

INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI

A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
313/761-4700 800/521-0600

Simulation of the Equilibrium Constant of the Weak Electrolyte
Solution using the Subspace Sampling Method

by

Chwen-Yang Shew

A dissertation submitted to the Graduate Faculty in Chemistry in partial
fulfillment of the requirements for the degree of Doctor of Philosophy,
The City University of New York

1995

UMI Number: 9605662

UMI Microform 9605662

Copyright 1995, by UMI Company. All rights reserved.

**This microform edition is protected against unauthorized
copying under Title 17, United States Code.**

UMI

**300 North Zeeb Road
Ann Arbor, MI 48103**

Approval Page

This manuscript has been read and accepted for the Graduate Faculty in Chemistry in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy

5/30/95
Date

Pamela A. Mulls
Chair of Examining Committee

5/30/95
Date

Michael Pye
Executive Officer

W. R. J. L.
Max Zien

Ronald J. L.

Supervisory Committee

THE CITY UNIVERSITY OF NEW YORK

Abstract

Simulation of the Equilibrium Constant of the Weak Electrolyte Solution using
the Subspace Sampling Method

by

Chwen-Yang Shew

Advisor: Pamela Mills

The Monte Carlo method is a powerful tool to compute the thermodynamic and structural properties of condensed phase problem. During my dissertation research, we developed a novel Monte Carlo method --- the Subspace Sampling Methods. The method combines the free energy difference method with some other Monte Carlo methods. The method divides configuration space into two subspaces. The two subspaces are sampling using two processes. One is a transfer process, which manages the crossing process from one subspace to another in an analogous manner to the acceptance ratio method of free energy difference calculation. The other is the move process, which is the transition in the individual subspace. In the simulation of the double well potential, the results demonstrate that the method can really facilitate the barrier crossing with a high efficiency. The convergence is more rapid than that of any other method. Due to the versatility of the method, it can be employed to simulate both continuous and discontinuous potential systems. The equilibrium constant and the density function (or other thermodynamic properties) is simulated simultaneously. In other words, the method is applicable to investigate the equilibrium properties of chemical reaction systems which are quite difficult for other methods. Recently, we carried out the simulation for the divalent weak electrolyte system based on the Sticky Electrolyte Model. The model is composed of the primitive model for strong electrolyte solutions and a sticky molecular potential. The high energy

barrier in this model impedes ion pairs association to form molecules. Although, standard Monte Carlo algorithms cannot simulate the system efficiently, the Subspace Sampling Method proved amenable to compute the equilibrium constant and pair correlation function. We compare the results of the simulation with those of analytic solutions (MSA/HNC and HNC). The simulation equilibrium constant has a similar behavior to that obtained from the analytic solution. However, the magnitude of the simulated equilibrium constants are one order greater. The simulated contact correlation function is smaller. The value is more close to the result of MSA/HNC than that of using HNC alone.

Acknowledgments

I am grateful to my mentor Professor Pamela Mills for her generous support and pertinent guidance which made the research possible, to my dissertation committee --- Professor Ronald Dickman of Lehman College, Professor Max Diem of Hunter College, Professor Godfrey Gumbs of Hunter College and Professor John Lombardi of the City College for their enthusiastic assistance and extensive discussion.

I'd like to acknowledge my parents for their constant encouragement. I am indebted to all the colleagues in this laboratory and school for their great help and opinion. They are ---- Ali, Barry, Chin-Chuan, Christopher, Eda, Hye-Rim and Melinda.

Furthermore, I wish to thank Professor G. Quigley, Mrs. Edna Bornstein and Mr. Waldmar Cieniewicz for their assistance in the computer facility. I also appreciate Professor J. Dannenberg to allow me to use his computers.

At last but not the least, I am offering the dissertation in memory of my brother Chiung-Yang.

TABLE OF CONTENTS

Page

CHAPTER 1

Review of Monte Carlo methods	1
Introduction	2
Monte Carlo Method	3
Importance Sampling	5
Markov Process	8
Computing the interaction in Simulation ---	
Periodic Boundary Condition and Minimum Image	12
Metropolis Monte Carlo Method	14
Monte Carlo Methods Including Force Biased	16
Free Energy Difference	21
Grand Canonical Ensemble	30
Isobaric-Isothermic Ensemble	35
Microcanonical Ensemble	37
Weak Electrolyte System	38
Reference	41

CHAPTER 2

A MONTE CARLO METHOD TO SIMULATE SYSTEMS WITH	
BARRIERS --- SUBSPACE SAMPLING METHOD	47
Abstract	48
Introduction	48

Methodology	49
Results and Discussion	51
Conclusions	57
Reference	59
Figure Captions	60

CHAPTER 3

THE SUBSPACE SAMPLING METHOD --- A MONTE CARLO APPROACH FOR SIMULATING THE SINGLE PARTICLE DENSITY FUNCTION AND THE EQUILIBRIUM CONSTANT FOR SYSTEMS DESCRIBED BY MULTIPLE HAMILTONIANS	72
Abstract	73
Introduction	73
Subspace Sampling Method	75
SSM and the Free Energy Difference Method	77
Modifications of SSM	82
Conclusions	87
References	89
Figure Captions	90

CHAPTER 4

MONTE CARLO SIMULATION OF THE PAIR CORRELATION FUNCTION AND THE EQUILIBRIUM ASSOCIATION CONSTANT OF THE STICKY ELECTROLYTE MODEL

USING THE SUBSPACE SAMPLING METHOD	102
Abstract	103
Introduction	103
Potential Model	106
Simulation Method	107
Computational Details	109
Results and Discussion	111
Conclusion	121
References	124
Figure Captions	126
 APPENDIX	
FORTRAN CODE FOR THE SUBSPACE SAMPLING	
METHOD	147
 BIBLIOGRAPHY	169

Lists of Tables

Page

Chapter 2

Table 1 Parameter sets of S.S.M method and acceptance probability	71
---	----

Chapter 4

Table 1 Comparison of simulated and exact average number of molecule	142
--	-----

Table 2 Simulated equilibrium constant of different simulation cell and initial concentration with and without dipolar interaction	143
---	-----

Table 3 Calculation of thermodynamics equilibrium constant for the system without dipolar interaction	145
--	-----

Table 4 Comparison the simulation using Ewald Summation and Minimum Image method	146
---	-----

<u>List of Illustrations</u>	Page
Chapter 1	
Figure 1 Connections of created model, analytic solution and simulation	44
Figure 2 Illustration of periodic boundary condition and minimum image	45
Figure 3 Flow chart of Metropolis Monte Carlo method	46
Chapter 2	
Figure 1 Double well potential for the barrier crossing problem	62
Figure 2 Flow chart for subspace sampling method	63
Figure 3 Comparison of simulated and exact density function	64
Figure 4 Convergence rate of density function for step size with different P	65
Figure 5 Convergence rate of density function for P with different step size	66
Figure 6 Convergence rate of density function for c with different p	67
Figure 7 Flow chart of Enhanced S.S.M method	68
Figure 8 Convergence rate of density function for Enhanced S.S.M method	69
Figure 9 Convergence rate of density function for Enhanced S.S.M method incorporating with Smart Monte Carlo method	70
Chapter 3	
Figure 1 Schematic of the two potentials used to demonstrate how to divide the subspace	92
Figure 2 Comparison of simulated and exact density function for two potentials	93
Figure 3 Convergence rate of density function for model (a) in Fig.1	94
Figure 4 Convergence rate of equilibrium constant for model (a) in Fig.1	95
Figure 5 Convergence rate of density function for model (b) in Fig 1	96
Figure 6 Convergence rate of equilibrium constant for model (b) in Fig.1	97

Figure 7 Potential function used for the comparison of different methodologies	98
Figure 8 Convergence rate of density function for various methodologies	99
Figure 9 Convergence rate of equilibrium constant for different methodologies	100
Figure 10 Convergence rate of density function for different methodologies testing with $\chi\chi(m)$	101

Chapter 4

Figure 1 Illustration of Sticky Electrolyte model	129
Figure 2 Comparison of simulated and exact density function of one A and one B system	130
Figure 3 Convergence rate of density function and equilibrium constant with different simulation parameter sets for one A and one B system	131
Figure 4 Comparison of the convergence rate of average molecule number by using S.S.M and Metropolis method for one A and one B system	132
Figure 5 Comparison of the convergence rate of average molecule number and energy by using S.S.M and Metropolis method for many particle system ...	133
Figure 6 Comparison of simulated ion-ion correlation function and that of Debye-Hückel theory	134
Figure 7 Plot of simulated equilibrium constant vs. the inverse of the volume of simulation cell	135
Figure 8 Comparison of the limiting simulated equilibrium constant and the result of the analytic solution	136
Figure 9 Linear relationship of $\ln K$ to the square root of the ionic concentration	137
Figure 10 Convergence rate of average molecule number and energy when the system includes the dipolar interaction	138
Figure 11 Simulated site-site correlation function and ion-ion correlation function with dipolar interaction included	139

Figure 12 Comparison of simulated equilibrium constant with and without dipolar interaction	140
Figure 13 Illustration of the influence of the dipolar interaction on contact correlation function	141

Chapter 1

Review of Monte Carlo methods

Introduction

This project develops and applies a novel Monte Carlo methodology to simulate chemical reaction systems, with a particular emphasis on the weak electrolyte system. The potential model which has been used to describe the weak electrolyte solution is the Sticky Electrolyte Model (1). The potential function employed in the Sticky Electrolyte Model contains two discrete regions defined by the distance between opposite charges. For short distances, the ion pair is subject to a molecular potential described by a square well. At larger distances the ions interact through the Coulomb potential. The analytic solution based on this model has been solved by Rasaiah and his co-workers (1-8). However, there is no appropriate simulation result available to evaluate the accuracy of the analytical solution. The small width of the molecular potential and the high activation barrier to molecular association and dissociation render the standard Monte Carlo methods impractical. The traditional Metropolis method rarely samples the tiny molecular potential. Furthermore, the energy barrier impedes the association of the ion pair resulting in an extremely slowly convergence rate.

Standard Monte Carlo methods that existed prior to 1990 were notoriously inefficient in facilitating the barrier crossing of discontinuous potentials. Any simulation method for the chemical reaction system must be able to establish the equilibrium distribution of the reactants and products numbers. The method should also have an efficient way to sample the configuration space so that the mechanical and structural properties can be calculated. No simulation method satisfied the above criteria. Therefore, we were motivated to develop a new method to solve the problem.

This manuscript is organized as follows. The first chapter reviews the recent developments in Monte Carlo methodologies. Chapter Two introduces a novel simulation method --- the Subspace Sampling Method and demonstrates its application to the barrier

crossing problem. Several methodologies of Subspace Sampling Method will also be compared. In Chapter Three, a further investigation about the applicability of this novel method is discussed. Various one dimensional potentials are used to demonstrate that the method can simulate the free energy difference (i.e. equilibrium constant) and the density function (or other thermodynamic properties) at the same time. In other words this method is capable of investigating the equilibrium properties of the chemical reaction system. In the last chapter, the simulation of the weak electrolyte system will be shown. The method is extended to the three dimensional many body problem. After the ion pair associates, it is modeled as either a dipolar molecule or a hard core molecule. The simulation results are used to compare with those of analytic solution. The computer code for the subspace sampling method is reported in the appendix.

Monte Carlo Method

The Monte Carlo method is a numerical method to solve an integral by using a sequence of random variables. In statistical physics and chemistry, the Monte Carlo method is typically employed to simulate the thermodynamic and structural properties of many-body condensed-matter systems, using random numbers generated by a computer (9).

Monte Carlo simulations occupy a central role in theoretical and experimental physical chemistry. The Monte Carlo approach typically proceeds by simulating an experimentally accessible parameter which is used either to test the accuracy of analytical solutions or to examine complex models of various systems. An example of a Monte Carlo simulation in physical chemistry that illustrates these ideas is the study of the strong electrolyte solution.

Figure 1-1 shows the importance of Monte Carlo method and its connection with experimental results and theoretical work (10,11,12). One experimental measurement that one wishes to understand, for example, is the pair correlation function, which can be obtained from x-ray scattering data. The analytical approach to the pair correlation function

uses Debye-Hückel theory. This approach employs two approximations: a high idealized model and simplifications of the true potential of mean force. The typical role of the Monte Carlo simulation is to test the model and/or test the severity of the mathematical approximations inherent in the analytical method. In Debye-Hückel theory, the model used is the primitive model where the aqueous solvent is described by a uniform dielectric constant and the solute molecules are point charges. Poisson's equation is constructed to describe the electrostatic charge distribution surrounding a central ion. Poisson's equation is then simplified in two ways: in the first, the Boltzmann distribution replaces the true statistical mechanical charge distribution. In addition, the exponential term of the Boltzmann distribution is then expanded to first order to obtain the so-called linear Poisson-Boltzmann equation. (If spherical geometry is employed, this is Debye-Hückel theory).

The above two assumption can only be true in low concentration solution. However, when the results of the analytic solution are inconsistent with experimental data, it is difficult to determine if this inconsistency is a consequence of the model itself or of the approximations inherent in the analytic solution. Simulation plays an important role at this stage because the Monte Carlo simulation makes no approximations. Therefore, it provides the opportunity to examine the accuracy of the analytic solution and the feasibility of the model (12). It is also useful to compute the thermodynamic properties (9, 12).

Although the Monte Carlo method is a powerful tool to investigate the structural and thermodynamic properties of the solution state, there still exists some problem for the simulation. One disadvantage of the method is its restriction to a finite system. For example, we cannot obtain the distribution function throughout all space which is necessary to simulate the thermodynamic properties of the system beyond the region of the finite simulation cell. The other disadvantage of the method is the sophisticated computational procedure required. The analytic solution has an explicit mathematical formula which is easier for calculation and results in simplified expressions for the distribution functions and thermodynamic coefficient. In spite of its disadvantages, the Monte Carlo method is

extremely useful for evaluating many features of the exact statistical mechanical properties of the liquid state.

In the subsequent sections the details of the Monte Carlo method is developed. The basic idea of the Monte Carlo method is based on importance sampling in statistics. In the next section we will discuss the notion of importance sampling.

Importance sampling

Importance sampling biases the sampling toward the most significant component of the integrand. Due to importance sampling, a complex integral can be converted to an arithmetic average with an estimation error that can, ideally, be reduced to zero (9).

If we consider the function $f(x)$, when x is in the range $[a,b]$, the integral I is given by

$$I = \int_a^b f(x) dx. \quad (1-1)$$

The average value of $f(x)$ is evaluated according to the mean-value theorem of calculus as follows:

$$\langle f(x) \rangle = \frac{I}{b-a}. \quad (1-2)$$

When the integral is computed by selecting n points randomly in the interval $[a,b]$ on a uniform distribution, the integral is expressed as

$$I = \int_a^b f(x) dx \approx \frac{b-a}{n} \sum_i^n f(x_i). \quad (1-3)$$

The error of the arithmetic summation is inversely proportional to the square root of the total sampling number n , according to the central limiting theorem (9).

In importance sampling, the points x_i are chosen using a weighing probability $P(x)$ instead of a uniform distribution. The probability distribution $P(x)$ obeys the normalization:

$$\int_a^b P(x) dx = 1; \quad P(x) > 0 \quad (1-4)$$

So the integral I is shown below:

$$I = \int_a^b f(x) dx = \int_a^b \frac{f(x)}{P(x)} P(x) dx, \quad (1-5)$$

This probability distribution is incorporated into the integral according to:

$$\langle f(x) \rangle \approx \frac{1}{n} \sum_i^n \frac{f(x_i)}{P(x_i)}. \quad (1-6)$$

where the variance of the average $\langle f(x) \rangle$ is given by

$$\sigma^2 = \int_a^b \left[\frac{f(x)}{P(x)} \right]^2 P(x) dx - \left[\int_a^b \frac{f(x)}{P(x)} P(x) dx \right]^2. \quad (1-7)$$

We select a $P(x)$ which makes σ a minimum. This is achieved when the $P(x)$ is chosen as:

$$P(x) = f(x) / \int_a^b f(x) dx. \quad (1-8)$$

Thus, σ reduces to zero in eqn. (1-7). In other words, ideally, there is no error existent for the integral I. Therefore, the errors are highly reduced for the finite simulation. The arithmetic summation in equation (1-6) is exactly equal to the equation (1-1) divided by (b-a). The above examples directly give us the sense what the importance sampling is and how the sampling methods affect the deviation of integral.

Importance sampling is applied to physical-chemical systems through sampling of the lower energy states more than the higher energy states. In statistical mechanics, the average value of any mechanical property A is computed from the partition function. According to the importance sampling, the integrals can be converted as follows:

$$\begin{aligned} \langle A \rangle &= \frac{\int A(x) \exp(-\beta V(x)) dx}{\int \exp(-\beta V(x)) dx} \\ &\approx \sum_{i=1}^n \frac{A(x_i) \exp(-\beta V(x_i))}{P(x_i)} / \sum_{i=1}^n \frac{\exp(-\beta V(x_i))}{P(x_i)}. \end{aligned} \quad (1-9)$$

Obviously, the above equation will be simplified when the weighing probability is chosen as:

$$P(x_i) = \frac{\exp(-\beta V(x_i))}{Z} \quad (1-10)$$

where the x_i is the spatial configuration and Z is the partition function. The average value of A is then obtained from:

$$\langle A \rangle = \frac{1}{n} \sum_{i=1}^n A(x_i). \quad (1-11)$$

Once again, the integral can be evaluated from an arithmetic summation when the weighing probability in equation (1-10) is the same as the equilibrium distribution of energy states.

In fact, the weighing probability is the essence of importance sampling. The contribution of different regions of configuration space to the total integral is dependent on the weighing probability. The central need of the simulation is the appropriate weighing probability. In the subsequent section, we will demonstrate how the weighing probability is achieved using a Markov Process.

Markov Process

The Monte Carlo simulation is a sequence of transitions among various states. The weighing probability (i.e. the equilibrium distribution of the energy states) is generated with such a successive sampling of configuration space. Each transition in the Monte Carlo simulation is determined by the transition probability of a Markov process. Before any further discussion we must define the Markov processes (9, 13, 14).

The Markov process is one in which the transition probability from state x_{n-1} to state x_n is associated only with previous state but not with any earlier state. The transition probability is expressed as

$$w(x_n, t_n | x_{n-1}, t_{n-1}; \dots; x_1, t_1) = w(x_n, t_n | x_{n-1}, t_{n-1}). \quad (1-12)$$

The transition probability is related to the state at time t_{n-1} only and retains no memory of its preceding history. There are four restrictions for the transition probability:

- (i) Connectivity or ergodicity : there exists $w(x, x') \neq 0$ for the whole configuration,
- (ii) Positivity : $w(x, x') \geq 0$ for all x and x' ,
- (iii) Conservation : $\sum_{x'} w(x, x') = 1$ for all x ,
- (iv) $\sum_{x'} w(x, x') P(x') = P(x)$ for all x .

The equation in condition (iv) plays a crucial role by connecting the weighing probability and the transition probability. Condition (iv) shows that the limiting distribution arrives at the weighing probability (equilibrium distribution), as we expected, after a sequence of Markovian transitions. In other words, the transition probability is the prerequisite condition to access the weighing probability. In order to access the weighing probability, we must compute the transition probability during the Monte Carlo simulation.

The formal derivation of restriction (iv) arises from the definition of a Markov process and the Chapman-Kolmogorov equation (11, 13). According to equation (1-12) the probability that a system is found in state x_n at time t_n is

$$P(x_n, t_n; \dots; x_1, t_1) = w(x_n, t_n | x_{n-1}, t_{n-1}) P(x_{n-1}, t_{n-1}; \dots; x_1, t_1). \quad (1-13)$$

The above probability is expanded to

$$P(x_n, t_n; \dots; x_1, t_1) = w(x_n, t_n | x_{n-1}, t_{n-1}) w(x_{n-1}, t_{n-1} | x_{n-2}, t_{n-2}) \dots \dots \dots w(x_2, t_2 | x_1, t_1) P(x_1, t_1). \quad (1-14)$$

The transition probability for $n=2$ is

$$w(x_2, t_2 | x_1, t_1) = \frac{P(x_2, t_2; x_1, t_1)}{P(x_1, t_1)} = \frac{P(x_2, t_2; x_1, t_1)}{\sum_{x_2} P(x_2, t_2; x_1, t_1)} \quad (1-15)$$

$$P(x_1, t_1) = \sum_{x_2} P(x_2, t_2; x_1, t_1). \quad (1-16)$$

The Chapman-Kolmogorov equation originates from the following equation:

$$P(x_3, t_3; x_1, t_1) = \sum_{x_2} P(x_3, t_3; x_2, t_2; x_1, t_1); \quad t_3 \geq t_2 \geq t_1 \quad (1-17)$$

By expanding the distribution probability P by using the definition of transition probability then the above equation becomes

$$w(x_3, t_3 | x_1, t_1) P(x_1, t_1) = \sum_{x_2} w(x_3, t_3 | x_2, t_2) w(x_2, t_2 | x_1, t_1) P(x_1, t_1). \quad (1-18)$$

After we cancel out the $P(x_1, t_1)$ equation (1-18) is simplified to

$$w(x_3, t_3 | x_1, t_1) = \sum_{x_2} w(x_3, t_3 | x_2, t_2) w(x_2, t_2 | x_1, t_1). \quad (1-19)$$

Due to the definition of equation (1-15) equation (1-19) should be equal to

$$\frac{P(x_3, t_3; x_1, t_1)}{P(x_1, t_1)} = \sum_{x_2} w(x_3, t_3 | x_2, t_2) \frac{P(x_2, t_2; x_1, t_1)}{P(x_1, t_1)} \quad (1-20)$$

that is

$$P(x_3, t_3; x_1, t_1) = \sum_{x_2} w(x_3, t_3 | x_2, t_2) P(x_2, t_2; x_1, t_1). \quad (1-21)$$

Let $P(x') = P(x_3, t_3; x_1, t_1)$, $P(x) = P(x_2, t_2; x_1, t_1)$ and $w(x', x) = w(x_3, t_3 | x_2, t_2)$. Equation (1-21) turns out to be the restriction (iv) of the transition probability, $\sum_{x'} w(x, x') P(x') = P(x)$. Since the transition probability is the key factor of the Monte Carlo simulation the subsequent discussion will show how to find the transition probability.

The explicit expression for the transition probability can be derived from the more restrictive Master Equation which is

$$\frac{dP(x, t)}{dt} = - \sum_{x'} w(x, x') P(x, t) + \sum_{x'} w(x, x') P(x', t). \quad (1-22)$$

The above equation is the rate of change of state x . When the system is in the stationary state the incoming rates and outgoing rates are equal and the solution of equation (1-22) is

$$\sum_{x'} w(x, x') P(x) = \sum_{x'} w(x', x) P(x'). \quad (1-23)$$

Replacement of the restriction (iii) to left hand side of equation (1-23) also leads to the result of restriction (iv). Using the Detailed Balance or Microscopic Reversibility, the solution for equation (1-23) is

$$\begin{aligned} w(x, x') P(x) &= w(x', x) P(x') \\ \text{or } \frac{w(x, x')}{w(x', x)} &= \frac{P(x')}{P(x)} = \frac{\exp(-\beta V(x'))}{\exp(-\beta V(x))}. \end{aligned} \quad (1-24)$$

One possible choice of the solution in equation (1-24) which was developed by Metropolis and his coworkers (15) and has been used widely, is

$$\begin{aligned} w(x, x') &= T(x, x') \frac{P(x')}{P(x)} & \text{if } V(x') > V(x) \\ &= T(x, x') & \text{if } V(x') < V(x) \end{aligned} \quad (1-25)$$

$$w(x, x) = 1 - \sum_{x' \neq x} w(x, x')$$

where the $T(x, x')$ is the trial probability of the transition and $T(x, x') = T(x', x)$ here. The trial probability must obey the criterion of conservation, i.e. $\sum_{x'} T(x, x') = 1$.

In summary, the weighing probability (or distribution probability) in importance sampling can be found by computer simulation as long as the transition probability for each transition is established. After a sequence of Markovian transitions, the weighing

probability is achieved. One of the most important steps in the Monte Carlo method is the computation of the transition probability. In order to enhance the efficacy of the Monte Carlo simulation, several modifications of the Metropolis transition probability have been proposed to include the trial probability. This subject will be continued in the latter sections after the discussion of the conventions of the computer simulation.

The transition probability of equation (1-25) is related to the interaction of configuration space. The energy difference of a transition determines the magnitude of the transition probability. From the viewpoint of the calculation of the interaction energy, especially for the many body interaction, two conventions for computer simulation need to be mentioned at this point --- Periodic Boundary Conditions and Minimum Image.

Computing the interactions in the simulation

Generally speaking, the current computer simulation method employs periodic boundary conditions to calculate the finite system (9, 12). Periodic boundary conditions are to confine the particles in a finite cell. The cell represents the system we are simulating. The particles surrounding the cell are considered to replicate the central cell through the whole space as shown in figure 2. Whenever the particle moves, its periodic image will move with the same way. If a particle drifts out the cell, another particle enters from the opposite side. Periodic boundary conditions reduce the surface effect of the simulation cell (9, 12, 23). The region near the edge is sampled more often than the case without the periodic boundary condition.

When periodic boundary conditions are employed, the computation of the interaction energy uses another convention --- minimum image (16). Minimum image eliminates the arbitrary boundary due to the periodic boundary condition. The pair interaction is based on the distance between the minimum images. It is clear in figure 2 that only those image

particles locate in the dashed cell interact with the particle resident in the center. When the particle moves, the dashed cell will accompany the particle.

The minimum image convention affects how accurately the finite cell simulates the infinite system, particularly when the potential contains long-range forces. In systems containing only short range forces, the spurious long-range correlation introduced by the minimum image convention does not affect the decay of the pair correlation function. The pair correlation function approaches unity before the image correlations occur. In systems containing long range forces, however, the tail part of the correlation function extends into the image region. In order to compensate for the tail part of the Coulomb potential, Ewald (17,18) and Valleau (19) have introduced various modifications into the calculation of the energy.

In Valleau's work the minimum image convention is still applied but the particle number and cell size are varied to maintain a constant concentration. The system energy is computed for each cell size. Then, extrapolation is applied to find the limiting value of the system with infinite size. The extrapolation properties are equivalent to the result of infinite system.

Unlike the minimum image convention, the Ewald approach includes the images from all periodic cells in the interaction energy. The summation of the interactions of these periodic cells corresponds to the calculation of the lattice energy. Recently, a fit of the complex Ewald summation with an analytic form has been reported (20). The analytic formula of the Ewald summation is more feasible to implement in a computer algorithm than the traditional equation with the complementary error function, especially when the force is calculated.

An alternative way to modify the long range electrostatic interaction is on the basis of self-consistent periodic boundary condition (9, 21, 22). It has been applied to investigate the polyionic DNA and counter ion interaction (21, 22). The DNA is modeled as an infinite cylinder. The axis of cylinder passes through the center of all periodic cells. The cells

surround the DNA cylinder with fixed radius and length. In order to include the contribution of small ions outside the central cell the external charge density must be taken into account so that the external potential from the periodic cell can be estimated. In this method, the external charge density is approximated by the average charge density of the central cell. The external potential is recomputed until no further change is observed. Convergence is evaluated from the variation of the charge density during the iteration. The minimum image convention is still applied here to the particle in the central simulation cell.

The periodic boundary condition and minimum image really convert the study of an infinite system to a finite system. However, the finite size cell inhibits the occurrence of long wavelength fluctuation (12). Thus, certain systems are not easily simulated such as the gas-liquid critical point and the long wavelength phonon of a solid. In spite of these limitations, the Monte Carlo simulation has successfully been used to explore a variety of areas since Metropolis and his coworkers proposed this idea in 1954. The Metropolis method forms the basis of all subsequent methodologies (15). The following section discusses details of the Metropolis method.

Metropolis Monte Carlo Method

In figure 3 the flow chart shows how the computer algorithm is organized in the Metropolis method (15, 23). The initial configuration is chosen arbitrarily but must be an energetically feasible choice, i.e. there should be no overlap for hard sphere particles. A particle is picked randomly or sequentially and moved to a trial position. The trial position is selected randomly within a cubic box based on the step size. The step size is the adjustable parameter of the Metropolis method. The new configuration of the system is generated from the trial position of the moved particle and the change of the configuration energy is computed by using the minimum image or some other convention.

The transition probability is analyzed from the microscopic reversibility or detailed balance as in the previous discussion and shown as follows.

$$w(r_i, r_j) = \text{Min} \left\{ 1, \frac{\Pi_j}{\Pi_i} \right\} \quad (1-26)$$

where $\Pi_i = \exp(-\beta V(r_i))$ and $\Pi_j = \exp(-\beta V(r_j))$. The transition probability determines whether the trial move is accepted. If it is accepted the particle moves to the new position. Otherwise, the particle returns to its original position. For each iteration, the mechanical properties are computed. The histogram of various thermodynamic parameters is established and the average value is obtained from the cumulative result.

In contrast to equation (1-25), the trial probability, which is the probability of attempting a move of a particle from r_i to r_j , is not shown in equation (1-26) because of cancellation. The trial probability here is

$$T(r_i, r_j) = \frac{1}{\delta_x \delta_y \delta_z} = T(r_j, r_i) \quad (1-27)$$

where the δ_x , δ_y and δ_z are the step sizes in the x, y and z directions, respectively. This illustrates explicitly the inclusion of the step size in the simulation.

The step size is a decisive factor that determines the rate of convergence of the simulation. Generally speaking, if the step size is too small, most trial moves are accepted because of the small change in the configuration energy. This results in a high acceptance/rejection ratio of trial moves and slows down the sampling of the entire configuration space. On the other hand, if the step size is too large the configuration energy change is great resulting in a low acceptance/rejection ratio and slow convergence. The rule of thumb in selecting the step size is to choose the step size to produce an acceptance ratio around 50 percent (24).

The Metropolis algorithm is strictly applied to the canonical ensemble. Consequently, there are several limitations imposed by this method. The first problem with the Metropolis method is the random walk, which is a time-consuming process (24). The second one is that the method can not be used to estimate the free energy (or partition function) of the system because importance sampling always samples the low energy states more than higher energy states (9, 12, 24). In order to compute the free energy (or partition function), the entire configuration space has to be counted thoroughly. Importance sampling is not able to sample the space completely without a tremendously long iteration. The third problem is that the method has a low crossing probability for systems with a high energy barrier, i.e. the so-called bottle-neck problem (25).

In the past two decades several algorithms have been developed to solve these problems. Recently, there have been a great deal of modified Monte Carlo methods proposed to scrutinize more complicated systems. The major difference used by these modified methods is the incorporation of a bias into the trial probability to sample the higher density probability region. The bias typically takes into account the forces acting on the particle at each step in the simulation. In the next section, the modified Monte Carlo methods that include force factors will be reviewed.

Modified Monte Carlo methods including the force

The Smart Monte Carlo method was developed by Friedman and his coworkers in 1978 (24). The method biases the displacement of particles in the direction of higher statistical probability. The random displacement in the Metropolis method is replaced by the motion of Brownian Dynamics (11, 26). The displacement from an initial position r_i to a final position r_f is

$$\Delta r = r_f - r_i = D\beta F_i \Delta t + R = \beta F_i A + R \quad (1-28)$$

where $A=D\Delta t$, D is the diffusion coefficient; F_i is the force exerted on the particle i due to all other particles in the system and R is the random displacement which is expressed as a gaussian distribution in Brownian Dynamics (11, 24, 26):

$$W(R) = (4\pi A)^{-3/2} \exp \left(-\frac{R^2}{4A} \right). \quad (1-29)$$

The trial probability T_{ij} is related to the probability that allows a particle to move from r_i to r_j under the force F_i when the random force is R . Since the r_i , r_j and F_i are specified, the probability to pick a certain R corresponds to equation (1-29). Combination of equation (1-28) and (1-29) gives the trial probability as :

$$T_{ij} = (4\pi A)^{-3/2} \exp \left[-\frac{(\Delta r - \beta A F_i)^2}{4A} \right]. \quad (1-30)$$

From microscopic reversibility or detailed balance, the ratio of the transition probabilities is obtained as

$$\frac{w_{ij}}{w_{ji}} = \frac{T_{ji} * \Pi_j}{T_{ij} * \Pi_i}. \quad (1-31)$$

The acceptance probability is defined by

$$w_{ij} = \min \left\{ 1, \frac{T_{ji} * \Pi_j}{T_{ij} * \Pi_i} \right\}. \quad (1-32)$$

The selection of T_{ij} is arbitrary as long as the T_{ij} is conservative, i.e. $\sum_i T_{ij} = 1$. The explicit form of the trial probability is

$$T_{ij} = C_s * \exp \left[-\frac{(\Delta r - \beta A F_i)^2}{4A} \right] \quad (1-33)$$

where C_s is a normalization constant and only a function of A so it is canceled in equation (1-32). In the Smart Monte Carlo method, the force exerted on the particle biases its motion to low energy states.

Berne et. al. reported a method similar to the bias method by expanding the potential around r_i to first order (27-29). Therefore,

$$\begin{aligned} T_{ij} &= C_f * \exp \{ -\lambda_i \beta \nabla_{r_i} V(r_i) * \delta r_i \} & \text{when } \delta r_i \in D \\ &= 0 & \text{otherwise} \end{aligned} \quad (1-34)$$

where λ_i is an adjustable constant and domain D is the region which the particle can move around in one step. δr_i is the displacement of the particle in the domain and $-\frac{1}{2} \Delta \leq \delta x_i, \delta y_i, \delta z_i \leq \frac{1}{2} \Delta$ where Δ is the step size of the Metropolis method. Actually, the normalization constant C_f in the trial probability is a function of force (or the potential gradient) and step size Δ . T_{ij} and T_{ji} have different normalization constants.

Smart Monte Carlo and Force Bias methods incorporate the sense of a physical trajectory to bias the motion of particles to low energy states. For example, during the simulation of a particle in a simple harmonic potential, the particle gradually steps down to the bottom of the potential when using Metropolis method. In the Smart Monte Carlo and Force Bias methods the displacement of a particle depends on the force exerted on the particle. The greater the force exerted on the particle, the larger the displacement. Hence, the particle is dragged down to the bottom more rapidly than the Metropolis' random walk and stays in there longer because of a smaller force near the bottom.

Cao and Berne extended and modified the algorithm of Smart Monte Carlo to explore the barrier crossing of the double well potential (27-29). Barrier crossing has been a

difficult task for computer simulations because of the low crossing rate even for the efficient Smart Monte Carlo or Force Bias methods. The idea is to follow Smart Monte Carlo. They implement one variable λ_i into the Force Bias method so the trial probability is selected according to :

$$T_{ij} = (4\pi A)^{-3/2} * \exp \left[-\frac{(\Delta r - \lambda_i \beta A F_i)^2}{4A} \right] \quad (1-35)$$

where λ_i has three possibilities. If $\lambda_i=0$ the situation returns to the random walk it becomes Smart Monte Carlo for $\lambda_i=1$. When the following case is chosen

$$\lambda_i = \frac{V(r_i)''}{|V(r_i)'|}, \quad (1-36)$$

the direction of motion is determined by the curvature of the potential, in addition to the force. If the particle resides in the position above the inflection point, $\lambda_i=-1$, the particle moves in the opposite direction against the force and crosses the barrier. When the particle is located in the position below the inflection point, the method is equivalent to the Smart Monte Carlo and the motion is biased to the direction of the force. The method is powerful for the double well potential but the application is limited to those continuous potentials with inflection points only.

Dynamic Monte Carlo method is another method that also originates from the idea of Brownian Dynamics (30). The recurrence formula of the displacement is similar to that of Smart Monte Carlo, that is

$$r_i^{t+1} - r_i^t = -\gamma \nabla V(r_i^t) + \alpha R \quad (1-37)$$

where γ and α are two adjustable parameters and R is random displacement referred to as the gaussian distribution. Instead of searching the trial probability, the method expresses

the Boltzmann distribution (or weighing) probability in terms of the parameter γ and α .

Consider the integrated form of the Chapman-Kolmogorov equation (in equation 1-21):

$$P(r, t + \Delta t) = \int w(r, t + \Delta t | r - \Delta r, t) P(r - \Delta r, t) d\Delta r. \quad (1-38)$$

Expanding equation (1-38) around the position r to second order and using the results of Brownian Dynamics ($\langle \Delta x_k \rangle = -\gamma V(x)$; $\langle \Delta x_j \Delta x_k \rangle \approx \alpha^2 \delta_{kj}$ (11)) yields

$$\begin{aligned} P(r, t + \Delta t) &= \int w(r, t + \Delta t | r - \Delta r, t) \left[P(r, t) + \sum_k \frac{\partial P(r, t)}{\partial x_k} \Delta x_k \right. \\ &\quad \left. + \sum_k \sum_j \frac{\partial^2 P(r, t)}{\partial x_j \partial x_k} \Delta x_j \Delta x_k \right] d\Delta r \\ &= P(r, t) + \sum_k \frac{\partial P(r, t)}{\partial x_k} \langle \Delta x_k \rangle + \sum_k \sum_j \frac{\partial^2 P(r, t)}{\partial x_j \partial x_k} \langle \Delta x_j \Delta x_k \rangle \\ &= P(r, t) + \sum_k \left[\frac{\partial}{\partial x_k} (\gamma P(r) \frac{\partial V(r)}{\partial x_k}) + \frac{1}{2} \frac{\partial^2}{\partial x_k^2} (\alpha^2 P(r)) \right]. \end{aligned} \quad (1-39)$$

For the stationary solution the distribution function of the weighing probability at time t is equal to that at time $t + \Delta t$. Therefore,

$$\frac{\partial}{\partial x_k} (\gamma P(r) \frac{\partial V(r)}{\partial x_k}) + \frac{1}{2} \frac{\partial^2}{\partial x_k^2} (\alpha^2 P(r)) = 0$$

and

$$P(r) = C_d \exp(-2\gamma V(r)/\alpha^2) \quad (1-41)$$

where γ should be sufficiently small to ensure that the second order Taylor expansion as a function of position r is adequate. Consequently, the following equality can be obtained.

$$kT = \alpha^2 / 2\gamma. \quad (1-42)$$

Equation (1-41) shows the connection between α and γ . In the simulation, the γ is selected and then the recurrent displacement is computed according to (1-37). In comparison with the force bias or Smart Monte Carlo methods, the trial move of Dynamic Monte Carlo does not need to incorporate the trial probability because the sequential motion is generated by the Markov process, equation (1-39) and the equilibrium distribution probability has been included in equation (1-40) .

The efficacy of the Metropolis method has been improved by including the physical trajectory to modify the slow convergence of the random walk. In addition, the Metropolis method has been expanded to compute non-mechanical variables (free energy simulation). In fact, the Monte Carlo method has more advantages for the free energy difference simulation than Molecular Dynamics (12). Although the molecular dynamics simulation contains time, which is not explicitly included in the Monte Carlo method, it is difficult to transfer the particle between two hamiltonian systems in the same configuration space using molecular dynamics. The subject of next section is free energy difference simulations using Monte Carlo Methods.

Free Energy Difference

Although the canonical Monte Carlo simulation can not efficiently calculate free energies, free energy differences can be attained in the canonical ensemble (31,32). In the free energy difference calculation two systems are involved. One is a reference system with a well-known free energy. The other is the new system. The method must be able to sample both systems in one simulation.

The basic principle of the free energy difference method arises from microscopic reversibility or detailed balance. If the sampling process proceeds between two energy systems (two hamiltonians) that span the same configuration space, the method mimics the essence of the Markov process except the transition involves a switching between the two

hamiltonians. The transition probability for transferring between these hamiltonians is based on the Metropolis method:

$$w(r^1, r^2) \exp(-\beta V(r^1)) = w(r^2, r^1) \exp(-\beta V(r^2)). \quad (1-43)$$

If we multiply by Q_1/Q_1 and Q_2/Q_2 on left and right hand sides respectively then integrate, equation (1-43) becomes:

$$\langle w(r^1, r^2) \rangle_{Q_1} = \langle w(r^2, r^1) \rangle_{Q_2}. \quad (1-44)$$

The ratio of the partition functions (Q_2/Q_1) is equal to the ratio of the average acceptance probability from system 1 to system 2 and that from the opposite direction. The method is known as the acceptance ratio (or probability ratio) method. In other words, the method counts how many times a particle transfers from one system to the other. If the two potentials have a small overlap, the method is limited because long computer times are needed to sample both potential functions. This results in a large statistical fluctuation.

In the Metropolis method, the transition probability is described by the Metropolis function $w(x) = \min \{ 1, x \}$ where $x = \exp[-\beta(V(r_j) - V(r_i))]$. Bennet altered the method to replace the weighing probability with the fermi function, which is $f(x) = 1/(1 + \exp(x))$ (31). The fermi distribution function extends the sampling regime so that the particle does not stay in the high probability density region all the time. If we compare the Fermi function with the Metropolis function, the former one climbs to 1 continuously when $V(r_j) - V(r_i)$ is less than zero instead of jumping to 1 suddenly like the Metropolis method. However, by choosing the Fermi function, the ratio of the acceptance probabilities is not directly equal to the ratio of partition functions. After a tedious statistics argument the correction obtained is as follows:

$$\ln (Q_2/Q_1) = C - \ln (n_2/n_1) \quad (1-45)$$

where n_1 and n_2 are the occupancy of the V^1 and V^2 regions respectively. The C is a constant and is the crucial factor of the method. According to the equation (1-44) and the analysis of the optimization of the free energy difference estimation, one more equation can be obtained:

$$\frac{Q_2}{Q_1} = \frac{\langle f(V^2 - V^1 + C) \rangle_1}{\langle f(V^1 - V^2 - C) \rangle_2} \exp(+C). \quad (1-46)$$

After implementing a series of simulations with different C values, we can graph equations (1-45) and (1-46) versus C , and the intersection point of the two function leads to the free energy difference.

Although this method has been used when the two potentials have only a small density overlap, the fundamental notion remains to confine two hamiltonians in the same configuration region. For those cases where the hamiltonians do not have the same size and the regime of overlap is small, Voter proposed a modified method (33).

In Voter's work, he expanded the hamiltonians to arbitrary sizes. For example, he considered two harmonic oscillators with different force constants in the same space. One has a minimum at position $r^1_{\min}$. The minimum of the other harmonic potential is at $r^2_{\min}$. This displacement D is defined as $D = r^2_{\min} - r^1_{\min}$. In this case the overlap in the density functions is obviously small because the minimums of the two potentials are away from each other. The ratio of the partition functions is similar to that of the acceptance ratio method:

$$Q_2 \int \frac{w(r^2, r^1) \exp(-V^2(r^2))}{Q_2} dr = Q_1 \int \frac{w(r^1, r^2) \exp(-V^1(r^1))}{Q_1} dr \quad (1-47)$$

where r^1 (or r^2) is the particle position at r but it belongs to potential system 1 (or 2). Implementing D into the equation (1-47) and changing the variable on the left hand side becomes

$$\begin{aligned} Q_2 \int \frac{w(r^1, r^2-D) \exp(-V^2(r^1))}{Q_2} dr \\ = Q_1 \int \frac{w(r^2, r^1+D) \exp(-V^1(r^1+D))}{Q_1} dr \end{aligned} \quad (1-48)$$

and

$$\frac{Q_2}{Q_1} = \frac{\langle w(r^2, r^1+D) \rangle_1}{\langle w(r^1, r^2-D) \rangle_2} \quad (1-49)$$

where the $\langle w(r^2, r^1+D) \rangle_1, \langle w(r^1, r^2-D) \rangle_2$ are the average acceptance probability under the system 1 and 2, respectively. When $D=0$, Voter's method reduces to the acceptance ratio method (31, 32). From equation (1-49) we can obtain the ratio of the partition functions as in the acceptance ratio method. The particle is switched to another potential system for each move by including the distance D . In Voter's work, two potential systems do not need to reside in the same confined region. By shifting the coordinate with the displacement D in equation (1-49), the probability density of two potentials produce a good overlap in the relative coordinate system.

In both the acceptance ratio method and Voter's method, a potential switching move is engaged. The trial move can be the random walk. The next step is to switch the particle to the other potential system and determine whether the process is accepted. In a different approach, Valleau applied a non-Boltzmann distribution to the free energy difference problem, which is known as the umbrella sampling method (16, 34). Again, there are two systems: one is the reference system with potential $V_0(r)$ at temperature T_0 and the other

system with potential $V(r)$ at temperature T . The free energy difference of these two ensembles is

$$\begin{aligned}
 \frac{A}{kT} - \frac{A_0}{kT_0} &= -\ln \frac{\int \exp[-V/kT + V_0/kT_0] \exp(-V_0/kT_0) dr}{\int \exp(-V_0/kT_0) dr} \\
 &= -\ln \langle \exp(-V/kT + V_0/kT_0) \rangle_0 \\
 &= -\ln \langle \exp(-\Delta V^*) \rangle_0
 \end{aligned} \tag{1-50}$$

where $V^* = V/kT$ is the reduced energy. The free energy difference is converted to the canonical ensemble average of the reference system. If we replace the variable r by ΔV^* the equation is simplified to a one dimensional integral problem, i.e.

$$\frac{A}{kT} - \frac{A_0}{kT_0} = -\ln \int_{-\infty}^{\infty} f_0(\Delta V^*) \exp(-\Delta V^*) d\Delta V^* \tag{1-51}$$

where $f_0(\Delta V^*)$ is the distribution function of ΔV^* in reference state. The equation shows the integration of configuration space is converted to ΔV^* space. The conversion factor is $f_0(\Delta V^*)$. The key point is to find the $f_0(\Delta V^*)$. Valleau defined a distribution function, which is a non-Boltzmann distribution, as

$$p(r) = \frac{w(r) \exp(-V_0(r)/kT_0)}{\int w(r) \exp(-V_0(r)/kT_0) dr} \tag{1-52}$$

where $w(r) = W(\Delta V^*)$ is any positive weighting function. The iteration process is still used as in the Metropolis method except the acceptance probability follows equation (1-52), which does not obey the Boltzmann distribution. Any mechanical property Θ is calculated as

$$\langle \Theta \rangle_0 = \int \Theta f_0(\Delta V^*) d\Delta V^* = \frac{\int (\Theta/w) w \exp(-V_0/kT_0) dr}{\int (1/w) w \exp(-V_0/kT_0) dr} = \frac{\langle \Theta/w \rangle_w}{\langle 1/w \rangle_w} \quad (1-53)$$

The average $\langle \Theta \rangle_0$ is carried out over the distribution in equation (1-51) which is equivalent to the $f_0(\Delta V^*)$. Actually, equation (1-53) implies that

$$f_0(\Delta V^*) = \frac{f_w(\Delta V^*)/W(\Delta V^*)}{\langle 1/w \rangle_w} \quad (1-54)$$

where $W(\Delta V^*)$, that is $w(r)$, is searched by trial and error until the $f_w(\Delta V^*)$ is as wide and uniform as possible so that the sampling process can, uniformly, cover all the configuration space, including the small overlap region of two potentials. The function $w(r)$ connects the distribution function of the configuration space with the ΔV^* space for the new potential system for which the free energy is calculated. It is also the bridge between the reference system and new system (equation 1-52).

In umbrella sampling, $w(r)$ is guessed and $f_w(\Delta V^*)$ and $\langle 1/w \rangle_w$ are computed during the simulation with the non-Boltzmann weighing probability (equations 1-24 and 1-52). A satisfactory guess for $w(r)$ (or $W(\Delta V^*)$) speeds up the convergence as well. The $f_0(\Delta V^*)$ is computed by the histogram during the simulation. Consequently, the $f_0(\Delta V^*)$ is found and inserted back into the equation (1-51) to compute the free energy. However, the searching for the appropriate $W(\Delta V^*)$ is a laborious task. It is an obvious disadvantage of the method.

The above free energy difference methods are best for systems with similar potentials. When the shapes of two potentials are widely different, several intermediate potentials must be included to produce the best estimation. The free energy difference of any adjacent states are evaluated by the umbrella sampling. The total free energy difference is obtained by summation of all intermediates.

An important method to compute free energy difference uses Zwanzig's perturbation theory (11, 35). However, the major problem of this method is that it can only be used to compute small free energy differences around 1.5 Kcal/mol (32). For many biochemical and chemical problems this limitation is too restrictive.

In the perturbation approach, the free energy difference between two systems is estimated from the system energy of any intermediate state $E(r;\lambda)$ where

$$\begin{aligned} E(r;\lambda) &= (\lambda-1)E_0(r) + \lambda E_1(r) \\ &= E_0(r) + \lambda(E_1(r)-E_0(r)) = E_0(r) + \lambda\Delta E^{(1)}(r) \end{aligned} \quad (1-55)$$

and $\Delta E^{(1)}(r)=E_1(r)-E_0(r)$. When λ is small, we can treat $\lambda\Delta E^{(1)}$ like a perturbation term.

The approximation made here is the linear relationship in equation (1-55).

The free energy is exactly equal to

$$\Delta A = \int_0^1 \left[\frac{\partial A(\lambda)}{\partial \lambda} \right] d\lambda \quad (1-56)$$

where

$$A(\lambda) = -kT \ln Z(\lambda) \quad (1-57)$$

and

$$Z(\lambda) = \int \exp[-\beta E(r,\lambda)] dr = Z(0) \langle \exp(-\beta \lambda \Delta E^{(1)}(r)) \rangle_0. \quad (1-58)$$

From equation (1-58), $\frac{\partial A(\lambda)}{\partial \lambda}$ in equation (1-56) is equal to $\langle \frac{\partial E(r,\lambda)}{\partial \lambda} \rangle$. After combining

equations (1-55) and (1-58), the free energy difference is

$$\Delta A = \int_0^1 \left\langle \frac{\partial E(r, \lambda)}{\partial \lambda} \right\rangle d\lambda = \int_0^1 \langle \Delta E^{(1)}(r) \rangle d\lambda \approx \sum_{i=1}^N \langle \Delta E(r) \rangle_i \Delta \lambda_i. \quad (1-59)$$

In the simulation, a series of λ values are chosen and $\langle \Delta E(r) \rangle_i$ are calculated. From a plot of $\langle \Delta E(r) \rangle_i$ versus $\Delta \lambda_i$, the final ΔA is computed by integration. In the algorithm, $E(r, \lambda_i)$ is evaluated using either Monte Carlo or Molecular Dynamics for each λ_i and $E(r, \lambda_{i+1})$ is also computed by only changing λ_i to λ_{i+1} in each move. The energy difference $\Delta \langle \Delta E^{(1)}(r) \rangle_i = \langle E(r, \lambda_{i+1}) \rangle - \langle E(r, \lambda_i) \rangle$ is obtained from the simulation.

Another free energy method based loosely on the perturbation approach is the expanded ensemble method, recently developed by Lyubartsev et. al. (36). This method is used to evaluate the free energy difference relative to the ideal gas of which the reciprocal temperature β is zero, i.e. the temperature approaches infinity. They inserted several substates with the temperature between $\beta_0=0$ to the expected temperature β . The partition function for each temperature is

$$Z_m = \frac{1}{N!} \int \prod d\mathbf{r}_i \exp(-\beta_m V\{\mathbf{r}_i\}) \quad (1-60)$$

where $\mathbf{r}_i$ is the position of the i th particle, and V is the hamiltonian of the system. Consequently, a new ensemble is created as

$$Z = \sum_{m=0}^M Z_m \exp(h_m) \quad (1-61)$$

where h_m are a set of constants related to the temperature. The probability P_m of the system which can exist at temperature β_m is

$$P_m = Z_m \exp(h_m) / Z. \quad (1-62)$$

Therefore,

$$\frac{P_m}{P_k} = \frac{Z_m}{Z_k} \exp(h_m - h_k) = \exp(-\beta_m F_m + \beta_k F_k + h_m - h_k) \quad (1-63)$$

When $k=M$ and $m=0$ the free energy F_M is

$$\beta_M F_M = -\ln(Z_M) = -\ln(P_M/P_0) + h_M - h_0 - \ln(Z_0). \quad (1-64)$$

The free energy at different temperatures depends on the constant h_M . To compute the free energy, both move processes and temperature switching processes, which are selected randomly, are included in the computer algorithm. The former process follows the Metropolis method to sample the position of the particles. In the temperature switching process, the positions of the particles are fixed and the temperature is varied randomly to a new one. The transition probability is $\min\{1, \exp[(\beta_k - \beta_m)H\{r_i\} + h_m - h_k]\}$ based on the equation (1-60) and (1-61). The P_m and average F_m are estimated from the histogram to determine the set of constants $\{h_m\}$. The optimal values of h_m are unknown in the beginning so that we have to guess them initially. After a short run, $\{h_m\}$ is adjusted by inserting the average F_m and P_m into equation (1-64). Convergence is achieved self-consistently.

This method appears most useful for calculating free energy difference in the canonical ensemble. Furthermore, the thermodynamic properties at various temperatures can also be computed at the same time. The scope of this application is still being investigated. In general, the free energy is not readily obtained in the canonical Monte Carlo simulation. However, the novel free energy difference calculations enable the use of the canonical ensemble to obtain non-mechanical variables. In addition to the canonical ensemble, other ensembles have also been used in the past few decades to simulate other thermodynamic properties. In the next topic a variety of ensembles are reviewed.

Grand Canonical Ensemble

Grand canonical ensemble simulations (GCMC) use the grand canonical ensemble to simulate an open system (9, 22, 37-39). In this ensemble three parameters, (μ, v, T) , are held to be constants, where μ is the chemical potential of the system and v is the volume. In the grand canonical ensemble, the particle number N is not fixed as in the simulations using canonical ensemble. In the implementation of the GCMC method, the average particle number is obtained by insertion and deletion processes. These two processes adjust the particle number until equilibrium is achieved. The move process is also included in the algorithm to sample the configuration space for a fixed particle number.

If the move process is selected, we repeat the canonical Monte Carlo procedure to sample the configuration space. However, the transition probability for insertions and deletions should relate to the partition function of the grand canonical ensemble, that is

$$\begin{aligned}\Xi &= \sum_N \exp(\mu N/kT) \int dr \frac{1}{N! \Lambda^{3N}} \exp(-V_N(r)/kT) \\ &= \sum_N \frac{\exp(\mu N/kT)}{N!} Z(N, v, T)\end{aligned}\quad (1-65)$$

where $Z(N, V, T)$ is the partition function of canonical ensemble when there are N particles confined in a cell with volume v and temperature T . $V_N(r)$ is the interaction of the system with N particles and $\Lambda = (h^2/2\pi m kT)^{1/2}$. The probability $\Pi(V_N(r), N)$ when N particles are in the system with the energy $V_N(r)$ is

$$\Pi(V_N(r), N) = \frac{\exp(\mu N/kT) \exp(-V_N(r)/kT)}{N! \Lambda^{3N} \Xi} \quad (1-66)$$

According to microscopic reversibility or detailed balance the transition probability for insertion is

$$\begin{aligned} T(N, N+1) \Pi(V_N(r), N) w(N, N+1) \\ = T(N+1, N) \Pi(V_{N+1}(r), N+1) w(N+1, N) \end{aligned} \quad (1-67)$$

where $T(N, N+1)$ and $T(N+1, N)$ are the trial probabilities of insertions and deletions, respectively. The insertion process adds one particle into the cell. The deletion process removes a particle from the system. The mathematical expressions for these two processes are

$$\begin{aligned} T(N, N+1) &= v && \text{(insertion)} \\ T(N+1, N) &= N+1 && \text{(deletion)}. \end{aligned} \quad (1-68)$$

Consequently, equation (1-67) is rearranged to

$$\frac{w(N, N+1)}{w(N+1, N)} = \frac{T(N+1, N) \Pi(V_{N+1}(r), N+1)}{T(N, N+1) \Pi(V_N(r), N)}. \quad (1-69)$$

If we combine equations (1-66), (1-68) and (1-69), one possible solution of equation (1-69) is

$$\begin{aligned} w(N, N+1) &= \min \left\{ 1, \frac{v}{\Lambda^{3N}} \exp(\beta\mu) \exp(-\beta(V_{N+1}(r) - V_N(r))) \right\} \\ &= \min \left\{ 1, \frac{1}{N} \exp[B - \beta(V_{N+1}(r) - V_N(r))] \right\} \end{aligned} \quad (1-70)$$

where

$$B = \beta(\mu - \mu_{\text{ideal}}) + \ln(N) = \beta\mu + \ln\left(\frac{v}{\Lambda^3}\right). \quad (1-71)$$

Clearly, the value of B determines the value of μ . B can be interpreted as the excess chemical potential. For the deletion process the transition probability is

$$w(N, N-1) = \min \{ 1, N \exp[-B - \beta(V_{N-1}(r) - V_N(r))] \}. \quad (1-72)$$

The insertion and deletion transition probabilities are equivalent to the following ensemble:

$$\Xi = \sum_N \frac{1}{N!} \exp(NB) \int dr \exp(-V_N(r)/kT). \quad (1-73)$$

In many cases, it is more illustrative to replace the chemical potential with the activity to create a constant $[a, v, T]$ ensemble (9, 22). The activity a is defined as

$$a = \frac{\exp(\mu/kT)}{\Lambda^3 N}. \quad (1-74)$$

The transition probability of insertion is modified to :

$$w(N, N+1) = \min \left\{ 1, \frac{a v}{N} \exp(-\beta(V_{N+1}(r) - V_N(r))) \right\} \quad (1-75)$$

and the deletion process is

$$w(N, N-1) = \min \left\{ 1, \frac{N}{a v} \exp(-\beta(V_{N-1}(r) - V_N(r))) \right\}. \quad (1-76)$$

One important application of grand canonical Monte Carlo is the simulation of the strong electrolyte system with solute stoichiometry as $M_{v+}X_{v-}$ where M is the cation and X

is the anion (39). Valleau generalized the GCMC method to any ionic species with different charges on the basis of the restricted primitive model. The transition probabilities are adapted to the ionic system where electroneutrality must be maintained. For this system, $q_+v_+ + q_-v_- = 0$, where q_+ and q_- are the charges of the cation and the anion respectively. Let i and j be the states before and after insertion of an electrically neutral strong electrolyte respectively so $N_j^+ = N_i^+ + v_+$ and $N_j^- = N_i^- + v_-$. If the probability of selecting and insertion or deletion processes is equal to P then the trial probability of one insertion step is

$$T_{ij} = \frac{P}{v^V} v_+! v_-! \quad (1-77)$$

where $v = v_+ + v_-$. In this procedure we place v_+ cations and v_- anions into the arbitrary positions of the cell. $v^V/v_+! v_-!$ is the total possible choices for this specific insertion step. The deletion process has to consider the total possible ways to pick cations and anions. For cations there are $N_j^+! / N_i^+!$ choices to this deletion procedure. Therefore, the trial probability is

$$T_{ji} = \frac{P N_i^+! N_i^-! v_+! v_-!}{N_j^+! N_j^-!} \quad (1-78)$$

The probability of a certain state is

$$P(V_N(r), N) = \frac{\exp[(\mu_+ N_i^+ + \mu_- N_i^-)/kT] \exp(-V_N(r)/kT)}{\Lambda^{3N_i^+} \Lambda^{3N_i^-} \Xi} \quad (1-79)$$

where $N = N_i^+ + N_i^-$. From equation (1-77)- (1-79) the transition probabilities can be expressed as the following ratio:

$$\frac{w_{ij}}{w_{ji}} = \frac{N_i^+! N_i^-!}{N_j^+! N_j^-!} \exp[B \cdot (V_j - V_i)/kT] \quad (1-80)$$

where

$$B = \beta(\mu - \mu_{\text{ideal}}) + \ln(N^{V+} N^{V-}) = \beta\mu + \ln\left(\frac{v}{\Lambda_+^{3v+} \Lambda_-^{3v-}}\right). \quad (1-81)$$

The transition probability can be simplified to

$$\begin{aligned} w_{ij} &= \min\left\{1, \frac{w_{ji}}{w_{ij}}\right\} && \text{for insertion} \\ w_{ji} &= \min\left\{1, \frac{w_{ij}}{w_{ji}}\right\} && \text{for deletion.} \end{aligned} \quad (1-82)$$

In the grand canonical ensemble the average particle number is computed. Moreover, the activity coefficient is readily obtained by the following relationship:

$$v \ln \gamma_{\pm} = \beta(\mu - \mu_{\text{ideal}}). \quad (1-83)$$

In the canonical ensemble, extraction of the activity coefficient is a laborious process (21, 40-42). The activity coefficient is related to the osmotic coefficient. According to the Gibbs-Duhem relation in the McMillan-Mayer model (43, 44)

$$M d \ln \gamma_{\pm} + d[M(1-\phi)] = 0 \quad (1-84)$$

where M is the concentration of solute, $\gamma_{\pm}$ is the activity coefficient and ϕ is the osmotic coefficient. In order to describe the properties of the liquid state, McMillan and Mayer showed that the equation of imperfect gas theory corresponds rigorously to the solution of nonelectrolytes (11, 43, 44). The solvent is considered as a dielectric continuum. The

restricted primitive model is consistent with McMillan-Mayer theory except the opposite ions of electrolyte have the same size and the particles behave like hard spheres when the distance of two ions is less than their sum of radii.

The formula to calculate osmotic pressure arises from the pressure equation of the McMillan-Mayer model (41, 42).

$$\phi^{-1} = \frac{U}{3NkT} + \frac{2\pi M}{3} \sum_i \sum_j x_i x_j g_{ij}(\sigma_{ij}) \sigma_{ij}^3 \quad (1-85)$$

where σ_{ij} is the sum of the radii of two different species, $g_{ij}(\sigma_{ij})$ is their contact correlation function and x_i is the mole fraction of species i . The deviation of the simulated $g_{ij}(\sigma_{ij})$ is always the limitation for the accuracy of the computed activity coefficients in the canonical ensemble (39). This limitation is relaxed in the grand canonical ensemble where the activity coefficient is specified.

Isobaric-Isothermal Ensemble

Due to the definition of the (N,P,T) ensemble, the volume of the system fluctuates. The algorithm has some similarity to the grand canonical ensemble except that the volume of the system, instead of the particle number, is varied randomly (45). The move process is still used in order to sample the configuration space. If the process of sampling the volume is selected, the reduced positions α_i of particles are the same but the box length L is changed. The reduced position is defined as

$$\alpha_i * L = r_i \quad (1-86)$$

where r_i is the real position of the i th particle. The partition function is

$$\Delta = L(2\pi m\beta/h^2)^{3N/2}(1/N!) \int dv \exp(-\beta p v) \int dr^N \exp(-\beta V(r^N)) \quad (1-87)$$

where L is a constant, v is the volume of the system and $V(r^N)$ is the total interaction of the N particles in the system. In terms of the reduced position, equation (1-87) gives rise to the probability distribution, which is proportional to

$$\Pi_j(v) \propto \exp(-\beta p v - \beta V(\alpha^N, L) + N \ln v) \quad (1-88)$$

where $N \ln v$ comes from the argument dr^N when it is transformed to reduced coordinates in equation (1-87) and $V(\alpha^N, L)$ is equal to $V(r^N)$. If microscopic reversibility or detailed balance is employed, the transition probability for the change in volume becomes

$$w_{ij} = \min\{1, \exp(-\beta p \Delta v - \beta \Delta V(\alpha^N, L) + N \Delta \ln v)\}. \quad (1-89)$$

where i and j represent the volume of the two different states; Δv and $\Delta \ln v$ are the volume and logarithm volume change, respectively; and $\Delta V(\alpha^N, L)$ is the energy difference of the transition. In order to calculate the interaction, the reduced position must be multiplied by the side length. Equation (1-89) can be simplified to

$$w_{ij} = \min\{1, \exp(-\beta \omega)\} \quad (1-90)$$

$$\text{with } \omega = p \Delta v + \Delta V(\alpha^N, L) - (N \Delta \ln v) / \beta. \quad (1-91)$$

In equation (1-90), the transition probability has the same form as the Metropolis method. From this method, the enthalpy and volume changes are obtained naturally. The

deviation of the simulation is larger than other ensembles because the volume change leads to larger fluctuations.

Microcanonical Ensemble

In the Microcanonical Ensemble the particle number N , the volume of the system V and total energy of the system E are held constant (9, 46). The partition function in phase space is

$$\Omega = \sum_i \delta(H(p_i, r_i) - E) \quad (1-92)$$

where $H(p_i, r_i)$ is the hamiltonian of the system; p_i and r_i are particles' momentum and position, respectively. Mechanical properties are given by

$$\langle A \rangle = \frac{1}{\Omega} \sum_i A_i \delta(H(p_i, r_i) - E). \quad (1-93)$$

In the Monte Carlo simulation using the microcanonical ensemble, the configuration with the required energy is found first. Any transition can produce only two results. The new configuration either has the same energy or does not. The transition probability is then either 0 or 1. The random walk produces a very slow convergence because we need only the configuration with energy E . A modified method to enhance convergence samples the constant energy surface instead of moving through all of configuration space.

Creutz included the kinetic energy (Demon Energy E_D) to force the system to fluctuate in its constant energy surface (46). Equation (1-92) simply becomes

$$\Omega = \sum_r \sum_{p_i} \delta(E_D(p_i), V(r_i) - E). \quad (1-93)$$

Therefore, the random move is still useful to produce a new configuration. Although the procedure changes the configuration interaction energy (potential energy), the demon energy will compensate for the difference as long as E_D is large enough. Otherwise, the new configuration results in a negative kinetic energy. In such a situation, the transition is discarded in the simulation. In other words the new kinetic energy has to be computed after the transition, that is $E_D(p_j) = E_D(p_i) - [V(r_j) - V(r_i)]$. When the potential difference is less than zero the transition is accepted or the $E_D(p_i)$ decides whether the transition is accepted. If $V(r_j) - V(r_i)$ is greater than zero only those demon energies which are greater than the difference can force the particles to move.

The probability for the demon energy is chosen from an exponential probability distribution (c.f. Maxwell distribution):

$$P(E_D) \propto \exp(-\beta E_D). \quad (1-94)$$

The temperature is easily found from the distribution by plotting the logarithm $P(E_D)$ versus E_D . The slope gives β . The selection of the initial demon energy requires care. For example, a large negative demon energy results in all of the transitions forbidden because $V(r_j) - V(r_i)$ is not enough to provide a positive kinetic energy. Consequently, the demon energy absorbs all the energy. Generally speaking, the initial demon energy is set to be zero.

From the preceding discussion, we can see that in each ensemble there is a natural set of thermodynamic and structural properties that are readily simulated. However, all the simulation methods have their own limitations. The further development of various ensemble methods is still an important subject.

Weak Electrolyte System

For the past two decades, the strong electrolyte system has been extensively studied using Monte Carlo simulation techniques. The strong electrolyte solution has cations, anions and solvent molecules. In the McMillan-Mayer model the solvent is treated implicitly as a continuum medium described by a dielectric constant. Thus, the system reduces to a two component problem. In the weak electrolyte solution, the molecules do not dissociate completely. Some weak electrolytes have a large association constant so that molecules are the dominant constituent of the solution; only a number of ions participate in the interaction and contribute to the stability of the system.

Rasaiah et al developed the "Sticky Electrolyte Model" in 1985 to investigate the weak electrolyte solution (1-8). The "Sticky Electrolyte Model", which is based on the restricted primitive model, was applied to diatomic weak electrolytes. In this model, a square well potential is used to describe the molecular potential when the distance of two opposite ions is less than the sum of their radii. Beyond the molecular region, the particles behave like Coulombic ions. The expression of the potential is

$$\begin{aligned}
 V_{AA}(r) &= V_{BB}(r) = \infty & r < \sigma \\
 V_{AB}(r) &= \frac{e^2}{\epsilon r} & r > \sigma \\
 &= \epsilon_1 & 0 < r < L-w/2 \\
 &= -\epsilon_2 & L-w/2 < r < L+w/2 \\
 &= \epsilon_1 & L+w/2 < r < \sigma \\
 &= 0 & r > \sigma
 \end{aligned} \tag{1-95}$$

where w is the width of square well potential; ϵ is the dielectric constant; ϵ_1 is the activation barrier of solvent repulsion; ϵ_2 is the depth of the square well potential. According to the above potential, the distance distinguishes between molecules and ions. One more criterion for the molecule is the "steric saturation". This means each atom A or B belongs to, at

most, one molecule. Moreover, the charge of each atom does not disappear when the ions form the molecule. In other words they become a dipolar molecule.

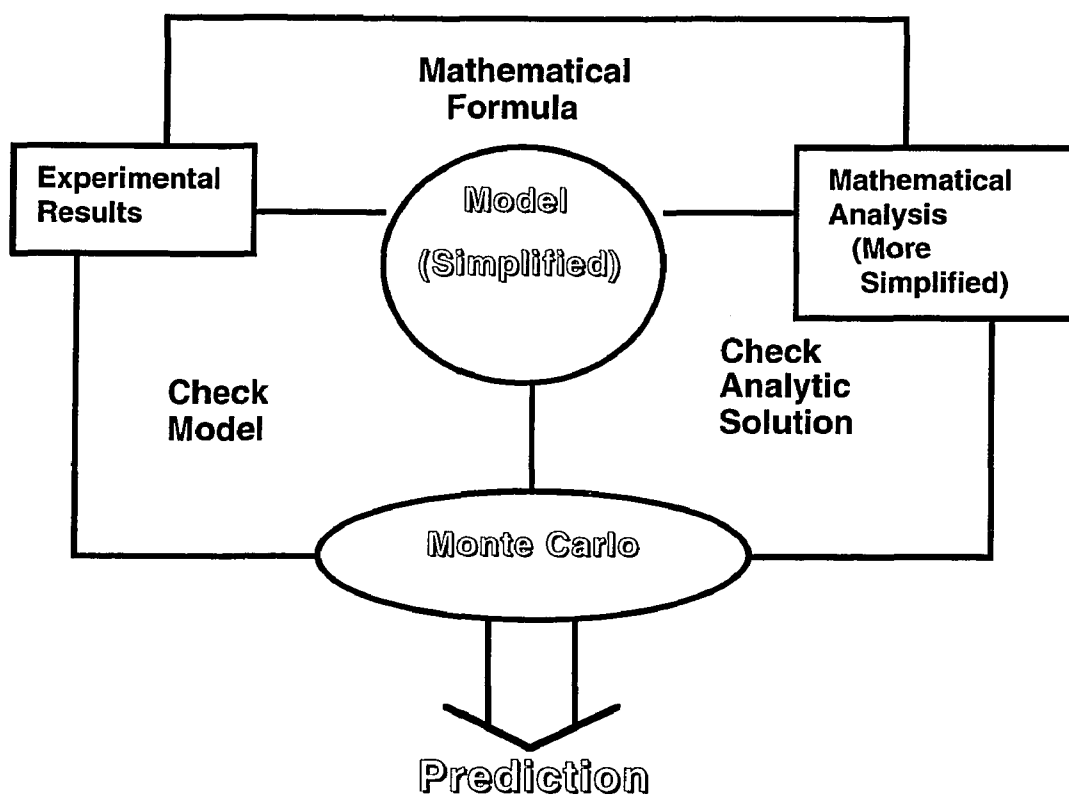
Analytical solutions for the Sticky Electrolyte Model have been developed by Rasaiah and co-workers. However, the evaluation of the accuracy of these solutions, and consequently the accuracy of the models, has been impeded by the lack of a Monte Carlo method that can simulate the weak electrolyte solution. The subject of this thesis is the development of such a methodology and the application of the methodology to study the weak electrolyte solution using the Sticky Electrolyte Model.

References

- (1) Lee, S.H.; Rasaiah, J.C.; Cummings, P.T. *J.Chem. Phys.* **1985**, 83, 317.
- (2) Rasaiah, J.C.; Lee, S. H. *J. Chem. Phys.* **1985**, 83, 5870.
- (3) Rasaiah, J.C.; Lee, S.H. *J. Chem. Phys.* **1985**, 83, 6396.
- (4) Rasaiah, J.C.; Lee, S.H. *J. Chem. Phys.* **1987**, 86, 983.
- (5) Rasaiah, J.C ; Zhu, J.; Lee, S.H. *J. Chem. Phys.* **1989**, 91, 495.
- (6) Rasaiah, J.C ; Zhu, J.; Lee, S.H. *J. Chem. Phys.* **1989**, 91, 505.
- (7) Rasaiah, J.C ; Zhu, J. *J. Chem. Phys.* **1991**, 94, 3141.
- (8) Rasaiah, J.C. *Int. J. Thermodynamics.* **1990**, 11, 1.
- (9) Heermann, D.W. *Computer Simulation Methods*, 2nd edition, Springer-Verlag **1990**.
- (10) Hecht, C.E. *Statistical Thermodynamics and Kinetic Theory* , W. H. Freeman and Company, New York, **1990**.
- (11) McQuarrie, D.A. *Statistical Mechanics*, HarperCollins Publishers, New York, **1976**.
- (12) Allen, M.P.; Tildesley, D.J. *Computer Simulation of Liquid*, Oxford Science Publications, Cambridge, **1987**.
- (13) Risken, H. *The Fokker-Planck Equation*, Springer, Berlin, New York, 1984
- (14) Van Kampen, N.G. *Stochastic Process in Physics and Chemistry*, North-Holland, Amsterdam, New York, **1981**.
- (15) Metropolis, N.; Rosenbluth, A.W; Rosenbluth, M.N.; Teller, A.H. and Teller, E. *J. Chem. Phys.* **1953**, 21, 1087.
- (16) Berne, B.J.; ed. *Statistical Mechanics, Part A: Equilibrium Techniques*, Plenum Press, New York, **1977**.
- (17) Ewald, P.P., *Ann. Phys. (Leipzig)* **1921**, 64, 253.
- (18) Brush, S.G.; Sahlin, H.L.; Teller, E. *J. Chem. Phys.* **1966**, 45, 2102.
- (19) Valleau, J.P.; Cohen, L.K. *J. Chem. Phys.* **1980**, 72, 5935.
- (20) Adams, D.J.; Dubey, G.S. *J. Comp. Phys.* **1987**, 72, 156.

- (21) Mills, P; Anderson, C.F.; Record. M.T., Jr. *J. Phys. Chem.* **1985**, 89, 3984.
- (22) Mills, P; Anderson, C.F.; Record. M.T., Jr. *J. Phys. Chem.* **1986**, 90, 6541.
- (23) Wood, W.W.; Parker, F.R. *J. Chem. Phys.* **1957**, 27, 720.
- (24) Rossky, P.J.; Doll, J.D.; Friedman, H.L. *J. Chem. Phys.* **1978**, 69, 4628.
- (25) Cao, J; Berne, B.J. *J. Chem. Phys.* **1990**, 92, 1980.
- (26) Ermak, D.L. *J. Chem. Phys.* 1975, 62, 4189; **1975**, 62, 4197.
- (27) Rao, M.; Berne, B. J. *J. Chem. Phys.* **1979**, 71, 129.
- (28) Rao, M.; Pangali,D; Berne, B.J. *Mol. Phys.* **1979**, 37, 1773.
- (29) Pangali, C.; Rao, M.; Berene, B.J. *Chem. Phys. Lett.* **1978**, 55, 413.
- (30) Kotelyanskii, M.J.; Suter, U.W. *J. Chem. Phys.* **1992**, 96, 5383.
- (31) Bennett, C.H. *J. Comput. Phys.* **1976**, 22, 245.
- (32) Beveridge, D.L.; Dicapua, F.M. *Annu. Rev. Biophys. Chem.* **1989**, 18, 431.
- (33) Voter, A. F. *J. Chem. Phys.* **1985**, 82, 1890.
- (34) Torrie, G.M; Valleau, J.P. *J. Compt. Phys.* **1977**, 23, 187.
- (35) Zwanzig, R.W. *J. Chem. Phys.* **1954**, 22, 1420.
- (36) Lyubartsev, A.P.; Martsinovski, A. A.; Shevkunov, S.V.; Vorontsov-McDonald, P.N. *J. Chem. Phys.* **1992**, 96, 1776.
- (37) Adams, D.J. *Mol. Phys.* **1974**, 28, 1241.
- (38) Adams, D.J. *Mol. Phys.* **1975**, 29, 307.
- (39) Valleau, J.P.; Cohen, L.K. *J. Chem. Phys.* **1980**, 72, 5942.
- (40) Guggenheim, E.A. *Thermodynamics*, North-Holland Physics Publishing, New York, 1967.
- (41) Rasaiah, J.C; Freedman, H.L. *J. Chem. Phys.* **1972**, 56, 248.
- (42) Rasaiah, J.C. *J. Solution. Chem.* **1973**, 2, 301.
- (43) Friedman, D. *Statistical Mechanics*, Prentice-Hall, Inc, New Jersey, **1985**.
- (44) McMillan, W.G. and Mayer, J. E. *J. Chem. Phys.* **1945**, 13, 276.
- (45) McDonald, I.R. *Mol. Phys.* **1972**, 23, 41.

(46) Creutz, M. *Phys. Rev. Lett.* **1983**, 50, 1411.



**Correlation Function
of
Strong Electrolyte**

**Debye-Hückel Theory
(Low Concentration)**

Primitive Model :

- (1) Hard Sphere with Charge
- (2) Dielectric Constant of Solvent
- (3) Interaction
 - Coulomb Potential
 - Hard Sphere Repulsion

Figure 1: Connection of model, analytic solution and computer simulation

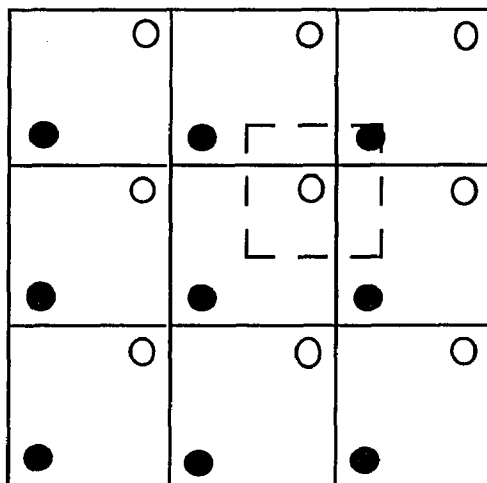


Figure 2: Periodic boundary condition and minimum image

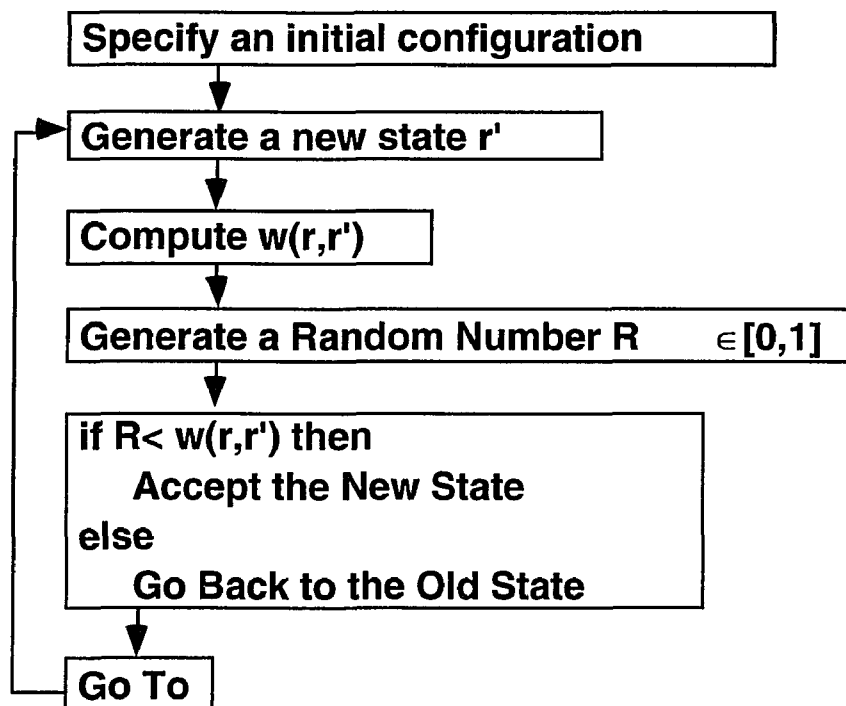


Figure 3: Flow chart of Metropolis Monte Carlo method

Chapter 2

A Monte Carlo Method to Simulate Systems with Barriers-- Subspace Sampling Method

Published in *J. Phys. Chem.* **1993**, 97, 13824

(Reprinted by Permission of the Journal of Physical Chemistry)

Abstract

We present the Subspace Sampling Method (SSM), a novel Monte Carlo methodology for simulating systems with large barriers. The Subspace Sampling Method partitions configuration space into subspaces. These subspaces are sampled using either random moves within subspaces or inducing transitions between subspaces to facilitate barrier crossing. Several adjustable parameters characterize SSM including the step size, the percentage of trial barrier crossings and the frequency of acceptance of barrier crossings. In order to examine the accuracy and efficacy of the method we compute the single-particle distribution function for a double well potential. The computed distribution function is independent of the adjustable parameters and converges to the exact function. For the one-dimensional potential, the Subspace Sampling Method provides an accurate and efficient method for simulating barrier crossing.

Introduction

The Metropolis Monte Carlo method is exceedingly useful to explore the thermodynamic and structural properties of dense systems (both homogeneous and heterogeneous) where the interparticle potential functions contain no discontinuities or barriers (1-3). In systems with barriers or discontinuities, such as weak electrolyte solutions or systems involving chemical reactions, the Metropolis Monte Carlo method fails to efficiently sample configuration space and consequently fails to converge to accurate equilibrium properties. Alternative Monte Carlo methodologies exist to simulate solutions with strong specific interactions, many accessible configurations or large barriers (4-9). These methods use alternative sampling functions that force the Markov chain to sample regions of configuration space that are rarely sampled in the Metropolis scheme.

We propose another method, the Subspace Sampling Method (SSM). SSM is motivated in part by the free energy difference methods of Bennett and Voter (11,12). Free energy differences are evaluated from the ratio of the classical partition functions representing two systems with different Hamiltonians. This ratio is obtained from the ensemble averages of the potential energy differences of the two systems. A switching function is used to compute potential energy differences from a Metropolis walk that only includes transitions from one system to the other. In the subspace sampling method we define subspaces from the classical potential energy function and then choose either to sample the subspace or to induce transitions among the subspaces. Using the subspace sampling method, we obtain an algorithm that can be used to compute distributions as well as estimate the free energy differences between the subspaces. With the proper choice of the subspaces, SSM can be readily applied to systems with large barriers or discontinuities.

In this paper, we discuss the method and its application to the simulation of a single particle in a symmetric, double-well potential with a large barrier. We compare our results and methodology directly with published work using the force-bias and antforce-bias methods (4,8). We show that the subspace sampling method is comparable to the antforce-bias method. We then introduce modifications into the method that enhance barrier crossing and produce more efficient sampling of subspaces.

Methodology

The basis of the Subspace Sampling Method is the division of configuration space into different regions (subspaces) depending upon the explicit form of the potential function. We then distinguish between transitions due to random moves within one subspace and transitions that result from moves among subspaces. Transition probabilities for random moves within subspaces follow the standard Metropolis sampling method. Transition probabilities for the transfer from one subspace to another are derived below.

We call transitions within one subspace a "move" and transitions between subspaces "transfers".

A transition from one subspace (space A, state i) to another subspace (space B, state j) occurs by 1) selecting the transfer process with probability P_{AB} , 2) selecting a random position in the subspace B with probability $1/\delta_B$, and 3) accepting the transition with probability W_{ij} . Consequently, the probability of finding a particle in subspace B, state j, at a time t, if the particle was initially in subspace A, state i is:

$$P_{AiBj} = P_{AB} \frac{1}{\delta_B} W_{ij} \quad (1)$$

Similarly, the probability of a transition to the subspace A from subspace B is given by:

$$P_{BjAi} = P_{BA} \frac{1}{\delta_A} W_{ji} \quad (2)$$

At equilibrium, the condition of microscopic reversibility gives:

$$\exp(-E_i/kT) P_{AiBj} = \exp(-E_j/kT) P_{BjAi} \quad (3)$$

and therefore

$$\frac{W_{ij}}{W_{ji}} = \exp(-(E_j - E_i)/kT) \frac{\delta_B}{\delta_A} \frac{P_{BA}}{P_{AB}} \quad (4)$$

In our application, we choose $P_{BA} = P_{AB}$. In addition, the *a priori* probability of locating the molecule at a position in the one subspace is obtained from the volume of the Monte Carlo cell occupied by that subspace, i.e.

$$\delta_B = V_B \text{ and } \delta_A = V_A \quad (5)$$

The transition probabilities W_{ij} and the algorithm we use are as follows.

Transfer from subspace A to B: We select the subspace transition with probability P_{AB} and randomly choose a particle from subspace A. Then we pick a trial position randomly from subspace B. The probability of accepting the transfer is

$$W_{ij} = c \frac{V_B}{V} \quad \text{if } E_j < E_i \quad (6)$$

or

$$W_{ij} = c \frac{V_B}{V} \exp(-(E_j - E_i)/kT) \quad \text{if } E_j > E_i$$

where c is a constant. The only constraint on the choice of c is that the product $c \frac{V_B}{V} \leq 1$.

Transfer from subspace B to A: This transfer is identical to the transfer from subspace A to B with the roles of A and B reversed.

Random Move: We select the random move process with probability P , choose one particle and attempt a trial move within a distance Δ . If the trial position is outside the selected subspace, the particle is returned to the subspace using periodic boundary conditions. The probability of accepting the trial move is in accordance with the Metropolis transition probability.

Results and Discussion

Double-Well potential

We test the accuracy and convergence of this algorithm by examining the single particle distribution function for a double-well potential. The potential function used is:

$$V(x) = \frac{(x^2 - a^2)^2}{2a^2} \quad |x| \leq L/2 \quad (7)$$

and

$$= \infty \quad |x| \geq L/2$$

and all parameters and units are chosen to conform to the simulation of Cao and Berne, viz. $L=12$, $a=4.0$, $m=1.0$, and $\beta=1$ (8). Our potential function differs somewhat from that of Cao and Berne's in that we apply a cutoff at L and have a finite length Monte Carlo cell. As shown in Figure 1, we divide configuration space into two subspaces A and B, corresponding to the two wells. A schematic of the algorithm we use is shown in Figure 2. We first decide a) to implement a random move of a particle while remaining in the subspace of the particle, or b) to transfer a particle from subspace A to subspace B, or c) to transfer a particle from subspace B to subspace A. The transition probabilities for this system are as follows:

If the particle is resident in subspace A, then the probability of a transition from subspace A, state i to subspace B, state j is:

$$P_{AiBj} = \begin{cases} P_{AB} c \frac{V_B}{V} & \text{if } E_{Bj} < E_{Ai} \\ P_{AB} c \frac{V_B}{V} \exp(-(E_{Bj}-E_{Ai})/kT) & \text{if } E_{Bj} > E_{Ai} \end{cases} \quad (8)$$

If the particle is not resident in subspace A, then

$$P_{AiBj} = 0 \quad (9)$$

The probability of a random move within subspace A is

$$P_{AiAj} = \begin{cases} P & \text{if } E_{Aj} < E_{Ai} \\ P \exp(-(E_{Aj}-E_{Ai})/kT) & \text{if } E_{Aj} > E_{Ai} \end{cases} \quad (10)$$

where we apply periodic boundary conditions at the subspace boundary.

To complete the transitions probabilities we have

$$P_{AiAi} = 1 - \sum_{i \neq j} P_{AiAj} - \sum_i P_{AiBj} \quad (11)$$

Convergence Criteria

We follow convergence by examining the fluctuation of the single particle density distribution in the manner suggested by Cao and Berne (8). Following m transitions, the cumulative average in the fluctuation of the distribution is

$$\chi^2(m) = \int [\rho(x,m) - \rho_{\text{exact}}(x)]^2 dx \quad (12)$$

To calculate $\rho(x,m)$ we construct a one-dimensional histogram and count the number of times a particle falls within a particular division of the histogram. We chose the grid defining our histogram to be 12/100 units. $\rho(x,m)$ is given by:

$$\rho(x,m)\Delta x = \sum_{i=1}^{100} \frac{N_i}{N_T} h_i \quad (13)$$

where N_i is the number of times a particle is found in the i th slot, N_T is the total number of moves of the simulation and Δx is the grid size. h_i is a step function, i.e., $h_i = 1$ if x is in the i th histogram slot, otherwise $h_i = 0$.

The evaluation of $\chi^2(m)$ is obtained from:

$$\chi^2(m) = \sum_{i=1}^{100} \left(\frac{N_i}{N_T} - \rho_{\text{exact}}(x_i)\Delta x \right)^2 \frac{1}{\Delta x} \quad (14)$$

where

$$\rho_{\text{exact}}(x_i)\Delta x = \sum_{i=1}^{100} \frac{100 \exp\left(\frac{-V(x_{\text{mid},i})}{kT}\right) h_i}{\int_{-6}^6 \exp\left(\frac{-V(x)}{kT}\right) dx} \Delta x \quad (15)$$

and $x_{\text{mid},i}$ is the midpoint of Δx corresponding to the i th slot.

Number of Parameters

The Subspace Sampling algorithm contains three parameters: the step size Δ , the probability of choosing a random move P , and the constant c in the acceptance probability. We chose $P_{AB}=P_{BA}$; consequently $P_{AB}=(1-P)/2$. However, as seen from the transition probabilities (equation 8), P and c are not independent. Rather, P and c influence the transition probability through the product $c\frac{(1-P)}{2}$.

Accuracy of the Simulation Method

We evaluate the accuracy of the simulation method by computing the single particle distribution function for a variety of values for Δ , P and c . The parameters used in a representative set of simulations and the percentage of the trial and accepted moves and transfers for the three possible transitions are shown in Table 1. In Figures 3a and 3b we show representative distribution functions corresponding to two sets of parameters. In both cases, the computed distribution function is equivalent to the exact function following 10^5 moves ($\chi^2(10^5)=10^{-4}$). In all the simulations, the computed distribution function converges to the exact function, although the rate of convergence depends upon the parameter set.

Convergence Rate of the Distribution Function

Although the computed distribution function is independent of the choice of the parameters, the convergence rate of the distribution can be a sensitive function of the parameters. In Figures 3a and 3b, the computed distribution function is shown following 10^4 , 4×10^4 and 10^5 moves. For the parameter set $c=1.0$, $\Delta=3/2$ and $P=0.7$ (Figure 3a), $\rho(x,m)\Delta x$ (hereafter referred to simply as $\rho(x)$) converges rapidly. However, when c is reduced to 0.2 (Δ and P remain fixed), the exact distribution function is not reproduced until after 10^5 moves.

In order to compare systematically the convergence rate of the distribution function for a variety of parameter sets, we compute the fluctuation in the distribution according to equation 15. In Figure 4 we show $\chi(m)$ as a function of step size for fixed c (1.0) and for $P=0.1$ (A), 0.5 (B) and 0.9 (C). We restrict our range of step sizes to $\Delta \leq 3/2$ in order to maintain the percentage of accepted trial moves to total moves above 50%. The rate of convergence of $\rho(x)$ depends upon both the step size and the choice of P . When P is small ($P=0.1, 0.5$), convergence rate of $\rho(x)$ is fairly insensitive to step size. For small values of P , barrier crossings are attempted frequently. Consequently, configuration space is sampled primarily through the transitions, which are independent of step size. As P increases, random moves within one subspace are attempted more frequently and the step size begins to influence the convergence rate. At $P=0.5$, the simulation using a small step size ($\Delta=3/32$) converges slowly. As P further increases ($P=0.9$), the convergence rate becomes a sensitive function of step size. At higher P , rapid convergence is observed when the fraction of accepted moves to total trial moves is in the range 50-80% (see Table 1). When the fraction of accepted moves is close to 100% (as occurs with $\Delta=3/32$), sampling of configuration space is inefficient resulting in a slowly converging distribution. In our simulations, the most rapidly converging $\rho(x)$ occurs for $3/8 < \Delta < 3/2$, which corresponds to a fraction of accepted moves/total moves of 50-85%.

In Figure 5 we show $\chi(m)$ for a fixed step size and variable P . For any choice of Δ , the convergence rate depends upon the choice P . In simulations with large step sizes, the most rapidly converging distribution occurs for large P . At lower values of P , one samples the transfers more often than the corresponding random moves. To speed convergence, an efficient random sampling of the subspace on each side of the barrier is necessary. Excessive sampling of barrier crossings actually impedes the rate of convergence. However, in simulations with very small step sizes ($3/32$), the slowest converging $\rho(x)$ occurs for large P . In these simulations, barrier crossing compensates for the small step size and facilitates the random sampling of the states on both sides of the barrier. However,

for all values of P , convergence is slower for the small step sizes ($3/32$) as compared with simulations using larger values of Δ . Barrier crossing cannot compensate entirely for inefficient sampling of the individual subspaces.

The choice of c dramatically affects the convergence of $\rho(x)$ when P is large while influencing the convergence rate only minimally when P is low. Figures 6a and 6b show similar convergence rates even though c is varied from 0.2 to 2.0. Although low values of P correspond to simulations dominated by barrier crossings, the fraction of accepted barrier crossings (as determined by c) does not significantly influence the rate of convergence of $\rho(x)$. However, when $P=0.9$, transfers are rarely attempted and the acceptance probability for transfers significantly influences the rate of convergence of $\rho(x)$. When $P=0.9$, only 10% of the trial transitions involve barrier crossings. Furthermore, when $c=0.2$ only 1% of the trial barrier crossings are accepted. This low value results in such few barrier crossings that significant density builds in one subspace rendering $\chi(m)$ larger (see Figure 3b). However, when $c=1.0$, 7% of the trial barrier crossings are accepted. This percentage is sufficiently large to insure that all subspaces are sampled frequently enough to prevent biasing of one subspace and consequently a slow convergence of $\rho(x)$.

Enhanced Method

The optimum choice of parameters corresponds to an enhanced, but equal, sampling of the subspaces. In order to minimize ineffective barrier crossings, we modified the algorithm to include a random move whenever a transfer across the barrier is disallowed (see schematic, Figure 7). As expected, this Enhanced Method improves convergence somewhat and the rate of convergence is, in general, less sensitive to the choice of the parameters than the Subspace Sampling Method (see Figure 8). However, when optimal parameters were chosen, the Subspace Sampling Method converged as quickly as the enhanced scheme.

Smart Monte Carlo with Enhanced SSM

For transitions involving random moves, we replace the standard Metropolis method with Smart Monte Carlo (7). Barrier crossings are selected as in the Enhanced Scheme. Figure 9 demonstrates that the convergence rate of $\rho(x)$ is less sensitive to the parameter set than the comparable rate using the unmodified Subspace Sampling Method. The most rapid convergence rate of $\rho(x)$ is observed with a $\Delta=3/8$ and for all values of P less than 0.9. With the larger value of P , too few barrier crossings are attempted and the convergence rate is slower as in the unmodified SSM. The convergence rate displays a slight sensitivity to the step sizes chosen. However, even with variations in the step size over the same range as chosen previously, the convergence rate of $\rho(x)$ is comparable to or better than the corresponding rate obtained from the unmodified SSM.

Comparison with other techniques

Cao and Berne compared the rate of convergence of $\rho(x)$ for one particle in the double well using the no-bias, force-bias and anti-force bias with and without variable step size (see Figures 1-3, ref 8). Following 10^5 moves, the no-bias method (the standard Metropolis Monte Carlo method, with a fixed step size) and the force-bias method converged very slowly. The anti-force bias method with and without a variable-step size, and the no-bias method using a variable step size converged rapidly ($\chi(10^5) < 0.025$). This convergence is similar to that using the Subspace Sampling Methods and slightly slower than the Enhanced SSM/Smart Monte Carlo. In general, rapid convergence is facilitated by efficient random sampling within each subspace and independent barrier crossings. The anti-force bias, variable step-size methods, as well as the Subspace Sampling Methods, fulfill these criteria.

Conclusions

These one-dimensional, single particle simulations demonstrate that the Subspace Sampling Methods are capable of reproducing the exact partition function for the double well potential. These methods require three adjustable parameters: the step size (Δ), the trial probability (P) and the multiplicative constant (c). Although the single particle distribution function is independent of the choice of parameter set, the convergence rate of the distribution function can be sensitive to the choice of P and c . This sensitivity is reduced when the Enhanced Method/Smart MC algorithm is used. In general, rapid convergence is achieved when two criteria are met. The first criterion requires efficient sampling of the subspaces. This can be optimized by varying the step size and choosing the Δ that corresponds to a percentage of accepted moves to total attempted moves in the range of 50-90%. The second criterion requires a reasonable frequency of successful barrier crossings. The frequency of accepted barrier crossings (expressed as a percentage of accepted crossing/total transitions) can be as low as 5% to insure an even sampling of the subspaces.

The Subspace Sampling Method converges as rapidly as the anti-force-bias and variable-step Metropolis method when the optimum set of parameters is chosen. The advantage of SSM over the anti-force bias method is its applicability to systems containing barriers as well as discontinuities. We are examining the applicability and utility of the Subspace Sampling Method to more complex systems containing many particles, three dimensions and asymmetric potential functions.

References

- (1) Metropolis, N.; Rosenbluth, A.W.; Rosenbluth, M.N.; Teller, A.H.; Teller, E. *J. Chem. Phys.* **1953**, 21, 1087.
- (2) Berne, B.J.; ed. *Statistical Mechanics, Part A: Equilibrium Techniques*, Plenum Press, New York, **1977**.
- (3) Ciccotti, G.; Frenkel, D.; McDonald, I.R.; eds. *Simulation of Liquids and Solids, Molecular Dynamics and Monte Carlo Methods in Statistical Mechanics* North-Holland Physics Publishing, **1987**.
- (4) Pangali, C.; Rao, M.; Berne, B.J. *Chem. Phys. Lett.* **1978**, 3, 413.
- (5) Rao, M.; Berne, B. *J. Chem. Phys.* **1979**, 71, 129.
- (6) Rao, M.; Pangali, C.; Berne, B.J. *Mol. Phys.* **1979**, 37, 1773.
- (7) Rossky, P.J.; Doll, J.D.; Friedman, H.L. *J. Chem. Phys.* **1978**, 69, 4628.
- (8) Cao, J. ; Berne, B.J. *J. Chem. Phys.* **1990**, 92, 1980.
- (9) Kotelyanskii, M.J.; Suter, U. W. *J. Chem. Phys.* **1992**, 96, 5383.
- (10) Bennett, C.H. *J. Computational Physics* **1976**, 22, 245.
- (11) Voter, A.F. *J. Chem. Phys.* **1985**, 82, 1890.

Figure Captions

Figure 1--The double well potential used in the simulations. The potential is used to define the subspaces as shown.

Figure 2--The flowchart used to implement the Subspace Sampling Method.

Figure 3--Comparison of the convergence of the single particle distribution function with the exact function. The solid line corresponds to the exact distribution. The symbols refer to the computed distribution following: $\circ$: 10^4 ; Δ : 4×10^4 ; $+$: 10^5 moves.

Figure 4--Convergence of $\rho(x)\Delta x$ for variable Δ corresponding to three choices of P . In all figures $c=1$. The step sizes plotted are: $\circ$: $3/2$; Δ : $3/4$; $+$: $3/8$; $\diamond$: $3/16$; $\square$: $3/32$.

Figure 5--Convergence of $\rho(x)\Delta x$ for variable P corresponding to three choices of step size. In all figures $c=1$. The P values plotted are: $\circ$: 0.0 ; $\circ$: 0.1 ; Δ : 0.3 ; $+$: 0.5 ; $\diamond$: 0.7 ; $\square$: 0.9 .

Figure 6--Convergence of $\rho(x)\Delta x$ for variable c corresponding to three choices of P . In all figures $\Delta=3/8$. The c values plotted are: $+$: 0.2 ; $\circ$: 1.0 ; Δ : 2.0 .

Figure 7--The flowchart used to implement the Enhanced Subspace Sampling Algorithm. The cells bordered by the double box are the modifications of the Subspace Sampling Method.

Figure 8--Convergence of $\rho(x)\Delta(x)$ using the Enhanced Subspace Sampling Method as a Function of Step Size for three values of P . The step sizes show are: Δ : $3/2$; $+$: $3/4$; $\circ$: $3/8$; $\diamond$: $3/16$.

Figure 9--Convergence of $\rho(x)\Delta(x)$ using the Smart Monte Carlo/Enhanced Subspace Method as a function of P for selected choices of Δ and c . The values of P are: $\circ$: 0.1 ; Δ : 0.3 ; $+$: 0.5 ; $\diamond$: 0.7 ; $\square$: 0.9 .

Table 1. The parameters for a representative set of simulations using the Subspace Sampling Method is tabulated. The percentage of transitions that were trial and attempted transfers from subspace B to subspace A, transfers from A to B and moves are shown.

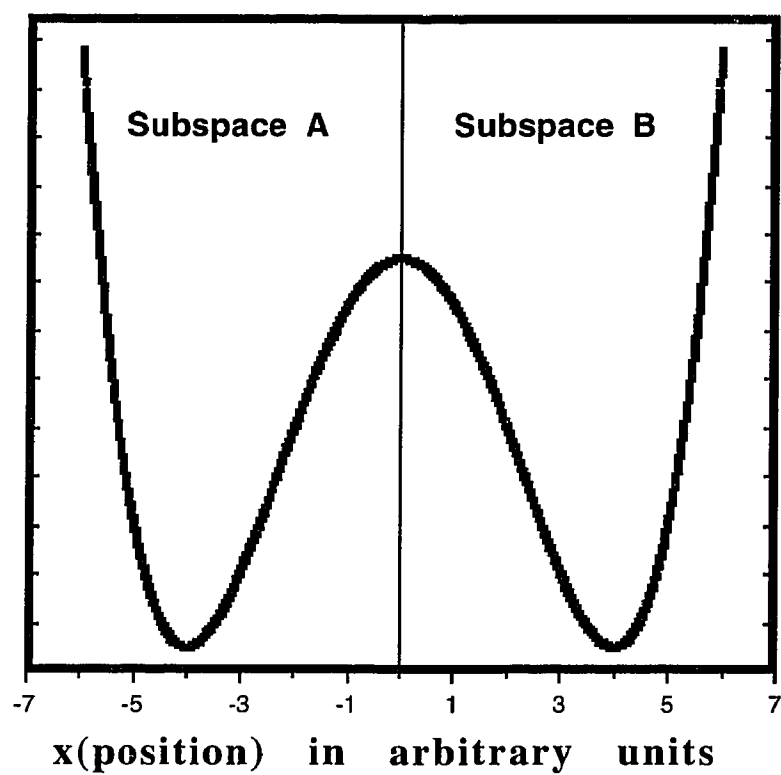


Figure 1

Flow Chart For the Subspace Sampling Algorithm

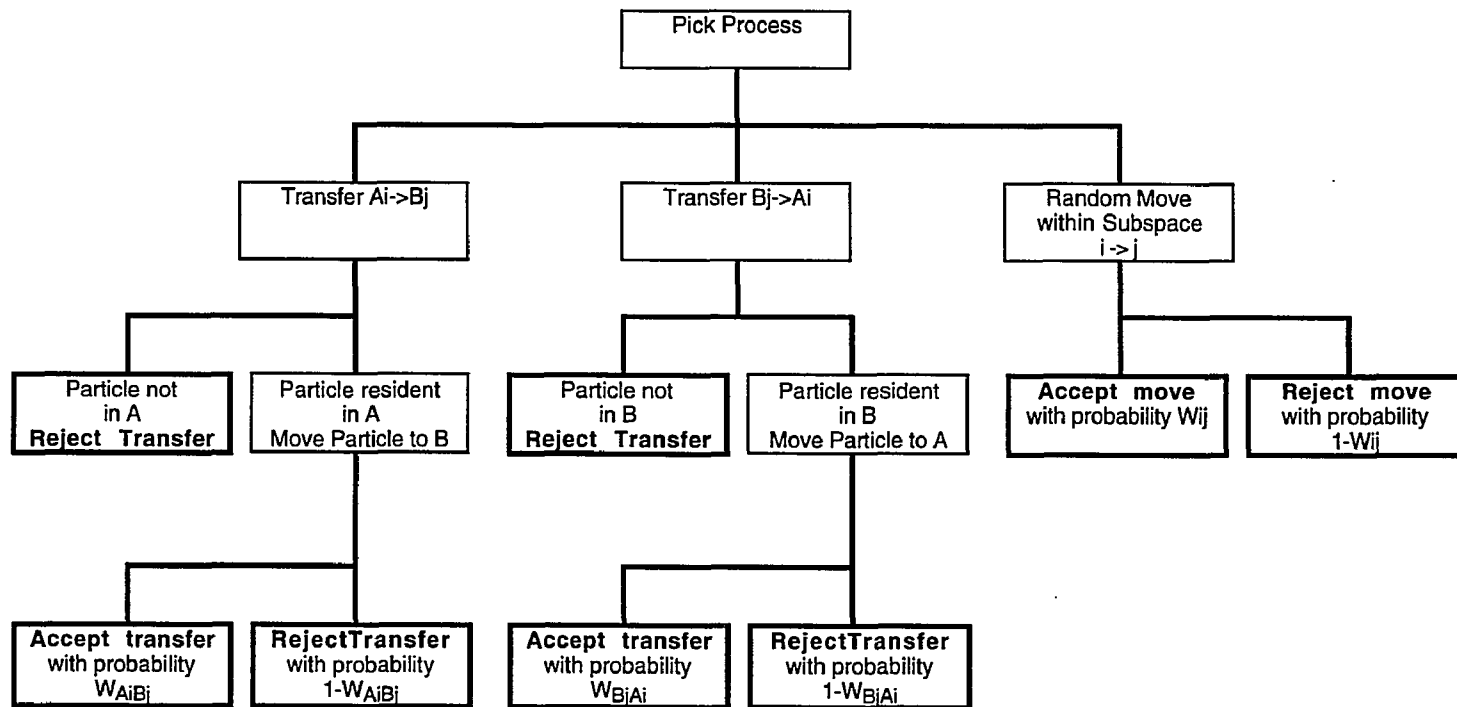


Figure 2

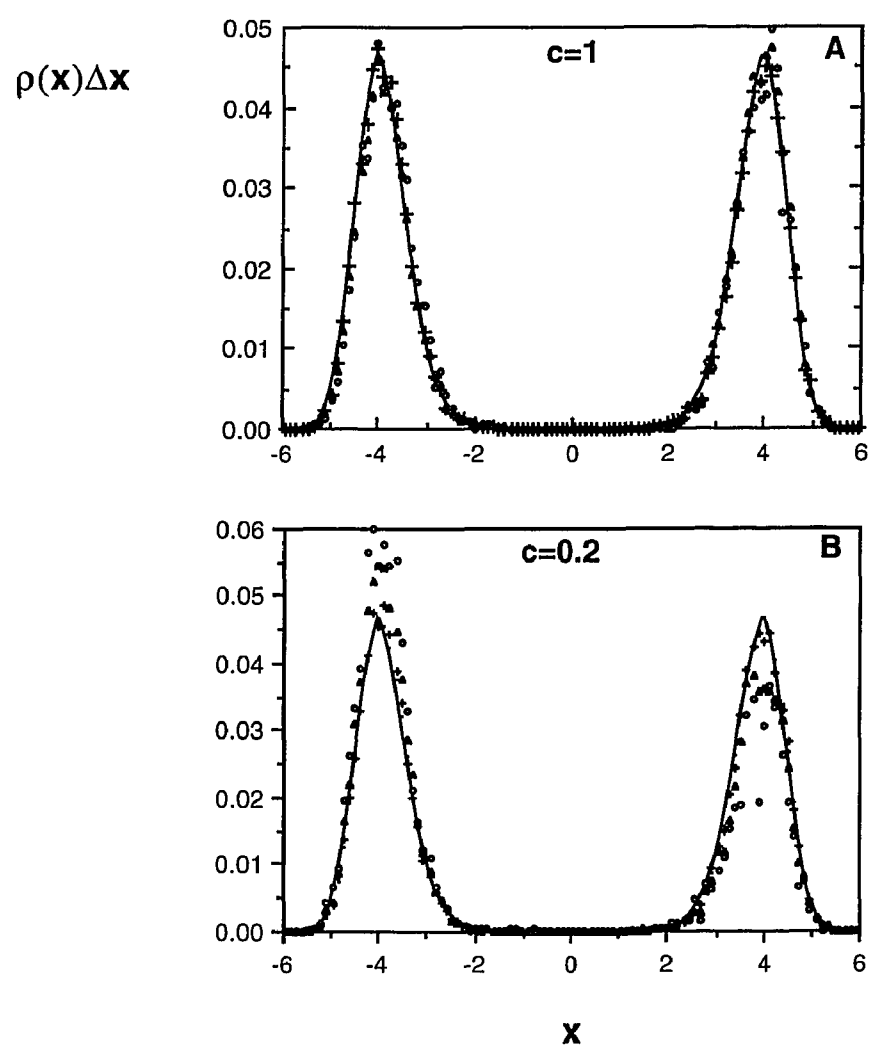


Figure 3

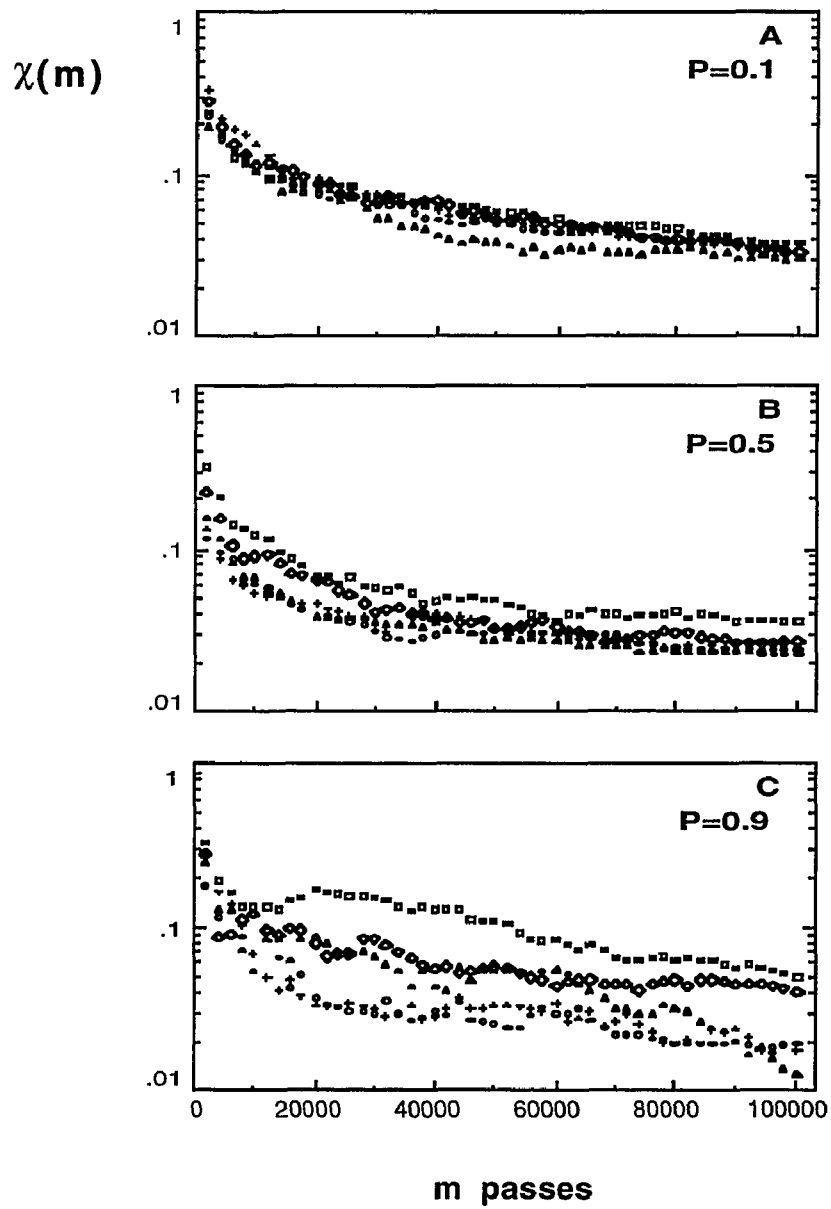


Figure 4

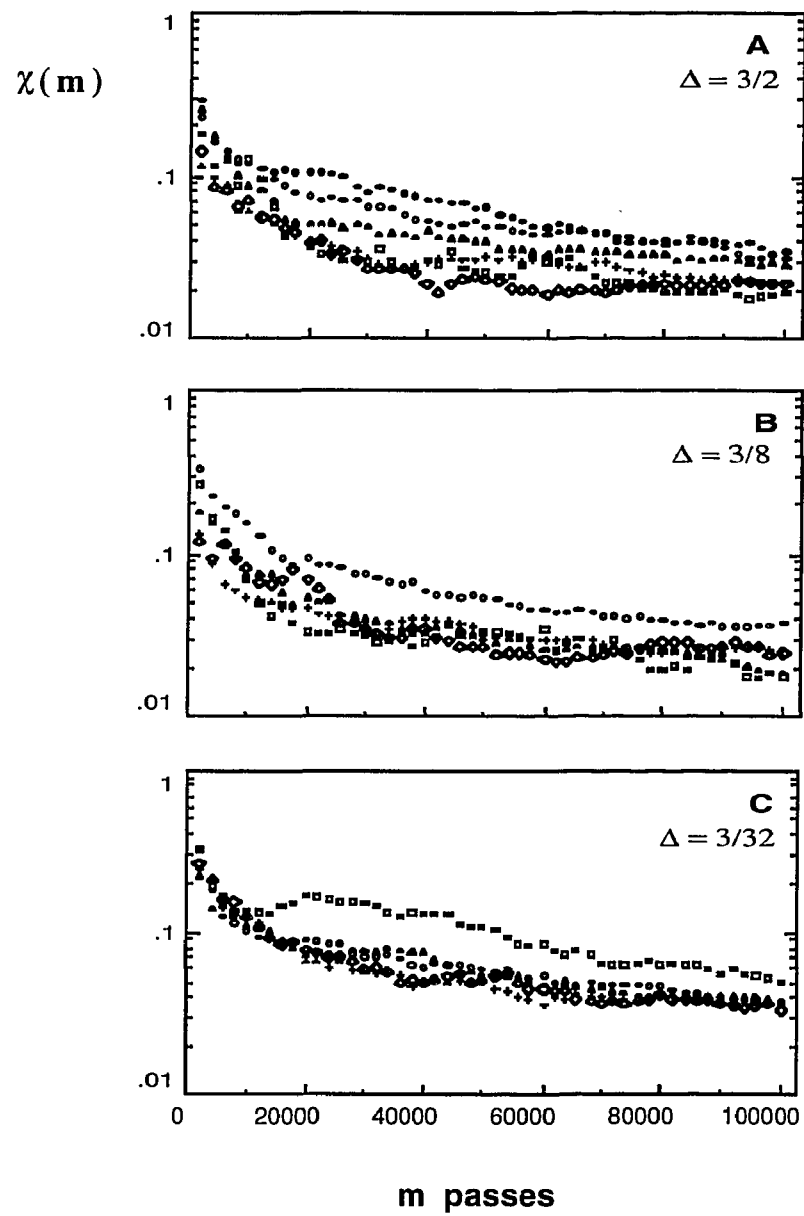


Figure 5

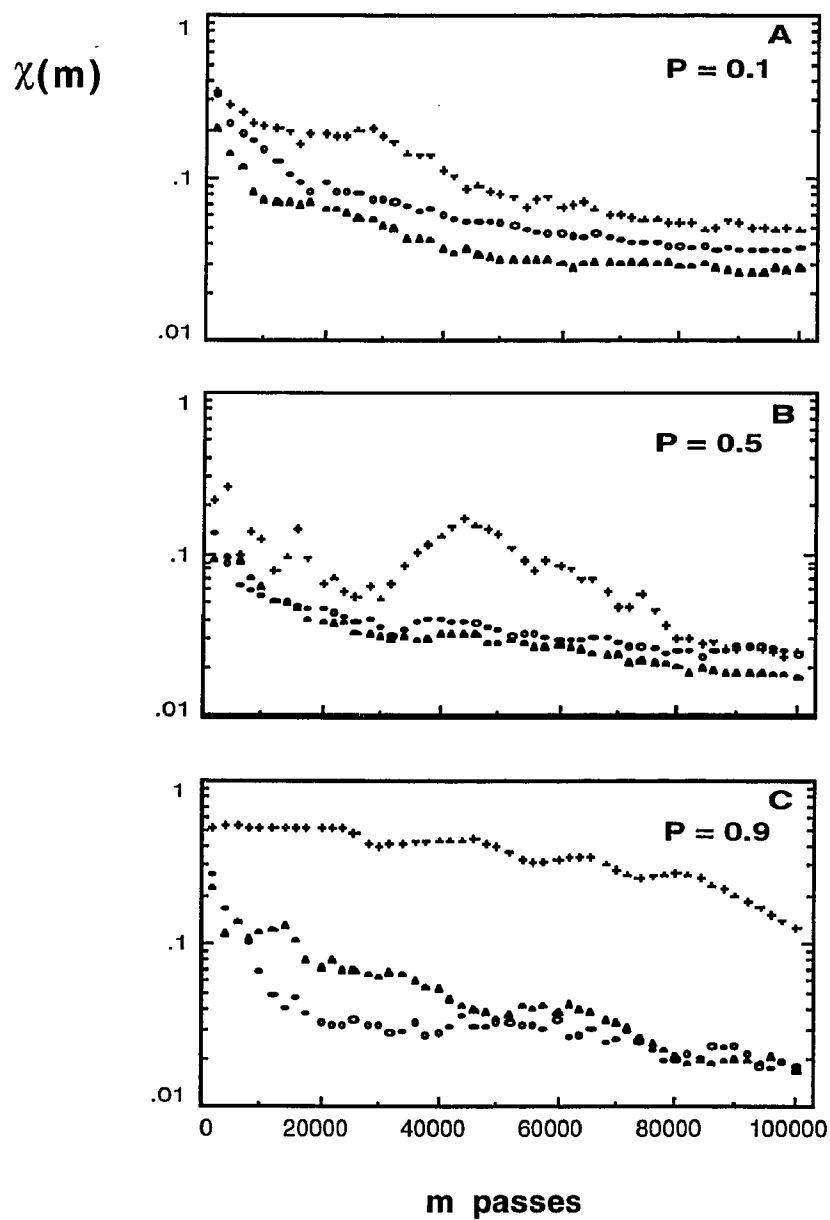


Figure 6

Flow Chart For the Enhanced Subspace Sampling Algorithm

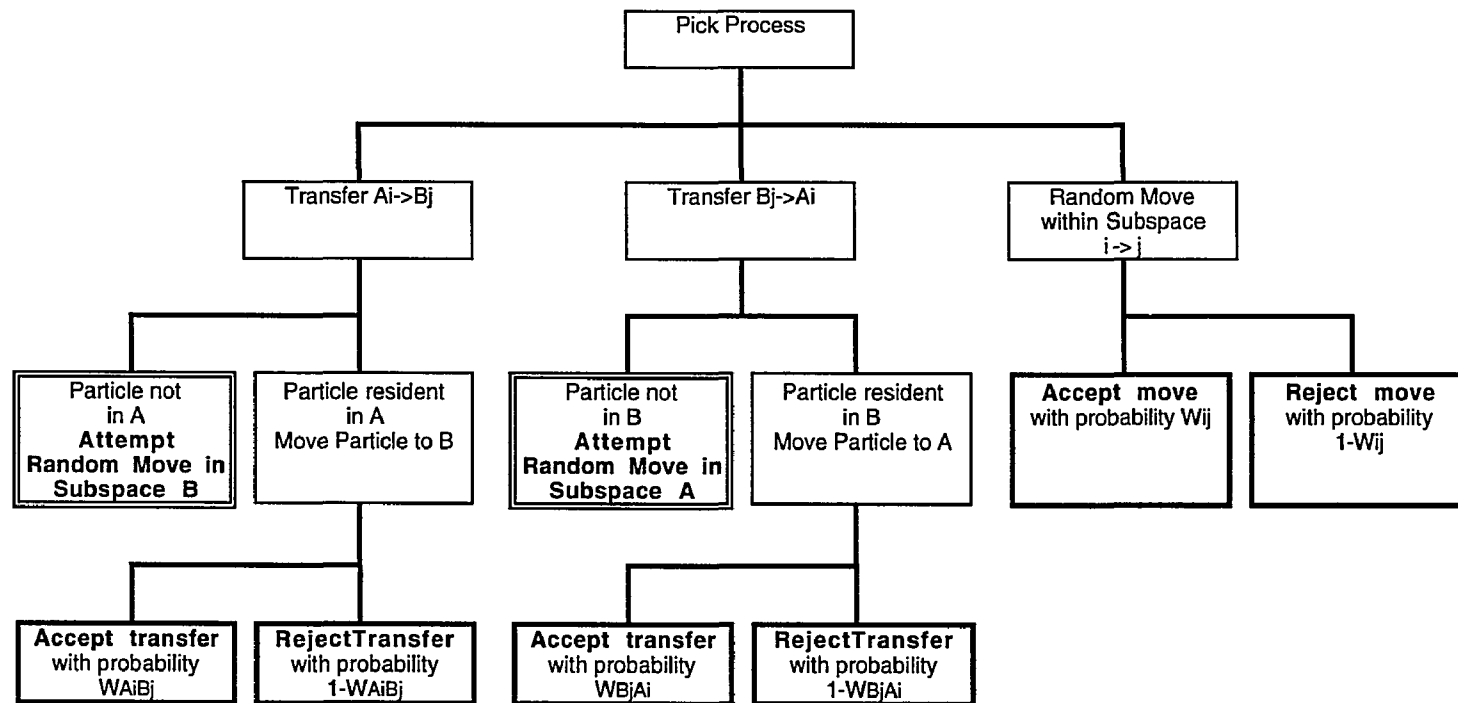


Figure 7

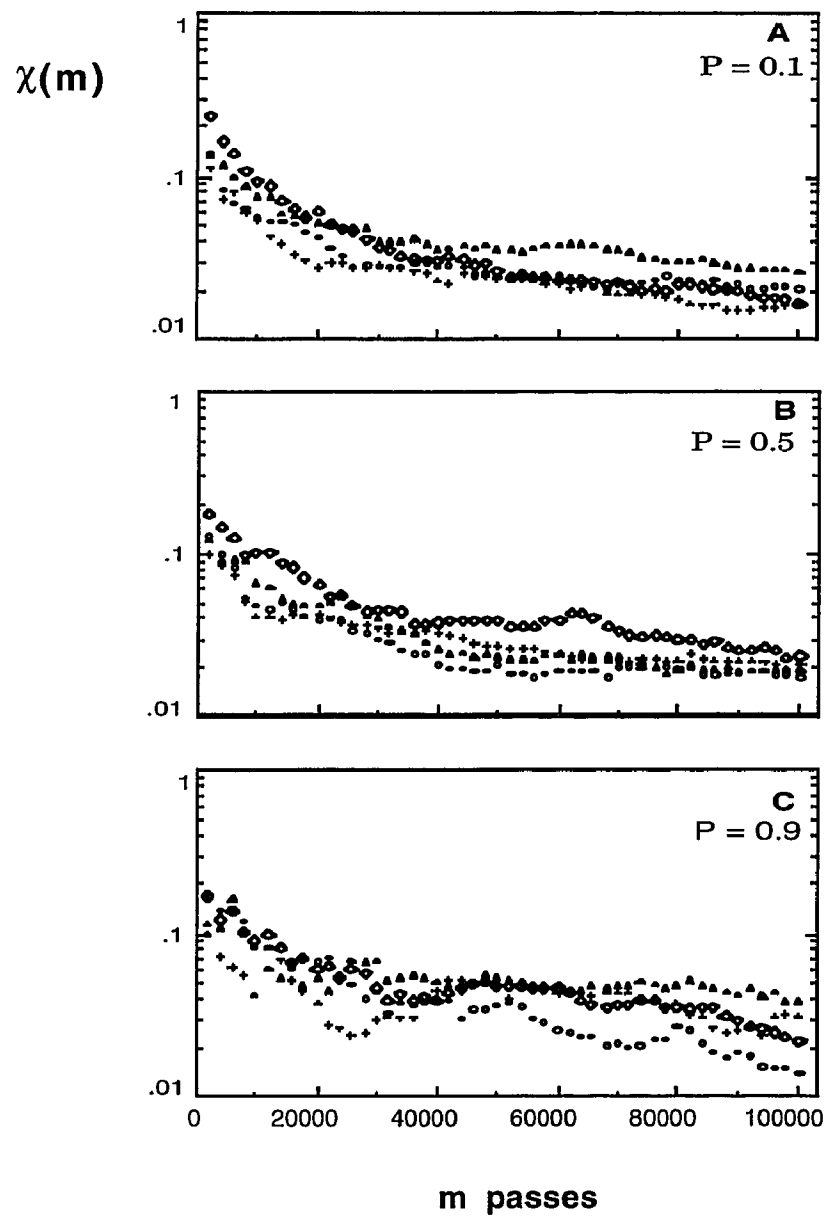


Figure 8

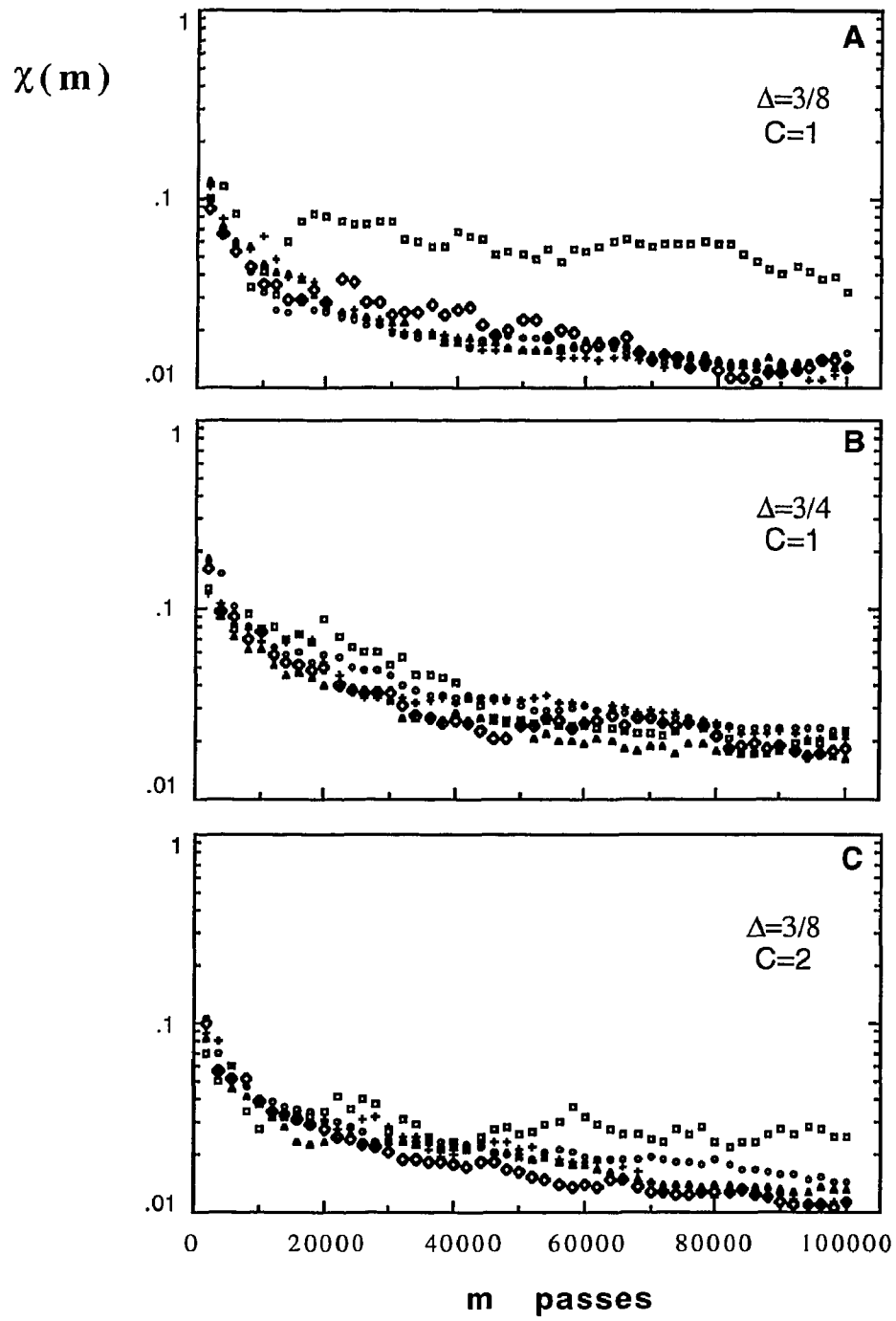


Figure 9

Table 1. Parameters, Trial and Acceptance Probabilities for a Selected Set of Simulations Using the Subspace Sampling Method

Parameters			Percentage of Trial and Accepted Transfers and Moves					
Step Size	c	P	Trial Transfer of B to A	Accepted Transfer of B to A	Trial Transfer of A to B	Accepted Transfer of A to B	Trial Random Moves	Accepted Random Moves
3/2	1.	0.0	49.94	7.08	50.06	7.06	0.00	0.00
3/2	1.	0.1	44.72	6.96	45.21	6.89	10.07	50.62
3/2	1.	0.5	25.23	6.97	25.06	7.01	49.70	50.25
3/2	1.	0.9	4.97	6.93	5.04	6.81	90.00	50.64
3/2	0.2	0.7	15.05	1.40	14.93	1.41	70.02	50.22
3/8	1.	0.1	44.97	6.97	45.23	6.93	9.80	85.63
3/8	1.	0.5	25.11	7.04	25.09	7.05	49.80	85.58
3/8	1.	0.9	4.99	6.98	5.04	6.90	89.97	85.74
3/32	1.	0.1	45.09	7.14	44.99	7.15	9.93	96.63
3/32	1.	0.5	25.07	6.82	24.91	6.86	50.03	96.51
3/32	1.	0.9	5.00	6.82	5.01	6.78	89.99	96.49
3/8	2.	0.1	44.88	14.10	45.16	14.01	9.96	85.39
3/8	2.	0.5	25.10	14.17	24.86	14.30	50.04	85.61
3/8	2.	0.9	5.02	14.05	4.99	14.14	90.00	85.64
3/16	2.	0.1	45.23	14.14	44.80	14.28	9.98	92.68
3/16	2.	0.5	25.19	14.08	24.98	14.19	49.83	92.92
3/16	2.	0.9	4.97	13.98	4.99	13.93	90.04	92.88
3/8	1.5	0.0	49.72	10.42	50.28	10.30	0.00	0.00
3/8	1.5	0.1	44.99	10.69	44.91	10.71	10.10	86.04
3/8	1.5	0.5	25.18	10.48	24.99	10.55	49.83	85.48
3/8	1.5	0.9	5.03	10.45	4.99	10.53	89.98	85.45
3/8	0.02	0.1	45.10	0.15	44.94	0.14	9.96	85.42
3/8	0.02	0.5	24.99	0.14	25.05	0.14	49.96	85.41
3/8	0.02	0.9	5.03	0.10	5.07	0.08	89.89	85.77

Chapter 3

The Subspace Sampling Method--

**A Monte Carlo Approach for Simulating the Single Particle Density
Function and the Equilibrium Constant for Systems Described by Multiple
Hamiltonians**

Abstract

We compute the single particle density function and the equilibrium constant using the Subspace Sampling Method (SSM) for three one-dimensional potentials. SSM is a versatile Monte Carlo methodology that partitions configuration space into subspaces. Such partitioning facilitates the sampling of regions of configuration space separated by potential barriers or described by different Hamiltonians. We show that SSM can be readily used to compute simultaneously the density function and the equilibrium constant for the subspaces. Furthermore, alternative sampling schemes can be included in the SSM methodology to enhance convergence and minimize the need to search for an optimum parameter set.

Introduction

The well-known difficulties of the standard Metropolis method to model accurately systems with barriers have motivated a variety of alternative methodologies (1-10). We recently introduced the Subspace Sampling Method in order to simulate systems with discontinuous potentials, and large barriers (1). The Subspace Sampling Method (SSM) is an approach that partitions the potential function into subspaces. Mechanical variables, as well as the relative occupancy of the subspaces, can be simulated by choosing between two types of transitions: moves within the subspace and transfers among the subspaces. SSM is a versatile methodology into which any transition probabilities for the random sampling of the individual subspaces can be incorporated. In addition, large barriers can be overcome by the forced transfer across the potential barrier thereby enhancing sampling of the entire configuration space.

We tested the accuracy and efficiency of our methodology by computing the single particle distribution function in a double-well potential (1). We compared our results with those using the anti-force bias and variable step Monte Carlo methods (5). These latter

methodologies belong to a general class of alternative sampling schemes that modify the random walk by biasing the trial displacements according to the potential function (5-10). Originally, biasing methodologies selected displacements in the direction of the force in order to sample the low energy states more efficiently (6-8). The anti-force-bias approach selects some displacements against the force to increase the probability of crossing a potential barrier. The variable step Monte Carlo method alters the step size to sample inaccessible regions of configuration space (5). We showed that the Subspace Sampling Method, using the standard Metropolis walk for the subspaces, was, at least, comparable to the anti-force-bias and variable step size methods. For some SSM parameter sets, the convergence rate of the single particle distribution function was even more rapid (1).

The Subspace Sampling Method appeared to be a promising method for simulating systems with barriers. However, SSM requires a choice of three parameters that must be optimized to achieve a convergence rate comparable to other methods. Consequently, we sought to exploit the versatility of the method to determine if the incorporation of other sampling schemes could both speed convergence and minimize the need to search for optimum parameters.

A significant feature of the Subspace Sampling Method is the partitioning of configuration space into subspaces. This suggests that SSM can be used to compute the equilibrium constant or relative occupancy of two regions of configuration space represented by different Hamiltonians. Thus, SSM belongs to a second class of sampling methodologies where the Metropolis walk is coupled to different sampling distributions (J-walking) (11,12) or a different system (free energy difference) (13,14). The J-walking (jump-walking) method of Frantz, et al. introduces trial displacements chosen from a distribution at a higher temperature interspersed among the standard Metropolis walk (11,12). The increase in temperature facilitates trial "jumps" across energy barriers. Frantz, et al. demonstrated that J-walking was an efficient algorithm to ascend the one-dimensional

double-well potential. The free energy difference methods of Bennett and Voter were designed primarily to obtain Helmholtz free energy differences between two canonical ensembles represented by different Hamiltonians (13,14). The original development defined Hamiltonians that spanned the same configuration space (13). Consequently, one ensemble can act as a reference state and free energies relative to the reference state can be computed within the canonical ensemble. The extension of this method to systems containing overlapping or non-overlapping regions of configuration space provides a general method to compute equilibrium constants and transition state theory rate constants within the canonical ensemble (14).

In this paper, we apply SSM to a variety of single particle potential functions containing barriers, deep potential wells and discontinuous potentials. In the first section, we review the Subspace Sampling Methodology. We then compute the single particle distribution function ($p(x)$) and the equilibrium constant (K) for two potentials containing barriers and discontinuities. We show that SSM simultaneously and efficiently reproduces both the distribution function and the equilibrium constant. However, the convergence rates demonstrate a sensitivity to the choice of parameter set. Therefore, we examine the sensitivity of the parameter set to convergence rates of both $p(x)$ and K using a harmonic-hyperbolic potential for three SSM methodologies: 1) the single-particle enhanced method, 2) SSM with J-walking and 3) SSM with smart Monte Carlo. We show that SSM with the smart Monte Carlo algorithm again simultaneously simulates both the distribution function and the equilibrium constant with minimum sensitivity to the parameter set.

Subspace Sampling Method

The Subspace Sampling Method arbitrarily divides configuration space into subspaces. Transitions between two states in configuration space are separated into two classes: transitions between states within one subspace (called moves) and transitions between subspaces (called transfers). The choice of subspace is typically dictated by the

potential function. Regions separated by large potential barriers or discontinuities are assigned different subspaces to facilitate sampling of configuration space.

Transitions between subspaces

The sampling of configuration space for two subspaces is initiated by randomly selecting a process from 1) a transfer from subspace A to B with probability P_{AB} , 2) a transfer from subspace B to A with probability P_{BA} , or 3) a random move with probability P . We choose $P_{AB}=P_{BA}$ and consequently $P=1-2P_{AB}$. When the transfer process is selected, for example from subspace A to B, we then pick a particle from subspace A and a trial position randomly from subspace B. The probability of accepting the transfer from subspace A to subspace B is

$$\begin{aligned} W_{AiBj} &= c \frac{V_B}{V} && \text{if } E_{Bj} < E_{Ai} \\ W_{AiBj} &= c \frac{V_B}{V} \exp(-\beta(E_{Bj}-E_{Ai})) && \text{if } E_{Bj} > E_{Ai} \\ W_{AiBj} &= 0 && \text{if no particle exists in subspace A} \end{aligned} \quad (1)$$

where c is a constant, and E_{Bj} and E_{Ai} are the potential energies of subspace B, state j and subspace A, state i , respectively and $\beta = 1/kT$. In this work, we select $V_A=V_B=\frac{1}{2}V$, where V is the total volume of the Monte Carlo cell and V_A and V_B are the volumes of the individual subspaces. Thus the underlying matrix for the trial moves is symmetric and the acceptance probability for the transfer process is reduced to the Metropolis expression (2).

Transitions within one subspace

Transitions within one subspace are implemented according to the standard Metropolis algorithm. The probability of accepting a move from state i to state j within one subspace can be represented as:

$$W_{ij} = \text{Min} \{ 1, q \} \quad (2)$$

where

$$q = \frac{T_{ji}}{T_{ij}} \exp(-\beta(E_j - E_i)) \quad (3)$$

and

$$T_{ij} = 1, T_{ji} = 1. \quad (4)$$

Periodic boundary conditions are imposed for each subspace.

The Subspace Sampling Method contains three adjustable parameters: P —the probability of selecting a move within a subspace, c —a constant that modulates the acceptance probability of a transfer and Δ —the maximum step size for the trial displacement in the random move. In the following section, we examine the accuracy, efficiency and sensitivity to the parameter set of the Subspace Sampling Method.

SSM and the Free Energy Difference Method

The free energy difference method corresponds to the Subspace Sampling Method when $P=0$. To evaluate the convergence properties and accuracy of SSM for several parameter sets, we compute $\rho(x)$ and K for the two asymmetric potentials shown in figure 1. In both models chosen, there are two subspaces corresponding respectively to the regions $-6 < x < 0$ and $0 < x < 6$ enclosed within a cell of length 12. We employ a dimensionless unit system throughout this paper ($\beta=1$). A detailed description of the models is as follows:

Model (a)--A two-step potential function

The two-state potential function (figure 1a) is:

$$V_A(x)=0 \quad -6 < x < 0 \quad (5)$$

$$V_B(x)=\epsilon_2 \quad 0 < x < 6$$

The lower energy level is set to be zero. We considered one higher energy level of ϵ_2 equal to 2.

Model (b)--An activation barrier

The potential has a flat barrier (figure 1b) which simulates a one dimensional activation complex.

$$V_A(x)=0 \quad -6 < x < -3 \quad (6)$$

$$V_A(x)=2 \quad -3 < x < 0$$

$$V_B(x)=2 \quad 0 < x < 3$$

$$V_B(x)=0 \quad 3 < x < 6$$

Computation of the density function, equilibrium constant and evaluation of convergence

In all of our simulations, we compute the single particle density function for a variety of parameter sets and compare the computed density with the exact solution. We compute the density function according to:

$$\rho(x,m)\Delta x = \sum_{i=1}^{100} \frac{N_i}{N_T} h_i \quad (7)$$

where N_i is the number of times a particle is found in the i th slot, N_T is the total number of moves of the simulation and Δx is the grid size. h_i is a step function, i.e., $h_i = 1$ if x is in

the i th histogram slot, otherwise $h_i = 0$. The exact solution for the density function is computed from:

$$\rho(x) = \exp(-V(x)) / \int \exp(-V(x)) dx \quad (8)$$

where for model a:

$$\begin{aligned} \rho(x) &= \frac{1}{6(1+\exp(-2))} & -6 < x < 0 \\ &= \frac{\exp(-2)}{6(1+\exp(-2))} & 0 < x < 6 \end{aligned} \quad (9)$$

and for model b:

$$\begin{aligned} \rho(x) &= \frac{1}{6(1+\exp(-2))} & -6 < x < -3 \text{ and } 3 < x < 6 \\ &= \frac{\exp(-2)}{6(1+\exp(-2))} & -3 < x < 3. \end{aligned} \quad (10)$$

We use the square roote of the average fluctuation of $\rho(x,m)$ to evaluate the convergence of $\rho(x,m)$ following m passes according to the method of Cao and Berne (5):

$$\chi(m) = \text{sqrt} \left[\int [\rho(x,m) - \rho_{\text{exact}}(x)]^2 dx \right]. \quad (11)$$

For the asymmetric potential functions, $\chi(m)$ only weakly represents the convergence in the entire distribution function. $\chi(m)$ is dominated by the regions containing maximum particle density. In order to examine the convergence of the regions of low particle density, we define a fractional deviation of the particle density according to:

$$\chi\chi(m) = \text{sqrt} \left[\int \left[\frac{\rho(x,m) - \rho_{\text{exact}}(x)}{\rho_{\text{exact}}(x)} \right]^2 dx \right]. \quad (12)$$

Results and Discussion

Accuracy of SSM

We compare the computed single particle distribution function with the corresponding exact distribution for the two potentials (figure 2). As shown in figure 2, the Subspace Sampling Method reproduces the exact distribution function for these potentials. The parameters for each calculation are given in the figure caption. These parameter sets produce a $\chi(10^5)$ around 0.01. Although SSM is capable of reproducing the exact distribution for all parameter sets considered, the convergence rate of $\rho(x)$ can be a sensitive function of the parameter set.

Convergence Properties of SSM

In figures 3-4, we show $\chi(m)$ and $K(m)$ for different parameter sets corresponding to model a when $\epsilon_2=2$. The equilibrium constant is evaluated according to:

$$\langle K \rangle = \frac{\int_0^6 \rho(x,m) dx}{\int_{-6}^0 \rho(x,m) dx} \quad (13)$$

$\langle K \rangle$ represents the ratio of the occupancies of the two subspaces. In model a, the equilibrium constant is not related to the step size, therefore only the values of c and P affect the convergence rate of K . As shown in figures 3-4, the convergence rates of $\rho(x)$ and K exhibit opposite trends: the convergence rate of $\rho(x)$ is facilitated by a large P whereas K converges more rapidly with a small P . For $P=0.9$, K fails to converge within 10^6 passes. On the other hand, the free energy difference method ($P=0.0$) produces the most rapidly convergent K , regardless of the parameter set chosen. However, $\rho(x)$ computed using the free energy difference method fails to converge within 10^6 passes.

To examine the properties of the Subspace Sampling Method in greater detail, we compute $\rho(x)$ and K for model b. The subspace division in this model does not follow the

potential discontinuities. Rather, we symmetrically partition the higher energy state. The equilibrium constant for model b is computed according to:

$$\langle K \rangle = \frac{\int_{-3}^3 \rho(x,m) dx}{\int_{-6}^0 \rho(x,m) dx + \int_0^6 \rho(x,m) dx} \quad (14)$$

where $K(\text{model b}) = K(\text{model a})$. The simulation of $\rho(x)$ for this model converges rapidly ($\chi(10^6) < 0.005$) and with little sensitivity to the parameter set (Figure 5). In this simulation, the free energy difference method ($P=0.0$) produces a rapidly convergent $\rho(x)$ in contrast with the step potential. This is a consequence of the subspace definition. At low values of P , barrier crossing is the primary means of sampling configuration space. In model b, efficient sampling of the subspaces occurs for small values of P because the low energy state is contained in both subspaces. At large values of P , efficient sampling of the subspaces also occurs because the high energy state is sampled both through barrier crossing and through the random move. However, the ability to accurately and efficiently sample the high energy state is very sensitive to the step size as exhibited by the calculation of the equilibrium constant.

For large step sizes, the equilibrium constant converges to the exact value following 10^6 passes (figure 6a) for all values of P . As with the step potential, low values of P (facilitated transfers) produces the most rapid convergence. However, unlike the simulation of K for the step potential, even for high values of P , K converges within 10^6 passes. The insensitivity of the convergence rate of K to the parameter set (for a fixed Δ) is a consequence of an increase in the acceptance of barrier crossings. The convergence properties of K are determined primarily by barrier crossings (as evidence by rapid convergence of K at $P=0.0$). In model b, sampling of the high energy state is facilitated by an increased acceptance of barrier crossings because the particle originates partially in the

high energy state. Consequently, when the particle originates in the high energy state, transfers from one subspace to the other occur with certainty.

Transfers between the two energy states are inefficient with a small step size as demonstrated by the simulated values of K (figure 6b). The small step size results in more sampling of the high energy state than is appropriate. (K is defined as the ratio of the occupancy of the high energy state to the low energy state). This illustrates the careful balance between individual subspace sampling and barrier crossings. With the small Δ , the probability of transferring between the two energy states is reduced (as compared with large Δ), consequently barrier crossings are less efficient and K converges slowly.

The simulations of barrier crossings demonstrate that for rapid convergence of both $\rho(x)$ and K we need a method that efficiently samples the individual subspaces with frequent barrier crossings. The free energy difference method only uses barrier crossing and therefore inefficiently samples the subspaces. With a proper choice of Δ , P and c , the Subspace Sampling Method can balance barrier crossing with subspace sampling. The relationships among barriers crossing and subspace sampling can be seen from the parameters K and $\chi(m)$. K indicates the efficiency of transferring between the two energy states while $\chi(m)$ reflects the convergence of the high density region of the distribution function. Figures 5 and 6 demonstrate that two parameter sets produce the same $\rho(m)$ but different convergence rates of K . The optimization of the parameter set to produce rapid convergence rates of both $\rho(x)$ and K can be time consuming. In the following simulations, we demonstrate a more efficient subspace sampling method that minimizes the need for three parameters.

Modifications of SSM

The versatility of SSM can be exploited to include more efficient sampling schemes within the SSM algorithm. In this section, we incorporate three methodologies into SSM:

1) a single-particle enhanced method that increases subspace sampling, 2) smart Monte Carlo to improve subspace sampling and 3) a modified J-Walking approach to enhance barrier crossing coupled to smart Monte Carlo for subspace sampling. We examine the distribution function and the equilibrium constant for the potential shown in figure 7a. We choose this potential, rather than the step potentials, because the distribution function is non-uniform in both subspaces. This places a greater demand on the sampling scheme to efficiently sample all regions of configuration space.

Methodologies Considered

Enhanced Method for Single Particle Simulations

For one particle simulations, many trial transfers are rejected ($W_{AiBj} = 0$) because there is no particle resident in the initiating subspace. We enhance the efficiency of our one particle simulations by randomly moving the particle within subspace B when a transfer from subspace A to B is chosen but no particle is resident in A. As shown previously, this enhancement results in slightly faster convergence of the one particle distribution function. We note that this modification is unnecessary for sufficiently large systems.

J-Walking

We explore an alternative set of transfer probabilities which is a modification of the J-walking method of Frantz, et. al. (11,12). The probability of accepting a transfer from subspace A, state i to subspace B, state j is:

$$W_{AiBj} = \min [1, q] \quad (15)$$

where

$$q = \exp(-\beta(E_{Bj} - E_{Ai})) \frac{T_{BjAi}}{T_{AiBj}} \quad (16)$$

$$T_{AiBj} = \frac{\exp(-\beta E_{Bj})}{Q_B} \quad (17)$$

$$T_{BjAi} = \frac{\exp(-\beta' E_{Ai})}{Q_A}$$

$$T_{AiBj} = 0 \quad \text{if no particle exists in subspace A.}$$

In our one particle, one-dimensional systems, the partition functions Q_A and Q_B are evaluated analytically at the temperature T' ($\beta' = 1/kT'$) and are given by

$$Q_A = \int \exp(-\beta' E_A(x_A)) dx_A \quad (17)$$

$$Q_B = \int \exp(-\beta' E_B(x_B)) dx_B$$

where the integration is over the space spanned by the potential function.

In the development of the J-walking method, Frantz, et al. use a higher temperature ($\beta' < \beta$) to jump over the potential barriers. These jumps are interspersed among the standard Metropolis walks. We implement transfers in the following manner: we select a transfer process according to the standard SSM probability $(1-P)/2$. We then select the position of the particle after transferring according to the distribution in equation 17. As shown in the previous section, efficient sampling of the entire configuration space is a result of a careful balance of individual subspace sampling and accepted barrier crossings. In systems with non-uniform potentials, this balance is affected by the trial distribution from which the move or transfer is selected. Consequently, this modified J-walking method provides an alternative distribution from which to select the position of the transferred particle. The temperature T' acts as an adjustable parameter to optimize convergence.

Smart Monte Carlo

We incorporate the smart Monte Carlo method into the transition probabilities for the random move within one subspace by modifying the trial probability (T_{ij}) in equation 3 according to

$$T_{ij} = \left[\frac{1}{4\pi\Delta^2} \right]^{1/2} \exp \left(-\frac{[x_j - x_i - \beta A f_i]^2}{4\Delta^2} \right). \quad (18)$$

In equation 18, f_i is the force on the particle in the initial state i , $x_j - x_i$ is the trial displacement of the particle, A is an adjustable parameter and $\Delta = A^{1/2}$. Reflecting walls are used for the subspaces in the smart Monte Carlo approach.

Potential Function

A quadratic potential replaces the uniform energy level of model b to represent the harmonic potential of an associated molecule (figure 7a). Explicitly, the potential is:

$$\begin{aligned} V_A(x) &= \frac{8}{9} (x+3)^2 & -6 < x < 0 \\ V_B(x) &= 16 - \frac{96}{x+6} & 0 < x < 6. \end{aligned} \quad (19)$$

Also shown in figure 7b is the distribution function (computed for the parameter set $c=2$, $P=0.9$, $\Delta=3/2$) compared with the exact distribution. The exact density distribution is obtained from:

$$\rho(x) = \frac{\exp(-V(x))}{Q} \quad (20)$$

where Q is the partition function of the system given by:

$$Q = \int_{-6}^0 \exp\left(-\frac{8}{9} (x+3)^2\right) dx + \int_0^6 \exp\left(-\left(16 - \frac{96}{x+6}\right)\right) dx \quad (21)$$

The integration with the harmonic potential in equation 21 is estimated using a Taylor series expansion for the integral. Therefore,

$$\int_{-6}^0 \exp\left(-\frac{8}{9}(x+3)^2\right) dx = \frac{3}{\sqrt{8}} [F_1(\sqrt{8}) - F_1(-\sqrt{8})] \quad (22)$$

where

$$F_1(u) = \sum_{n=1}^N \frac{(-1)^{n+1} u^{2n-1}}{(2n-1)(n-1)!} \quad (23)$$

and $u = \sqrt{\frac{8}{9}}(x+3)$.

The integration with hyperbolic potential in equation 21 is evaluated according to:

$$\int_0^6 \exp\left(-\left(16 - \frac{96}{x+6}\right)\right) dx = 96 \exp(-16) \int_8^{16} \frac{\exp(t)}{t^2} dt = 96 \exp(-16) [F_2(16) - F_2(8)] \quad (24)$$

where

$$F_2(t) = -\frac{\exp(t)}{t} + \ln(t) + \sum_{n=1}^N \frac{t^n}{(n)(n!)} \quad (25)$$

and $t = \frac{x+6}{96}$. Convergence in both integrations required at least forty terms in the series expansions.

For all parameter sets studied, and for all methodologies considered, SSM is capable of reproducing the exact function. However, as noted previously, the rate of convergence of $\rho(x)$ and K are dependent upon the methodology and the parameter set.

Results and Discussion

In figure 8, we compare the $\chi(m)$ for the three alternative subspace sampling methodologies corresponding to several parameter sets. In all three methodologies, the single particle enhanced method is incorporated into SSM. The methodologies considered

are 1) SSM, 2) SSM incorporating smart Monte Carlo for the random move, and 3) SSM incorporating the J-walking distributions for the transfer process and smart Monte Carlo for the random move. The convergence rate of $\rho(x)$ using SSM is sensitive to the choice of P and Δ (figure 8a). Such sensitivity is diminished when the smart Monte Carlo algorithm is used for the random move (figure 8b). Similar convergence rates are observed with the J-walking algorithm. However, J-walking introduces another parameter (β') that influences the convergence properties. We examined a variety of temperatures and found that $\beta'=1.25\beta$ produces the optimum convergence in both $\rho(x)$ (figure 8c) and K (figure 9c).

Different trends are observed for the equilibrium constant. The convergence rate of K is accelerated and the sensitivity to the parameter set is extremely reduced when the smart Monte Carlo or J-Walking methods are incorporated in the subspace sampling methodologies (figures 9). In figures 9 we show the convergence rate of K for one set of parameters (see figure caption). Other parameters chosen exhibit the same convergence properties.

Although the smart Monte Carlo and J-Walking methods exhibit similar convergence properties of the equilibrium constant, small differences in the convergence rates of the particle density are observed. Figure 10 shows that the subspace sampling methodologies demonstrate different sensitivities to the parameter set when convergence of the entire distribution function is examined. Rapid convergence in $\rho(x)$ can be observed when only $\chi(m)$ is considered because $\chi(m)$ is dominated by regions of high particle density. $\chi\chi(m)$ equally weighs the regions of high and low particle density and is thus a measure of the convergence properties of both regions. Figure 10 demonstrates that only the subspace sampling method with the smart Monte Carlo algorithm efficiently samples both high and low density regions, independent of the parameter set.

Conclusions

The single particle density for three asymmetric, continuous and discontinuous potential functions were studied in this paper. For all potential functions, the Subspace Sampling Method and its variants reproduce the exact solution. The disadvantage of the SSM is the number of adjustable parameters that must be optimized to speed convergence of the simulation. Furthermore, the optimum choice of these parameters depends upon the variable desired. Frequent barrier crossings result in a rapidly convergent equilibrium constant but come at the expense of the convergence of the distributions. Free energy difference methods ($P=0.0$ in the standard subspace sampling methodology), for example, exhibit in the most accelerated convergence in K but the slowest convergence in $\rho(x)$.

Optimum convergence of both $\rho(x)$ and K occurs when the subspaces are efficiently sampled between each barrier crossing. Consequently, the most efficient method we find uses the smart Monte Carlo algorithm for the individual subspace sampling (9). The efficiency of this algorithm minimizes the sensitivity of the convergence rates of both $\rho(x)$ and K to the subspace sampling parameter set. Efficient sampling of the entire configuration space of one subspace appears to produce efficient barrier crossings that further accelerate the convergence rate of K . This is in contrast to the Metropolis method where the lower energy states are sampled more frequently than in smart Monte Carlo and consequently transitions to a higher energy state (across a barrier) occur less often.

The Subspace Sampling Method is a general methodology for facilitating sampling of configuration space that contains barriers, discontinuities and wells in the potential function. An appropriate selection of the transition probability for inter- and intra-subspace sampling results in rapid convergence of the simulation variables. In the subsequent paper, we extend the Subspace Sampling Method to multiple particles and to three dimensions to simulate the weak electrolyte system using the Sticky Electrolyte Model (15).

References

- (1) Shew, C.-Y., Mills, P.A. *J. Phys. Chem.* **1993**, 97, 13824.
- (2) Metropolis, N.; Rosenbluth, A.W.; Rosenbluth, M.N.; Teller, A.H.; Teller, E. *J. Chem. Phys.* **1953**, 21, 1087.
- (3) Berne, B.J.; ed. *Statistical Mechanics, Part A: Equilibrium Techniques*, Plenum Press, New York, **1977**.
- (4) Ciccotti, G.; Frenkel, D.; McDonald, I.R.; eds. *Simulation of Liquids and Solids, Molecular Dynamics and Monte Carlo Methods in Statistical Mechanics* North-Holland Physics Publishing, **1987**.
- (5) Cao, J. ; Berne, B.J. *J. Chem. Phys.* **1990**, 92, 1980.
- (6) Pangali, C.; Rao, M.; Berne, B.J. *Chem. Phys. Lett.* **1978**, 3, 413.
- (7) Rao, M.; Berne, B. *J. Chem. Phys.* **1979**, 71, 129.
- (8) Rao, M.; Pangali, C.; Berne, B.J. *Mol. Phys.* **1979**, 37, 1773.
- (9) Rossky, P.J.; Doll, J.D.; Friedman, H.L. *J. Chem. Phys.* **1978**, 69, 4628.
- (10) Mezei, M.; Bencsath, K.A.; Goldman, S.; Singh, S. *Molecular Simulations* **1987**, 1, 87.
- (11) Frantz, D.D., Freeman, D.L., Doll, J.D. *J. Chem. Phys.* **1990**, 93, 2769.
- (12) Frantz, D.D., Freeman, D.L., Doll, J.D. *J. Chem. Phys.* **1992**, 97, 5713.
- (13) Bennett, C.H. *J. Computational Physics*, **1976**, 22, 245.
- (14) Voter, A.F. *J. Chem. Phys.* **1985**, 82 , 1890.
- (15) Lee, S.H.; Rasaiah, J.C.; Cummings, P.T. *J. Chem. Phys.* **1985**, **83**, 317.

Figure Captions

Figure 1. A schematic of the two potential functions. The Monte Carlo cell is a one-dimensional unit of length 12. (Dimensionless units are used throughout.) The subspace definitions are indicated with the dashed line.

Figure 2. The single particle density function for the two models. The computed $\rho(x)$ following 10^5 passes is given by o and the exact functions are given by the solid line. The parameters for the computations correspond to: model (a) $c=1$, $P=0.9$, $\Delta=3/2$, $\epsilon_2=2$; model (b) $c=2$, $P=0.7$, $\Delta=3/2$, $\epsilon_2=2$.

Figure 3. The convergence rates of the single particle density function for model (a) and for three parameter sets. The parameter sets shown are A: $c=2$, $\Delta=3/2$, B: $c=2$, $\Delta=3/8$ and C: $c=1$, $\Delta=3/2$ and for all figures: $\square$: $P=0.1$, $\circ$: $P=0.5$, Δ : $P=0.9$, $\bullet$: $P=0.0$.

Figure 4. The convergence rates of the equilibrium constant expressed as a deviation from the exact function for model (a) and for the two parameter sets corresponding to figure 3A and 3C.

Figure 5. The convergence rates of the single particle density function for model (b) and for two parameter sets. The parameter sets shown are A: $c=2$, $\Delta=3/2$ and B: $c=2$, $\Delta=3/8$ and for all figures : $\square$: $P=0.1$, $\circ$: $P=0.5$, Δ : $P=0.9$, $\bullet$: $P=0.0$.

Figure 6. The convergence rates of the equilibrium constant expressed as a deviation from the exact function for model (b) and for two parameter sets. The parameter sets shown are given in figure 5.

Figure 7. a. The potential function used for the comparison of the three Subspace Sampling Methodologies given in the text. b. The simulated (o) and exact (solid line)

single particle distribution function for this potential computed using SSM with the Metropolis sampling scheme and corresponding to the parameter set $c=2$, $P=0.9$, $\Delta=3/2$.

Figure 8. The convergence rates of the single particle density function for the three methodologies. Figure 8a: the Metropolis method with subspace sampling, $c=2$, $\Delta=3/8$, and $\square$: $P=0.1$, Δ : $P=0.3$, $\circ$: $P=0.5$, $\blacksquare$: $P=0.7$, $\bullet$: $P=0.9$. Figure 8b: the smart Monte Carlo with subspace sampling, $c=2$, $\Delta=3/8$, and $\square$: $P=0.1$, Δ : $P=0.3$, $\circ$: $P=0.5$, $\blacksquare$: $P=0.7$, $\bullet$: $P=0.9$. Figure 8c: J-walking with subspace sampling, $\square$: $P=0.1$, Δ : $P=0.3$, $\circ$: $P=0.5$, $\blacksquare$: $P=0.7$, $\bullet$: $P=0.9$, $\Delta=3/8$, $\beta_T=1.25\beta$; $\beta=1$.

Figure 9. The convergence rates of the equilibrium constant for three methodologies. The parameter sets correspond to figure 8.

Figure 10. The convergence rates of the single particle density as represented by $cc(m)$. Figure A corresponds to SSM, Figure B SSM with smart Monte Carlo, and Figure C SSM, J-Walking and smart Monte Carlo. The parameter sets correspond to figure 8.

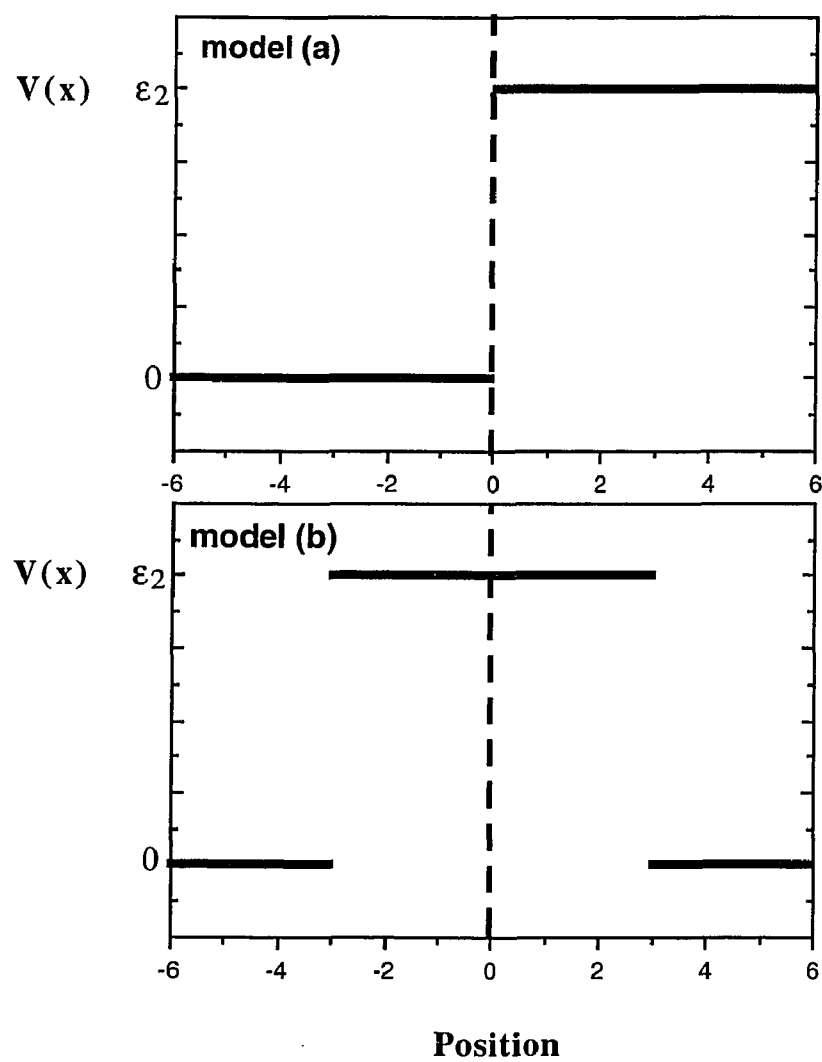


Figure 1

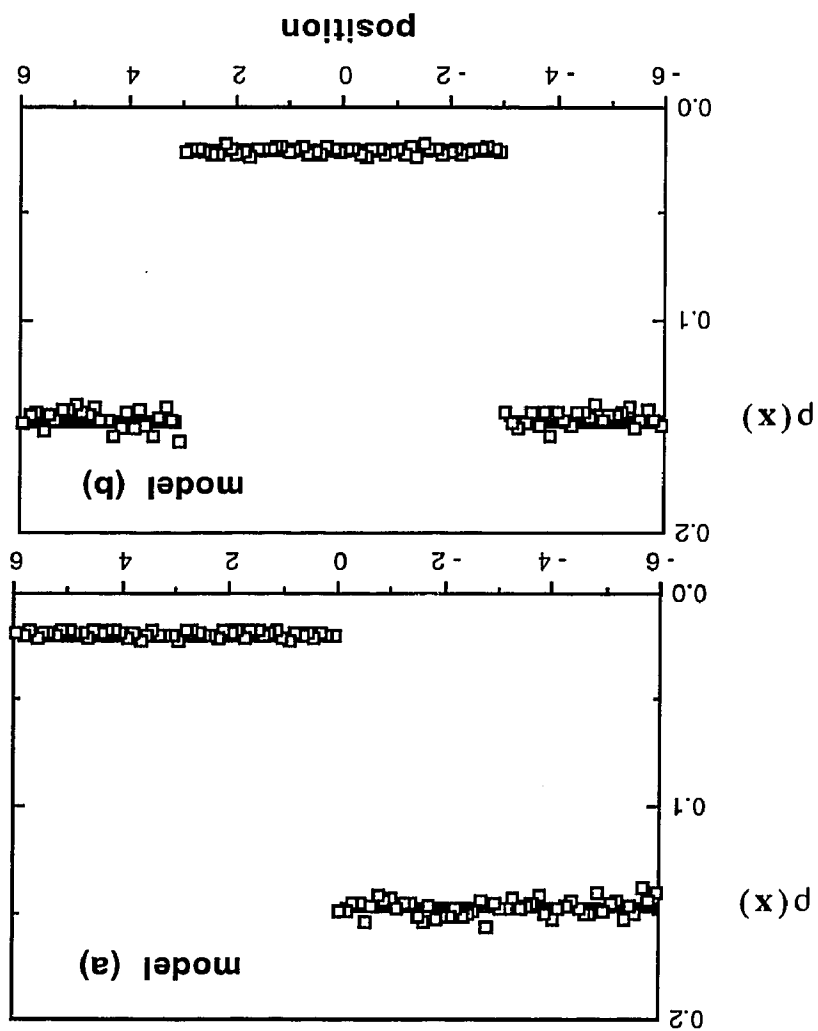


Figure 2

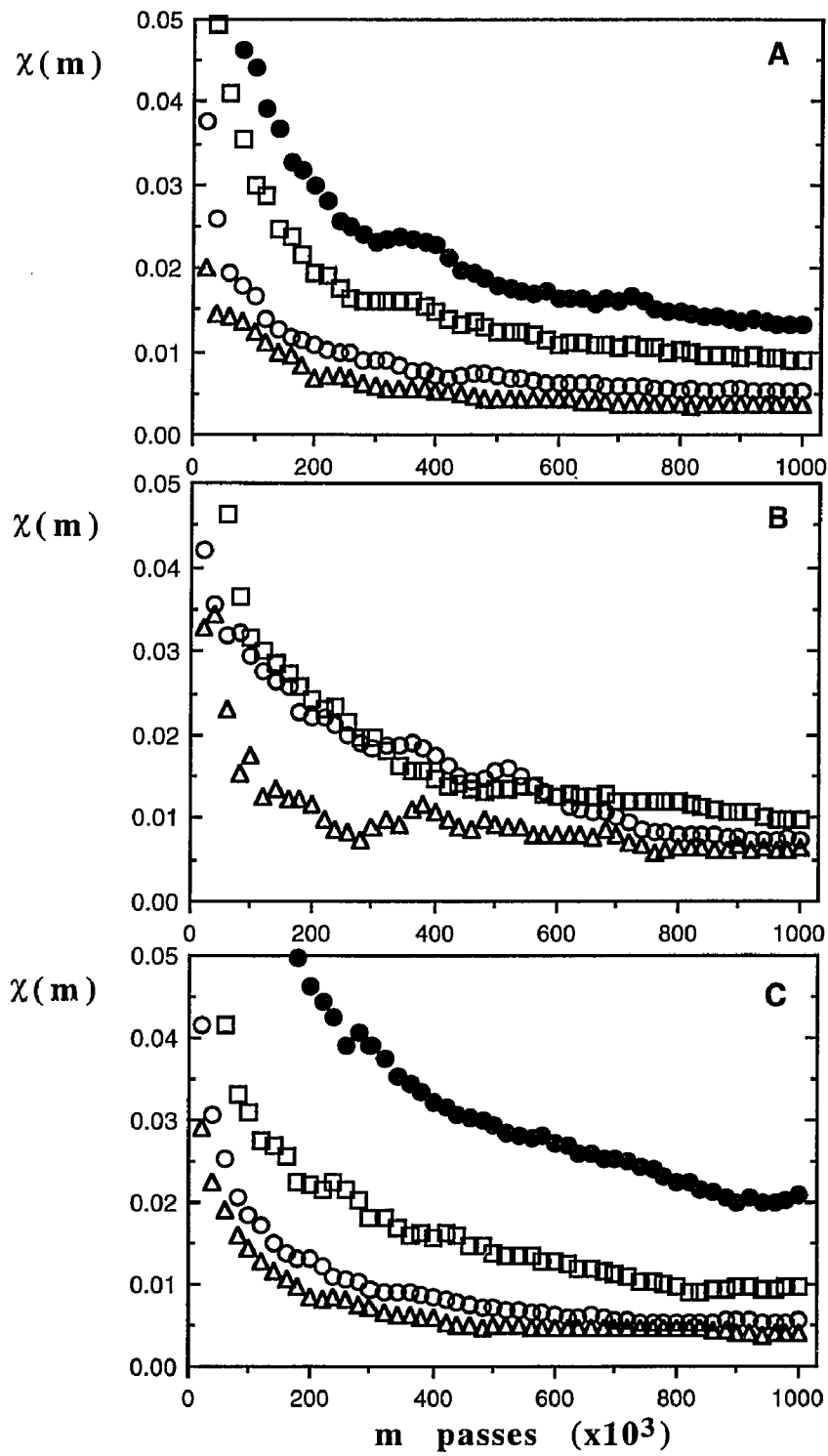


Figure 3

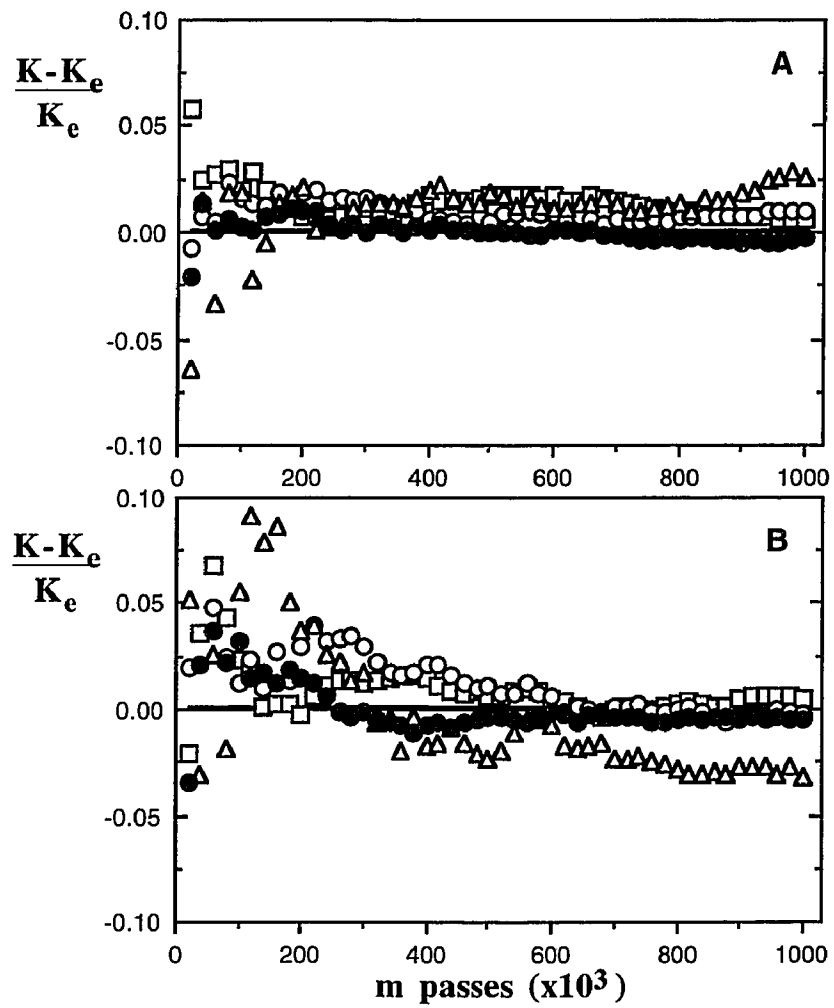


Figure 4

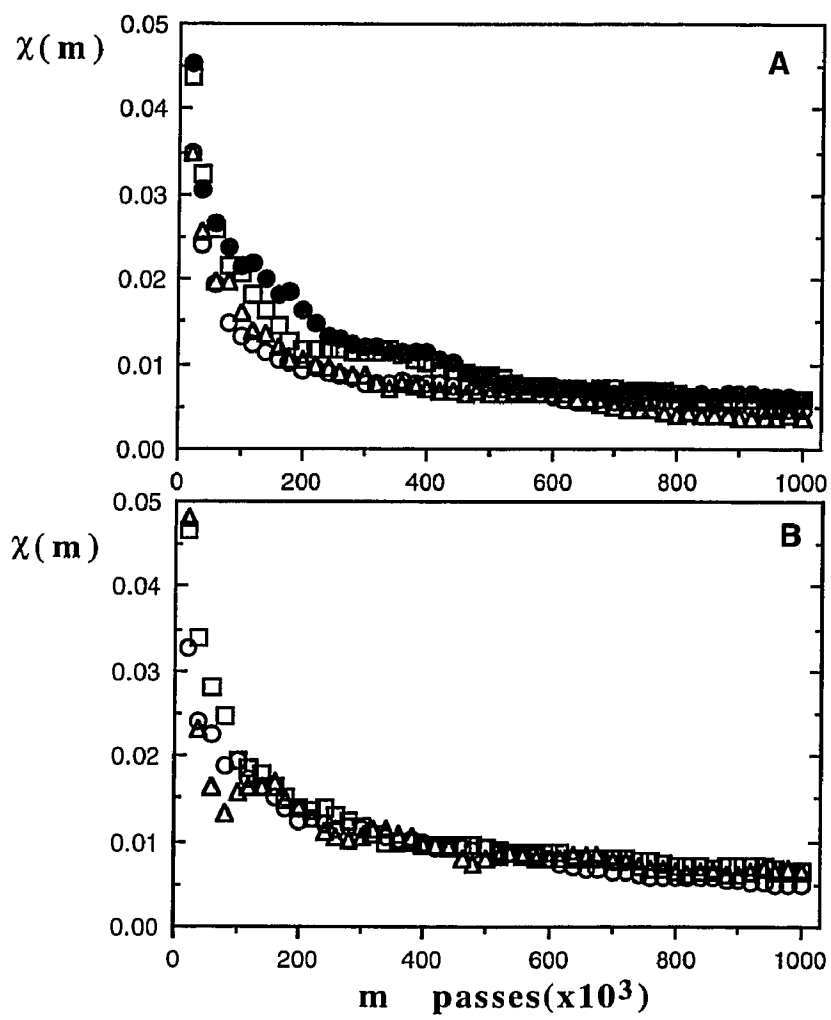


Figure 5

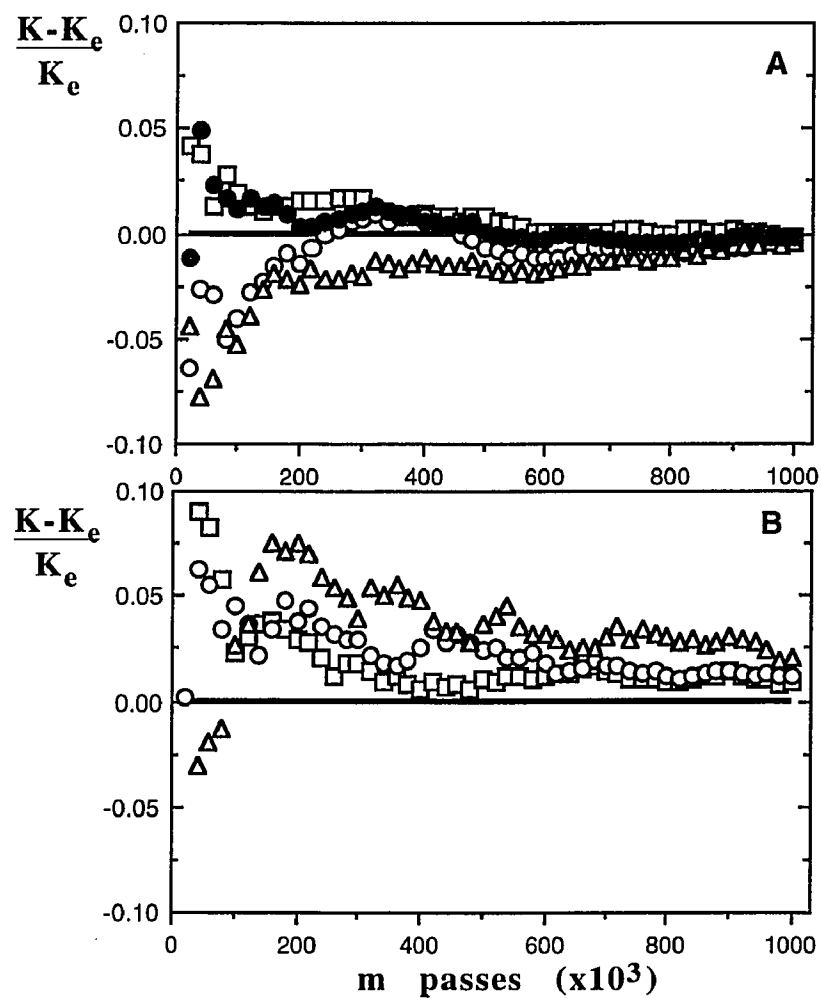


Figure 6

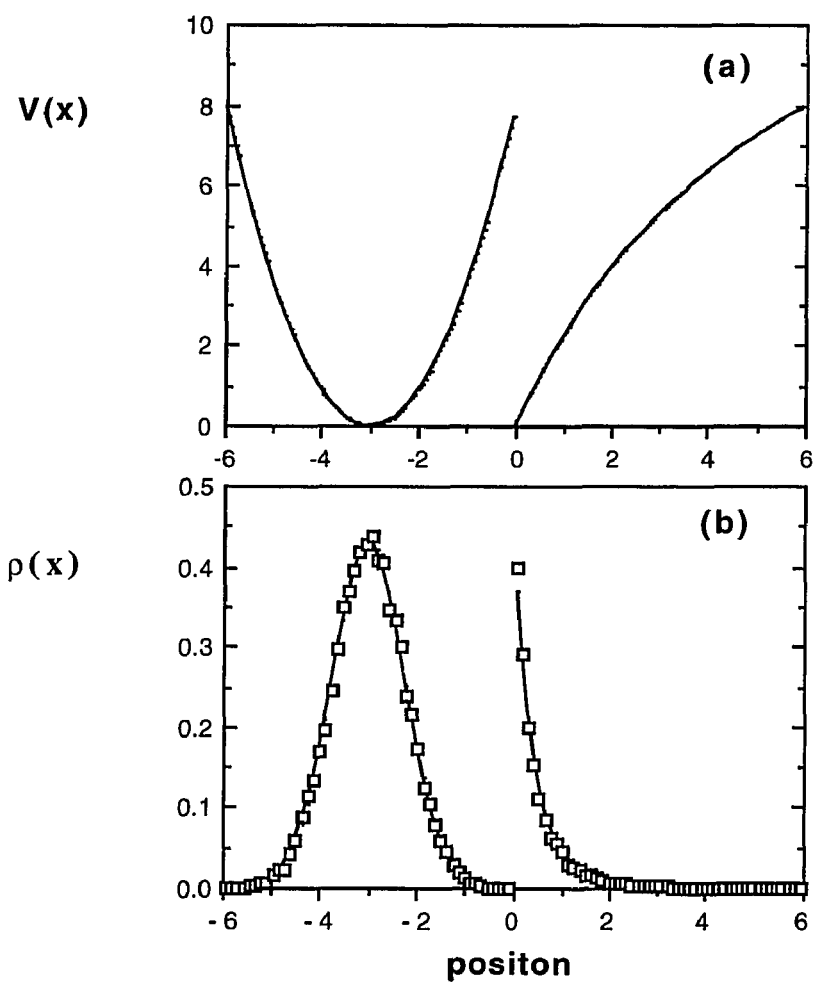


Figure 7

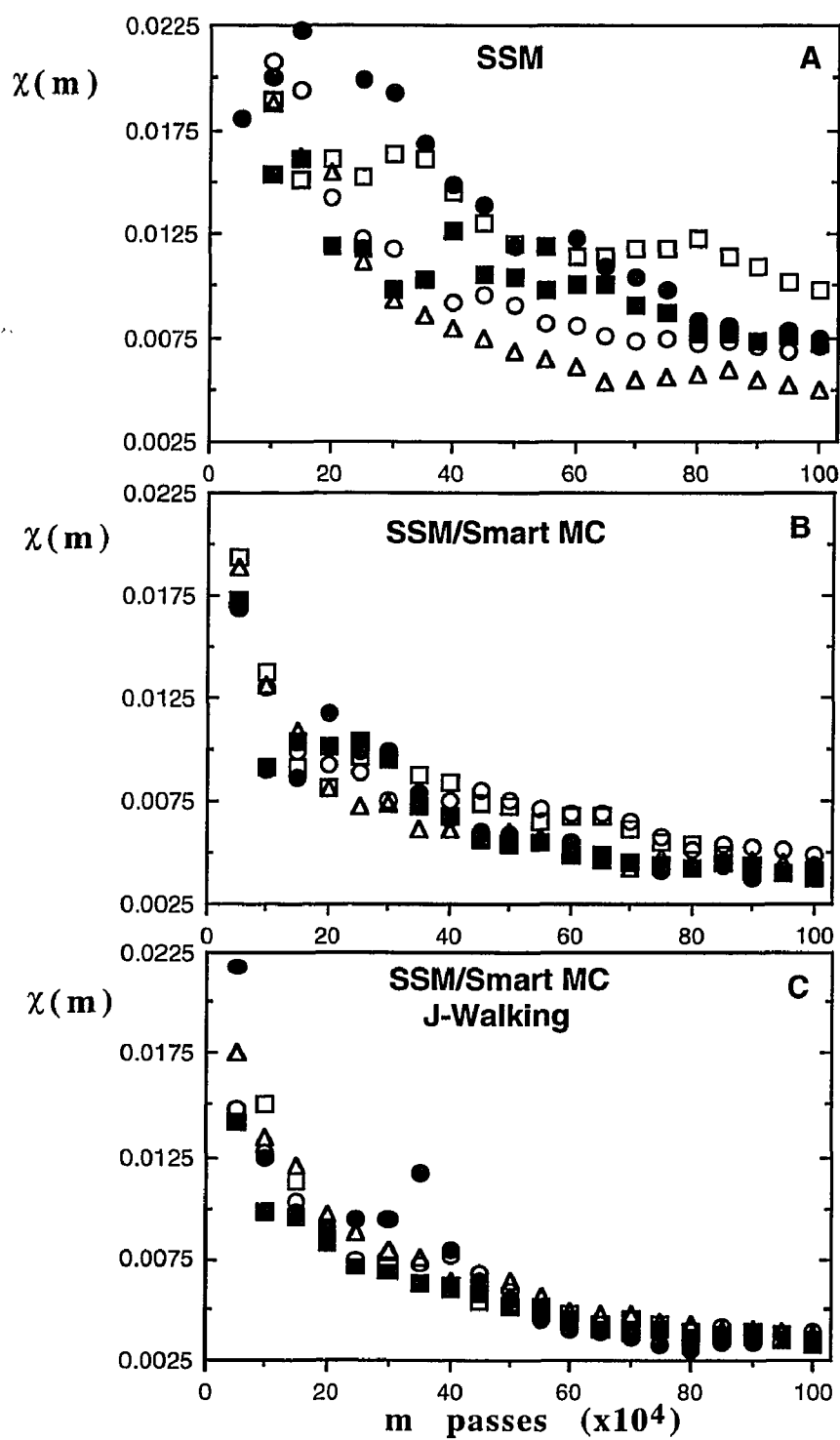


Figure 8

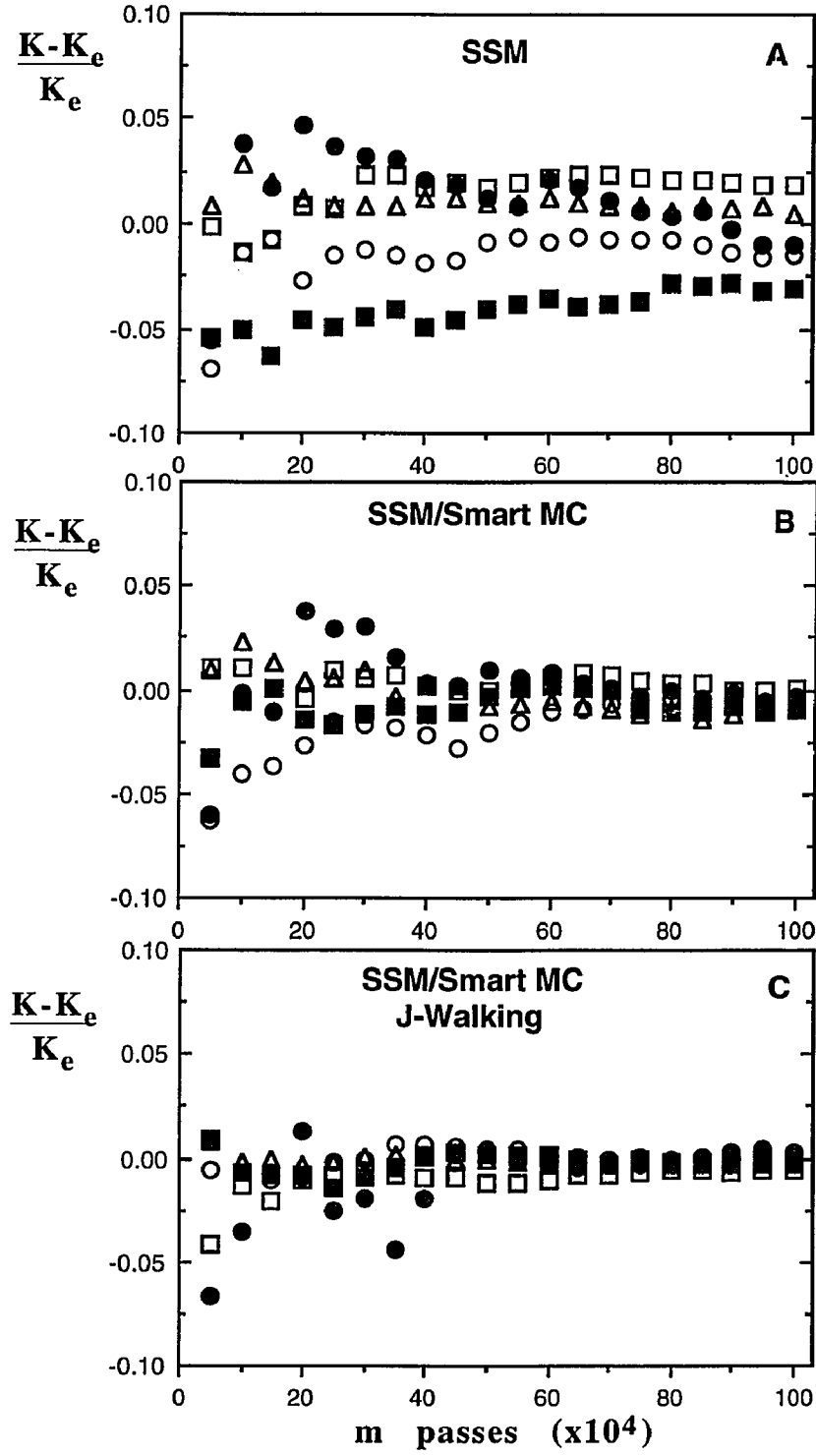


Figure 9

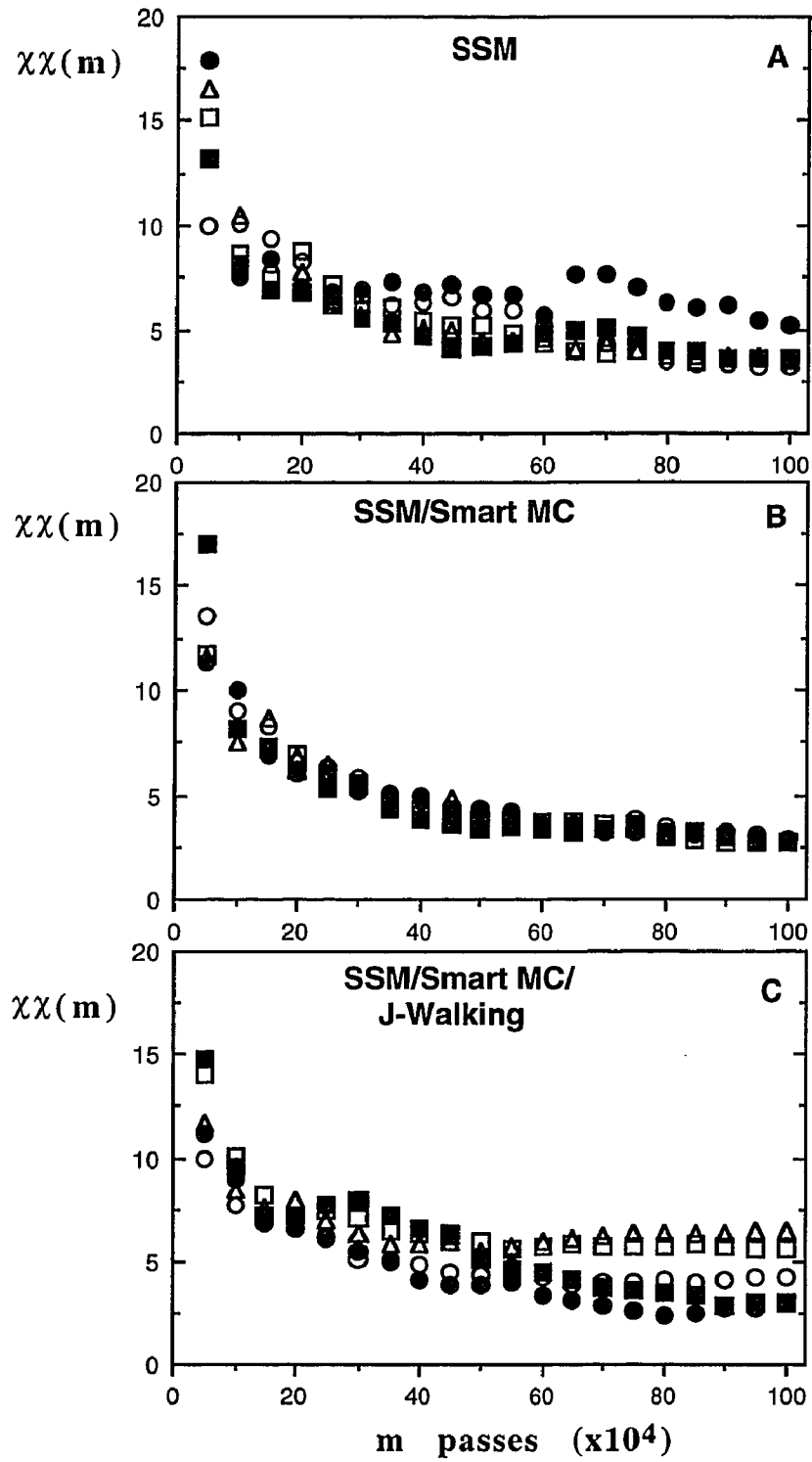


Figure 10

Chapter 4

Monte Carlo Simulations of the Pair Correlation Function and the Equilibrium Association Constant of the Sticky Electrolyte Model Using the Subspace Sampling Method

Abstract

We extend the Subspace Sampling Method (SSM) to three-dimensions to simulate the thermodynamic and structural properties of the weak electrolyte solution. We demonstrate that the Subspace Sampling Method is particularly well-suited to investigate chemical reaction systems. The weak electrolyte system is simulated using the Sticky Electrolyte Model (1) with and without dipolar interactions. For low concentrations of electrolyte (0.0625M), the simulated equilibrium association constant is insensitive to the inclusion of the dipolar interactions when the depth of the molecular potential (ϵ_2/k_B) is 4000 K and the width is 0.42 Å. The internal energy of the molecule is the dominate source of stabilization energy; the dipolar interactions contribute only minimally to stabilize the system. The parameters chosen in these simulations are consistent with those used in the analytical solutions (1,2). However, the equilibrium constant computed from the Monte Carlo simulations and extrapolated to an infinite cell size is approximately one order of magnitude greater than that of the hypernetted-chain/mean sphere approximation (HNC/MSA) (1) or the HNC approximation (2). The simulated site-site correlation function is lower than both approximation methods when the dipolar interaction is included but resembles more closely the results of the HNC/MSA calculations (1).

Introduction

Statistical mechanical studies of chemical reactions have introduced various degrees of "stickiness" to emulate the chemical association of both nonionic and ionic fluids (1-14). For simple models of chemical association, analytical solutions for the thermodynamic properties of these reacting systems have been obtained using the mean spherical approximation (MSA), MSA along with the Percus-Yevick (PY) or hypernetted chain (HNC) equations (1-8). One such simple model of the weak electrolyte solution is the

Sticky Electrolyte Model (SEM) (1-8). In addition, various realistic improvements of SEM have been considered including symmetric and asymmetric hard spheres (1-8).

The Sticky Electrolyte Model partitions configuration space into two regions determined by the distance between two ions with opposite charges. In the region where the inter-ionic distance is greater than the hard sphere distance of closest approach (σ), the ionic interactions consist of charged hard spheres immersed in a dielectric continuum described by a uniform dielectric constant (ϵ). The interactions in this region thus follow the restricted primitive model (RPM). Chemical association occurs when the inter-ionic distance is less than s . In this region, the potential is described by a square well of depth ϵ_2 and width w . The formation of a molecule must obey "steric saturation" (1), i.e., each atom A or B can belong to one molecule only. After an ion pair forms a molecule, the charge remains on each atom of the molecule. Consequently, molecules are modeled as dipoles and thus can interact with other molecules and ions through a dipolar mechanism.

The inclusion of "stickiness" in SEM results in a reduction of the number of free ions below that predicted from the restricted primitive model (RPM) at corresponding concentrations of ions (1-8). Analytical solutions for SEM have been achieved using MSA for the ionic interactions and PY or HNC for the molecular state. Initially, the HNC equations were not applied to the ionic state because the HNC results for dilute solutions (less than 0.05 M) of strong electrolytes produced high charge densities near the ions as compared with the Debye-Hückel theory (1,15). Hence, MSA was chosen to describe the behavior of the electrostatic interactions in the weak electrolyte system. However, for the molecular part of the SEM potential, the HNC equation produced the most realistic average number of molecules. The MSA/PY approximation, applied to the molecular state, resulted in an unrealistic negative average number of molecules for divalent weak electrolytes.

Although the HNC/MSA approximation resulted in reasonable pair correlation functions and equilibrium constants for the weak electrolyte solution studied, the accuracy

of HNC/MSA was uncertain. In their subsequent work, Lee et al. used the HNC equation alone for both the molecular and electrostatic parts of the Sticky Electrolyte Model (2). The HNC results for the equilibrium constant were almost identical with the MSA/HNC method but the contact correlation function of the HNC method is much greater at low electrolyte concentrations. When the initial electrolyte concentration is 0.1M the contact correlation function is 2.6 for HNC/MSA and 5.93 for HNC for the 2-2 electrolyte (2). The difference between the two methods becomes smaller only when initial concentration is high ($>1.5\text{M}$) (2). Furthermore, in the HNC results trimer and tetramer clusters of dipolar molecules and ions were formed that were absent in the MSA/HNC results (2).

The Monte Carlo (MC) method is an extremely versatile approach that can be applied to test the accuracy of approximate analytical theories and to generalize to systems of arbitrary charge and geometry. Monte Carlo simulations have been used to obtain the structural and thermodynamic properties of aqueous solutions of strong electrolytes (15-19) or heterogeneous solutions of strong electrolytes interacting with planes (membranes) (20), charged spheres (micelles) (21,22) or rods (nucleic acids) (23-28). These simulations of ionic solutions typically employ the restricted primitive model or variants of the model, e.g. modifications of the dielectric constant to account for the dielectric discontinuity due to the macromolecules (25,26). The application of Monte Carlo methods to study chemical reactions is impeded by the deep potential well and the discontinuous potential needed to model the molecular interactions. Deep potential wells produce a bottleneck in the implementation of the standard Monte Carlo algorithm resulting in an inefficient sampling of configuration space. Force-bias and anti-force bias methods, which overcome the bottleneck problem, cannot be implemented with discontinuous potential functions (31).

The recently introduced Subspace Sampling Method (SSM) overcomes the bottleneck problem and thus appears well-suited to simulate the thermodynamic and structural properties of the Sticky Electrolyte Model (32,33). SSM is formally equivalent to the

reaction ensemble method (34) and bond-bias method (35) which were shown to accurately reproduce the equilibrium properties for some simple potential functions. In the previous paper, we showed that SSM can accurately reproduce both the equilibrium constant and the pair correlation functions for simple one-dimensional potentials. To simulate both thermodynamic and structural properties, the method must be able to sample the configuration space of the reactant and product and concomitantly generate the probability distribution for each of the components. In our previous work we demonstrated the Subspace Sampling Method is a powerful method that can satisfy both criteria (33).

In this work we show how the Subspace Sampling Method using the Sticky Electrolyte Model can be applied to simulate the association of hard sphere divalent ions ($A^{2+} + B^{2-} \rightleftharpoons AB$). In the first section, we compare the simulated results with the exact solution for a system containing one A and one B. This reveals the accuracy and validity of the method. In the second section, we carry out the simulation for an arbitrary number of ion pairs. The simulated equilibrium constant and correlation function will be shown. In the final section, we incorporate dipolar interactions into the SEM and compare our results with the analytical solutions.

Potential Model

The weak electrolyte potential of this work is based on the Sticky Electrolyte Model (1). The interaction of an AB ion pair is divided into two components depending upon the distance between the individual ions. If the distance is greater than the sum of radii of the ions (σ) the interaction energy follows Coulombs law (eqn 1). The A and B particles are then envisioned as individual ions. Otherwise, when the distance is less than σ , the AB pair forms a molecule and a square well potential is used to describe the molecular energy. Figure 1 shows the AB pair potential. Two ions with the same charge obey the restrictive primitive model. The potential model is therefore:

$$U(r_{ij}) = \frac{q_i q_j}{4\pi\epsilon\epsilon_0 r_{ij}} \quad r_{ij} > \sigma \quad i, j = A^+ \text{ or } B^- \quad (1)$$

$$U(r_{ij}) = \infty \quad r_{ij} < \sigma \quad i = A^+ \text{ or } B^- \quad (2)$$

$$U(r_{AB}) = -\epsilon_2 \quad d_{\min} < r_{AB} < d_{\max} \quad (3)$$

$$U(r_{AB}) = \infty \quad d_{\max} < r_{AB} < \sigma \text{ or } r_{AB} < d_{\min} \quad (4)$$

where s is the sum of radii of A (r_A) and B (r_B) and $r_A = r_B = 2.1 \text{ \AA}$, r_{ij} is the interatomic distance between two unassociated ions, r_{AB} is the interatomic distance between two oppositely charged associated ions, ϵ is the dielectric constant of the solvent, ϵ_2 is the depth of the square potential and $d_{\min}$ and $d_{\max}$ determine the bond distance (w) of molecule. The hard sphere model is applied to all pair potentials involving an AB molecule (AB-AB, A^+ -AB, and B^- -AB). In a subsequent section, the hard sphere model is modified to include dipolar interactions among AB molecules in order to compare the Monte Carlo simulation results directly with the analytical approximations for the Sticky Electrolyte Model (1-8).

Method

In order to extend the Subspace Sampling Method to three dimensions and multiple particles, the transition probability is modified to include the number of particles of the different species and the variations in the subspace volumes. Here, we refer to the transfer processes in the previous papers (32, 33) as "Association" and "Dissociation" processes to distinguish the direction of reaction (or transfer process). From microscopic reversibility, the transition probabilities for association and dissociation in the canonical ensemble have the form:

$$\begin{aligned} P_{\text{associate}} &= \frac{1}{N_A} \frac{1}{N_B} \frac{1}{\delta_{\text{mol}}} \frac{1}{V} W_{\text{associate}} \exp\left(-\frac{E_i}{kT}\right) \\ &= P_{\text{dissociate}} \frac{1}{N_{AB}} \frac{1}{\delta_{\text{ion}}} \frac{1}{\delta_{\text{ion}}} W_{\text{dissociate}} \exp\left(-\frac{E_i}{kT}\right). \end{aligned} \quad (5)$$

$P_{\text{associate}}$ ($P_{\text{dissociate}}$) is the probability we select the association (dissociation) process. When the association process is selected we randomly pick one cation and one anion from the N_A and N_B pool of cations and anions, respectively, and force the formation of a molecule. After the association occurs, the first particle is placed randomly in the space of the system with probability $\frac{1}{V}$. Then, the bond length of the two particles is selected from a nonuniform distribution (weighted by r^2) so that the possible volume in which a molecule can reside is $d_{\text{mol}} = \frac{4}{3} \pi (d_{\text{max}}^3 - d_{\text{min}}^3)$. The solid angle (θ, ϕ) , which positions particle B relative to particle A in the molecule, is chosen from a uniform distribution of $\cos\theta$ in the range $[-1, 1]$ and a uniform distribution of ϕ between 0 and 2π . On the left hand side of the equality of equation 5 the first five arguments represent the trial probability of the association process. $W_{\text{associate}}$ is the acceptance probability and $\exp(-\frac{E_i}{kT})$ is the Boltzmann factor.

The same method is applied to obtain the probabilities of the dissociation process. In this process a molecule is selected from the N_{AB} pool of molecules and dissociated. The dissociated cation and anion are randomly placed in the Monte Carlo cell. Thus, δ_{ion} is equal to the volume of the cell (V).

From equation (5), the ratio of the association and dissociation acceptance probability is obtained. One solution for the acceptance probabilities is:

$$W_{\text{associate}} = C_{ij} \frac{N_A N_B}{N^2} \frac{\delta_{\text{mol}}}{V} M(E_i \rightarrow E_j) \quad (6)$$

$$W_{\text{dissociate}} = C_{ij} \frac{N_A N_B}{N^2} \frac{\delta_{\text{mol}}}{V} M(E_i \rightarrow E_j) \quad (7)$$

where

$$\begin{aligned} M(E_i \rightarrow E_j) &= \text{Min} \left\{ 1, \exp\left(-\frac{E_j - E_i}{kT}\right) \right\} \\ M(E_j \rightarrow E_i) &= \text{Min} \left\{ 1, \exp\left(-\frac{E_i - E_j}{kT}\right) \right\}. \end{aligned} \quad (8)$$

In equations 6-8, N is the total number of particles A and B and C_{ij} and C_{ji} are constants and chosen to be equal. We obtain the value of the constants by maximizing the acceptance probability according to:

$$C_{ij}=C_{ji}=\{\text{Max}[\frac{N_A N_B}{N^2} \frac{\delta_{\text{mol}}}{V} M(E_i-E_j), \frac{N_{AB}}{N^2} \frac{\delta_{\text{ion}}}{V} M(E_i-E_j)]\}^{-1}. \quad (9)$$

The acceptance probabilities in equations 6 and 7 are in a more generalized form from what we chose before (32,33). In our previous work, we studied two subspaces with the same volume and with a fixed number of particles. Thus, the maximum value of the constants (C_{ij} and C_{ji}) in the acceptance probabilities cancelled the volume factor of the subspaces and the acceptance probabilities reduced to the Metropolis expressions. In this work, the volume and the number of particles in the two subspaces differ; consequently, the maximum values of the constants have contributions from these volume factors and particle numbers.

By choosing values of C that maximize the acceptance probabilities (equations 6, 7, 9), the simulation parameters are reduced from C , P and Δ to P and Δ in comparison with the previous papers (32,33), where Δ is the step size and P is the probability a move process is selected ($P=1-P_{\text{association}}-P_{\text{dissociation}}$). The current formula for computing the acceptance probability simplifies the search for the optimum parameter set.

Computational details

In this first section, we compare our results with the standard Metropolis method (36). In the Metropolis method, a cation (A) or anion (B) is selected randomly and a trial position is computed. After each random move, we search all inter-ionic distances to determine if the random move results in the formation or dissociation of an uncharged AB molecule. Once the molecules are distinguished from individual ions, the configuration energy is computed according to equations 1-4. In the simulations using both SSM and the

Metropolis method periodic boundary conditions and the minimum image conventions are applied.

In the Subspace Sampling Method of this work the Smart Monte Carlo method is implemented for the move process involving ions. This is the most efficient algorithm for sampling configuration space (see 32, 33). For random moves involving molecules, a random walk is assumed. In the last part of this work (including the dipolar interaction and Ewald potential (37, 38)) we employ the random walk for all move processes to simplify the time consuming calculation of the force exerted on each particle.

The properties computed in these simulations are the equilibrium association constant (K), the pair correlation function for the oppositely charged ions (g_{AB}) and, in the simulation with one A and one B ion, the local probability density $\rho(r)$. The association equilibrium constant K is computed as :

$$K = \frac{\langle n_{AB} \rangle}{\langle n_A \rangle \langle n_B \rangle} * L_A * V \quad (10)$$

where $\langle n_{AB} \rangle$, $\langle n_A \rangle$ and $\langle n_B \rangle$ are the average number of molecules, ions A and ions B respectively. V is the volume of the system in the units of dm^{-3} and L_A is Avogadro's number.

In the simulation of the association of one A cation and one B anion, the local probability density is computed. The probability density is given by

$$\rho(r) = \frac{1}{4\pi r^2 \Delta r} \left(\sum_{i=1}^{\eta} h_i \frac{M_i}{M} \right)$$

where M_i is the number of times the distance between A and B falls within the i th histogram slot during the simulation, M is the total number of passes of the simulation, Δr is the grid size of the histogram chosen to be $0.5\text{\AA}$ and η is the number of divisions of the

histogram. The step function h_i is unity when the distance between the two particles corresponds to the i th histogram slot, otherwise $h_i = 0$. The units of $\rho(r)$ are $\text{\AA}^{-3}$.

For the many body simulations of the Sticky Electrolyte Model, the ionic correlation function g_{AB} is

$$g_{AB} = \frac{V}{4\pi r^2 \langle n_B \rangle} \left\langle \frac{dn_B}{dr} \right\rangle = \frac{V}{4\pi (r + \Delta r/2)^2 \langle n_B \rangle} \left(\frac{\langle \Delta n_B(r) \rangle}{\Delta r} \right) \quad (11)$$

where Δr is the width of the histogram used to calculate the correlation function. $\langle \Delta n_B(r) \rangle$ is the average number of anions (B) found between r and $r + \Delta r$ away from a cation (A). Δr is selected as $0.6 \text{\AA}$.

Results and Discussion

Test of the SSM Approach--One A and One B case

We use this simplified system to test the accuracy and the rate of convergence of the SSM method. The rate of convergence is determined from the cumulative deviation of the particle density and equilibrium association constant (see equation 8, Figure 6 and Figure 9 of the previous paper (33)). The parameters used for the potential (Figure 1) are: $\epsilon_2 = 5kT$, $d_{\max} = 7 \text{\AA}$, $d_{\min} = 3 \text{\AA}$, $\Delta = 12.5 \text{\AA}$, $\sigma = 7 \text{\AA}$ and the bond distance $L = 5.0 \pm 2.0 \text{\AA}$. In Figure 2, the simulated probability density ($\rho(r)$) is compared with the exact solution for this potential. The exact partition function is calculated by numerical integration using Simpson's method. The three dimensional integration is simplified to two dimensions before integration. In the expansion series, at least 100 terms are used to guarantee the convergence for the integration. Following 3×10^6 moves, the simulation results reproduce the exact distribution. The rate of convergence of the probability density and the average number of the uncharged molecules (AB) are shown in Figure 3 as a function of the parameter P . The convergence of the probability density and the average number of uncharged molecules (N_m) are sensitive to the simulation parameters. A low value of P , which corresponds to

sampling the association and dissociation processes, produces a rapid convergence in the molecule number but gives a slower rate of convergence for the density of probability. For the larger P the rates of convergence are reversed. These observations are consistent with our previous results where the rates of convergence of the equilibrium constant and the particle density displayed opposite trends (33). The smaller P is close to the free energy difference method, which has a faster rate of convergence for the simulation of the ratio of the partition function of two subspaces but a slower rate for the probability density. However, the convergence of the molecule number is established after one million passes, independent of the choice of P . Thus, the greater P more efficiently builds up the probability density.

We compared the rate of convergence of the average molecule number computed using SSM and the Metropolis algorithm for a potential with a deeper well. In the potential shown in Figure 1, we increased the magnitude of ϵ_2 to $10kT$. For this potential function, the Metropolis method always exhibits a slow and fluctuating convergence, regardless of the step size (Figure 4). The standard error of the mean from the Metropolis method, in Figure 4, is 0.067 but the corresponding standard error from S.S.M. is only 0.0039. SSM exhibits a faster rate of convergence and a smaller standard error.

In Table 1 we compile the results of several simulations for various potential parameters using the Subspace Sampling Method. For all choices of the potential, SSM reproduces the average particle number extraordinarily accurately. The accuracy of these results, as compared with the Metropolis method, are a consequence of a high acceptance probability for the transfer processes. For example, in Figure 4, the acceptance probability of the transferring processes between the two subspaces is as great as 3.5% for S.S.M. method. On the other hand, it is merely around 0.02% for Metropolis method.

The Subspace Sampling Method accurately and efficiently reproduces the average molecule number and the pair correlation for this one ion pair systems. We generalize the method to an arbitrary number of particles in the next section.

Simulation of the Sticky Electrolyte Model without dipolar interactions.

For the simulation of a realistic model of the divalent ion association, we used the potential in Figure 1 with the following parameters: $\epsilon_2=13.4228kT$, $d_{min}=1.89\text{\AA}$, $d_{max}=2.31\text{\AA}$, $\sigma=4.2\text{\AA}$, $\epsilon=78$ and the bond length $L=2.1\pm0.21\text{\AA}$. The conditions are chosen to be the same as the MSA/HNC approximation for the divalent weak electrolyte ions with bond distance $L=\sigma/2\pm0.05\sigma$ (1). The potential energy is given in equations 1-4. The initial concentrations (C_{int}) correspond to those using the MSA/HNC approximation. C_{int} represents the concentration of the binary molecules in the system with no dissociated ions. The ion pair concentration represents the concentration of unassociated cations.

The Subspace Sampling Method exhibits rapid convergence in the average number of molecules and in the total energy of the system. As shown in Figure 5, the average molecule number and energy converge after 1×10^6 moves for the system initially with 12 divalent cations and anions. For this system, the Metropolis method fails to achieve convergence in either the energy or the molecule number following 8×10^6 moves. Thus S.S.M. produces a more accurate and efficient method to calculate the equilibrium constant from the average number of molecules.

The results in Figure 5 show that the convergence in the system energy is coincident with convergence in the molecule number (N_{mol}). The system energy cannot converge before the N_{mol} is convergent. This demonstrates that the simulation method used for the weak electrolyte solution must achieve an accurate and efficient distribution of the particle number in order to calculate the system energy and other mechanical properties.

The pair correlation function

We computed the pair correlation function for the dissociated ions and the equilibrium association constant for several initial concentrations and several sizes of the system. In Figure 6, we show two pair correlation functions corresponding to two different cell sizes with the same initial molecular concentration. In the region near the inter-ionic contact point, the pair correlation functions are identical. However, the tail part of the correlation functions show a slight dependence on particle number (cell size). In each of these simulations, the total number of dissociated ions differs as a consequence of the finite size of the cell. To estimate the accuracy of the simulated g_{AB} , we compared the calculated g_{AB} with the corresponding Debye-Hückel correlation function. In Figure 6 the two solid curves represent two Debye-Hückel correlation functions with ionic concentrations of 0.003M and 0.02M for the upper and lower curves, respectively. The Debye-Hückel correlation functions bracket the calculated pair correlation functions. The ionic concentration in the simulations varied between 0.00339 and 0.00447 M. The similarity between the Debye-Hückel and the simulated correlation functions demonstrate that the ions of the weak electrolyte solution still behave like the restrictive primitive model. However, exact agreement between the Debye-Hückel and calculated pair correlation function does not occur because of the finite cubic cell used in the simulations. In the finite cubic cell, the tail part of the pair correlation function is truncated in comparison with the infinite system. Consequently, the correlation function of the finite system exhibits a dependence on the cell size.

Estimation of the Equilibrium Constant

We used the SSM method, in conjunction with an extrapolation method, to estimate the equilibrium association constant for the Sticky Electrolyte Model. In Table 2, the equilibrium constants for calculations that vary in the initial molecule concentrations and the system sizes (different total particle numbers, N_{tot}) are reported. The equilibrium constants apparently depend on the particle number (size of the system) as well as the initial

concentration of the system. The sensitivity of the equilibrium constant and the internal energy to the size of the system may be a consequence of the cut-off of the long range Coulomb potential. In simulations of the strong electrolyte solution, the internal energy of the system decreased as the size of the system increased (15, 16, 19). To determine the value of the equilibrium constant for an infinitely large system (at fixed initial concentration), we examined the limiting behavior of K when the size of the system (the total number of particles) increases as shown in Figure 7. A quadratic polynomial function is used to fit the data points and the limiting value of K is obtained by extrapolation to an infinite box size ($1/V \rightarrow 0$). The extrapolated values of the equilibrium constant from these plots are sensitive to the initial concentration.

The Thermodynamic Equilibrium Constant

The dependence of the association constant on initial molecule concentration is shown in Figure 8. The increase in K as the concentration decreases exhibits similar dependence as the K computed using the MSA/HNC equation. However, the SSM values for K exceed those of the analytical solution. This is presumably a consequence of the inclusion of dipolar terms in the configuration energy in the MSA/HNC equation. We examine the influence of these dipolar terms on the structural properties of the weak electrolyte solution in the subsequent section.

The dependence of K on concentration in both the MSA/HNC and SSM approaches, demonstrates the inability of both of these methods to produce a true thermodynamic equilibrium constant. In order to estimate the thermodynamic equilibrium constant from the SSM method we evaluated the activity coefficient of the relevant species. The thermodynamic equilibrium constant (K_T) can be expressed in terms of the simulated equilibrium constant K as:

$$K_T = \frac{\gamma_{AB}}{\gamma_{\pm}^2} \frac{N_{AB}}{N_A N_B} * V * L_A = \frac{\gamma_{AB}}{\gamma_{\pm}^2} * K \quad (12)$$

where γ_{AB} is the activity coefficient of the associated species and set equal to unity, $\gamma_{\pm}$ is the mean ionic activity coefficient of the dissociated species, L_A is Avogadro's number and K is in units of L/mol. Although the $\gamma_{\pm}$ in equation 12 is dominated by the Coulombic interactions among the dissociated ions (as in the strong electrolyte solution), $g_{\pm}$ exhibits a dependence on the simulation cell size. Figure 9 shows that $\ln K$ is a linear function of the square root of ionic concentration. However, for each of the initial concentrations the slope of the fitted linear equation varies. It is, thus, not possible to separate the cell-size dependence of $\gamma_{\pm}$ from its concentration dependence and obtain a true limiting value for $\gamma_{\pm}$.

To obtain an estimate of $\gamma_{\pm}$, we used the Debye-Hückel activity coefficients corresponding to the ionic concentration obtained from the limiting value of K for the three lowest values of C_{int} . For $C_{int}=0.5625M$, the ionic concentration is outside the Debye-Hückel limit. Consequently, we used the activity coefficient from the simulations of Valleau et al. for divalent ions in the restricted primitive model. Table 3 compares the simulated K and K_T (13). The error estimates reported in Table 3 are computed from the standard error of the mean number of particles obtained from block averages. Each block includes 100,000 steps. There are 80 blocks for each simulation. The small K_T for $C_{int}=0.5625 M$ arises from the slow convergence of K as a function of cell size which limits the accuracy of the extrapolated K . However, Figure 7 indicates that the maximum value of K_T is 8600 for $C_{int}=0.5625M$. We thus estimate K_T to be approximately 8400.

An alternative approach to obtain the thermodynamic equilibrium constant uses the pair correlation function. We attempted to fit the simulated pair correlation functions with the corresponding function from Debye-Hückel theory in order to generate an infinite dilution function. The infinite dilution ionic concentration estimated by direct extrapolation is approximately 0.524 times that of the concentration computed from the infinite dilution pair correlation function (i.e. by fitting with Debye-Hückel theory) for all initial concentrations. The number 0.524 is equal to the volume ratio of a sphere imbedded in a

cubic box to the box itself. The difference arises from the spherical symmetry in Debye-Hückel theory as compared with the cubic dimensions of the Monte Carlo cell. Moreover, the fluctuation in the pair correlation functions impedes the accuracy of the fit to the Debye-Hückel pair correlation function. Thus, the extrapolation of calculated equilibrium constant to an infinitely large cell size in conjunction with the Debye-Hückel activity coefficients yields a more accurate estimate of the thermodynamic equilibrium constant.

The simulation including the dipolar interaction

We modified the potential to include interactions among the molecules and between molecules and ions. The pair potentials included in the configuration energy are given by equations 1-4. In addition, the pair potential for the interaction between two AB molecules is given by:

$$U_{ij} = \frac{q_{Ai}q_{Bj}}{4\pi\epsilon_0\epsilon_r r_{AiBj}} + \frac{q_{Aj}q_{Bi}}{4\pi\epsilon_0\epsilon_r r_{AjBi}} \quad (13)$$

where r_{AiBj} is the distance between the cation in the i th molecule and the anion in the j th molecule. (The interaction energy between the cation and anion in the same molecule is not included in the electrostatic component of the configuration energy. This interaction energy is included only in the molecular part of the potential function.) The pair potentials for the interaction between an AB molecule and an unassociated ion (C) is given by

$$U_{ij} = \frac{q_{Ai}q_{Cj}}{4\pi\epsilon_0\epsilon_r r_{AiCj}} + \frac{q_{Bi}q_{Cj}}{4\pi\epsilon_0\epsilon_r r_{BiCj}} \quad (14)$$

The molecule energy of the AB molecule is still determined by the square well potential (equations 3-4).

In Figure 10 we show the rate of convergence of the average number of molecules and the system energy of 16 A and 16 B particles in the cell, thus the initial concentration is equal to 0.0625M. Convergence occurs after 3 million passes. We used the random walk to

sample configuration space for the move process. We computed the cation-anion site-site correlation function, which includes all inter-ionic and inter-molecular positive and negative charges except for those charges identified as positioned on the same dipolar molecule. Figure 11a shows this site-site correlation function for various total particle numbers in the system while the concentration is fixed at 0.0625M. The correlation function is independent of the cell size over the range of distances studied ($r < 24.2 \text{ \AA}$). In Figure 11b the ion pair correlation function is shown. The simulated correlation function is also independent of the total particle number (cell size). This differs significantly from the correlation function calculated without inclusion of the dipolar interaction. The Debye-Hückel correlation functions from two different ion concentrations are compared with simulated result in Figure 11b. The simulated ion pair correlation function behaves almost identically to the Debye-Hückel function.

The simulated equilibrium constants are shown in Figure 12 and compared with the case without dipolar interaction when $C_{\text{int}} = 0.0625 \text{ M}$. The limiting value of the equilibrium constant computed with the dipolar interaction appears to converge to that computed from the Sticky Electrolyte Model without dipolar interactions. However, for a finite cell size, the inclusion of the dipolar interaction reduces the computed equilibrium constant below that computed without the dipolar interaction. Furthermore, when the cell size is small, the total number of molecules is small (for fixed C_{int}) and the equilibrium constant is substantially below that of the SEM without dipolar interactions. When the total number of particles increases for fixed C_{int} (by an increase in the cell size) the number of molecules increases more rapidly than the number of ions. The internal energy of the molecule, modeled in these simulations as a square well, contributes substantially to the total energy when the system has a greater number of total particles. Consequently, the dipolar contribution to the total energy becomes less significant as the number of molecules increases. Hence the equilibrium constant computed with the dipolar interaction approaches that computed without the dipolar interaction.

Comparison with the Analytical Results

We compared the simulated and analytical pair correlation function at the contact point between oppositely charged particles. The contact density is computed from the pair correlation function (equation 11) where all cation-anion distances are included except for the distances originating from an individual molecule. Thus, the site-site correlation function includes cation-anion correlations between free ions, between one free ion and the opposite charge located on a dipolar molecular and between opposite charges of two different dipolar molecules. We denote this site-site correlation function g_{+-} to distinguish it from the correlation function composed only of free cations and anions (g_{AB}).

The site-site correlation function for the dipolar interaction (Figure 11a) is approximately 1.9, which is smaller than the result ($g_{+-}(\sigma)=2.5$) of the analytic solution (MSA/HNC) when the initial concentration is 0.0625 M. Both values for $g_{+-}(\sigma)$ are substantially below the value from the simulation without the dipolar interaction. The significant reduction of $g_{+-}(\sigma)$ with the dipolar interaction is primarily a consequence of the correlation between cations (or anions) of one molecule with the oppositely charged ions of another molecule or a free ion. Included in the computation of the correlation functions are all correlations between oppositely charged ions except for those ion pairs defined as molecules. Because of the large equilibrium association constant the major species in these simulations are the molecules. Consequently, the site-site correlation function is dominated by correlations between oppositely charged ends of different molecules or between molecules and single ions.

In this work, the reduction of the contact site-site correlation function $g_{+-}(\sigma)$ from the MC simulations, as compared with the analytical solutions, is consistent with the corresponding increase in the simulated equilibrium constant above that of the analytical K . Examining the interaction in Figure 13, the site-site correlation function is least favorable for the dipolar-dipolar interaction due to their steric hindrance and repulsive force. The ion-

ion interaction has the lowest energy and has no steric hindrance. The ion-dipole interaction is between the above two cases. The greater equilibrium K produces more molecules and an increased number of dipolar-dipolar interactions. Consequently, the site-site correlation function should be smaller than the correlation function of the lower K .

The molecular correlation function is the primary determinant of the equilibrium constant. Both the Hypernetted Chain/Mean Sphere Approximation and the Hypernetted Chain approximation produce lower equilibrium constants than the MC simulations. The greater K from the MC simulations, both with and without dipolar interactions, may arise because 1) the HNC approximation is not appropriate for the molecular potential, and/or 2) a δ -function is employed to describe the molecular correlation function in the HNC results. The Hypernetted Chain approximation has led to better results than the PY approximation for the molecular potential of extremely short range interactions (1,2). However, HNC is still not completely consistent with our simulation results. On the other hand, the contact correlation function of MSA/HNC method more closely resembles the MC results than that of HNC alone. MSA/HNC seems consistent with the dynamical behavior of liquid state as well. The frequent association and dissociation processes of the weak electrolyte in solution may impede the formation of trimers or tetramers (which were observed in the HNC results) and lead to a smaller correlation function as observed in the MC simulations.

Although MSA is not appropriate for the ionic correlation function (16) of a strong electrolyte solution it may be useful for the dipolar-dipolar interaction which is demonstrated in the work from Rasaiah's laboratory (1-8). Since there was no appropriate simulation result available for dilute solutions of weak electrolytes, the accuracy of the correlation function from the MSA calculations was estimated for extremely high concentrations of weak electrolytes by comparison with Monte Carlo results for the dipolar liquid (2-4). This comparison assumes that the high concentration, weak electrolyte system behaves similarly to the dipolar liquid as long as the dielectric constant of the two systems

are similar. For high concentrations of electrolyte, the MSA site-site correlation function is similar to the correlation function of the dipolar liquid suggesting that MSA is a reasonably accurate approximation for the high concentration, weak electrolyte solution when dipolar interactions are included. The similarity between the MC and the MSA/HNC results for low concentrations of the weak electrolyte solution suggests that the MSA/HNC approximation is fairly accurate over a wide range of electrolyte concentrations.

The Monte Carlo approach, in addition to the other statistical mechanical approaches to the weak electrolyte solution, approximate the long-range interactions of the Coulomb potential to some extent. Consequently, the equilibrium constant is sensitive to the size of the system simulated. Inclusion of the Ewald potential can enhance the contribution from long-range interactions. However, the Ewald potential contributes minimally when the ion pair concentration is low (16) as in all our simulations of the weak electrolyte solution. The inclusion of the dipolar terms in the SEM also contain a long-range potential. We incorporated the Ewald potential into our simulations with the dipolar terms and found that inclusion of the Ewald potential did not alter the computation of K or the pair potential (Table 4) presumably because the molecular potential dominates the system energy. Consequently, the use of the Ewald potential did not limit the cell-size dependence of K nor did it improve the accuracy of the pair correlation function. An accurate estimation of an infinite dilution equilibrium constant from a simulation with a finite cell size requires an accurate calculation of the tail part of the pair correlation function. The Ewald potential does not facilitate an accurate estimate of the tail of the pair correlation function at the ionic or dipolar concentrations simulated in this study.

Conclusions

We performed Monte Carlo simulations of the divalent weak electrolyte solution using the Sticky Electrolyte Model. We demonstrated the ability of the Subspace Sampling Method (SSM) to simulate efficiently and accurately the Sticky Electrolyte Model. The

basis of SSM is the separation of the transferring processes (i.e. association and dissociation processes) from the move processes. The transferring processes have a low acceptance probability for the potential model used in this study. To enhance the total number of transferring processes, the constant c in the transferring transition probability is maximized. This directly facilitates the transferring processes between two subspaces in a manner analogous to the free energy difference method (39-40) and results in the efficient computation of the average number of molecules (or the equilibrium constant) in each subspace. By maximizing the transferring constant, the SSM parameter set that must be optimized is reduced to two parameters (P and Δ). These parameters were then optimized for efficient computation of the correlation function or probability density in addition to the equilibrium constant. We show that SSM is a Monte Carlo methodology that can be employed to accurately and efficiently compute both the correlation function and the equilibrium constant for the SEM which contains a deep and narrow potential well.

We used the optimum SSM parameter set to compute the equilibrium constant for the Sticky Electrolyte Model with and without dipolar interactions. The simulated equilibrium constant is dependent upon the system size and/or total number of particle. The limiting value of the equilibrium constant is obtained by extrapolating the computed K to an infinitely large cell size. The thermodynamic equilibrium constant is estimated from the appropriate activity coefficients. We compared our simulated equilibrium constant with the results of the MSA/HNC and HNC approximation (2). In agreement with the MSA/HNC or HNC results, our computed equilibrium constants decreased with increasing concentration of weak electrolyte. However, when the dipolar interaction between ion pairs was excluded, the Monte Carlo values of the equilibrium constant are one order of magnitude greater than the corresponding MSA/HNC or HNC solutions. Furthermore, the simulated contact site-site correlation function $g_{AB}(\sigma)$ is much higher than the analytical solution.

The difference between the MC and MSA/HNC site-site correlation function is a consequence of the dipolar interactions between ion pairs. When dipolar interactions are included in the potential function, the site-site correlation function is dominated by the dipolar molecules due to the large K . The Monte Carlo equilibrium constant exceeds that of the MSA/HNC or the HNC alone calculation. As a consequence of the increase in the number of dipolar molecules in the Monte Carlo simulations, there is a corresponding reduction in the site-site correlation functions at the contact distance below either the MSA/HNC or HNC predictions. However, the MC contact correlation function is more consistent with the MSA/HNC result.

At finite concentrations, the equilibrium constant computed with the dipolar potential is lower than that of the equilibrium constant computed without the dipolar terms. However, the infinite box size equilibrium constant appears to be insensitive to the inclusion of the dipolar interactions. As the size of the system increases (at fixed weak electrolyte concentration) the molecular potential (the square well) dominates the total energy of the system as compared with the coulomb energy for the dipolar and ionic interaction. This results in a limiting equilibrium constant that appears independent, within computational error, of the dipolar terms.

References

- (1) Lee, S. H.; Rasaiah, J. C. and Cummings, P. T. *J. Chem. Phys.* **1985**, *83*, 317.
- (2) Rasaiah, J. C.; Lee, S. H.; *J. Chem. Phys.* **1985**, *83*, 5870.
- (3) Rasaiah, J. C.; Lee, S. H.; *J. Chem. Phys.* **1985**, *83*, 6396.
- (4) Rasaiah, J. C.; Lee, S. H.; *J. Chem. Phys.* **1987**, *86*, 983.
- (5) Rasaiah, J. C.; Zhu, J.; Lee, S. H. *J. Chem. Phys.* **1989**, *91*, 495.
- (6) Rasaiah, J.C.; Zhu, J.; Lee, S.H. *J. Chem. Phys.* **1989**, *91*, 505.
- (7) Rasaiah, J.C.; Zhu, J. *J. Chem. Phys.* **1991**, *94*, 3141.
- (8) Rasaiah, J.C., *Int. J. Thermophysics*, **1990**, *11*, 1.
- (9) Herrer, J. N.; Blum, L. *J. Chem. Phys.* **1991**, *94*, 5077.
- (10) Pizio, O. A. *J. Chem. Phys.* **1984**, *100*, 548.
- (11) Cummings, P. T. and Stell, F. *Mol. Phys.* **1984**, *51*, 253.
- (12) Cummings, P. T. and Stell, F. *Mol. Phys.* **1985**, *55*, 33.
- (13) Andersen, H. C. *J. Chem. Phys.* **1973**, *59*, 4714.
- (14) Hfye, J. S. and Olaussen, K. *Physica A* **1980**, *104*, 435.
- (15) Card, D. N. and Valleau, J. P. *J. Chem. Phys.* **1970**, *52*, 6232.
- (16) Valleau, J. P.; Cohen, L. K. and Card, D. N. *J. Chem. Phys.* **1980**, *72*, 5942.
- (17) Larsen, B. *Chem. Phys. Letters* **1974**, *27*, 47.
- (18) Rogde, S. A. *Chem. Phys. Letters* **1983**, *103*, 133.
- (19) Sorensen, T. S. *J. Chem. Soc. Faraday Trans.* **1990**, *86*, 1815.
- (20) Valleau, J. P.; Ivkov, R. and Torrie, G. M.; *J. Chem. Phys.* **1991**, *95*, 520, (and references therein).
- (21) Linse, P.; Gunnarsson, G.; Jönsson, B. *J. Phys. Chem.* **1982**, *85*, 413.
- (22) Linse, P.; Jönsson, B. *J. Chem. Phys.* **1983**, *78*, 3167.
- (23) Bratko, D.; Vlachy, V. *Chem. Phys. Lett.* **1982**, *90*, 434.
- (24) LeBret, M.; Zimm, B. H. *Biopolymers* **1984**, *23*, 271.

- (25) Murthy, C. S.; Bacquet, R. J.; Rossky, P.J. *J. Phys. Chem.* **1985**, *89*, 701.
- (26) Mills, P. A.; Anderson, C. F.; Record, M. T., Jr. *J. Phys. Chem.* **1985**, *89*, 3984.
- (27) Mills, P. A.; Anderson, C. F.; Record, M. T., Jr. *J. Phys. Chem.* **1986**, *90*, 6541.
- (28) Bacquet, R. J.; Rossky, P. J. *J. Phys. Chem.* **1988**, *92*, 3604.
- (29) Sharp, K. A.; Jean-Charles, J.; Honig, B. *J. Phys. Chem.* **1992**, *96*, 3822.
- (30) Sitkoff, D.; Sharp, K. A.; Honig, B.. *J. Phys. Chem.* **1994**, *97*, 1978.
- (31) Cao, J. and Berne, B. *J. Chem. Phys.* **1990**, *92*, 1980.
- (32) Shew, C.-Y. and Mills, P. *J. Phys. Chem.* **1993**, *97*, 13824.
- (33) Shew, C.-Y. and Mills, P. *J. Phys. Chem.* (submitted).
- (34) Smith, W. R. and Triska, B. *J. Chem. Phys.* **1994**, *100*, 3019.
- (35) Tsangaris, D.M.; de Pablo, J.J. *J. Chem. Phys.* **1994**, *101*, 1477.
- (36) Metropolis, N.; Rosenbluth, A. W.; Rosenbluth, M. N.; Teller, A. H.; Teller, E. *J. Chem. Phys.* **1953**, *21*, 1087.
- (37) Brush, S. G.; Sahlin, H. L.; Teller, E. *J. Chem. Phys.* **1966**, *45*, 2102.
- (38) Adams, D. J. Duber, G. S. *J. Comp. Phys.* **1987**, *72*, 156.
- (39) Voter, A.F. *J. Chem. Phys.* **1985**, *82* , 1890.
- (40) Bennet, C.H. *J. Comp. Phys.* **1976**, *22*, 245.
- (41) Rasaiah, J. C.J. *J. Chem. Phys.* **1972**, *56*, 3071.

Figure Captions

Figure 1. The AB pair interaction energy (equation 1, text) for the Sticky Electrolyte Model.

Figure 2. The simulated probability density for one A and one B when $\epsilon_2=5kT$, $d_{min}=3\text{\AA}$, $d_{max}=7\text{\AA}$, $\sigma=7\text{\AA}$, $\epsilon=78$, $p=0.3$, passes=3 million. O : simulated result; + : exact solution.

Figure 3(A). The chi-square test for convergence of the probability density computed during the simulation with the same parameters in Figure 2 for several choices of p. O : $p=0.1$; + : $p=0.5$; $\square$: $p=0.9$.

Figure 3(B). The cumulative deviation of the average number of molecules during the simulation. Symbols and the parameters are the same as in Figure 3a.

Figure 4. Comparison between the rate of convergence of the average number of molecules for the Subspace Sampling Method and the Metropolis Method when $\epsilon_2=10kT$, $d_{min}=3\text{\AA}$, $d_{max}=7\text{\AA}$, $\sigma=7\text{\AA}$, $\epsilon=78$. + : Metropolis Method $\Delta=25\text{\AA}$; $\blacktriangle$: Subspace Sampling Method for $p=0.1$ and $\Delta=12.5\text{\AA}$.

Figure 5. The comparison of the average number of molecules and the system energy for a total electrolyte concentration $[AB]=0.5625M$ (12A and 12 B particles) and parameters $\sigma=4.2\text{\AA}$, $d_{max}=2.31\text{\AA}$, $d_{min}=1.89\text{\AA}$ and $L=\sigma/2\pm 0.05\sigma=2.1\pm 0.21\text{\AA}$ computed using the Subspace Sampling Method and the Metropolis Method. + : Metropolis Method with $\Delta=12.5\text{\AA}$; O: Subspace Sampling Method for $p=0.3$ and $\Delta=3.125\text{\AA}$.

Figure 6. The simulated AB pair correlation function by Subspace Sampling Method when the initial concentration is 0.0625M and the square well potential parameters correspond to figure 5. o: 8A and 8B total particles in the system; $\blacktriangle$: 24A and 24B

maximum particles. The solid curves represent two different Debye-Hückel correlation functions. The concentrations are 0.003 M for the upper curve and 0.02 M for the lower one.

Figure 7. The plot of the equilibrium constant vs. the inverse of the volume of the Monte Carlo cell for three fixed electrolyte concentrations and the same parameters as figure 5 to obtain the limiting value of K. The initial concentrations correspond to: O : 0.00984M; Δ : 0.0625M; $\square$: 0.5625M.

Figure 8. Comparison of the equilibrium constant of the Monte Carlo simulation with the analytical solution at four initial electrolyte concentrations. O : results from the MC simulation; $\blacksquare$: results from the HNC/MSA approximation.

Figure 9. The linear relationship of $\ln K$ to the square root of the ionic concentration when $C_{int}=0.00984$ M.

Figure 10. The rate of convergence of the average number of molecules and system energy for the simulation including the dipolar interaction at an initial concentration of 0.0625M (16 A and 16 B particles) with $\sigma=4.2\text{\AA}$, $d_{max}=2.31\text{\AA}$, $d_{min}=1.89\text{\AA}$ and $L=2.1\pm0.21\text{\AA}$.

Figure 11(A). The site-site correlation function with the dipolar interactions included for two cell sizes using the S.S.M./Random Walk algorithm when $C_{int}=0.0625$ M. $\bullet$: 16 A and 16B particles; $\square$: 32A and 32B particles.

Figure 11(B). The ion-ion pair correlation function for the same parameter set as in Figure 11a. The solid curves represent the Debye-Hückel correlation functions shown in Figure 6.

Figure 12. Comparison of the simulated equilibrium constant with and without dipolar interactions for different total particle numbers (cell size) when $C_{int}=0.0625$ M. O: without dipolar interactions; Δ : dipolar interactions included.

Figure 13. Schematic of the of dipole-dipole, ion-dipole and ion-ion interactions used to rationalize the site-site correlation function in terms of energy and steric hindrance. The energy of each model is 0.208 kT for (a); 3.25 kT for (b); and 6.84 kT for (c) when the center-to-center distance between two particles in a molecule is 0.42 Å.

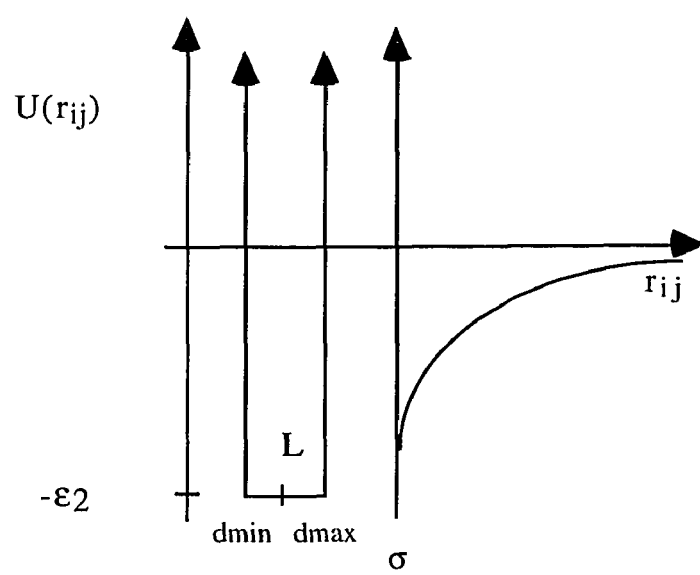


Figure 1

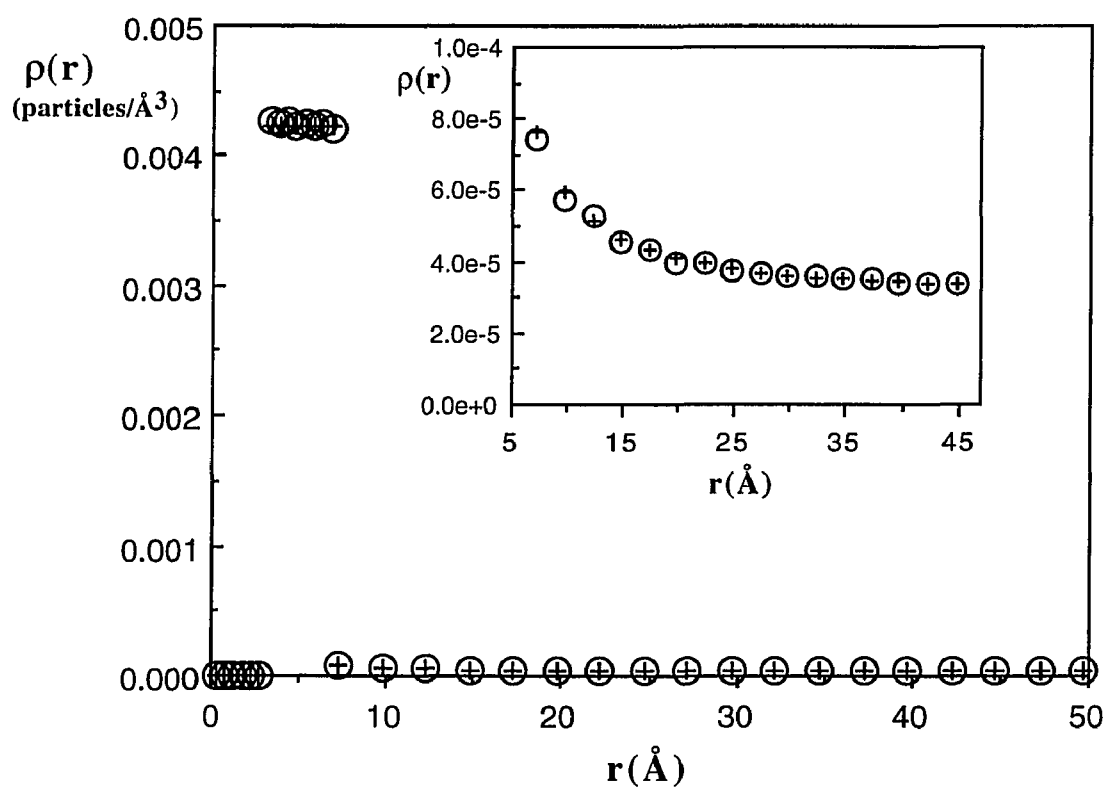


Figure 2

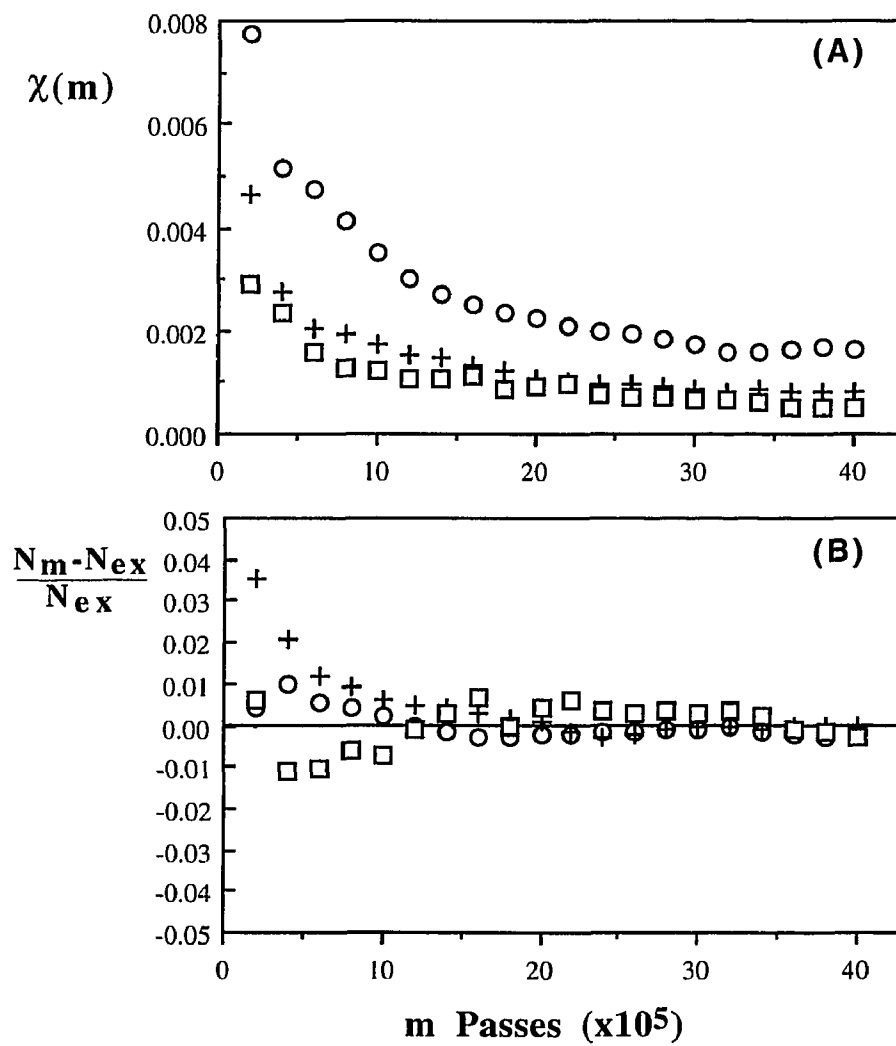


Figure 3

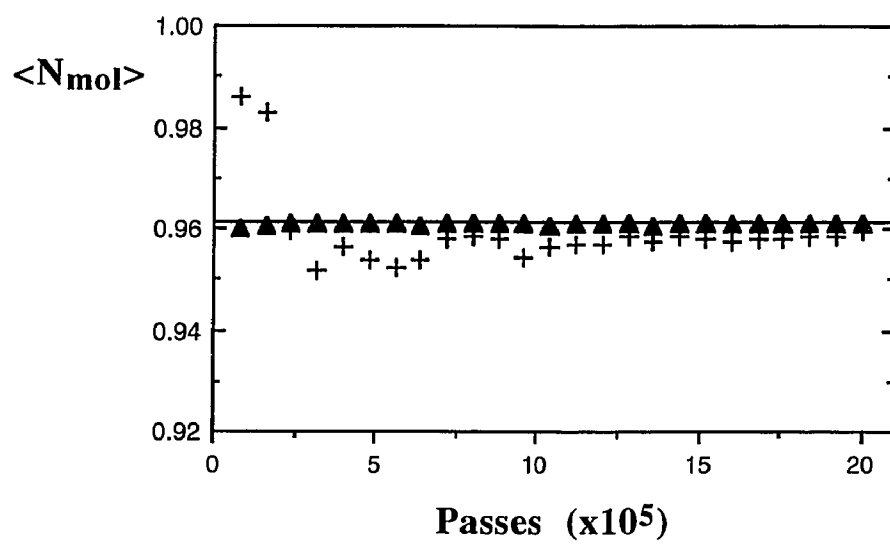


Figure 4

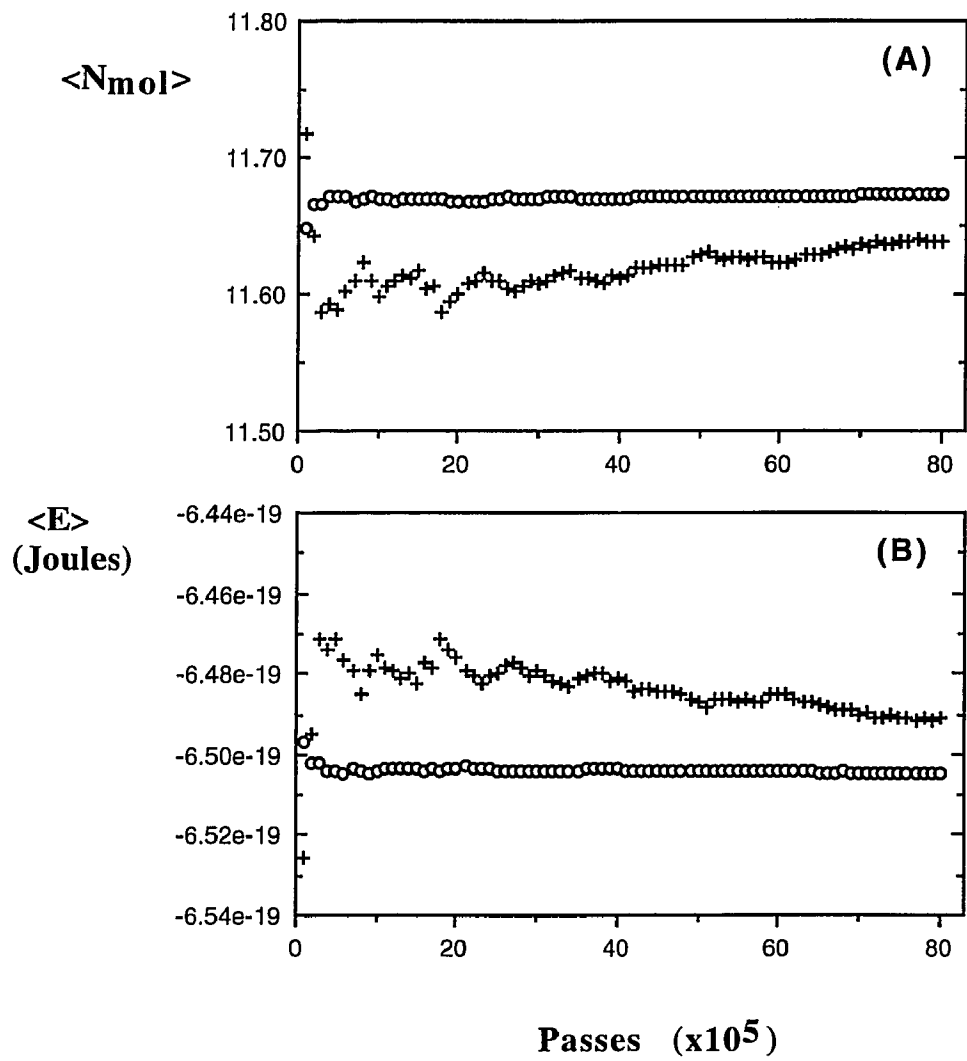


Figure 5

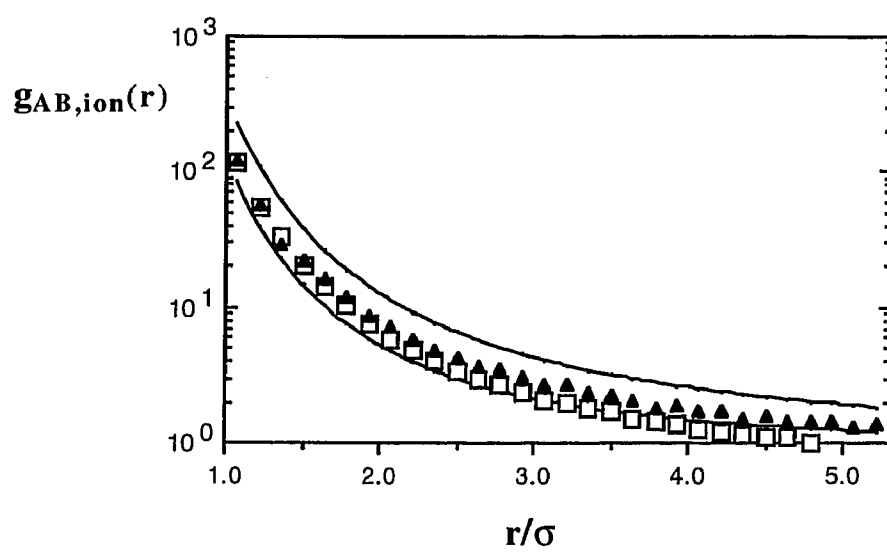


Figure 6

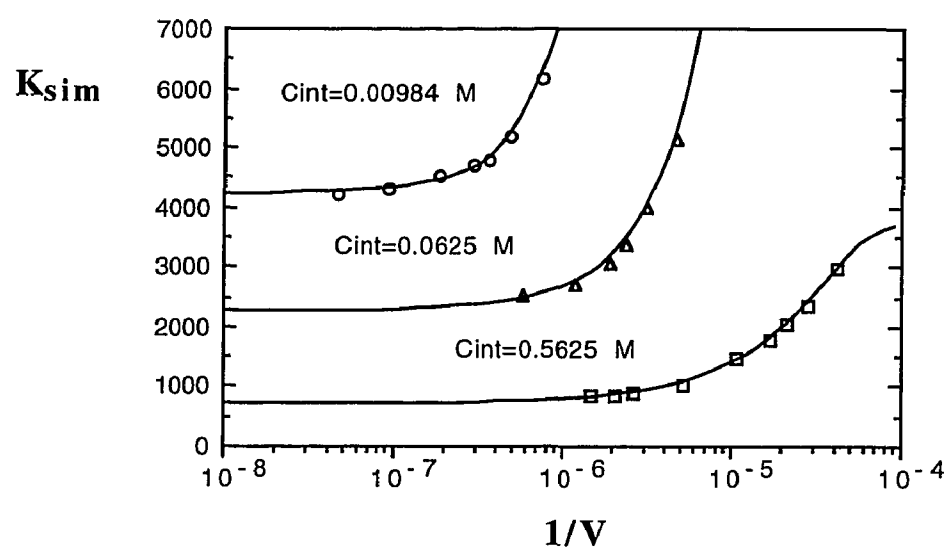


Figure 7

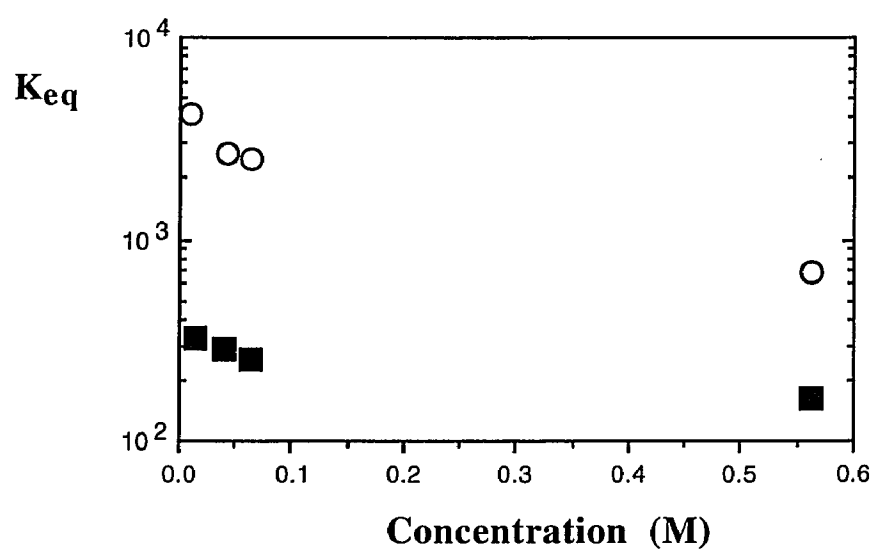


Figure 8

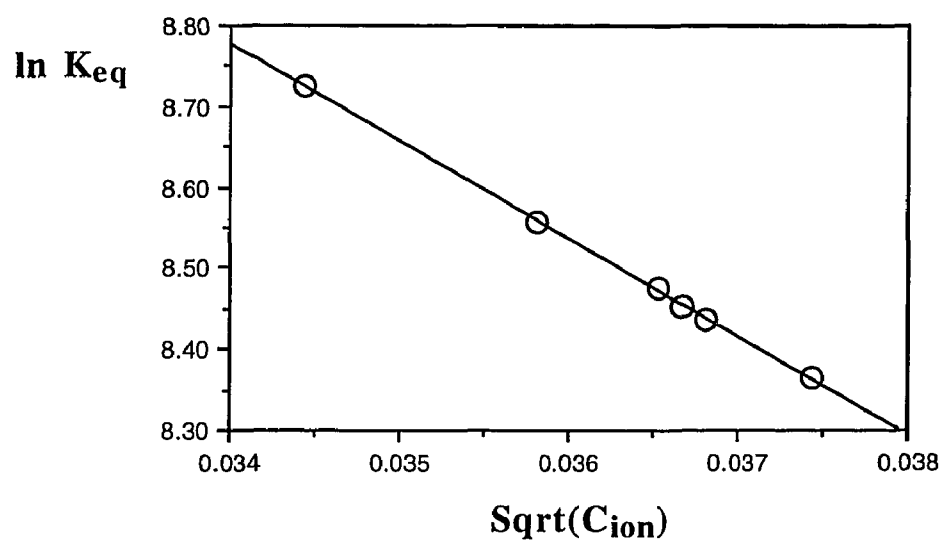


Figure 9

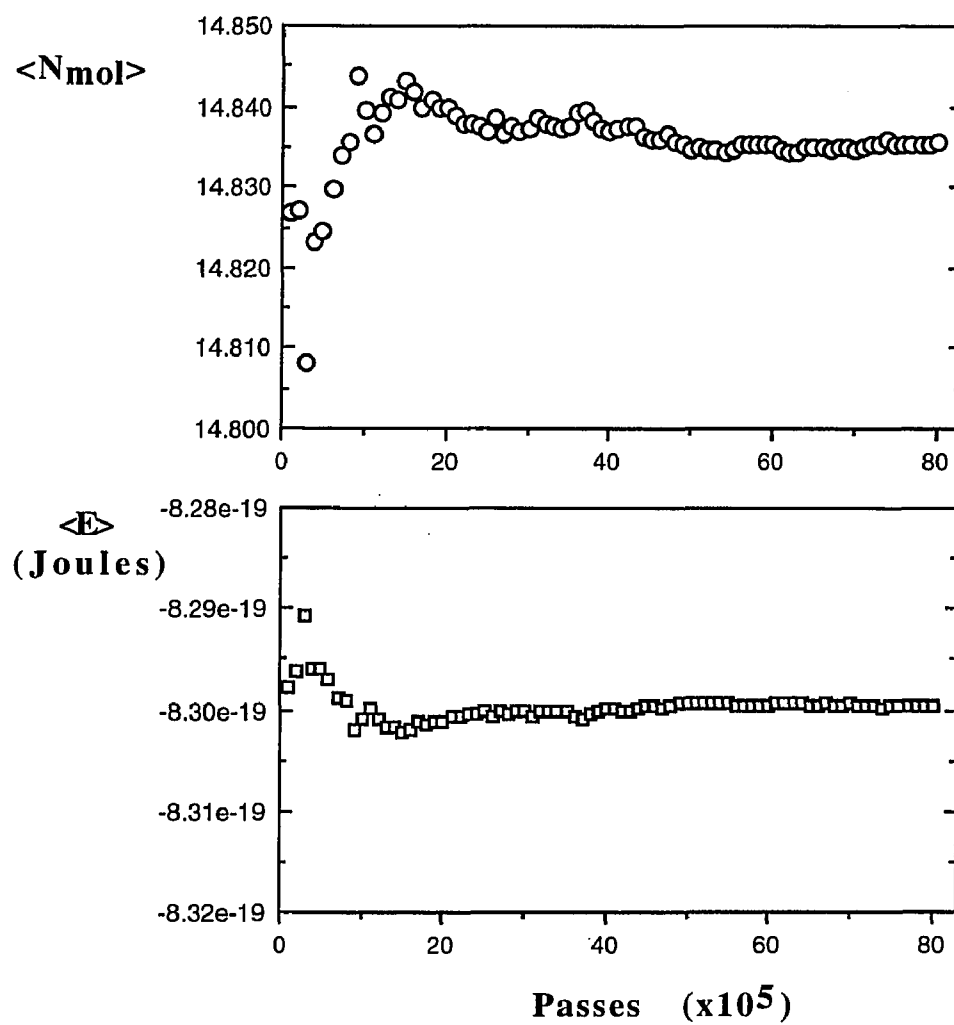


Figure 10

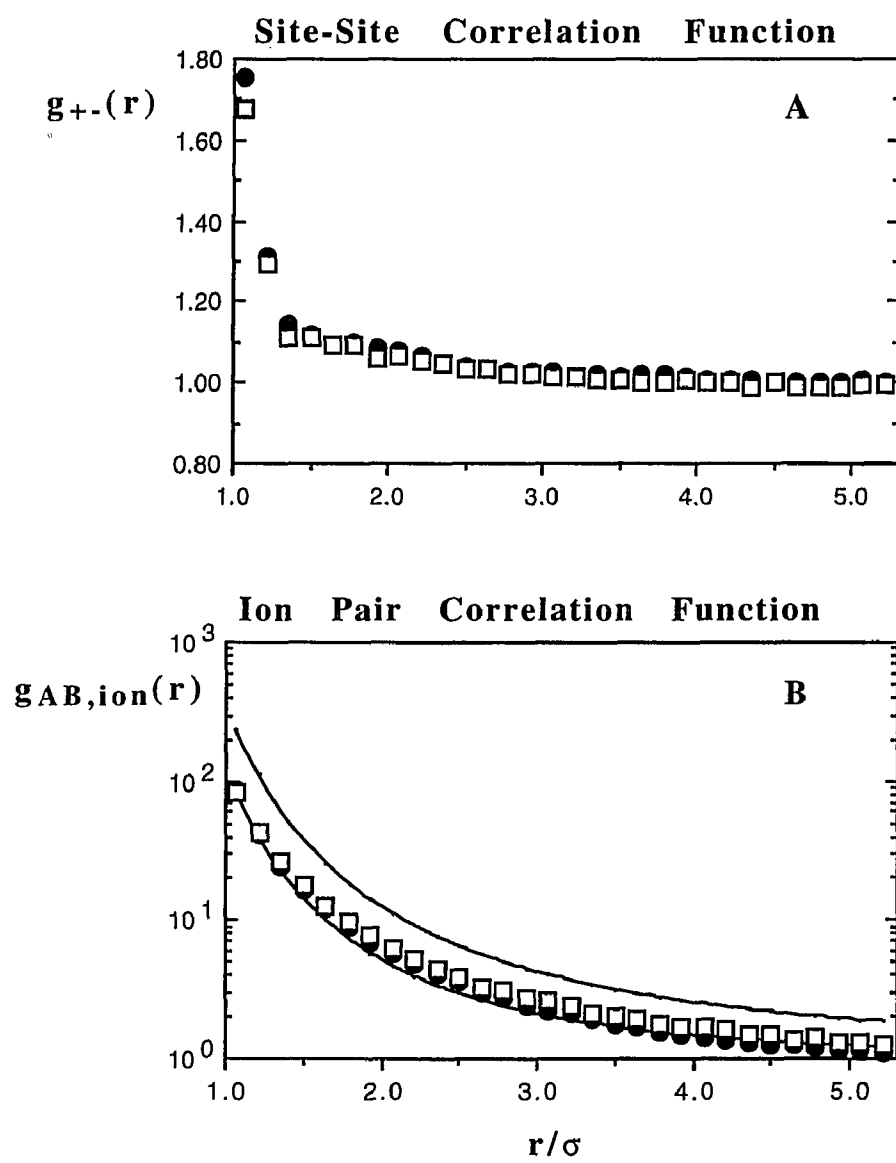


Figure 11

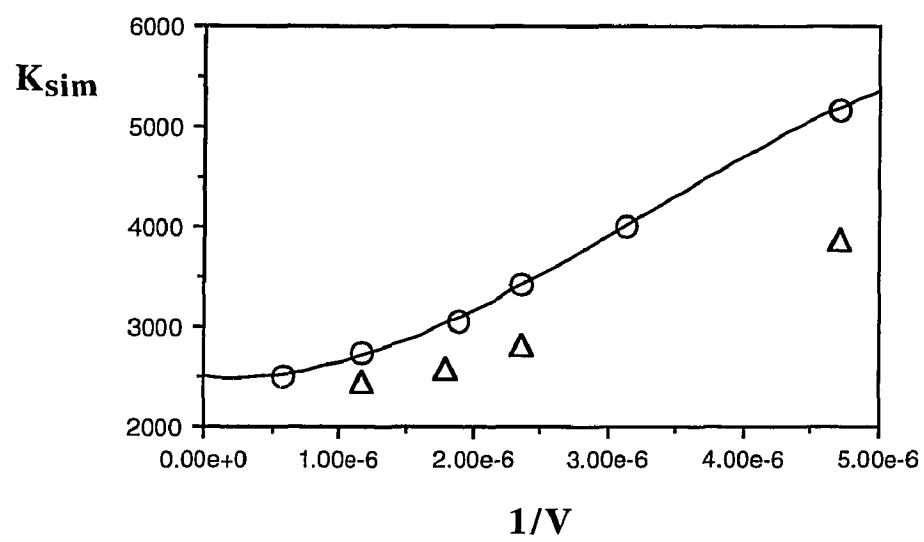


Figure 12

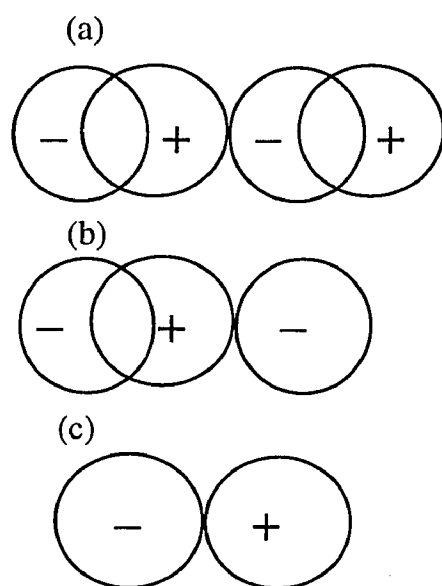


Figure 13

Table 1
Comparison of the Average Number of Molecules
Computed for the Simulation of One Cation and One Anion with
the Exact Value

$\epsilon_2(\text{kT})$	w (Å)	$\langle N_{\text{mol}} \rangle$	exact $\langle N_{\text{mol}} \rangle$
0.5	4	0.00185 (0.00094)	0.001838
2.5	4	0.0134 (0.0024)	0.13420
10	4	0.9610 (0.0039)	0.96093
0.5	11	0.0158 (0.0027)	0.01584
2.5	11	0.1064 (0.0059)	0.10640
5	11	0.5912 (0.0069)	0.59162
10	11	0.9953 (0.0015)	0.99537

Table 2
Calculated Equilibrium Constant and Average Energy
for Simulations of Different Cell Size and Particle Number
when $L = \frac{\sigma}{2} \pm 0.05\sigma$

Table 2a
Equilibrium Constant and Average Energy
Computed without Dipolar Interaction

Conc.(M)	P	Δ (Å)	N_{total}	$\langle K \rangle$	$-\langle E \rangle / N_{\text{total}} k T$
0.00984	0.5	50/12	16	6157 (295)	5.981 (0.012)
	0.5	50/16	24	5196 (219)	5.897 (0.011)
	0.5	50/16	32	4784 (216)	5.881 (0.013)
	0.3	50/12	40	4694 (174)	5.869 (0.012)
	0.2	50/6	64	4509 (157)	5.855 (0.012)
	0.2	50/6	128	4297 (180)	5.819 (0.014)
	0.3	50/12	256	4227 (179)	5.830 (0.014)
0.0625	0.1	50/12	16	5157 (310)	6.396 (0.007)
	0.5	50/12	24	3983 (258)	6.369 (0.008)
	0.5	50/16	32	3386 (175)	6.334 (0.007)
	0.5	50/16	40	3041 (187)	6.311 (0.009)
	0.2	50/24	64	2713 (140)	6.288 (0.008)
	0.2	50/12	128	2507 (141)	6.265 (0.010)
0.5625	0.5	50/16	16	2948 (214)	6.597 (0.003)
	0.1	50/16	24	2346 (126)	6.586 (0.003)
	0.5	50/16	32	2024 (166)	6.566 (0.005)
	0.3	50/12	40	1757 (115)	6.557 (0.004)
	0.3	50/24	64	1453 (94)	6.539 (0.004)
	0.2	50/12	128	1035 (61)	6.505 (0.005)
	0.2	50/12	256	884 (64)	6.487 (0.007)
	0.2	50/12	330	857 (57)	6.481 (0.006)
	0.2	50/16	448	822 (53)	6.475 (0.007)

Table 2b
Equilibrium Constant and Average Energy
Computed with Dipolar Interaction

Conc.(M)	P	Δ (Å)	N_{total}	$\langle K \rangle$	$-\langle E \rangle / N_{\text{total}} k T$
0.0625	0.2	50/12	16	3856 (164)	6.375 (0.005)
	0.2	50/12	32	2803 (131)	6.303 (0.005)
	0.2	50/16	42	2570 (334)	6.280 (0.021)
	0.2	50/16	64	2441 (95)	6.268 (0.007)

Table 3
Comparison of the extrapolated equilibrium constant
and the thermodynamic equilibrium constant

C_{int} (M)	K_{limiting}	γ_{+-}^2	K_T
0.00984	4200 (300)	0.4926*	8400 (600)
0.0425	2700 (200)	0.3153*	8600 (600)
0.0625	2300 (200)	0.2722*	8400 (700)
0.5625	700 (50)	0.0866†	8000 (600)

*Computed from the Debye-Hückel limiting law.

†Obtained from Monte Carlo simulations (Ref. 13)

Table 4
Comparison of the Simulation Results Using the Ewald
Potential with the Minimum Image Method for the Coulomb
Interaction for $C_{int} = 0.0625 \text{ M}^*$

Max. NO. of Ions (N_{tot})	Average Values	With Dipolar Interactions		Without Dipolar Interactions	
		Ewald	MI	Ewald	MI
$N_{tot}=32$	$\langle N_{mol} \rangle$	14.836	14.909	14.878	14.937
	$-\langle E_{tot} \rangle / N_{tot} kT$	6.303	6.331	6.316	6.334
	$-\langle E_{coulomb} \rangle / N_{tot} kT$	0.089	0.077	0.065	0.068
$N_{tot}=64$	$\langle N_{mol} \rangle$	29.512	29.578	29.577	29.635
	$-\langle E_{tot} \rangle / N_{tot} kT$	6.268	6.284	6.277	6.288
	$-\langle E_{coulomb} \rangle / N_{tot} kT$	0.078	0.081	0.074	0.073

* Parameters Used: $\epsilon_2=13.4228kT$, $L=2.1\pm0.21\text{\AA}$ and $T=298 \text{ K}$

Appendix

Fortran Code for the Subspace Sampling Method

```
c
c Simulation of the Equilibrium Constant of Divalent Weak Electrolyte c
c --- By three dimensional Subspace Sampling Method c
c c Array x, y, z : Position of particles c
c Array ion : Original charge of particle (+2 or -2) c
c Array q : Current status of particle (+2,-2 for ions and 0 for c
c the particles on molecule) c
c Array mol : Specify the patnership of each particle (0 for ions) c
c Array es : Bond distance of specific molecule (0 for ions) c
c Array fs, fsx2, fsy2, fsz2 : table for pair distance and vector c
c cccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccc c
c dimension sum(120),rad(3),pmss(3),pbpc(4),pq(120) c
c dimension qip(2),ipc(2),ipm(2),xm(2),ym(2),zm(2),iqm(2) c
c dimension fs(349,349) c
c dimension fsx2(349,349),fsy2(349,349),fsz2(349,349) c
c common/coord/x(400),y(400),z(400) c
c common/coord1/q(400),ion(400),mol(400),es(400) c
c common/ff/fs c
c common/fxyz/fsx2,fsy2,fsz2 c
c common/constant1/temp,bk c
c common/constant3/a,b,c,step c
c common/energy/u1,u2 c
c common/constants/ix,ibx c
c common/para/dmin,dmax,depth,req,r1,r2 c
c common/ba1/qip,ipc,ipm,xm,ym,zm,iqm c
c common/ba2/esm(2),xm(2),ym(2),zm(2),iqm(2) c
c common/sum/sum c
c common/mass/pmss c
c common/ex/pq c
c cccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccc c
c Read all the variabes for the simulation system c
c c depth : the depth of square well potential (=13.4228kT) c
c dmin : the left boud of square well potential (=1.89) c
c dmax : the right bound of square well potential (=2.31) c
c temp : 298 c
c bk : Boltzmann constant c
c r1,r2 : radii of cations and anions respectively (r1,r2=2.1) c
c..n : 2*n is the total number of particles c
c cccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccc c
c open(14,file="bivfile") c
c read(14,*)n,ix,ibx c
c read(14,*)depth,bk c
c do 3 i=1,2*n c
c read(14,*)x(i),y(i),z(i) c
```



```
c
c
c nfs_ndedn file : simulated molecule number
c nfs_ndede file : simulated system energy
c nfs_ndednp file : correlation function of opposite charged particles
c nfs_ndedpp file : correlation function of same charged particles
c
ccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccc
c
      open(16,file="nfs_ndedn")
      open(17,file="nfs_ndede")
      open(15,file="nfs_ndednp")
      open(18,file="nfs_ndedpp")
do 9998 num=1,nos
c
ccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccc
c
c Calculate the probability to pick Association, Dissociation and
c Move Process
c Select the process and Choose the particle(s)
c Select the S.S.M (const1.eq.1) or Enhanced S.S.M (const1.eq.2)
c
ccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccc
c
      const1=1
      call probab(num,n,pbpc(2),pbpc(3),pbpc(4),prob)
      ipc(1)=0
      ipc(2)=0
      qip(1)=0
      qip(2)=0
      call randb(xx,24)
c
c Choosing move process
c
      if(xx.le.pbpc(2)) go to 200
c
c Choosing association process
c
      if(xx.gt.pbpc(2).and.xx.le.pbpc(3)) then
        sum(61)=sum(61)+1
        call randb(xxb,4)
        qip(1)=-2.
        if(xxb.gt.0.5) qip(1)=2.
        call site(num,n,ipc(1),qip(1),yest)
if (const1.eq.1) then
  if(yest.eq.2) go to 750
elseif (const1 .eq.2) then
  if(yest.eq.2) go to 200
end if
      mptc=1
      call setup(num,mptc,ipc(1))
      qip(2)=-1*qip(1)
      call site(num,n,ipc(2),qip(2),yest)
      mptc=2
      call setup(num,mptc,ipc(2))
      go to 400
```


[illegible]

```

c      subroutine newtona(rdist)
common/para/dmin,dmax,depth,req,r1,r2
7      call rand(x1,6)
      call randb(x2,6)
      call rand(x3,6)
      xxt=sqrt(x1*x1+x2*x2+x3*x3)
      rdist=xxt*dmax
      if(rdist.lt.dmin.or.rdist.ge.dmax) go to 7
9      return
      end
c
c
c
c      Dissociation Process
c      - Separate two ions on the molecules and place in the space
c
c
c
c
c      subroutine dissociate(num,n,m1)
common/coord/x(400),y(400),z(400)
common/coord1/q(400),ion(400),mol(400),es(400)
common/constant3/a,b,c,step
common/para/dmin,dmax,depth,req,r1,r2
      m2=mol(m1)
      call move(x(m1),y(m1),z(m1),a,2.)
      dstp=0.5*a
      call rand(xxb,16)
      x(m2)=x(m2)+(2*xxb-1)*dstp
      call rand(xxb,16)
      y(m2)=y(m2)+(2*xxb-1)*dstp
      call rand(xxb,16)
      z(m2)=z(m2)+(2*xxb-1)*dstp
      call boxlim(x(m2),y(m2),z(m2))
      mol(m1)=0
      mol(m2)=0
      es(m1)=0
      es(m2)=0
      q(m1)=ion(m1)
      q(m2)=ion(m2)
      call fast(m1,n)
      call fast(m2,n)
      return
      end
c
c
c
c      Define the positon of the second particle of a molecule
c
c
c
c
c
c
c
c      subroutine defpos(m1,m2,rdist)
common/coord/x(400),y(400),z(400)
      call rand(xxb,4)
      rr2=xxb*6.283185308
      call randb(xxb,8)
      rr=2*xxb-1
      rr3=acos(rr)

```



```

c
c
c      Calculation of transition probability for different processes
c
c
c
c      subroutine accept( yes, ux, n, dmin, dmax, yesac, const)
c      common/constant1/temp, bk
c      common/coord1/q(400), ion(400), mol(400), es(400)
c      common/energy/u1, u2
c      common/constant3/a, b, c, step
c      double precision wd, cond
c      sm=0
c      do 7 i=1, 2*n
c         if (mol(i).ne.0) sm=sm+1
c      continue
c      st=n-sm/2
c      sion=(st+1)*(st+1)
c      t2n=n*n
c      u2=ux
c
c      Acceptance probability for Move process
c
c      if (const.eq.1) then
c         if(yesac.eq.11) then
c            call accm(yes, ux, tr, trf)
c            go to 91
c         end if
c         if(yesac.eq.12) pc32=1.
c      else if (const.eq.2) then
c         if(yesac.eq.1) then
c            if(u2.gt.u1.and.u2.lt.98) then
c               pct=(u1-u2)/bk/temp
c               pc32=exp(pct)
c            else
c               pc32=1.
c            end if
c         end if
c      end if
c
c      Acceptance probability for Association process
c
c      if(yesac.eq.2) then
c         tmax=dmax*dmax*dmax
c         tmin=dmin*dmin*dmin
c         ta=a*a*a
c         pc=4.*3.14159*(tmax-tmin)/3./ta
c         pc1=sion*pc/t2n
c         pc2=(sm/2.)/t2n
c         pct=(u1-u2)/bk/temp
c         if(u2.gt.u1.and.u2.lt.98) then
c            pc1=pc1*exp(pct)
c         end if
c         if(u1.gt.u2) pc2=pc2*exp(-pct)
c         constm=max(pc1, pc2)
c         const=1./constm
c         pc32=const*pc1

```

```

      end if
c
c      Acceptance probability for Dissociation process
c
      if(yesac.eq.3) then
          tmax=dmax*dmax*dmax
          tmin=dmin*dmin*dmin
          ta=a*a*a
          pc=4.*3.14159*(tmax-tmin)/3./ta
          pc2=(1.+sm/2.)/t2n
          pc1=st*st*pc/t2n
          pct=(u1-u2)/bk/temp
          if(u2.gt.u1.and.u2.lt.98) then
              pc2=pc2*exp(pct)
          end if
          if(u1.gt.u2) pc1=pc1*exp(-pct)
          constm=max(pc1,pc2)
          const=1./constm
          pc32=const*pc2
      end if
c
c Generating the Random Number and Carrying out the Monte Carlo method
c
      yes=1
      u2=ux
      if(u2.ge.99) go to 89
      if(u2.ge.u1) then
          go to 85
      else
          call rand(xx,24)
          if(xx.ge.pc32) then
              go to 89
          else
              u1=u2
              yes=1
              go to 91
          end if
      end if
85      call rand(xx,24)
      cond=pc32
      if(xx.ge.cond) then
          go to 89
      else
          u1=u2
      end if
      yes=1
      go to 91
89      yes=2
91      return
      end
c
c
c
c Calculation of transition probability
c
c when Smart Monte Carlo is chosen.
c
c Only for strutureless ion particles
c

```

```
c
      subroutine accm(yes,ux,tr,trf)
      common/constant1/temp,bk
      common/energy/u1,u2
      double precision q
         yes=1
         u2=ux
         if(u2.ge.99) go to 89
            q=(trf/tr)*exp(-(u2-u1)/temp/bk)
         if(q.lt.1) then
            go to 85
         else
            u1=u2
            yes=1
            go to 91
         end if
85      call rand(xx,24)
         if(xx.ge.q) then
            go to 89
         else
            u1=u2
         end if
         yes=1
         go to 91
89      yes=2
91      return
      end

c
ccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccc c
c                                                                                   c
c               Smart Monte Carlo Move                                           c
c                                                                                   c
ccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccc c
c
      subroutine qset(n,m1,tr,trf,yesq)
      common/coord/x(449),y(449),z(449)
      common/coord1/q(449),ion(449),mol(449),es(449)
      common/constant3/a,b,c,step
      call force(n,m1,fx,fy,fz)
         std=(50./step)**2
         ddx=fx*std
         ddy=fy*std
         ddz=fz*std
         scrf=sqrt(2.*std)
      call gauss(rfx,r1,scrf)
      call gauss(rfy,r1,scrf)
      call gauss(rfz,r1,scrf)
         tr=exp(-(rfx*rfx+rfy*rfy+rfz*rfz)/4./std)
         x(m1)=x(m1)+ddx+rfx
         y(m1)=y(m1)+ddy+rfy
         z(m1)=z(m1)+ddz+rfz
      call boxlim(x(m1),y(m1),z(m1))
         call fast(m1,n)
      call force(n,m1,fx,fy,fz)
         btx=-(ddx+rfx)-std*fx
         bty=-(ddy+rfy)-std*fy
         btz=-(ddz+rfz)-std*fz
```

```

      trf=exp(-(btX*btX+btY*btY+btZ*btZ)/4./std)
      yesq=1
9       return
      end

c
ccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccc
c
c   Calculation of force exerted on the partilce for Smart Monte Calo    c
c                                                                           c
ccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccc
c

      subroutine force(n,m1,fx,fy,fz)
      common/coord/x(449),y(449),z(449)
      common/constant3/a,b,c,step
      common/coordl/q(449),ion(449),mol(449),es(449)
      fx=0
      fy=0
      fz=0
do 10 i=1,2*n
      if(i.ne.m1.and.q(i).ne.0) then
          sdx=x(i)-x(m1)
          sdy=y(i)-y(m1)
          sdz=z(i)-z(m1)
          if (sdx.gt.a/2) sdx=sdx-a
          if (sdx.lt.-a/2) sdx=sdx+a
          if (sdy.gt.a/2) sdy=sdy-a
          if (sdy.lt.-a/2) sdy=sdy+a
          if (sdz.gt.a/2) sdz=sdz-a
          if (sdz.lt.-a/2) sdz=sdz+a
          sd=sqrt(sdx*sdx+sdy*sdy+sdz*sdz)
          sd3=sd*sd*sd
          tqg=-q(m1)*q(i)*7.138/sd3
          fx=fx+sdx*tqg
          fy=fy+sdy*tqg
          fz=fz+sdz*tqg
      end if
10     continue
      return
      end

c
ccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccc
c
c           Gaussian Randor love for Smart Monte Carlo                  c
c   Using " Polar Methdo "                                              c
c   Ref.: Heermann, D.W. Computer Simulation Methods in theoretical    c
c         physics., 2nd edition, Springer-Verlag (1990)                 c
c                                                                           c
ccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccc
c

      subroutine gauss(rf,r1,sclf)
4       call rand(xx,24)
          uu1=xx
          call randb(xx,24)
          uu2=xx
          v1=2.0*uu1-1.0
          v2=2.0*uu2-1.0
          s=v1*v1+v2*v2
          if(s.ge.1.0) go to 4

```


[illegible]

```

subroutine eng(u,n,const)
common/coord/x(400),y(400),z(400)
common/coord1/q(400),ion(400),mol(400),es(400)
common/ff/fs(349,349)
common/fxyz/fsx2(349,349),fsy2(349,349),fsz2(349,349)
common/constant3/a,b,c,step
common/para/dmin,dmax,depth,req,r1,r2
  u=0.0
  uion=0.0
  umol=0.0
  tmol=0.0
do 66 i=1,2*n-1
do 65 j=i+1,2*n
  ss=fs(i,j)
  mt=q(i)+q(j)
  if(mt.eq.4) d2=2*r1
  if(mt.eq.0) d2=4.2
  if(mt.eq.-4) d2=2*r2
  if(i.eq.mol(j)) then
    d2=dmin
    if(ss.lt.d2) uion=uion+100.
  else
    if(ss.ge.d2) then
      tqq=q(i)*q(j)
      if(tqq.eq.-4. .or. tqq.eq.4.) then
        if(const.eq.2) then
          sxa2=fsx2(i,j)
          sya2=fsy2(i,j)
          sza2=fsz2(i,j)
          ta=1/a
          ta2=ta*ta
          sx2=sxa2*ta2
          sy2=sya2*ta2
          sz2=sza2*ta2
          call engpdc(ss,dss,sx2,sy2,sz2)
        else
          dss=ss
        end if
        uion=uion+tqq*dss*2.9375e-20
      end if
    else
      uion=uion+100.
    end if
  end if
65 continue
  if(es(i).ne.0.and.es(i).le.2.31.and.es(i).ge.1.89) then
    umol=umol-depth
  else
    umol=umol+100
  end if
66 continue
  if(es(2*n).ne.0.and.es(2*n).le.2.31.and.es(2*n).ge.1.89) then
    umol=umol-depth
  else
    umol=umol+100
  end if
  u=uion+umol/2.

```



```

end if
    sx2=sx21*sx21
c
    aty=abs(y(i)-y(j))
    if(aty.gt.dsa) then
        sy21=aty-a
    else
        sy21=aty
    end if
    sy2=sy21*sy21
c
    atz=abs(z(i)-z(j))
    if(atz.gt.dsa) then
        sz21=atz-a
    else
        sz21=atz
    end if
c
    sz2=sz21*sz21
    ss=sqrt(sx2+sy2+sz2)
    return
end
c
ccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccc
c
c Establish the histogram of simulated thermodynamics properties          c
c                                                                           c
c cccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccc
c
    subroutine calculate(n,nos,num,mt4,numpt,const)
    dimension ns(120),tdps(120),tdps1(120)
    dimension stdps(120),stdps1(120)

    common/energy/u1,u2
    common/constant3/a,b,c,step
    common/index/ns,t6,t7
    common/ff/fs(349,349)
    common/sum/sum(120)
    common/coord1/q(400),ion(400),mol(400),es(400)
    common/ex/pq(120)
    double precision t6,t7,tm4
    tm4=0.
    do 95 i=1,2*n
        if(mol(i).ne.0) then
            tm4=tm4+1.
        end if
95      continue
        ttm4=2.*n
        tm4=tm4*0.5
        sum(2)=sum(2)+u1
        sum(1)=sum(1)+tm4*0.0001
        t6=t6+(n-tm4)
        t7=t7+tm4
c
c List the energy and molecule number distribution
c
    do 98 i7=1,10
        t21=(-i7*0.6-4)*1.0e-20

```

```

        t22=(-i7*0.6-5)*1.0e-20
        if(tm4.ge.(i7-1)*5+1 .and. tm4.le.i7*5) ns(i7)=ns(i7)+1
        if(u1.lt.t21 .and. u1.ge.t22) ns(i7+10)=ns(i7+10)+1
98      continue
c
c Histogram of Correlation fuction. ! Further calcualtion is necessary
c
        ws=0.6
do 114 i=1,2*n-1
do 113 j=i+1,2*n
    if (const.eq.1) then
        tspq1=q(i)
        tspq2=q(j)
    else
        tspq1=ion(i)
        tspq2=ion(j)
    end if
    if (tspq1.eq.2..and.tspq2.eq.-2.) then
        ss=fs(i,j)
        do 111 k=1,30
            pss=(k-1)*ws+4.2
            pss1=k*ws+4.2
            if(ss.ge.pss.and.ss.lt.pss1) tdps(k)=tdps(k)+1.
111        continue
        end if
        if (tspq1.eq.2..and.tspq2.eq.2.) then
            ss=fs(i,j)
            do 112 k=1,30
                pss=(k-1)*ws+4.2
                pss1=k*ws+4.2
                if(ss.ge.pss.and.ss.lt.pss1) tdps1(k)=tdps1(k)+1.
112            continue
        end if
113        continue
114    continue
c
c Output the results to files
c
116    if(mod(num,numpt) .eq.0) then
        sum(16)=sum(16)+sum(2)
        sum(17)=sum(17)+sum(1)
        write(16,*) sum(16)/float(num)
        write(17,*) sum(17)/float(num)
        sum(1)=0
        sum(2)=0
        totm=0.
    end if
c
    if(mod(num,numpt) .eq.0) then
        do 120 i=1,30
            write(15,*) tdps(i)/float(num)
            write(18,*) tdps1(i)/float(num)
120        continue
    end if
    return
end
c

```



```

common/xyz/fsx2(349,349),fsy2(349,349),fsz2(349,349)
common/ba1/qip(2),ipc(2),ipm(2),xm(2),ym(2),zm(2),iqm(2)
common/ba2/esm(2),xmol(2),ymol(2),zmol(2),iqmol(2)
  do 8 i=1,mptc
    m1=ipc(i)
    m2=ipm(i)
    x(m1)=xm(i)
    y(m1)=ym(i)
    z(m1)=zm(i)
    q(m1)=iqm(i)
    es(m1)=esm(i)
    mol(m1)=ipm(i)
    call fast(m1,n)
  if(m2.eq.0) go to 8
    x(m2)=xmol(i)
    y(m2)=ymol(i)
    z(m2)=zmol(i)
    q(m2)=iqmol(i)
    es(m2)=esm(i)
    mol(m2)=m1
  call fast(m2,n)
8 continue
return
end

```

Bibliography

Chapter 1

- (1) Lee, S.H.; Rasaiah, J.C.; Cummings, P.T. *J.Chem. Phys.* **1985**, 83, 317.
- (2) Rasaiah, J.C.; Lee, S. H. *J. Chem. Phys.* **1985**, 83, 5870.
- (3) Rasaiah, J.C.; Lee, S.H. *J. Chem. Phys.* **1985**, 83, 6396.
- (4) Rasaiah, J.C.; Lee, S.H. *J. Chem. Phys.* **1987**, 86, 983.
- (5) Rasaiah, J.C ; Zhu, J.; Lee, S.H. *J. Chem. Phys.* **1989**, 91, 495.
- (6) Rasaiah, J.C ; Zhu, J.; Lee, S.H. *J. Chem. Phys.* **1989**, 91, 505.
- (7) Rasaiah, J.C ; Zhu, J. *J. Chem. Phys.* **1991**, 94, 3141.
- (8) Rasaiah, J.C. *Int. J. Thermodynamics.* **1990**, 11, 1.
- (9) Heermann, D.W. *Computer Simulation Methods*, 2nd edition, Springer-Verlag **1990**.
- (10) Hecht, C.E. *Statistical Thermodynamics and Kinetic Theory* , W. H. Freeman and Company, New York, **1990**.
- (11) McQuarrie, D.A. *Statistical Mechanics*, HarperCollins Publishers, New York, **1976**.
- (12) Allen, M.P.; Tledesley, D.J. *Computer Simulation of Liquid*, Oxford Science Publications, Cambridge, **1987**.
- (13) Risken, H. *The Fokker-Planck Equation*, Springer, Berlin, New York, 1984
- (14) Van Kampen, N.G. *Stochastic Process in Physics and Chemistry*, North-Holland, Amsterdam, New York, **1981**.
- (15) Metropolis, N.; Rosenbluth, A.W; Rosenbluth, M.N.; Teller, A.H. and Teller, E. *J. Chem. Phys.* **1953**, 21, 1087.
- (16) Berne, B.J.; ed. *Statistical Mechanics, Part A: Equilibrium Techniques*, Plenum Press, New York, **1977**.
- (17) Ewald, P.P., *Ann, Phys. (Leipzig)* **1921**, 64, 253.
- (18) Brush, S.G.; Sahlin, H.L.; Teller, E. *J. Chem. Phys.* **1966**, 45, 2102.
- (19) Valleau, J.P.; Cohen, L.K. *J. Chem. Phys.* **1980**, 72, 5935.
- (20) Adams, D.J.; Dubey, G.S. *J. Comp. Phys.* **1987**, 72, 156.

- (21) Mills, P; Anderson, C.F.; Record, M.T., Jr. *J. Phys. Chem.* **1985**, 89, 3984.
- (22) Mills, P; Anderson, C.F.; Record, M.T., Jr. *J. Phys. Chem.* **1986**, 90, 6541.
- (23) Wood, W.W.; Parker, F.R. *J. Chem. Phys.* **1957**, 27, 720.
- (24) Rossky, P.J.; Doll, J.D.; Friedman, H.L. *J. Chem. Phys.* **1978**, 69, 4628.
- (25) Cao, J; Berne, B.J. *J. Chem. Phys.* **1990**, 92, 1980.
- (26) Ermak, D.L. *J. Chem. Phys.* 1975, 62, 4189; **1975**, 62, 4197.
- (27) Rao, M.; Berne, B. J. *J. Chem. Phys.* **1979**, 71, 129.
- (28) Rao, M.; Pangali, D; Berne, B.J. *Mol. Phys.* **1979**, 37, 1773.
- (29) Pangali, C.; Rao, M.; Berne, B.J. *Chem. Phys. Lett.* **1978**, 55, 413.
- (30) Kotelyanskii, M.J.; Suter, U.W. *J. Chem. Phys.* **1992**, 96, 5383.
- (31) Bennett, C.H. *J. Comput. Phys.* **1976**, 22, 245.
- (32) Beveridge, D.L.; Dicapua, F.M. *Annu. Rev. Biophys. Chem.* **1989**, 18, 431.
- (33) Voter, A. F. *J. Chem. Phys.* **1985**, 82, 1890.
- (34) Torrie, G.M; Valleau, J.P. *J. Comput. Phys.* **1977**, 23, 187.
- (35) Zwanzig, R.W. *J. Chem. Phys.* **1954**, 22, 1420.
- (36) Lyubartsev, A.P.; Martsinovski, A. A.; Shevkunov, S.V.; Vorontsov-McDonald, P.N. *J. Chem. Phys.* **1992**, 96, 1776.
- (37) Adams, D.J. *Mol. Phys.* **1974**, 28, 1241.
- (38) Adams, D.J. *Mol. Phys.* **1975**, 29, 307.
- (39) Valleau, J.P.; Cohen, L.K. *J. Chem. Phys.* **1980**, 72, 5942.
- (40) Guggenheim, E.A. *Thermodynamics*, North-Holland Physics Publishing, New York, 1967.
- (41) Rasaiah, J.C; Freedman, H.L. *J. Chem. Phys.* **1972**, 56, 248.
- (42) Rasaiah, J.C. *J. Solution. Chem.* **1973**, 2, 301.
- (43) Friedman, D. *Statistical Mechanics*, Prentice-Hall, Inc, New Jersey, **1985**.
- (44) McMillan, W.G. and Mayer, J. E. *J. Chem. Phys.* **1945**, 13, 276.
- (45) McDonald, I.R. *Mol. Phys.* **1972**, 23, 41.

(46) Creutz, M. *Phys. Rev. Lett.* **1983**, 50, 1411.

Chapter 2

- (1) Metropolis, N.; Rosenbluth, A.W.; Rosenbluth, M.N.; Teller, A.H.; Teller, E. *J. Chem. Phys.* **1953**, 21, 1087.
- (2) Berne, B.J.; ed. *Statistical Mechanics, Part A: Equilibrium Techniques*, Plenum Press, New York, **1977**.
- (3) Ciccotti, G.; Frenkel, D.; McDonald, I.R.; eds. *Simulation of Liquids and Solids, Molecular Dynamics and Monte Carlo Methods in Statistical Mechanics* North-Holland Physics Publishing, **1987**.
- (4) Pangali, C.; Rao, M.; Berne, B.J. *Chem. Phys. Lett.* **1978**, 3, 413.
- (5) Rao, M.; Berne, B. *J. Chem. Phys.* **1979**, 71, 129.
- (6) Rao, M.; Pangali, C.; Berne, B.J. *Mol. Phys.* **1979**, 37, 1773.
- (7) Rossky, P.J.; Doll, J.D.; Friedman, H.L. *J. Chem. Phys.* **1978**, 69, 4628.
- (8) Cao, J. ; Berne, B.J. *J. Chem. Phys.* **1990**, 92, 1980.
- (9) Kotelyanskii, M.J.; Suter, U. W. *J. Chem. Phys.* **1992**, 96, 5383.
- (10) Bennett, C.H. *J. Computational Physics* **1976**, 22, 245.
- (11) Voter, A.F. *J. Chem. Phys.* **1985**, 82, 1890.

Chapter 3

- (1) Shew, C.-Y., Mills, P.A. *J. Phys. Chem.* **1993**, 97, 13824.
- (2) Metropolis, N.; Rosenbluth, A.W.; Rosenbluth, M.N.; Teller, A.H.; Teller, E. *J. Chem. Phys.* **1953**, 21, 1087.
- (3) Berne, B.J.; ed. *Statistical Mechanics, Part A: Equilibrium Techniques*, Plenum Press, New York, **1977**.

- (4) Ciccotti, G.; Frenkel, D.; McDonald, I.R.; eds. *Simulation of Liquids and Solids, Molecular Dynamics and Monte Carlo Methods in Statistical Mechanics* North-Holland Physics Publishing, **1987**.
- (5) Cao, J. ; Berne, B.J. *J. Chem. Phys.* **1990**, 92, 1980.
- (6) Pangali, C.; Rao, M.; Berne, B.J. *Chem. Phys. Lett.* **1978**, 3, 413.
- (7) Rao, M.; Berne, B. *J Chem. Phys.* **1979**, 71, 129.
- (8) Rao, M.; Pangali, C.; Berne, B.J. *Mol. Phys.* **1979**, 37, 1773.
- (9) Rossky, P.J.; Doll, J.D.; Friedman, H.L. *J. Chem. Phys.* **1978**, 69, 4628.
- (10) Mezei, M.; Bencsath, K.A.; Goldman, S.; Singh, S. *Molecular Simulations* **1987**, 1, 87.
- (11) Frantz, D.D., Freeman, D.L., Doll, J.D. *J. Chem. Phys.* **1990**, 93, 2769.
- (12) Frantz, D.D., Freeman, D.L., Doll, J.D. *J. Chem. Phys.* **1992**, 97, 5713.
- (13) Bennett, C.H. *J. Computational Physics*, **1976**, 22, 245.
- (14) Voter, A.F. *J. Chem. Phys.* **1985**, 82 , 1890.
- (15) Lee, S.H.; Rasaiah, J.C.; Cummings, P.T. *J. Chem. Phys.* **1985**, **83**, 317.

Chapter 4

- (1) Lee, S. H.; Rasaiah, J. C. and Cummings, P. T. *J. Chem. Phys.* **1985**, 83, 317.
- (2) Rasaiah, J. C.; Lee, S. H.; *J. Chem. Phys.* **1985**, 83, 5870.
- (3) Rasaiah, J. C.; Lee, S. H.; *J. Chem. Phys.* **1985**, 83, 6396.
- (4) Rasaiah, J. C.; Lee, S. H.; *J. Chem. Phys.* **1987**, 86, 983.
- (5) Rasaiah, J. C.; Zhu, J.; Lee, S. H. *J. Chem. Phys.* **1989**, 91, 495.
- (6) Rasaiah, J.C.; Zhu, J.; Lee, S.H. *J. Chem. Phys.* **1989**, 91, 505.
- (7) Rasaiah, J.C.; Zhu, J. *J. Chem. Phys.* **1991**, 94, 3141.
- (8) Rasaiah, J.C., *Int. J. Thermophysics*, **1990**, 11, 1.
- (9) Herrer, J. N.; Blum, L. *J. Chem. Phys.* **1991**, 94, 5077.
- (10) Pizio, O. A. *J. Chem. Phys.* **1984**, 100, 548.

- (11) Cummings, P. T. and Stell, F. *Mol. Phys.* **1984**, *51*, 253.
- (12) Cummings, P. T. and Stell, F. *Mol. Phys.* **1985**, *55*, 33.
- (13) Andersen, H. C. J. *Chem. Phys.* **1973**, *59*, 4714.
- (14) Hfye, J. S. and Olaussen, K. *Physica A* **1980**, *104*, 435.
- (15) Card, D. N. and Valleau, J. P. *J. Chem. Phys.* **1970**, *52*, 6232.
- (16) Valleau, J. P.; Cohen, L. K, and Card, D. N. *J. Chem. Phys.* **1980**, *72*, 5942.
- (17) Larsen, B. *Chem. Phys. Letters* **1974**, *27*, 47.
- (18) Rogde, S. A. *Chem. Phys. Letters* **1983**, *103*, 133.
- (19) Sorensen, T. S. *J. Chem. Soc. Faraday Trans.* **1990**, *86*, 1815.
- (20) Valleau, J. P.; Ivkov, R. and Torrie, G. M.; *J. Chem. Phys.* **1991**, *95*, 520, (and references therein).
- (21) Linse, P.; Gunnarsson, G.; Jönsson, B. *J. Phys. Chem.* **1982**, *85*, 413.
- (22) Linse, P.; Jönsson, B. *J. Chem. Phys.* **1983**, *78*, 3167.
- (23) Bratko, D.; Vlachy, V. *Chem. Phys. Lett.* **1982**, *90*, 434.
- (24) LeBret, M.; Zimm, B. H. *Biopolymers* **1984**, *23*, 271.
- (25) Murthy, C. S.; Bacquet, R. J.; Rossky, P.J. *J. Phys. Chem.* **1985**, *89*, 701.
- (26) Mills, P. A.; Anderson, C. F.; Record, M. T., Jr. *J. Phys. Chem.* **1985**, *89*, 3984.
- (27) Mills, P. A.; Anderson, C. F.; Record, M. T., Jr. *J. Phys. Chem.* **1986**, *90*, 6541.
- (28) Bacquet, R. J.; Rossky, P. J. *J. Phys. Chem.* **1988**, *92*, 3604.
- (29) Sharp, K. A.; Jean-Charles, J.; Honig, B. *J. Phys. Chem.* **1992**, *96*, 3822.
- (30) Sitkoff, D.; Sharp, K. A.; Honig, B.. *J. Phys. Chem.* **1994**, *97*, 1978.
- (31) Cao, J. and Berne, B. *J. Chem. Phys.* **1990**, *92*, 1980.
- (32) Shew, C.-Y. and Mills, P. *J. Phys. Chem.* **1993**, *97*, 13824.
- (33) Shew, C.-Y. and Mills, P. *J. Phys. Chem.* (in press).
- (34) Smith, W. R. and Triska, B. *J. Chem. Phys.* **1994**, *100*, 3019.
- (35) Tsangaris, D.M.; de Pablo, J.J. *J. Chem. Phys.* **1994**, *101*, 1477.

- (36) Metropolis, N.; Rosenbluth, A. W.; Rosenbluth, M. N.; Teller, A. H.; Teller, E. *J. Chem. Phys.* **1953**, *21*, 1087.
- (37) Brush, S. G.; Sahlin, H. L.; Teller, E. *J. Chem. Phys.* **1966**, *45*, 2102.
- (38) Adams, D. J. Duber, G. S. *J. Comp. Phys.* **1987**, *72*, 156.
- (39) Voter, A. F. *J. Chem. Phys.* **1985**, *82* , 1890.
- (40) Bennet, C.H. *J. Comp. Phys.* **1976**, *22*, 245.
- (41) Rasaiah, J. C. *J. Chem. Phys.* **1972**, *22*, 245.