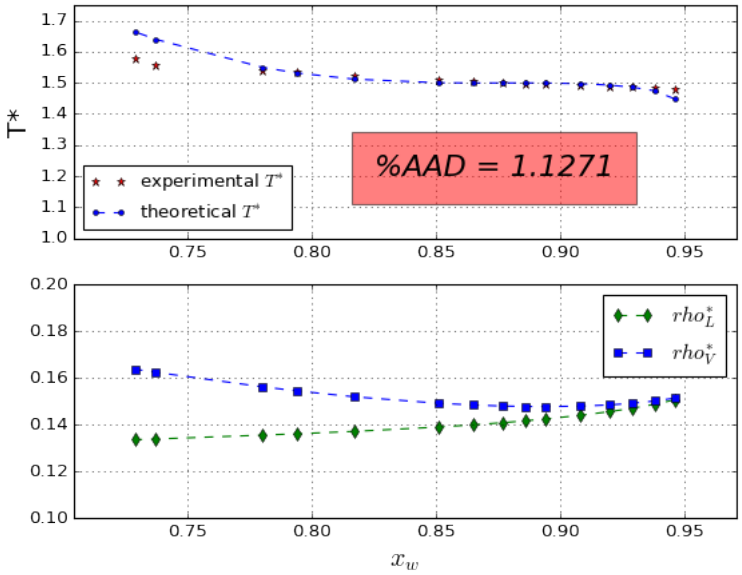


Figure 3.9: Prediction of equilibrium values for the nonlinear system (2.67) using the parameter on the table 3.3.

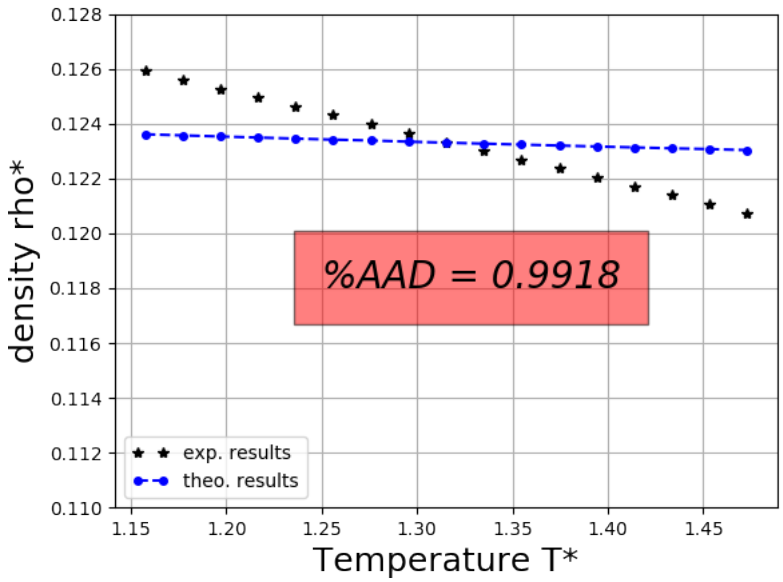


Figure 3.10: Density vs. Temperature for the ionic liquid using the parameters on table 3.3.

Chapter 4

Conclusions

To finish, this work focused on the application of the stochastic techniques of optimization, in order to determine the best parameters that can predict the observed experimental properties. The techniques presented in the first chapter were the most suitable to carry out the computations needed to reach the main objective. Additionally, based on the literature, the second chapter develops the computations to understand what is happening with the ionic liquids. The most important facts are:

- The use of stochastic techniques was totally necessary due to the complexity of the model equations. Any other method based on the derivatives can fail in giving the desired results, for example roots of equations or nonlinear equation systems. The study of the most common stochastic techniques, provides the advantages of reducing the zones where methods of Newtonian type can converge. This last combination of methods saves an important quantity of computational time.

- The fact that the fitting can not be done using at the same time the equilibrium temperature and the density data of the IL, means that the model is incomplete. From one hand, the hypothesis of a spherical model for the ions is possibly insufficient. The other, there exist more hydrogen bonds than those assumed in the water-water association.

- The obtained parameters present great physical meaning, this can be determined by the expected properties already mentioned in the last chapter. Among the effects of the found parameters, it was found that the molecules of study are in fact bigger than a water molecule, they have large range potential and energy of interaction, which allow the molecules associate easily between ions of the same type.

Appendix A

Here, we sketch an idea to proof the convergence of the Nelder and Mead method for finding a minimum.

The first important fact is that we are always considering, an initial simplex to be non degenerated. See the comments in the equation (1.2).

Definition 1. *The volume of the simplex (geometric figure) expressed in the equation* (1.3) *is given by the expression*

$$Vol = \frac{1}{n!}\,|det(V)| \tag{A.1}$$

where the matrix V is composed by the vertices of the simplex in the form:

$$V = (x_2 - x_1, x_3 - x_1, \ldots, x_{n+1} - x_1)\,. \tag{A.2}$$

Each column of V is formed by the difference between two vertices as is indicated in (A.2)

Proposition 1. *Suppose that the simplex S_{j+1} in the iteration $j+1$ (with volume Vol_{j+1}) is obtained from the simplex S_j in the iteration j (with volume Vol_j) by changing the vertex x_{n+1} by a vertex of the form $x_0 + \theta(x_0 - x_{n+1})$ then,*

$$Vol_{j+1} = |\theta| Vol_j \tag{A.3}$$

Proof. Noting the matrices:

$$V_j = (x_2 - x_1, x_3 - x_1, \ldots, x_{n+1} - x_1)$$

and

$$V_{j+1} = (x_2 - x_1, x_3 - x_1, \ldots, x_0 + \theta(x_0 - x_{n+1}) - x_1)$$

then:

$$\begin{aligned} det(V_{j+1}) &= (1+\theta) det\,(x_2 - x_1, x_3 - x_1, \ldots, x_0 - x_1) - \theta det(V_j) \\ &= \frac{1+\theta}{n} det\left(x_2 - x_1, x_3 - x_1, \ldots, \sum_{i=1}^{n} x_i - n x_1\right) - \theta det(V_j) \\ &= \frac{1+\theta}{n} det\left(x_2 - x_1, x_3 - x_1, \ldots, \sum_{i=1}^{n} x_i - n x_1 - \sum_{i=2}^{n} x_i + (n-1)x_1\right)^{0} \\ &\quad - \theta det(V_j) \end{aligned} \tag{A.4}$$

Then by the equation (A.4), we obtained the desired result. □

Proposition 2. *Suppose that the simplex S_{j+1} in the iteration $j+1$ is obtained from the simplex S_j by carrying out a reduction. Then,*

$$Vol_{j+1} = \sigma^n Vol_j \tag{A.5}$$

Proof. The solution is evident by observing that

$$\begin{aligned} V_{j+1} &= (\sigma(x_2 - x_1), \sigma(x_3 - x_1), \ldots, \sigma(x_2 - x_1)) \\ &= \sigma V_j \end{aligned} \tag{A.6}$$

Therefore, we obtain the desired result applying the properties of the determinant on (A.6) to have the volume required. □

The propositions 1 and 2 show that from an initial simplex non-degenerated, the followed simplexes in the algorithm are always non-degenerated. This fact partially shows that the search for the minimum is extensive and it is not carried out on a subspace of $\mathbb{R}^n$.

Bibliography

[1] Emile Aarts and P. van Laarhoven. Statistical cooling: A general approach to combinatorial optimization problems. *Philips Journal of research*, 40(4):193, 1985.

[2] J.D. Agudelo-Giraldo, E. Restrepo-Parra, and J. Restrepo. Monte carlo simulation of roughness effect on magnetic and magnetotransport behavior of la2/3ca1/3mno3/la1/3ca2/3mno3 bilayers. *Physics B: Condensed Matter*, 434:149 – 154, 2014.

[3] J. A. Barker and D. Henderson. Perturbation theory and equation of state for fluids. ii. a successful theory of liquids. *The Journal of Chemical Physics*, 47(11):4714–4721, 1967.

[4] J. A. Barker and D. Henderson. What is "liquid"? understanding the states of matter. *Rev. Mod. Phys.*, 48:587–671, Oct 1976.

[5] Claude J. P. Bélisle. Convergence theorems for a class of simulated annealing algorithms on $\mathbb{R}^d$. *Journal of Applied Probability*, 29(4):885–895, 1992.

[6] Claude J. P. Bélisle, H. Edwin Romeijn, and Robert L. Smith. Hit-and-run algorithms for generating multivariate distributions. *Mathematics of Operations Research*, 18(2):255–266, 1993.

[7] Dimitris Bertsimas and Santosh Vempala. Solving convex programs by random walks. *J. ACM*, 51(4):540–556, July 2004.

[8] Ihor O. Bohachevsky, Mark E. Johnson, and Myron L. Stein. Generalized simulated annealing for function optimization. *Technometrics*, 28(3):209–217, 1986.

[9] A. Boneh and A. Golan. Constraints' redundancy and feasible region boundedness by random feasible point generator (rfpg). In *3rd European Congress on Operations Research (EURO III)*, 1979.

[10] Tomáš Boublík. Hard-Sphere Equation of State. *The Journal of Chemical Physics*, 53(1):471–472, 1970.

[11] Richard H. Byrd, Humaid Fayez Khalfan, and Robert B. Schnabel. Analysis of a symmetric rank-one trust region method. *SIAM Journal on Optimization*, 6(4):1025–1039, 1996.

[12] Margarida F Cardoso, Romualdo L Salcedo, and S Feyo De Azevedo. The simplex-simulated annealing approach to continuous non-linear optimization. *Computers & chemical engineering*, 20(9):1065–1080, 1996.

[13] Walter G. Chapman, Keith E. Gubbins, George Jackson, and Maciej Radosz. New reference equation of state for associating liquids. *Industrial & Engineering Chemistry Research*, 29(8):1709–1721, 1990.

[14] Walter G. Chapman, George Jackson, and Keith E. Gubbins. Phase equilibria of associating fluids. *Molecular Physics*, 65(5):1057–1079, 1988.

[15] George B. Dantzig. A history of scientific computing. chapter Origins of the Simplex Method, pages 141–151. ACM, New York, NY, USA, 1990.

[16] Anton Dekkers and Emile Aarts. Global optimization and simulated annealing. *Mathematical Programming*, 50(1):367–393, 1991.

[17] M. Dorigo and L. M. Gambardella. Ant colony system: A cooperative learning approach to the traveling salesman problem. *Trans. Evol. Comp*, 1(1):53–66, April 1997.

[18] Russell C. Eberhart, James Kennedy, and Yuhui Shi. *Swarm intelligence*. The Morgan Kaufmann series in evolutionary computation. Morgan Kaufmann Publishers, 1st edition, 2001.

[19] A. Galindo, L. A. Davies, A. Gil-Villegas, and G. Jackson. The thermodynamics of mixtures and the corresponding mixing rules in the saft-vr approach for potentials of variable range. *Molecular Physics*, 93(2):241–252, 1998.

[20] Amparo Galindo, Alejandro Gil-Villegas, George Jackson, and Andrew N. Burgess. SAFT-VRE: Phase behavior of electrolyte solutions with the statistical associating fluid theory for potentials of variable range. *The Journal of Physical Chemistry B*, 103(46):10272–10281, 1999.

[21] Alejandro Gil-Villegas, Amparo Galindo, Paul J Whitehead, Stuart J Mills, George Jackson, and Andrew N Burgess. Statistical associating fluid theory for chain molecules with attractive potentials of variable range. *The Journal of chemical physics*, 106(10):4168–4186, 1997.

[22] W. K. Hastings. Monte carlo sampling methods using markov chains and their applications. *Biometrika*, 57(1):97–109, 1970.

[23] Abdel-Rahman Hedar and Masao Fukushima. Hybrid simulated annealing and direct search method for nonlinear unconstrained global optimization. *Optimization Methods and Software*, 17(5):891–912, 2002.

[24] John H. Holland. *Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence*. The MIT Press, 1992.

[25] Robert Hooke and T. A. Jeeves. "direct search" solution of numerical and statistical problems. *J. ACM*, 8(2):212–229, April 1961.

[26] George Jackson, Walter G. Chapman, and Keith E. Gubbins. Phase equilibria of associating fluids. *Molecular Physics*, 65(1):1–31, 1988.

[27] David E. Kaufman and Robert L. Smith. Direction choice for accelerated convergence in hit-and-run sampling. *Operations Research*, 46(1):84–95, 1998.

[28] J. Kiefer and J. Wolfowitz. Stochastic estimation of the maximum of a regression function. *Ann. Math. Statist.*, 23(3):462–466, 09 1952.

[29] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi. Optimization by simulated annealing. *Science*, 220(4598):671–680, 1983.

[30] Lloyd L. Lee. *Molecular thermodynamics of nonideal fluids*. Butterworth Publishers, Boston, 1988.

[31] Adrian S. Lewis and Michael L. Overton. Nonsmooth optimization via quasi-newton methods. *Mathematical Programming*, 141(1):135–163, 2012.

[32] M. Locatelli. Simulated annealing algorithms for continuous global optimization: Convergence conditions. *Journal of Optimization Theory and Applications*, 104(1):121–133, 2000.

[33] Alexandros Lymperiadis, Claire S. Adjiman, Amparo Galindo, and George Jackson. A group contribution method for associating chain molecules based on the statistical associating fluid theory (SAFT-γ). *The Journal of Chemical Physics*, 127(23):234903, 2007.

[34] G. A. Mansoori, N. F. Carnahan, K. E. Starling, and T. W. Leland. Equilibrium thermodynamic properties of the mixture of hard spheres. *The Journal of Chemical Physics*, 54(4):1523–1525, 1971.

[35] George Marsaglia. Choosing a point from the surface of a sphere. *Ann. Math. Statist.*, 43(2):645–646, 04 1972.

[36] Karl Menger. Eine dimensionstheoretische bemerkung von o. schreier. *Monatshefte für Mathematik und Physik*, 37(1):7–12, 1930.

[37] Huseyin Onur Mete and Zelda B. Zabinsky. Pattern hit-and-run for sampling efficiently on polytopes. *Operations Research Letters*, 40(1):6 – 11, 2012.

[38] Nicholas Metropolis. The beginning of the monte carlo method. *Los Alamos Science*, 15(584):125–130, 1987.

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

[40] Jorge Nocedal and Stephen Wright. *Numerical optimization*. Springer Science & Business Media, 2006.

[41] Felipe A. Perdomo and Alejandro Gil-Villegas. Predicting thermophysical properties of biodiesel fuel blends using the saft-vr approach. *Fluid Phase Equilibria*, 306(1):124 – 128, 2011. 20 years of the {SAFT} equation of stateRecent advances and challengesSymposium.

[42] Felipe A. Perdomo, Beatriz M. Millán, and José L. Aragón. Predicting the physical–chemical properties of biodiesel fuels assessing the molecular structure with the SAFT-γ group contribution approach. *Energy*, 72:274 – 290, 2014.

[43] Felipe A. Perdomo, Luis Perdomo, Beatriz M. Millán, and José L. Aragón. Design and improvement of biodiesel fuels blends by optimization of their molecular structures and compositions. *Chemical Engineering Research and Design*, 92(8):1482 – 1494, 2014. Green Processes and Eco-technologies – Focus on Biofuels.

[44] W.H. Press. *Numerical Recipes 3rd Edition: The Art of Scientific Computing*. Cambridge University Press, 2007.

[45] William H. Press and Saul A. Teukolsky. Simulated annealing optimization over continuous spaces. *Computers in Physics*, 5(4):426–429, 1991.

[46] L. Ramos-Rivera, D. Escobar, V. Benavides-Palacios, P.J. Arango, and E. Restrepo-Parra. Effect of annealing process on tin/tic bilayers grown by pulsed arc discharge. *Physica B: Condensed Matter*, 407(16):3248–3251, 2012. Frontiers of Condensed Matter V (FCM 2010).

[47] T. M. Reed and K. E. Gubbins. *Applied Statistical Mechanics*. McGraw-Hill Kogakusha, 1973.

[48] E. Restrepo-Parra, J.D. Agudelo-Giraldo, and J. Restrepo. Thickness and bilayer number dependence on exchange bias in ferromagnetic/antiferromagnetic multilayers based on la1-xcaxmno3. *Physica B: Condensed Matter*, 440:61 – 66, 2014.

[49] Herbert Robbins and Sutton Monro. A stochastic approximation method. *Ann. Math. Statist.*, 22(3):400–407, 09 1951.

[50] H. Edwin Romeijn and Robert L. Smith. Simulated annealing for constrained global optimization. *Journal of Global Optimization*, 5(2):101–126, 1994.

[51] John Shipley Rowlinson and FuL Swinton. *Liquids and Liquid Mixtures: Butterworths Monographs in Chemistry*. Butterworth-Heinemann, third edition edition, 1982.

[52] Jerome Sacks. Asymptotic distribution of stochastic approximation procedures. *Ann. Math. Statist.*, 29(2):373–405, 06 1958.

[53] Robert L. Smith. Efficient Monte-Carlo Procedures for Generating Points Uniformly Distributed Over Bounded Regions. *Operations Research*, 32(6):1296–1308, 1984.

[54] Virginia Torczon. On the convergence of pattern search algorithms. *SIAM J. on Optimization*, 7(1):1–25, January 1997.

[55] A. M. Turing. I.—computing machinery and intelligence. *Mind*, LIX(236):433–460, 1950.

[56] Rafael Arango Vélez, Felipe Jaramillo Ayerbe, Juan Carlos Riaño Rojas, and Flavio Prieto Ortiz. Caracterización del patrón capilaroscópico del pliegue ungueal en personas sin enfermedad evidente mediante técnicas de inteligencia artificial. prueba piloto. *Biosalud*, (6):85–95, 2007.

[57] J. H. Venter. An extension of the robbins-monro procedure. *Ann. Math. Statist.*, 38(1):181–190, 02 1967.

[58] Zelda B. Zabinsky and Robert L. Smith. *Encyclopedia of Operations Research and Management Science*, chapter Hit-and-Run Methods, pages 721–729. Springer US, Boston, MA, 2013.

[59] Zelda B. Zabinsky, Robert L. Smith, J. Fred McDonald, H. Edwin Romeijn, and David E. Kaufman. Improving hit-and-run for global optimization. *Journal of Global Optimization*, 3(2):171–192, June 1993.