

D.P. Landau
S.P. Lewis
H.-B. Schüttler
(Eds.)

Computer Simulation Studies in Condensed- Matter Physics XVI



Springer

SPRINGER PROCEEDINGS IN PHYSICS

- 87 **Proceedings
of the 25th International Conference
on the Physics of Semiconductors**
Editors: N. Miura and T. Ando
- 88 **Starburst Galaxies**
Near and Far
Editors: L. Tacconi and D. Lutz
- 89 **Computer Simulation Studies
in Condensed-Matter Physics XIV**
Editors: D.P. Landau, S.P. Lewis,
and H.-B. Schüttler
- 90 **Computer Simulation Studies
in Condensed-Matter Physics XV**
Editors: D.P. Landau, S.P. Lewis,
and H.-B. Schüttler
- 91 **The Dense Interstellar Medium
in Galaxies**
Editors: S. Pfalzner, C. Kramer,
C. Straubmeier, and A. Heithausen
- 92 **Beyond the Standard Model 2003**
Editor: H.V. Klapdor-Kleingrothaus
- 93 **ISSMGE**
Experimental Studies
Editor: T. Schanz
- 94 **ISSMGE**
Numerical and Theoretical Approaches
Editor: T. Schanz
- 95 **Computer Simulation Studies
in Condensed-Matter Physics XVI**
Editors: D.P. Landau, S.P. Lewis,
and H.-B. Schüttler
- 96 **Electromagnetics in a Complex World**
Editors: I.M. Pinto, V. Galdi,
and L.B. Felsen
- 97 **Fields, Networks,
Computational Methods and Systems
in Modern Electrodynamics**
A Tribute to Leopold B. Felsen
Editors: P. Russer and M. Mongiardo
- 98 **Particle Physics and the Universe**
Proceedings of the 9th Adriatic Meeting,
Sept. 2003, Dubrovnik
Editors: J. Trampetić and J. Wess
- 99 **Cosmic Explosions**
On the 10th Anniversary of SN1993
(IAU Colloquium 192)
Editors: J. M. Marcaide and K. W. Weiler
- 100 **Lasers in the Conservation of Artworks**
LACONA V Proceedings,
Osnabrück, Germany, Sept. 15–18, 2003
Editors: K. Dickmann, C. Fotakis,
and J.F. Asmus
- 101 **Progress in Turbulence**
Editors: J. Peinke, A. Kittel, S. Barth,
and M. Oberlack
- 102 **Adaptive Optics
for Industry and Medicine**
Proceedings
of the 4th International Workshop
Editor: U. Wittrock
- 103 **Computer Simulation Studies
in Condensed-Matter Physics XVII**
Editors: D.P. Landau, S.P. Lewis,
and H.-B. Schüttler
- 104 **Complex Computing-Networks**
Brain-like and Wave-oriented
Electrodynamical Algorithms
Editors: I.C. Gökna and L. Sevgi

Volumes 60–86 are listed at the end of the book.

D.P. Landau S.P. Lewis H.-B. Schüttler
(Eds.)

Computer Simulation Studies in Condensed-Matter Physics XVI

Proceedings of the Seventeenth Workshop
Athens, GA, USA, February 16–20, 2004

With 128 Figures and 23 Tables



Prof. Dr. Landau, David P.,
University of Georgia, Center
Simulational Physics
30602-2451 Athens, GA, U.S.A.
EMAIL: dlandau@hal.physast.uga.edu

Dr. Lewis, Steven P.,
University of Georgia, Center
Simulational Physics
30602-2451 Athens, GA, U.S.A.
EMAIL: lewis@hal.physast.uga.edu

Prof. Dr. Schüttler,
Heinz-Bernd, University of
Georgia, Center Simulational Physics
30602-2451 Athens, GA, U.S.A.
EMAIL: hbs@hal.physast.uga.edu

ISSN 0930-8989

ISBN-10 3-540-26564-3 Springer Berlin Heidelberg New York

ISBN-13 978-3-540-26564-1 Springer Berlin Heidelberg New York

Bibliographic information published by Die Deutsche Bibliothek

Die Deutsche Bibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data is available in the Internet at <<http://dnb.ddb.de>>.

This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilm or in any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provisions of the German Copyright Law of September 9, 1965, in its current version, and permission for use must always be obtained from Springer-Verlag. Violations are liable to prosecution under the German Copyright Law.

Springer is a part of Springer Science+Business Media.

springer.com

© Springer-Verlag Berlin Heidelberg 2006
Printed in The Netherlands

The use of general descriptive names, registered names, trademarks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

Production: SPI Publisher Services

Cover concept: eStudio Calamar Steinen

Cover production: *design & production* GmbH, Heidelberg

Printed on acid-free paper SPIN: 11399339 57/3100/SPI - 5 4 3 2 1 0

Preface

Over fifteen years ago, because of the tremendous increase in the power and utility of computer simulations, The University of Georgia formed the first institutional unit devoted to the use of simulations in research and teaching: The Center for Simulational Physics. As the international simulations community expanded further, we sensed a need for a meeting place for both experienced simulators and neophytes to discuss new techniques and recent results in an environment which promoted lively discussion. As a consequence, the Center for Simulational Physics established an annual workshop on Recent Developments in Computer Simulation Studies in Condensed Matter Physics. This year's workshop was the seventeenth in this series, and the continued interest shown by the scientific community demonstrates quite clearly the useful purpose that these meetings have served. The latest workshop was held at The University of Georgia, February 16–20, 2004, and these proceedings provide a “status report” on a number of important topics. This volume is published with the goal of timely dissemination of the material to a wider audience.

We wish to offer a special thanks to IBM and to SGI for partial support of this year's workshop.

This volume contains both invited papers and contributed presentations on problems in both classical and quantum condensed matter physics. We hope that each reader will benefit from specialized results as well as profit from exposure to new algorithms, methods of analysis, and conceptual developments.

Athens, GA, USA
October 2004

*D.P. Landau
S.P. Lewis
H.-B. Schüttler*

Contents

1 Computer Simulation Studies in Condensed Matter Physics: An Introduction <i>D.P. Landau, S.P. Lewis, H.-B. Schüttler</i>	1
<hr/>	
Part I Systems out of Equilibrium	
<hr/>	
2 Shake, Rattle or Roll: Things to do with a Granular Mixture on a Computer <i>D.C. Rapaport</i>	7
3 A New Method of Investigating Equilibrium Properties from Nonequilibrium Work <i>S. Yukawa</i>	19
4 Numerical Simulations of Critical Dynamics far from Equilibrium <i>B. Zheng</i>	25
<hr/>	
Part II Soft and Disordered Materials	
<hr/>	
5 Entropy Driven Phase Separation <i>R. Vink</i>	45
6 Supercooled Liquids under Shear: Computational Approach <i>R. Yamamoto, R. Temam, J. Dean, D. Grove, C. Chambers, K.B. Bruce, E. Bertino</i>	61
7 Optimizing Glasses with Extremal Dynamics <i>S. Boettcher, A.G. Percus</i>	74
8 Stochastic Collision Molecular Dynamics Simulations for Ion Transfer Across Liquid–Liquid Interfaces <i>S. Frank, W. Schmickler</i>	80

Part III Biological Systems

9 Generalized-Ensemble Simulations of Small Proteins	
<i>U.H.E. Hansmann</i>	87
10 A Biological Coevolution Model	
with Correlated Individual-Based Dynamics	
<i>V. Sevim, P.A. Rikvold</i>	90
11 An Image Recognition Algorithm for Automatic Counting	
of Brain Cells of Fruit Fly	
<i>T. Shimada, K. Kato, K. Ito</i>	95
12 Preferred Binding Sites of Gene-Regulatory Proteins	
Based on the Deterministic Dead-End Elimination Algorithm	
<i>R.G. Endres, T.C. Schultheiss, N.S. Wingreen</i>	100

Part IV Algorithms and Methods

13 Geometric Cluster Algorithm for Interacting Fluids	
<i>E. Luijten, J. Liu</i>	109
14 Polymer Simulations	
with a Flat Histogram Stochastic Growth Algorithm	
<i>T. Prellberg, J. Krawczyk, A. Rechnitzer</i>	122
15 Convergence of the Wang–Landau Algorithm	
and Statistical Error	
<i>C. Zhou, R.N. Bhatt</i>	136
16 Wang–Landau Sampling with Cluster Updates	
<i>M. Körner, M. Troyer</i>	142
17 Multibaric-Multithermal Simulations	
for Lennard–Jones Fluids	
<i>H. Okumura, Y. Okamoto</i>	146
18 A Successive Umbrella Sampling Algorithm	
to Sample and Overcome Free Energy Barriers	
<i>P. Virnau, M. Müller</i>	151

Part V Computer Tools

19 C++ and Generic Programming for Rapid Development of Monte Carlo Simulations	
<i>G. Brown, H.K. Lee, T.C. Schulthess</i>	157
20 Visualization of Vector Spin Configurations	
<i>R. Hihinashvili, J. Adler, S.-H. Tsai, D.P. Landau</i>	169
21 The BlueGene/L Project	
<i>D. Chen</i>	174

Part VI Molecules, Clusters and Nanoparticles

22 All-Electron Path Integral Monte Carlo Simulations of Small Atoms and Molecules	
<i>J. Shumway</i>	181
23 Projective Dynamics in Realistic Models of Nanomagnets	
<i>S.H. Thompson, G. Brown, P.A. Rikvold</i>	196
24 Cumulants for an Ising Model for Folded 1-d Small-World Materials	
<i>M.A. Novotny</i>	201
25 Embryonic Forms of Nickel and Palladium: A Molecular Dynamics Computer Simulation	
<i>Z. El-bayyari</i>	205

Part VII Surfaces and Alloys

26 Usage of Pattern Recognition Scheme in Kinetic Monte Carlo Simulations: Application to Cluster Diffusion on Cu(111)	
<i>C. Ghosh, A. Kara, T.S. Rahman</i>	215
27 Including Long-Range Interactions in Atomistic Modelling of Diffusional Phase Changes	
<i>D.R. Mason, R.E. Rudd, A.P. Sutton</i>	241
28 Br Electrodeposition on Au(100): From DFT to Experiment	
<i>S.J. Mitchell, M.T.M. Koper</i>	258

29 Simulation of ZnSe, ZnS Coating on CdSe Substrate: The Electronic Structure and Absorption Spectra	
<i>J.S. Kim, S.J. Yun, G.S. Lee</i>	265
30 Simulation of Islands and Vacancy Structures for Si/Ge-covered Si(001) Using a Hybrid MC-MD Algorithm	
<i>L. Nurminen, F. Tavazza, D.P. Landau, A. Kuronen, K. Kaski</i>	270
31 Spin-Polarons in the FM Kondo Model	
<i>M. Daghofer, W. Koller, A. Prüll, H.G. Evertz, W. von der Linden</i>	276
List of Contributors	283

Part I

Systems out of Equilibrium

1 Computer Simulation Studies in Condensed Matter Physics: An Introduction

D.P. Landau, S.P. Lewis, and H.-B. Schüttler

Center for Simulational Physics, The University of Georgia,
Athens, GA 30602-2451, USA

Computer simulation studies in condensed matter physics now play a fundamental role in many areas of investigation. The “status report” which is contained in this volume is the result of presentations and discussion that occurred during the 17th Annual Workshop at the Center for Simulational Physics. The texts of both longer, invited presentations as well as a number of contributed papers are included. The reader will find that the scope of simulational/computational studies is broad and that substantial potential for cross-fertilization of methods between different sub-fields exists.

The volume opens with three papers on systems that are not in equilibrium. The first, invited, presentation by Rapaport describes a molecular dynamics approach to the simulation of granular media. The simulation includes normal forces (always) and tangential forces (sometimes), and results are given for particle segregation in flowing, rotating, and vibrating media. Both quantitative measurements and visual images provide information about the behavior. Next, Yukawa introduces a new method to determine equilibrium properties by measuring non-equilibrium work. In tests on a Lennard-Jones system the method yields good results for the pressure but not for the free energy. The chapter ends with another invited paper in which Zheng describes extensive simulations of critical dynamics in the “far from equilibrium” regime. Monte Carlo generated data for the “short time” region are used to determine both equilibrium and dynamic critical exponents for simple models, including frustrated systems, of statistical physics.

Part II is centered on soft and disordered materials and begins with two invited presentations. Vink first describes a grand canonical Monte Carlo approach to colloid-polymer mixtures. Polymers are treated as large, rigid objects with fixed effective radius surrounded by a depletion zone. The mixture has zero energy but phase separation is driven by entropic effects. The phase diagram is determined using this new method. Yamamoto then examines the microscopic dynamics of supercooled liquids under shear. Molecular dynamics methods are used to examine a two-dimension liquid with soft-core interactions. The effect of the shear is to reduce the structural relaxation time and shear viscosity, but the microscopic dynamics remain isotropic. In the first of two contributed presentations Boettcher and Percus introduce a method, “extremal optimization”, as a local search heuristic. The power

of the method is demonstrated via application to spin glass ground states. This Part closes with a stochastic collision molecular dynamics study of ion transfer across liquid-liquid interfaces by Frank and Schmickler. The potential energy surface is calculated and the results are compared with those from transition-state theory and experiment.

Simulations of biological systems are contained in Part III which begins with an invited paper by Hansmann. He reviews the generalized ensemble method for protein simulations in which the standard Boltzmann sampling weight probability is modified so as to enable efficient escapes from local energy minima. He then shows that this approach allows for efficient all-atom simulations for small proteins with up to 35 amino acids. Next, Sevim and Rikvold present simulations of a biological coevolution model in which species are represented by genomes of bit strings and the procreation probability of each species is coupled to the populations of other species via a certain interaction matrix. They show that these interspecies correlations do not significantly affect the statistical behavior of this model eco-system. Then, Shimada, Kato and Ito describe a Monte Carlo algorithm for the counting of cells in high-resolution microscopy images of *Drosophila* brains. The algorithm is based on a random deposition of volume-excluded objects which approximate the imaged cell nuclei, using an appropriately constructed deposition probability function. Cell counts are given for samples containing on the order of 10^5 cells, with percent level counting accuracy. To complete this part, Endres, Schulthess and Wingreen show how energy minimization of phenomenological atomistic interaction potentials can be applied as an efficient tool to identify DNA binding sites of gene-regulatory proteins. Using a deterministic dead-end elimination algorithm they conduct a global search through DNA sequence and conformational space to predict likely binding site characteristics.

Algorithmic developments are featured in Part IV, beginning with an invited presentation by Luitjen and Liu who introduce a novel cluster Monte Carlo algorithm for the simulation of complex fluids. The algorithm employs rejection-free non-local geometric moves to achieve large reductions in the Monte Carlo correlation times. A substantial increase in efficiency, compared to standard local updating methods, is demonstrated for binary and simple Lennard-Jones fluid models. The next paper, also invited, by Prellberg, Krawczyk and Rechnitzer, describes how the computational efficiency of lattice polymer simulation algorithms can be substantially enhanced by incorporating the flat histogram approach into the pruned and enriched Rosenbluth (PERM) method. The method is illustrated by applications to self-avoiding random walk models and HP protein models. Several contributed papers follow. Zhou and Bhatt present a mathematical analysis of the convergence properties and statistical errors of the Wang-Landau (WL) algorithm. For an artificial ensemble, they prove rigorously that convergence for the density of states is achieved and show explicitly how the statistical error depends

on the WL modification factor. Körner and Troyer then describe a new algorithm combining cluster updates with Wang-Landau sampling, based on a series expansion of the partition function. The substantial reduction in Monte Carlo correlation times achieved with this algorithm is illustrated by applications to the one-dimensional XY model. In the contribution by Okumura and Okamoto, a new Monte Carlo approach for simulating continuum systems, referred to as the multibaric-multithermal algorithm, is discussed. It permits calculations of averages at all temperatures and pressures in a single run, and the effectiveness of the method is demonstrated in applications to the Lennard-Jones liquid. Lastly, Virnau and Müller have developed a novel implementation of the umbrella sampling method which allows effective sampling to overcome free energy barriers. They apply this method to the phase transition and the interfacial tension in a bead-spring model for hexadecane.

Part V contains three articles on several exciting new “computer tools” that will clearly be of future importance. First, in an invited presentation Brown et al. introduce the concept of “Generic programming” and provide an assessment of the potential benefits of such an approach. They describe the newly developed C++ Ψ -Mag toolset and use it to calculate energies in a wide variety of spin models. The toolset is also used to perform Metropolis Monte Carlo simulations of an anisotropic Heisenberg spin model. Hihinashvili et al. then describe the visualization of three-dimensional classical vector spins using the AViz atomistic visualization package developed at The Technion. Both the computer interface and typical results are shown. This Part ends with another invited paper as Chen presents a description of a massively parallel, 360 Teraflop design point parallel supercomputer currently under development. With a “system on a chip” design, the potential advantages are greater computing power, lower power consumption, and reduced cost per flop. The promise of Petaflop computing appears to be not very far off in the future.

Part VI treats molecules, clusters, and nanoparticles. An invited paper by Shumway detailed the development of a path-integral Monte Carlo method for doing ab initio all-electron simulations of atoms and molecules at finite temperature. He emphasizes the significant technical challenges, focusing especially on the sensitivity of the calculations to the nodal structure used in the fixed-node approximations. Preliminary results are shown for the light atoms up to Ne and for two small molecules: LiH and CH₄. Thompson, Brown, and Rikvold use projective dynamics to simulate the magnetization reversal for a three-dimensional model of an iron nano-pillar. This system permits richer dynamics than previous one-dimensional models, thus yielding more realistic simulations. Novotny then presents results for the fourth-order cumulant of a ferromagnetic Ising model on a one-dimensional lattice folded according to a finite number of random small-world bond connections. El-bayyari conducts constant-temperature molecular dynamics simulations for the geometry and

energetics of nickel and palladium nanoclusters. Empirical many-body classical potentials were used, and comparison was made to ab initio calculations.

The volume concludes with Part VII, containing a number of articles about surfaces and alloys. This part begins with two invited articles. First, Ghosh, Kara, and Rahman employ a novel pattern-recognition scheme within kinetic Monte Carlo simulations to study adatom cluster diffusion. Their studies of small copper clusters on Cu (111) find that certain “magic clusters” exhibit non-random, patterned diffusion pathways at room temperature, but at higher temperatures even these magic clusters revert to random diffusion. Microscopic details of the various diffusion processes are discussed. Next, Mason, Rudd, and Sutton study diffusional phase transformations in both Al-Cu and Al Cu-Mg alloys using an off-lattice atomistic kinetic Monte Carlo technique. They emphasize the importance of incorporating strain relaxation around vacancies and other defect structures, and describe an efficient way to include these effects by doing only partial relaxations during trial-move assessment and full relaxations after move selection. Their results are compared in detail to comparable rigid-lattice simulations. Mitchell and Koper perform a series of density functional theory calculations to parameterize a potential for use in off-lattice classical Monte Carlo simulations of bromine adsorbate phases on the Au (100) surface. Kim, Yun, and Lee apply density functional theory to study the electronic structure and optical absorption spectrum of a CdSe substrate coated with either ZnS or ZnSe under various coating conditions. Nurminen et al. implement a new hybrid MC-MD approach to investigate the formation and stability of islands and dimmer-vacancy structures of clean and Ge-covered Si (001). The simulations—done off-lattice and employing collective moves—yield results in good qualitative agreement with experiments. In the final article Daghofer et al. present Monte Carlo simulations of hole doping in the ferromagnetic Kondo model. They find that the doped holes do not phase separate, but rather form independent ferromagnetic spin-polarons.

2 Shake, Rattle or Roll: Things to do with a Granular Mixture on a Computer

D.C. Rapaport

Physics Department, Bar-Ilan University, Ramat-Gan, Israel
rapaport@mail.biu.ac.il

Abstract. Studies of granular media continue to produce surprising results that are unique to this class of matter. A particularly prominent feature of granular mixtures is the tendency to segregate into their individual components when externally driven in various ways, an effect that has consequences for industrial processing. This article describes the computer simulation of several different modes of segregation in flowing, rotating and vibrating granular media, in some cases reproducing known behavior, in others predicting effects that have yet to be observed experimentally.

2.1 Introduction

Granular media exhibit very different behavior from systems governed by thermodynamics and statistical mechanics. These differences can be so pronounced that the intuition gained from these theories proves to be of little help in trying to understand the mechanisms responsible for granular behavior. The tendency of noncohesive granular mixtures to segregate into individual species is one of the more conspicuous of these properties, and the fact that segregation occurs even when there are no apparent energetic or entropic advantages is what makes such behavior so fascinating. Since mixing and segregation are important processes, both in industry (ranging from pharmaceuticals and food processing to mining) and in nature, considerable effort has been invested in exploring the underlying causes of these phenomena. In some instances segregation is the desired effect, while in others it is something to be avoided. Given the absence of a general theory of granular matter that can serve as the basis for understanding segregation, much of the work in this field in recent years has been based on computer simulation, and it seems reasonable to assume that this situation will persist.

In this paper we begin with a brief overview of the simulation methodology and then survey a selection of recent results from simulations of various kinds. Some of this material has appeared in print over the past few years (more extensive bibliographies will be found in the papers cited), whereas other results have been extracted from unpublished simulations of systems presently under investigation. Certain aspects of the observed behavior correspond to well-known phenomena associated with real granular media, and

some of the systems are clearly of relevance from an industrial point of view. Other results are not so familiar, and some have yet to be observed in real materials. As originally presented, the lecture included interactive demonstrations of most of the examples, either in real time or as prerecorded animations from simulations requiring more extensive computation; the printed page is limited to static images.

A variety of different systems containing two distinct granular species have been treated by means of simulation in two or three dimensions. The examples described here include the following: vertical segregation in both sheared flow and vertical vibrated layers; horizontal segregation in vertically vibrated layers using a base with a sawtooth profile; axial and radial segregation in horizontal rotating cylinders; and, finally, horizontal segregation in horizontally vibrated layers. In some instances the segregating system undergoes spontaneous subdivision into two regions each having a much higher concentration (or even exclusively) one of the species, in other instances more complex segregation patterns emerge and evolve as the simulation progresses. Brief mention is also made of systems containing three granular species.

2.2 Simulation Methodology

The simulation methodology is based on the approach used for molecular dynamics (MD), with the particles now representing discrete grains rather than atoms. A variety of models have been employed in granular simulations over the years, based on both soft and (occasionally) hard particles; though referred to as “soft”, this characterization refers more to the continuous nature of the potential function used to prevent particle overlap during collision than to the amount of overlap itself, which is minimal (think in terms of golf balls). Soft particles are more widely used (including here) because of the simpler nature of the computations involved. In addition to this excluded-volume repulsive interaction, the particles are also subject to dissipative forces. The interactions fall into two classes, those between pairs of particles, and those between the particles and the container walls, which can be treated as special cases of the interparticle interactions; a schematic diagram showing the different contributions appears in Fig. 2.1. The simulations are carried out in two or three dimensions, depending on the nature of the problem.

Consider a pair of spherical granular particles i and j with diameters d_i and d_j , respectively. The repulsive force that acts between particles closer than a specified range [1] is

$$\mathbf{f}_v = \frac{48}{r_{ij}} \left[\left(\frac{d_{ij}}{r_{ij}} \right)^{12} - \frac{1}{2} \left(\frac{d_{ij}}{r_{ij}} \right)^6 \right] \hat{\mathbf{r}}_{ij}, \quad (2.1)$$

where $\mathbf{r}_{ij} = \mathbf{r}_i - \mathbf{r}_j$ is the particle separation, $r_{ij} = |\mathbf{r}_{ij}|$, and the effective diameter appearing in the interaction computations is $d_{ij} = (d_i + d_j)/2$. This

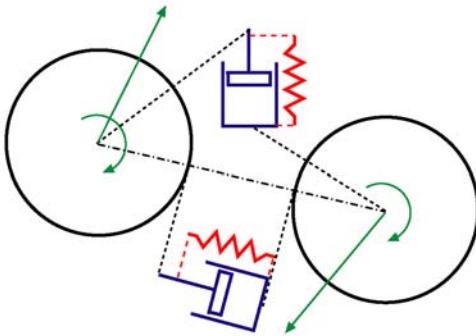


Fig. 2.1. Schematic form of the forces between grains; the normal forces are always present, but either or both of the tangential forces are optional, depending on the system

particular interaction is derived from the functional form of the Lennard-Jones potential; it acts whenever $r_{ij} < 2^{1/6} d_{ij}$ and is continuous at the cutoff point (although its derivative is not). Alternatives to this form of overlap interaction that are also in routine use include functions that depend on the overlap either linearly or to the 3/2-power, but the choice of function is of little consequence insofar as the bulk behavior is concerned. Note that because of the slight degree of softness, the particle diameter is not a precisely defined quantity.

What distinguishes the interactions used for granular media from those in atomistic MD is the presence of dissipative forces that act throughout the duration of each collision. The first of these is a normal (viscous) damping force

$$\mathbf{f}_d = -\gamma_n(\hat{\mathbf{r}}_{ij} \cdot \mathbf{v}_{ij})\hat{\mathbf{r}}_{ij}, \quad (2.2)$$

that depends on the component of the relative velocity of the particles $\mathbf{v}_{ij} = \mathbf{v}_i - \mathbf{v}_j$ in the direction between the particle centers. The factor γ_n is the normal damping coefficient, assumed to be the same for all particles. The total force parallel to \mathbf{r}_{ij} is $\mathbf{f}_n = \mathbf{f}_v + \mathbf{f}_d$.

Frictional damping can also act in the transverse direction at the point of contact of the particles. The relative transverse velocity of the particle surfaces at this point, allowing for particle rotation, is

$$\mathbf{v}_{ij}^s = \mathbf{v}_{ij} - (\hat{\mathbf{r}}_{ij} \cdot \mathbf{v}_{ij})\hat{\mathbf{r}}_{ij} - \left(\frac{d_i \boldsymbol{\omega}_i + d_j \boldsymbol{\omega}_j}{d_i + d_j} \right) \times \mathbf{r}_{ij}, \quad (2.3)$$

where $\boldsymbol{\omega}_i$ is the angular velocity of particle i . The sliding friction is then

$$\mathbf{f}_s = -\min(\gamma_s^{c_i c_j} |\mathbf{v}_{ij}^s|, \mu^{c_i c_j} |\mathbf{f}_n|) \hat{\mathbf{v}}_{ij}^s. \quad (2.4)$$

Here, $\gamma_s^{c_i c_j}$ is the sliding friction coefficient, whose value depends on the species c_i and c_j of the particles involved (the values may be the same). The

static friction coefficient $\mu^{c_i c_j}$ sets an upper bound to the sliding friction proportional to $|\mathbf{f}_n|$; in models of this kind there is no true static friction – at best it can be represented by a tangential restoring force that depends on the relative displacement occurring in the course of a contact,

$$\mathbf{u}_{ij} = \int_{(\text{coll})} \mathbf{v}_{ij}^s(\tau) d\tau \quad (2.5)$$

although it should be emphasized that this is not a strictly correct means for incorporating the effects of static friction. The particles are also subject to gravity.

Similar considerations apply to the interactions between the particles and the rough (nonslip) container walls. Walls play essential role in granular dynamics and a suitably simplified description is needed for the simulations. There are several methods for representing walls, for example, by constructing explicit rough, energy-absorbing walls out of (a large number of) constrained particles with similar properties to the mobile particles, or by representing the walls as geometrically smooth surfaces with similar frictional properties to the particles themselves. Depending on the problem, the actual wall shape may be flat, curved, or sawtooth-shaped; in addition, the wall may be vibrating, translating, or rotating. Periodic boundaries can be used where the presence of an explicit wall is inappropriate.

Other details of the simulation follow standard molecular dynamics procedures [2]. Neighbor lists are used to organize the force computations efficiently. The translational and rotational equations of motion are integrated using the leapfrog method. Because of the heavy computations involved in certain of the studies, parallel computing methods based on a spatial decomposition of the system can be used to distribute the workload across several coupled processors.

2.3 Chute Flow

The first example involves the flow of a binary granular mixture down an inclined chute – the work is described in detail in [3]. The base of the chute is rough, resulting in sheared flow. In the initial state the larger grains (all grains are made from the same material) are located close to the base, but as the simulation progresses they all gradually rise to the top of the layer. The example appearing here is based on a two-dimensional simulation in which the side boundaries of the system are periodic; Fig. 2.2 shows intermediate and late states of the run.

In order to explain the underlying mechanism it is necessary to consider the issue of particle packing: as the larger grains ascend, the space they once occupied is filled by smaller grains; for a larger grain to descend a vacancy of sufficient size must form, an unlikely event given that the cooperation of

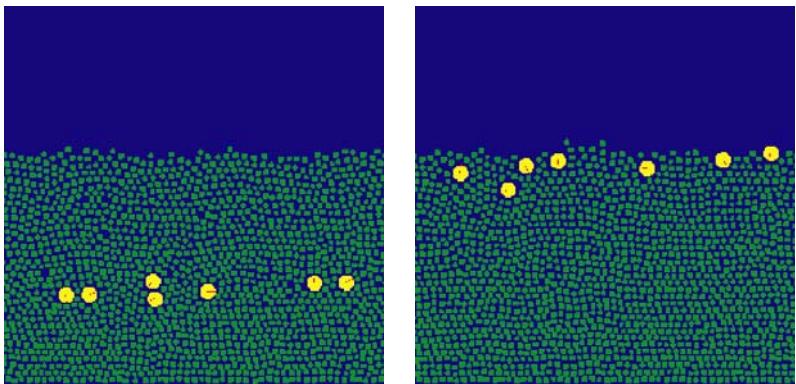


Fig. 2.2. Intermediate and late states of simulated chute flow; although the system is drawn horizontally, the actual slope is downward to the right

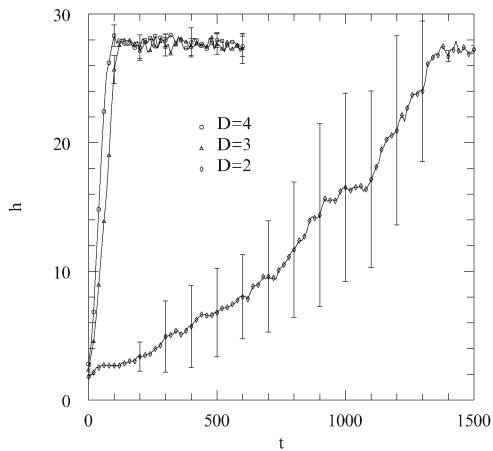


Fig. 2.3. Time dependence of the height of the large grains as a function of their relative size (D)

several of the smaller grains is required. The actual rate at which the large grains ascend depends on the size ratio and the slope angle; the dependence on the relative size is shown in Fig. 2.3. An “inverse grading” effect of this type may or may not be desirable in an industrial processing context.

2.4 Brazil-Nut Effect

An entirely different means of achieving size segregation is through vertical vibration. An initially mixed granular layer containing grains of two sizes will gradually segregate, with the larger grains moving to the upper region

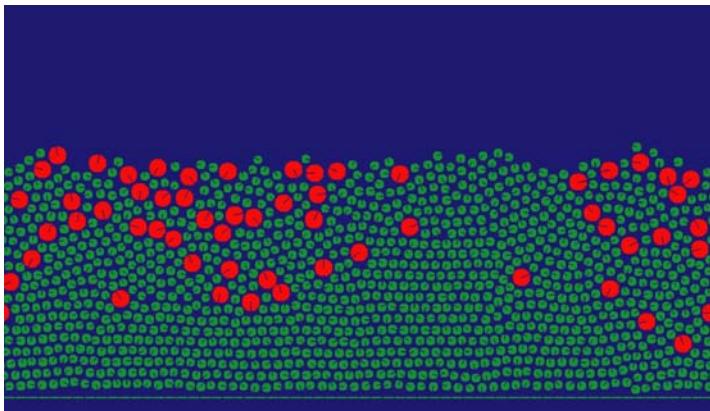


Fig. 2.4. Simulation of a vertically vibrated layer

of the layer. Even though the larger particles are heavier than the equivalent volume of smaller particles (assuming the same material) they still are driven upwards; such behavior is the exact opposite of what would occur with a solid body immersed in a fluid. This phenomenon is readily observed in the kitchen – in granulated coffee for example. The underlying mechanism is as before, with the smaller grains readily filling the space vacated by the ascending larger grains. Convective flows may also enhance the process under certain conditions, but these are not essential for obtaining segregation. The original simulations of this phenomenon [4] used Monte Carlo techniques; the present work, an example of which appears in Fig. 2.4, employs a dynamical approach.

2.5 Granular Ratchet

The previous two examples dealt with well-known segregation effects, so that the role of simulation was to confirm the ability of simplified granular models to reproduce the observed behavior. The next example is different, in that the phenomenon has yet to be observed experimentally.

It is known that a vertically oscillated granular layer (containing grains of just a single size), in which the base of the container has an asymmetric sawtooth profile, undergoes horizontal motion in addition to the vertical vibration. The direction and magnitude of the horizontal flow depend on the parameters of the system, such as the vibration frequency and the sawtooth size and shape. Simulations reveal [5] the presence of vertically stratified flow, and it is even possible to generate horizontal flows in opposite directions at different heights within the layer, as shown in Fig. 2.5.

The actual form of the flow profile varies considerably as the parameters are altered. Examples of how the flow changes with height within a layer for different frequencies and numbers of sawteeth appear in Fig. 2.6.

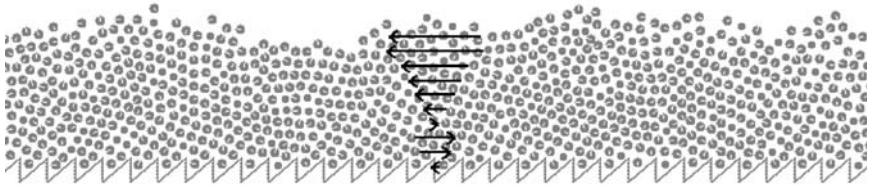


Fig. 2.5. Stratified horizontal flow in a vertically vibrated layer with a sawtooth base; arrows show the mean flow at each level

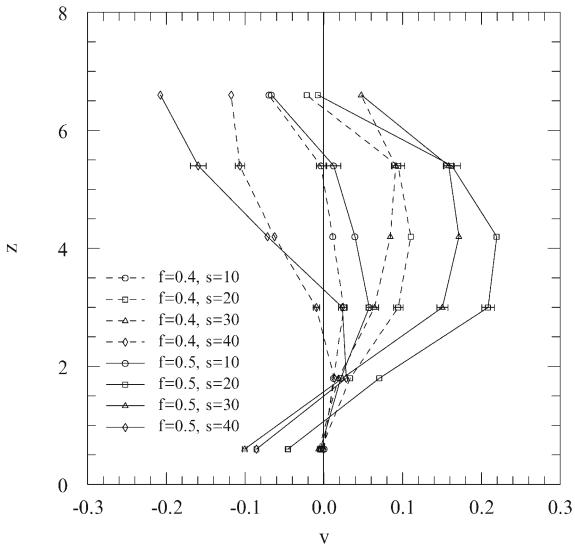


Fig. 2.6. Stratified flow profiles (velocity as a function of height z) for different frequencies (f) and numbers of sawteeth (s)

If this entirely unexpected behavior is now combined with the more familiar vertical segregation – the Brazil nut effect – observed with two grain sizes, then horizontal segregation of the two species is the predicted outcome. This is indeed what is observed in the simulations [6], with the detailed behavior depending once again on the parameters of the system. Figure 2.7 shows the simulation of a hypothetical separator after it has essentially completed its task; for computational convenience, grains exiting over the end walls, or through the slit at the bottom of the right-hand wall, are recycled as shown. Although the simulations have been carried out in two dimensions, a similar effect is observed in three dimensions – in a cylindrical container with sawteeth arranged in concentric rings so that the larger grains migrate outwards.

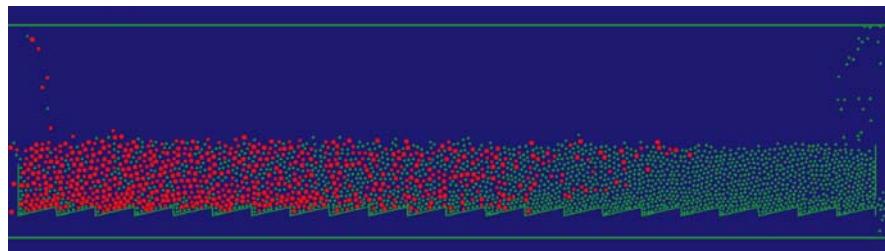


Fig. 2.7. Vertically vibrated system that produces horizontal size segregation

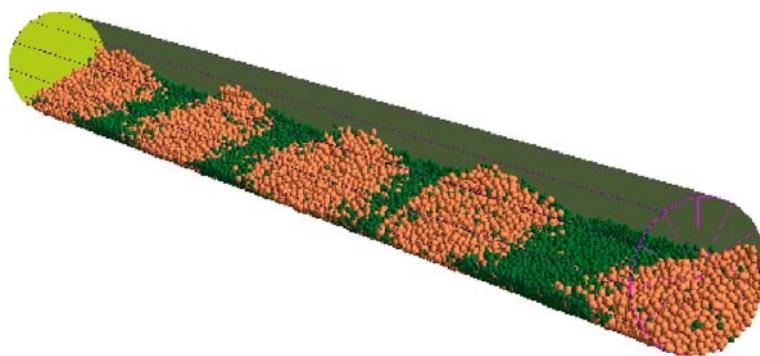


Fig. 2.8. Axial segregation in a rotating cylinder

2.6 Rotating Cylinder – Axial Segregation

There are some segregation phenomena that are inherently three-dimensional in nature. An example is the segregation into axial bands that occurs in a rotating horizontal cylinder partially filled with grains of two different kinds. A number of experimental studies of systems of this kind have been carried out, including some aided by the use of magnetic-resonance imaging (MRI) techniques that facilitate the examination of what actually occurs beneath the surface. The first simulations that systematically investigated this form of segregation were described in [1]. The two species of grains must differ in either size or frictional properties (or both), and over an extensive series of runs it became clear that there are certain requirements regarding the relative values of the friction coefficients appearing in the grain–grain and grain–boundary interactions in order to ensure segregation. Figure 2.8 shows the outcome of one such run.

In simulations of this type it is especially important to be able to visualize the detailed behavior. Figure 2.9 shows a set of images of a particular grain configuration, with views from above and below, as well as views showing just one species at a time. These help establish the nature of the segregation

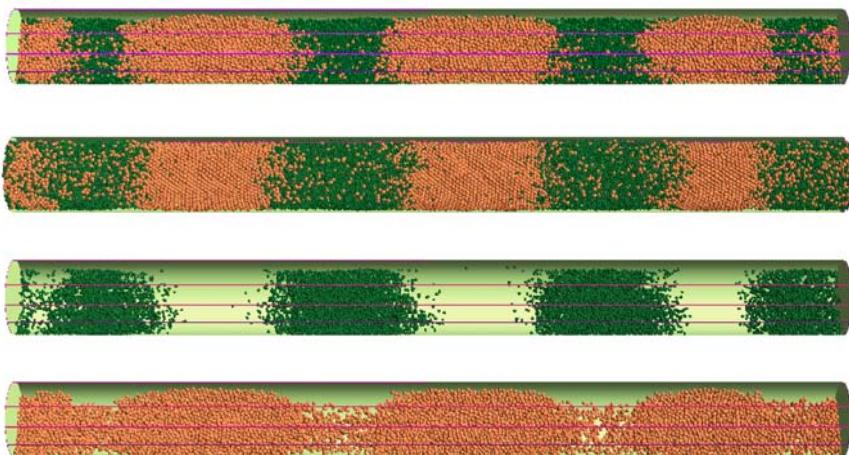


Fig. 2.9. Different views of a segregated state (see text)

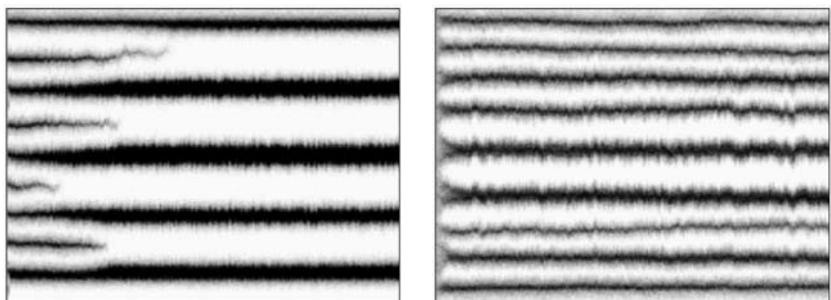


Fig. 2.10. Space–time plots showing band evolution for two of the rotating cylinder runs

that is taking place – here, for example, the separation of the smaller grains into distinct bands is complete.

Various kinds of behavior can be observed, including band merging (and consequent coarsening) and traveling bands. The entire history of a run, showing how and when bands merge, can be condensed into a single “space–time” plot, in which the horizontal scale represents time and the vertical scale the relative species concentration in slices along the cylinder axis; two examples are shown in Fig. 2.10. The underlying mechanism is not entirely understood, but the existence of different stability angles for the two grain species leads to a possible means for ensuring segregation. In the case of three kinds of grains, unpublished simulations reveal more complex band patterns,

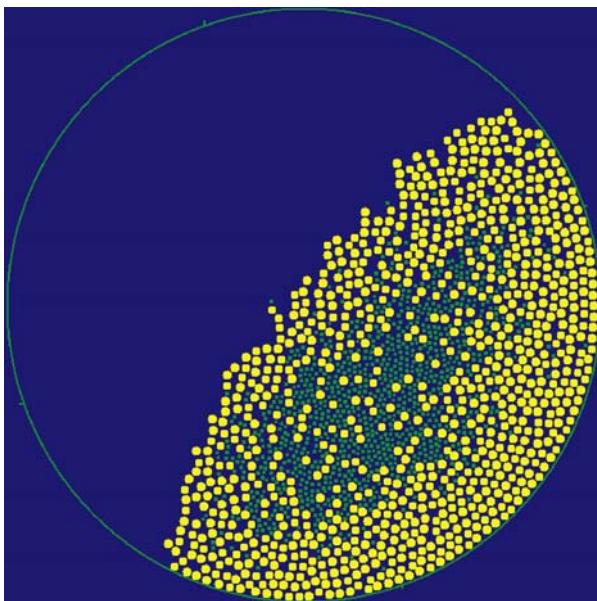


Fig. 2.11. Radial segregation in a rotating cylinder

with the intermediate-sized grains concentrated at the interfaces between small and large grains.

2.7 Rotating Cylinder – Radial Segregation

Another form of segregation known to occur in a horizontal cylinder involves the smaller grains congregating near the rotation axis and the larger grains on the outside – an essentially two-dimensional effect. Simulations (in two dimensions) display this behavior, but, unlike the other simulations discussed here, the model grains require the inclusion of the elastic restoring force mentioned earlier; the two consequences of this particular interaction are an increase in the surface slope and a reduction in the “fluidity” of the material, so that it becomes more solid-like with (sometimes intermittent) flows confined to the layers closest to the surface; while this characteristic of a true granular material is not fully reproduced, the effect is enhanced relative to models in which only the viscous forces are present. Figure 2.11 shows an example of the resulting behavior.

2.8 Horizontally vibrated layer

Yet another segregation effect is encountered in a very thin layer (essentially a monolayer) containing a granular mixture that is subjected to horizontal

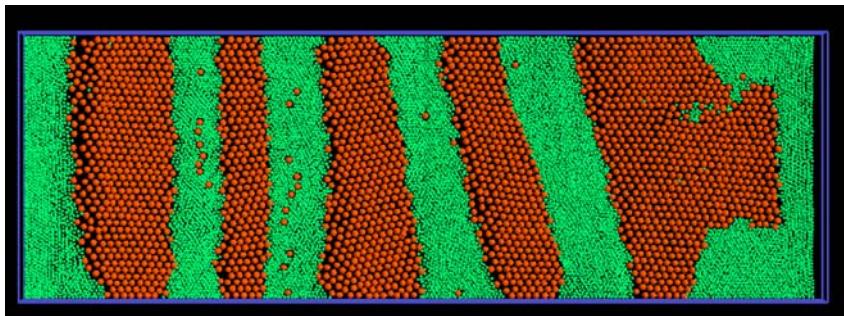


Fig. 2.12. Band formation in a horizontally vibrated layer

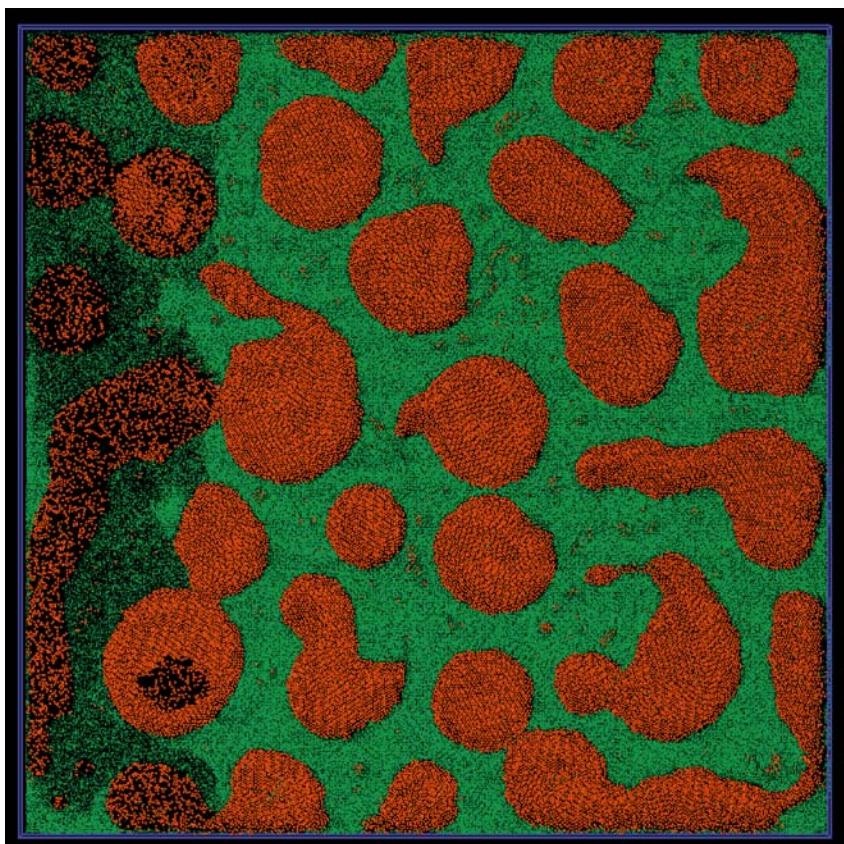


Fig. 2.13. Segregation patterns in a layer subjected to circular horizontal motion

side-to-side vibration. Beginning with an initially mixed state, the eventual outcome of the experiment is the formation of linear stripes perpendicular to the vibration direction [7]. Simulations of the corresponding model system

reveal analogous behavior; an example is shown in Fig. 2.12. The segregation patterns exhibit coarsening as the simulation progresses.

Two obvious extensions of this phenomenon, neither of them as yet studied experimentally, are the use of three granular species with the appearance of additional bands (similar to the rotating cylinder with three species) and the replacement of the side-to-side motion by circular motion (the container is not rotated however). The outcome in the latter case is shown in Fig. 2.13; here again the patterns exhibit coarsening over time, starting with small-scale structures that evolve into the forms such as those shown here.

2.9 Conclusion

Discrete-particle simulations of segregating granular mixtures in a variety of contexts have been described. It is apparent that systems of this kind exhibit a strong tendency to segregate spontaneously under suitable driving conditions; such behavior is very different from atomistic (energy-conserving) systems governed by thermodynamics. Where the corresponding experiments are available, the simulations are in broad agreement, at least qualitatively (quantitative comparison with experiment can be used to validate the models and tune the interaction parameters). While simulation is normally used to model experiment, in at least two of the examples presented here the situation is reversed, with the simulations predicting behavior yet to be observed in the laboratory; confirmation (hopefully) is eagerly awaited.

Acknowledgment

This work was partially supported by the Israel Science Foundation.

References

1. D.C. Rapaport: Phys. Rev. E **65**, 061306 (2002)
2. D.C. Rapaport: *The Art of Molecular Dynamics Simulation*. 2nd edn. (Cambridge University Press, Cambridge 2004)
3. D. Hirshfeld, D.C. Rapaport: Phys. Rev. E **56**, 2012 (1997)
4. A. Rosato, K.J. Strandburg, F. Prinz, R.H. Swendsen: Phys. Rev. Lett. **58**, 1038 (1987)
5. M. Levanon, D.C. Rapaport: Phys. Rev. E **64**, 011304 (2001)
6. D.C. Rapaport: Phys. Rev. E **64**, 061304 (2001)
7. T. Mullin: Phys. Rev. Lett. **84**, 4741 (1999)

3 A New Method of Investigating Equilibrium Properties from Nonequilibrium Work

S. Yukawa

Department of Applied Physics, School of Engineering, The University of Tokyo
7-3-1, Hongo, Bunkyo-ku, Tokyo 113-8656, Japan

Abstract. A new method of investigating thermal equilibrium natures from non-equilibrium work is proposed. This method is enable us to calculate thermal equilibrium values of physical quantities, even the free energy, with help of the information of nonequilibrium work. Applications for Lennard-Jones particle system are reported. For physical quantities such as pressure the new method shows good results. But the free energy is not good. The idea of improvement is also proposed.

3.1 Introduction

To study equilibrium properties by a computer, we often make a long time simulation for getting equilibrium states and calculate some physical quantities. A computation time for the equilibration process is usually wasted. In addition, we sometimes want to do a parameter study, then we have to repeat again the above procedures. In such case, wasted computational time increases. To avoid these waste computational time, some techniques have been proposed and used [8].

In this article another method for avoiding equilibration is proposed; this is based on the idea of nonequilibrium thermodynamics, that is, Jarzynski equality. The Jarzynski equality connects with nonequilibrium work and a difference of thermal equilibrium free energy in theoretically. But, in an application of the Jarzynski equality to an statistical object, unknown features still exist, even experimentally the equality was checked [3]. Thus we will check usability of the equality and clarify some problems through an application for Lennard-Jones (spline) particle system.

3.2 Method

To explain the present method, we consider a simple example; there is a cylinder containing gas, for simplicity, N mono-atomic molecules. The cylinder itself is subjected to a thermal bath which temperature is written as $T = 1/\beta$.

First a piston of cylinder is fixed, in which a volume is written as V_0 , and the gas is in thermal equilibrium state according to the temperature and its

density. Then we consider a process of changing a volume of gas by pushing the piston from V_0 to V_1 . When the process is sufficiently slowly, it can be regarded as quasi-static process. In this case a work $W[V_0 \rightarrow V_1]$ done during the process gives a difference of Helmholtz free energy F :

$$\langle W[V_0 \rightarrow V_1] \rangle = F(N, \beta, V_1) - F(N, \beta, V_0),$$

where $\langle \dots \rangle$ is an average over trials. It can be said that we can obtain the equilibrium information from the quasi-static work.

If the process is not sufficiently slow, what happens? We know an answer that the total work is larger than the difference of free energy:

$$\langle W[V_0 \rightarrow V_1] \rangle > F(N, \beta, V_1) - F(N, \beta, V_0).$$

That is nothing but the second law of thermodynamics. Additional work is regarded as an entropy production. This is an inequality. If we could write this inequality to be an equality, we get the information of equilibrium from nonequilibrium work. This is enabled by the Jarzynski equality.

In 1997, Jarzynski showed an equality relating the nonequilibrium work to the equilibrium free energy difference [1, 2]. The Jarzynski equality is written as follows:

$$\langle \exp(-\beta W[\lambda_0 \rightarrow \lambda_1]) \rangle = \exp(-\beta (F(\lambda_1) - F(\lambda_0))), \quad (3.1)$$

where β is an inverse temperature of environment and $W[\lambda_0 \rightarrow \lambda_1]$ is a work done during the process changing a parameter λ from λ_0 to λ_1 . The average is taken over initial thermal equilibrium states. It is remarkable that the equality is always valid independently of whether a intermediate state is equilibrium or nonequilibrium.

This equality is easily proved (see [1, 2]); requirements are that dynamics of system can describe a thermal relaxation to a thermal equilibrium state corresponding Hamiltonian with λ and that work can be defined with a parameter λ . It is also regarded as a kind of symmetry of distribution function of entropy production [5]. It is note that the weight before taking the trial ensemble average, $\exp(-\beta W)$ is proportional to the equilibrium Boltzmann weight. Thus we can obtain thermal expectation values of observable, in addition to the free energy difference from the Jarzynski equality.

In an actual computer simulation, we generate nonequilibrium processes of changing a parameter, sample the nonequilibrium work and calculate physical quantities. After computation, we obtain a free energy difference and thermal expectation values of physical quantities with the help of the Jarzynski equality.

3.3 Application for Lennard–Jones System

3.3.1 Model

Here we apply the present method to Lennard–Jones particle system; for computational advantages, we use not the original 12^{-6} model but a Lennard–Jones spline model. Two-body interacting potential is as follows:

$$\phi(r) = \begin{cases} 4\epsilon \left\{ \left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^6 \right\} & \text{for } 0 < r \leq r_c, \\ -a \frac{\epsilon}{\sigma^2} (r - r_m)^2 & \text{for } r_c < r \leq r_m, \\ 0 & \text{otherwise,} \end{cases} \quad (3.2)$$

where ϵ, σ are dimensional parameters represent energy and length scales, respectively. These two and a mass of particle m determine all physical dimensions. In an actual computer simulation, all of them are taken to be 1. r_c is a inflection point of Lennard–Jones potential, $r_c = \sqrt[6]{26/7}\sigma$. The first part of the above potential is an original one of Lennard–Jones 12^{-6} potential. The second part is a polynomial interpolation from the inflection point r_c to a vanishing point of potential r_m . Assuming a smooth continuation of interacting force, we can determine two parameters $a = (4536 2^{2/3} \sqrt[3]{7/13})/3211$ and $r_m = (55 \sqrt[6]{13/7})(18 2^{5/6})\sigma$. The above modification changes quantitative properties of the system. But qualitative feature is not changed.

Geometry of the system is a cubic box with fixed length, which volume is denoted as V . Periodic boundary conditions are imposed on all directions. Dynamics is Hamiltonian dynamics with Nosé–Hoover thermostat with temperature T [6]. A number of particles N is fixed, so we treat NVT ensembles. The thermodynamic operation for the Jarzynski process is chosen as increment of density which operation is achieved by introducing a new parameter α as a factor of σ ; if we set α taken to be 0, the system is identified as an ideal gas. For $\alpha = 1$, the original scale is revealed.

The actual simulation is done by the following way: Initially we prepare $N = V/\sigma^3$ particles with $\alpha = 0$, in which there is no interaction between them. Their velocities are taken from Maxwellian distribution with definite temperature T . Then we can easily set up an initial thermal equilibrium state. Thermodynamic operation for a Jarzynski calculation is done by changing α from 0 to 1 linearly in a period of t_{op} . Then a fractional density ρ , which is defined as a ratio of total cubic core volume $N\alpha^3\sigma^3$ to total volume V , is varying from 0 to 1 as α^3 . During such process, we measure work needed to change the parameter α .

Repeating these procedure, we average an exponentiated work and get a ratio of final partition function to initial one for every densities. In addition the exponentiated work itself is a Boltzmann weight, thus we can obtain thermal equilibrium quantities of observables for every densities.

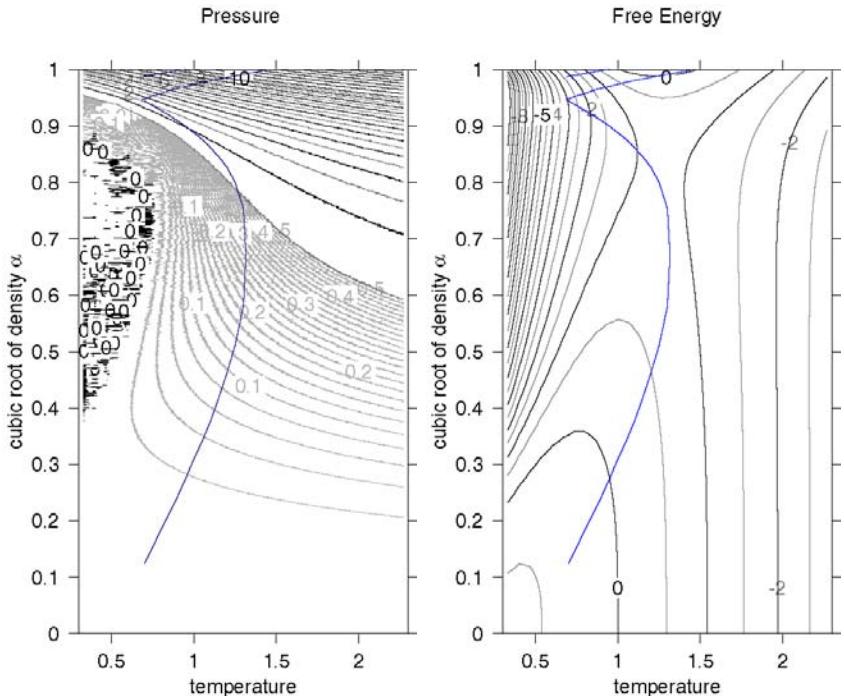


Fig. 3.1. (a) Global behavior of pressure and (b) Helmholtz free energy landscape: The horizontal axis represents the temperature and the vertical one does α , a cubic root of the fractional density. The result of $N = 8000$ system with $t_{\text{op}} = 100$ is shown. Arch-like curves are phase boundaries of the original Lennard-Jones system

3.3.2 Results

Here we see some results of calculation: Pressure, for instance of physical observable, and Helmholtz free energy are shown. First, global behavior of pressure in a temperature-density plane is presented in Fig. 3.1a. A horizontal axis represents the temperature of heat bath and a vertical one does the cubic root of fractional density, α . The pressure is calculated from the Virial theorem and reweighted with the exponentiated work. We use a $N = 8000$ system and an operation interval is taken to be $t_{\text{op}} = 100$. We take 10 samples for 30 different temperatures. Some isobars are shown in this figure.

For convenience, a phase boundary between gas, liquid and coexistence region of the original model is shown. For the present model, this curve is not equivalent to a real boundary. But some information can be obtained from this boundary, because quantitative difference is small. Coexistence region of the present model is observed as instable or metastable behavior of pressures.

To investigate dependence of the operation time t_{op} , we plot the pressure for several operation time in Fig. 3.2. Some thermal equilibrium values of

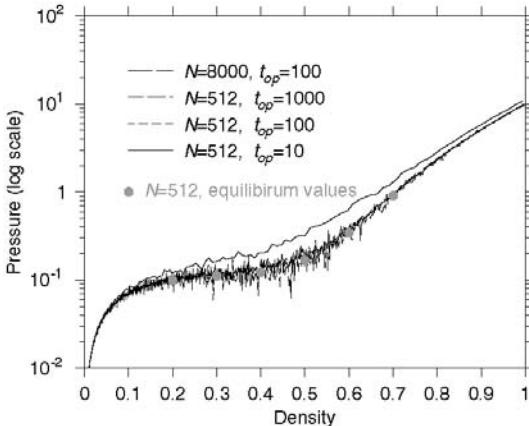


Fig. 3.2. Operation time dependence of the pressure: We show results of $N = 512$, $t_{\text{op}} = 10, 100, 1000$, $N = 8000$, $t_{\text{op}} = 1000$, and equilibrium calculation for the $N = 512$ system. Fluctuations of curve are almost equivalent to the order of statistical errors. For the equilibrium values, statistical errors are less than a symbol size

pressure are also shown in the same figure; the results are for systems with 512 and 8000 particles with a temperature 1. From the results of 512 and 8000 particles systems, dependence of the pressure on a system size is small at this temperature. The pressure of operation time 100 and 1000 are enough converged to equilibrium values, but the result of $t_{\text{op}} = 10$ is not good. Therefore the operation time 100 is sufficient for getting the equilibrium nature. It is remarkable because a relaxation time is about 10, which is estimated from a equilibrium calculation.

We can, of course, calculate the free energy difference. By choosing the initial state is trivial one, we obtain Helmholtz free energy itself. The result is shown in Fig. 3.1b. To check validity, we consider behavior of pressure from this profile; the pressure is proportional to a gradient of the free energy along the density axis. So the pressure near $\rho \sim 0, T = 1$ is negative from the profile. It does not agree with the equilibrium calculation. Thus this obtained free energy profile is not consistent with the actual profile.

3.4 Summary and Discussion

We have seen that the new method based on the Jarzynski equality is work for the Lennard-Jones system; the pressure reweighted with the exponential work shows good agreements of equilibrium expectation values. This calculation is more efficient than one of equilibrium simulations, because we can skip some initial relaxation processes.

From the theoretical basis of the Jarzynski equality, we cannot expect dependences of operation time. But in the present simulation, we observe the

dependence. Its origin lie in the deficient sampling; theoretically the equality requires all states, especially the states with exponentially small weight are quite important. This is also the main reason that the free energy shows a bad agreement. For the pressure this dependence is small, thus we can get good convergence.

To improve this point, we should use more fluctuating dynamics such as Monte Carlo rather than the molecular dynamics. There is a possibility that an artificial dynamics which is not appropriate for investigating equilibrium state is better than Monte Carlo method or molecular dynamics. Recently one possibility of improvement was proposed by Hukushima and Iba [7]. They used an extended ensemble with the Jarzynski equality.

Acknowledgment

The author is grateful to N. Ito, S. Miyashita, S. Todo, M. Kikuchi for valuable discussions. This work was partly supported by the Grant-in-Aid for Young Scientists (B, No. 14740229) and the Grant-in-Aid of Scientific Research (C, No. 15607003) of the Japanese Ministry of Education, Science, Sports and Culture. The part of computation in this work has been done using the facilities of the Earth Simulator Center, Japan Marine Science and Technology Center (JAMSTEC).

References

1. C. Jarzynski: Phys. Rev. Lett. **78**, 2690 (1997)
2. C. Jarzynski: Phys. Rev. E **56**, 5018 (1997)
3. J. Lipardt, S. Dumont, S.B. Smith, I. Tinoco Jr., C. Bustamante: Science **296**, 1832 (2002)
4. G.M. Wang, E.M. Sevick, E. Mittag, D.J. Searles, D.J. Evans: Phys. Rev. Lett. **89**, 05601 (2002)
5. G.E. Crooks: Phys. Rev. E **60**, 2721, (1999)
6. W.G. Hoover: Computational Statistical Mechanics (Elsevier, 1991)
7. K. Hukushima, Y. Iba: The Physical Society of Japan, 58th Annual Meeting 2003
8. For example, the nonequilibrium relaxation method: N. Ito: Physica A **196** 59, (1993)

4 Numerical Simulations of Critical Dynamics far from Equilibrium

B. Zheng

Zhejiang Institute of Modern physics, Zhejiang University
Hangzhou 310027, P.R. China

Abstract. Numerical simulations of critical dynamics far from equilibrium are reviewed. Based on the short-time dynamic scaling form, a dynamic approach is developed for numerical measurements of both dynamic and static critical exponents as well as the critical temperature. The method does not suffer from critical slowing down. Recent progress in frustrated systems, Kosterlitz–Thouless phase transitions, first order phase transitions and the fundamental deterministic dynamics is reported.

4.1 Introduction

In the past decade, many activities have been devoted to the non-equilibrium relaxation of critical dynamics. Traditionally, it is believed that universal dynamic scaling behavior exists only in the long-time regime of the dynamic evolution. The short-time dynamic behavior is supposed to depend on microscopic details. In 1989, however, with renormalization group methods Janssen, Schaub and Schmittmann derived a dynamic scaling form for the $O(N)$ vector model, which is valid up to the *macroscopic* short-time regime [1]. It is important that a new independent critical exponent must be introduced to describe the scaling behavior of the initial magnetization. This explains the anomalous behavior of the remanent magnetization in spin-glass dynamics [2].

To understand the short-time dynamic scaling behavior, it is essential to keep in mind that macroscopic initial conditions show up in the dynamic scaling form [1, 3]. For example, the power-law decay of the magnetization in the short-time regime observed in a dynamic relaxation starting from a completely ordered state [4, 5] is not “universal”, in the sense that it depends on the initial conditions. For a dynamic relaxation starting from a disordered state with a small initial magnetization, the magnetization does not decay, and rather shows an initial increase in the short-time regime [1, 3, 6].

Inspired and stimulated by these works, and the relevant experiments and theories in non-equilibrium dynamics, in the last years short-time critical dynamics has been systematically investigated with Monte Carlo methods [3, 6–10]. Simulations have been extended from simple spin models [3, 10, 11], to statistical systems with quenched disorder or frustration [12–15], XY models

and Josephson Junction arrays [16–23], quantum spin systems and lattice gauge theories [24–26], dynamic systems without detailed balance [27, 28], melting transitions [29–31] and fluid systems [32] as well as first-order phase transitions [33–38]. More complete list of the relevant references before 1998 can be found in [3]. All these results confirm the existence of a rather general dynamic scaling form in critical dynamic systems at macroscopic early times, and approximate scaling behavior in weak first-order phase transitions. Such a methodology should also be very interesting in experiments [39, 40]. The physical origin of the dynamic scaling behavior is the divergent or very large correlating time around the phase transition temperatures.

Actually, scaling behavior in non-equilibrium critical systems is not such a unique phenomenon in nature. For example, phase ordering dynamics and non-equilibrium critical dynamics share some similar features [41]. Spin-glass dynamics [2, 7, 13, 42, 43] as well as structural glass dynamics, different kinds of growth dynamics and aging phenomena in complex dynamic systems all may show certain scaling or quasi-scaling behavior.

What we emphasize is that the short-time dynamic scaling form not only is conceptually interesting, but also – more interestingly and importantly, provides new techniques for the measurements of both dynamic and static critical exponents as well as the critical temperature [4, 7, 9, 44], for an earlier review see [3]. Since the measurements are carried out in the short-time regime, *the dynamic approach does not suffer from critical slowing down*. Compared with those methods developed in equilibrium, e.g., the non-local cluster algorithms, the dynamic approach does study the original local dynamics and can be applied to disordered or frustrated systems. Furthermore, it is very difficult to numerically solve dynamic equations with a continuous time to the long-time regime, but the short-time dynamic approach works well [10].

In understanding the universal behavior of short-time critical dynamics, it is very important to distinguish the macroscopic and microscopic time scales. The dynamic scaling emerges only in the *macroscopic* short-time regime, after a time scale t_{mic} which is large enough in microscopic sense. t_{mic} is *not* universal. In some cases, t_{mic} is not so small. For accurate measurements of the critical temperatures and critical exponents, corrections to scaling must be taken into account in the short-time regime.

The purpose of this paper is to review numerical simulations of critical dynamics far from equilibrium in the past decade, and to report recent progress in frustrated systems, Kosterlitz–Thouless phase transitions and the fundamental deterministic dynamics. In Sect. 4.2, the short-time dynamic scaling form will be analyzed. In Sect. 4.3, the applications to second-order-like phase transitions, Kosterlitz–Thouless phase transitions and weak first-order phase transitions will be presented. In Sect. 4.4, the fundamental deterministic dynamics will be investigated. The final section contains some concluding remarks.

4.2 Short-Time Dynamic Scaling Form

4.2.1 Dynamic Scaling Form for Disordered Start

With renormalization group methods, Janssen, Schaub and Schmittmann [1] have shown that far from equilibrium, in a *macroscopic* short-time regime of the dynamical evolution, there already emerges universal scaling behavior. The relaxation process they considered is that a system initially in a high-temperature state, is suddenly quenched to the critical temperature T_c and then evolves with dynamics of model A. For an initial state with a non-zero but *small* magnetization, a generalized dynamic scaling form has been derived with an ϵ -expansion,

$$M^{(k)}(t, \tau, L, m_0) = b^{-k\beta/\nu} M^{(k)}(b^{-z}t, b^{1/\nu}\tau, b^{-1}L, b^{x_0}m_0), \quad (4.1)$$

where $M^{(k)}$ is the k th moment of the magnetization, t is the time of the dynamical relaxation, L is the lattice size, $\tau = (T - T_c)/T_c$ is the reduced temperature, and b is an arbitrary spatial rescaling factor. Here x_0 is a new independent critical exponent, the scaling dimension of the initial magnetization m_0 . It is interesting and important that the critical exponents β , ν and z in (4.1) are exactly those usually defined in equilibrium. These exponents can thus be extracted from the short-time critical dynamics.

Equation (4.1) provides the time evolution of the magnetization, with $k = 1$. Taking $b = t^{1/z}$ and assuming L sufficiently large, it gives for a small m_0

$$M(t, \tau, m_0) \sim m_0 t^\theta F(t^{1/\nu z} \tau), \quad \theta = \frac{(x_0 - \beta/\nu)}{z}. \quad (4.2)$$

Here $M(t, m_0 = 0) = 0$ has been used, and higher order terms of the small quantity $t^{x_0/z} m_0$ have been neglected. Exactly at the critical point ($\tau = 0$), (4.2) predicts a power-law behavior of the magnetization in the short-time region,

$$M(t) \sim t^\theta. \quad (4.3)$$

Up to now, analytical calculations and numerical simulations for a variety of statistical systems show that θ usually is positive, i.e. the magnetization undergoes an initial *increase*. This is a very prominent phenomenon in the short-time critical dynamics [1, 3, 6].

For small but non-zero τ , the power law behavior of the magnetization will be modified by the scaling function $F(t^{1/\nu z} \tau)$. Therefore, searching for the best power law behavior of the magnetization, we may determine the critical temperature in numerical simulations. To estimate the exponent $1/\nu$, we may take the derivative with respect to τ on both sides of (4.2), and obtain the logarithmic derivative of the magnetization

$$\partial_\tau \ln M(t, \tau, m_0)|_{\tau=0} \sim t^{-1/\nu z}. \quad (4.4)$$

Now we consider the case $m_0 = 0$. Using (4.1) for the second moment of the magnetization at the critical temperature gives

$$M^{(2)}(t) \sim t^{-2\beta/\nu z} M^{(2)}(1, t^{-1/z} L). \quad (4.5)$$

In the beginning of the time evolution the spatial correlation length is still small, even at the critical point. Thus it can be deduced that $M^{(2)}(t, L) \sim L^{-d}$. Taking this into account, (4.5) yields a power-law behavior

$$M^{(2)}(t) \sim t^{c_2}, \quad c_2 = \frac{(d - 2\beta/\nu)}{z}. \quad (4.6)$$

An analysis of the autocorrelation (for $m_0 = 0$) shows [45]

$$A(t) = \frac{1}{L^d} \sum_i < S_i(t) S_i(0) > \sim t^{-c_a}, \quad c_a = \frac{d}{z} - \theta. \quad (4.7)$$

In summary [3, 11], simulations of the dynamic process starting from a high temperature state with a small or zero initial magnetization, around the critical point, allow the critical temperature, the quantities θ , $1/\nu z$, c_2 and c_a to be measured and thus determination of the critical exponents θ , $1/\nu$, β/ν and z . The measurements are carried out in the macroscopic short-time regime, and therefore, do not suffer from critical slowing down. Since the spatial correlation length is small in the beginning of the time evolution, the finite size effect can be easily under controlled.

Compared with techniques developed in numerical simulations in equilibrium, the time t in the dynamic approach plays the role of the lattice size L in equilibrium. The dynamic scaling behavior holds only after a microscopic time scale t_{mic} , which is large enough in the microscopic sense but still small in the macroscopic time scale. This is just like that the finite size effect in equilibrium is negligible only for sufficiently large lattices.

4.2.2 Dynamic Scaling Form for Ordered Start

Another interesting and important process is the dynamic relaxation from a completely ordered state. The initial magnetization being exactly at its maximum point $m_0 = 1$, a scaling form

$$M^{(k)}(t, \tau, L) = b^{-k\beta/\nu} M^{(k)}(b^{-z} t, b^{1/\nu} \tau, b^{-1} L) \quad (4.8)$$

is expected. This scaling form looks the same as the dynamic scaling form in the long-time regime, however, it is now assumed already valid in the macroscopic short-time regime.

For the magnetization itself, $b = t^{1/z}$ yields

$$M(t, \tau) = t^{-\beta/\nu z} M(1, t^{1/\nu z} \tau). \quad (4.9)$$

This leads to the power-law behavior

$$M(t) \sim t^{-\beta/\nu z}, \quad (4.10)$$

at the critical point $\tau = 0$. For small but nonzero τ , the power-law behavior of the magnetization will be modified by the scaling function $M(1, t^{1/\nu z} \tau)$, thus allowing for determination of the critical temperature [3]. Taking the derivative with respect to τ on both sides of (4.9), gives the logarithmic derivative of the magnetization

$$\partial_\tau \ln M(t, \tau)|_{\tau=0} \sim t^{-1/\nu z}. \quad (4.11)$$

A Binder cumulant $U(t)$ can be constructed using the magnetization and its second moment. Finite size scaling shows that

$$U(t) = \frac{M^{(2)}}{(M)^2} - 1 \sim t^{d/z}. \quad (4.12)$$

In summary [3, 11], the short-time behavior of the dynamic relaxation starting from a completely ordered state is sufficient to determine all the critical exponents β , ν and z as well as the critical temperature. In practical simulations, these measurements of the critical exponents and critical temperature are usually better in quality than those from a relaxation process starting from a disordered state.

However, sometimes there are strong corrections to scaling, or, it may happen, e.g., for disordered systems, the completely ordered state is not known. Then we should either introduce new concepts and techniques, or stay with the dynamic process with a disordered initial state.

4.2.3 Corrections to Scaling

As mentioned already in Introduction, to describe the non-equilibrium dynamic behavior of certain dynamic systems, one may need to introduce corrections to scaling for the dynamic scaling forms in the preceding two subsections. Important examples are the 2D XY model and fully frustrated XY model which have been intensively studied in the past years [16–20, 22, 23].

Bray, Briant and Jervis have theoretically shown that there is a logarithmic correction for the 2D XY model in the dynamic process starting from a disordered state [17] (see also [46]). In the case with a logarithmic correction to scaling, standard measurements of the critical exponents without taking into account the correction to scaling could be correct only asymptotically in the limit $t \rightarrow \infty$. Therefore, it is very essential to clarify the logarithmic correction.

We first consider the case of $T < T_{KT}$, and a completely disordered initial state with $m_0 = 0$. Assuming a logarithmic correction for the non-equilibrium

spatial correlation length, from scaling analysis and finite size scaling analysis, e.g, the second moment should behave like [17]

$$M^{(2)}(t) \sim \left[\frac{t}{(1 + c \ln(t))} \right]^{(2-\eta)/z}. \quad (4.13)$$

Compared with the scaling behavior in the preceding subsections, here we have replace β/ν by $\eta/2$.

For the dynamic process quenched from a completely ordered state, there may also exist corrections to scaling in the early times, for example, the power-law corrections to scaling [19]

$$M(t) \sim t^{-\eta/2z} \left(1 + \frac{c}{t^b}\right). \quad (4.14)$$

In the cases of the simple Ising and Potts models, corrections to scaling are rather weak [3]. For the models with many meta-stable states such as systems with disorder, frustration or KT transitions, however, corrections to scaling could be strong. For accurate estimate of critical exponents, one needs to take into account corrections to scaling.

4.3 Applications of Short-Time Dynamic Scaling

4.3.1 The 2D *FFXY* Model

In this subsection, we study the second order phase transition in the 2D *FFXY* model [12].

The 2D *FFXY* model exhibits two kinds of phase transitions, i.e. the Kosterlitz–Thouless phase transition (*XY*-like) and the second order phase transition (*Ising*-like). Numerical simulations of the 2D *FFXY* model suffer severely from critical slowing down. Most of the recent works on the 2D *FFXY* model support that these two phase transitions take place at two different temperatures, however, it is not very clear whether the chiral degree of freedom of the *FFXY* model is in the same universality class of the *Ising* model [47–50].

The Hamiltonian of the 2D *FFXY* model can be written as

$$-\frac{H}{kT} = K \sum_{\langle ij \rangle} \cos(\theta_i - \theta_j + A_{ij}), \quad (4.15)$$

where θ_i is the angle of the spin (a planar unit vector) located on site i , A_{ij} determine the frustration and the sum is over the nearest neighbors on a square lattice. The order parameter for the second order phase transition is the staggered chiral magnetization defined as [51]

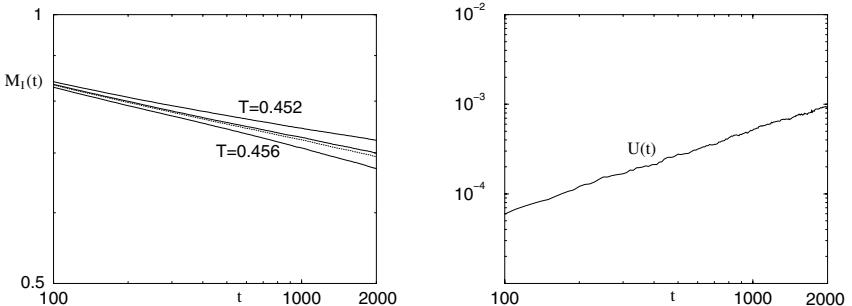


Fig. 4.1. The 2D FFFXY model with the Metropolis algorithm: (*Left*) the chiral magnetization $M_I(t)$ with $M_I(0) = 1$ in log–log scale. From above, the *solid lines* represent $M_I(t)$ at $T = 0.452$, 0.454 and 0.456 . The *dotted line* is at the critical temperature $T_c = 0.4545$. (*Right*) The Binder cumulant $U_I(t)$ at the critical point with $M_I(0) = 1$. Simulations are performed with a lattice size $L = 256$

$$M_I = \left\langle \frac{1}{L^2} \sum_r (-1)^{r_x + r_y} \sin \left[\sum_{\langle ij \rangle \in P_r} \sin(\theta_i - \theta_j + A_{ij}) \right] \right\rangle, \quad (4.16)$$

where r is the coordinate of the plaquette P_r .

In order to locate the critical temperature T_c , we have performed simulations with three different temperatures $T = 0.452$, 0.454 and 0.456 starting from an *ordered* initial state, i.e., the ground state. The average is taken over 2000 samples.

In Fig. 4.1 left, $M_I(t)$ is plotted in log-log scale. Based on (4.9), by searching for a curve with the best power law behavior, one can estimate the critical temperature. To avoid the effect of the t_{mic} , we carry out our measurements in the time interval [200,2000]. The resulting value $T_c = 0.4545(2)$ is very close to those reported in the recent references [48, 49, 52, 53]. The magnetization at T_c is also plotted in Fig. 4.1 left with a dotted line. From the slope of this curve, we find the critical exponent $\beta/\nu z = 0.0602(2)$. The quality of this measurement is very good. With β/ν in hand on can obtain a rigorous z or vice versa [3–5, 54, 55].

To extract the dynamical critical exponent z and the critical exponent $1/\nu z$, we measure the time-dependent Binder cumulant $U_I(t, L) = M_I^{(2)}/M_I^2 - 1$ and $\partial_\tau \ln M_I(t, \tau)$ at $T_c = 0.4545$. In Fig. 4.1 right, the Binder cumulant is plotted. The power law behavior is clearly seen.

In Table 4.1, all critical exponents are listed together with the results reported in literature. Our dynamic measurements support those from [53] but with extra new results for the dynamic exponents z and θ . The exponent ν of the 2D FFFXY model is different from that of the 2D Ising model by nearly 20% and it indicates that the chiral degree of freedom of the 2D FFFXY

Table 4.1. The exponents for the chiral degree of freedom of the 2D $FFXY$ model. For the Ising model, θ is taken from [9, 71], and z from [3, 5, 9, 71]

	ours [12]	[48] (1996)	[49] (1995)	[53] (1994)	[52] (1993)	Ising
T_c	0.4545(2)	0.451(1)	0.452(1)	0.454(2)	0.454(3)	
ν	0.81(2)	0.898(3)	1	0.813(5)	0.80(5)	1
$2\beta/\nu$	0.261(5)			0.22(2)	0.38(2)	0.25
z	2.17(4)					2.165(10)
θ	0.202(3)					0.191(3)

model is in a new universality class. Other exponents of the 2D $FFXY$ model look not much different from those of the 2D Ising model. From our numerical data, we observe that the values of the critical exponents are rather sensitive to the assumed or measured critical temperature T_c . This should be one of the main origins of different values of the exponents reported in the literature.

4.3.2 Kosterlitz–Thouless Transitions

In this subsection, we investigate the Kosterlitz–Thouless (KT) phase transitions in the 2D XY and $FFXY$ models [16, 21].

The Transition Temperature

The 2D XY model is defined by the Hamiltonian

$$-\frac{H}{kT} = K \sum_{\langle ij \rangle} \mathbf{S}_i \cdot \mathbf{S}_j, \quad (4.17)$$

where $\mathbf{S}_i = (S_{i,x}, S_{i,y})$ is a planar unit vector at site i . Large scale Monte Carlo simulations in equilibrium have been performed to understand the properties of the KT phase transition [56, 57]. The results support a KT singularity for the spatial correlation length

$$\xi(\tau) \sim \exp(b\tau^{-\nu}) \quad (4.18)$$

and for the susceptibility $\chi(\tau) \sim \xi^{2-\eta}(\tau)$. However, unconstrained four-parameter fits to the data do not yield completely satisfactory results [57]. η estimated from $\chi(\tau) \sim \xi^{2-\eta}(\tau)$ is above 0.7 and too big compared with the theoretical prediction $\eta = 0.25$. The temperatures for the available data of ξ and χ are still far from the transition temperature T_{KT} estimated to be around 0.89 to 0.90 (for details, see [57]). However, simulations in equilibrium with lower temperatures are very difficult.

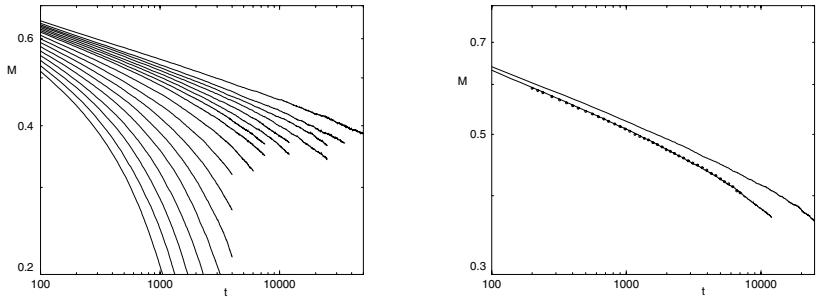


Fig. 4.2. The 2D XY model with the heat-bath algorithm: (*Left*) the magnetization $M(t)$. From above, the temperatures are 0.94, 0.95, 0.955, 0.96, 0.965, 0.97, 0.975, 0.98, 0.99, 1.00, 1.01, 1.02, 1.03, 1.04, 1.05, 1.06, 1.07. (*Right*) The scaling plot of the magnetization with a pair of temperatures (T_1, T_2) . The upper and lower solid lines correspond to temperatures $T_1 = 0.955$ and $T_2 = 0.965$. The circles are also the magnetization of $T_1 = 0.955$ but rescaled to have the best fit with that of $T_2 = 0.965$

We consider the dynamics relaxation starting from an *ordered initial state*, e.g., all $\mathbf{S}_i = (S_{i,x}, S_{i,y}) = (1, 0)$. Total samples for average is from 800 to 1200 for lattice size $L = 256$ and above 400 for $L = 512$.

The observable we measure is the magnetization defined as

$$M(t) = \frac{1}{L^d} \left\langle \sum_i S_{i,x}(t) \right\rangle. \quad (4.19)$$

From a general physical view point of the renormalization group transformation, the magnetization $M(t)$ is subject to a scaling form

$$M(t, \xi(\tau)) = t^{-\eta/2z} M(1, t^{-1/z} \xi(\tau)). \quad (4.20)$$

When the temperature is at T_{KT} (or below), i.e. $\tau = 0$, $\xi(\tau) \rightarrow \infty$ and $M(t)$ undergoes a power law decay $M(t) \sim t^{-\eta/2z}$. However, for $\tau > 0$, the power law behavior is modified by the scaling function $M(1, t^{-1/z} \xi(\tau))$. This fact can be used for the determination $\xi^z(\tau)$ and the exponent η/z . In Fig. 4.2 left, the magnetization is displayed for different temperatures.

From scaling collapse of two curves with a pair of temperatures (T_1, T_2) , we estimate the ratio ξ_1^z/ξ_2^z and η/z . Here ξ_1 and ξ_2 are the values of $\xi(\tau)$ at the temperatures T_1 and T_2 respectively. In Fig. 4.2 right, such a scaling plot is displayed for $(T_1, T_2) = (0.955, 0.965)$.

After extracting ξ_1^z/ξ_2^z for all temperatures, we fit the data to the exponential form in (4.18) and estimate the transition temperature T_{KT} , the exponent ν and parameter bz . The best results are given in Table 4.2 in comparison with those from simulations in equilibrium [16]. Our results are fitted

Table 4.2. The exponents and T_{KT} for the 2D XY model obtained in temperature interval $[T_1, T_2]$ in comparison with those in [57] (denoted by \dagger). T_{KT}^* and ν^* of [57] are from data of the susceptibility. The second and fourth column are the results with a fixed $\nu = 0.5$ as input

$[T_1, T_2]$	[0.94, 1.07]	[0.94, 1.07]	[0.98, 1.43] †	[0.98, 1.43] †
T_{KT}	0.8942	0.8926	0.8953	0.8914
bz	4.12	3.82	3.67	3.38
ν	0.48	0.5	0.47	0.5
T_{KT}^*			0.8871	0.8961
ν^*			0.57	0.5

from a relatively lower temperature interval [0.94, 1.07], and agree well with those obtained in a temperature interval [0.98, 1.43] in [57], in both cases of a unconstrained fit and a fit with a fixed $\nu = 0.5$. The estimate of T_{KT} is relatively stable. However, the unconstrained fit is not very stable.

Relaxation with Disordered Start

Now we study the dynamic relaxation starting from a disordered state for the 2D XY model [21]. We have performed the simulations up to $t = 10240$ Monte Carlo time steps. Samples for averaging are from 12000 to 24000.

To begin our investigation, we first perform simulations with a disordered initial state with a small non-zero magnetization m_0 . In Fig. 4.3 left, $M(m_0, t)$ is displayed on a log–log scale. From these curves, we can not detect a logarithmic correction. In a time interval [100, 1000], direct measurements of the slope yield the same exponents θ as with a power-law correction.

In Fig. 4.3 right, the auto-correlation with a disordered start is displayed on a log–log scale. Looking at the curves by eyes, they are not too far from a power-law behavior.

If one fits the curves with power-law corrections to scaling in the form like (4.14), however, the correction exponent b is rather small. According to the argument in [17], the corrections are logarithmic. In Fig. 4.3 right, dots represent the curves with the logarithmic corrections to scaling, and fit to the numerical data (solid lines) from rather early times. In Table 4.3, the resulting values of $(d - \eta)/z$ and $d/z - \theta$ are given.

The logarithmic correction is so strong such that the effective exponent obtained with a power-law fit would be correct only in the limit of $t \rightarrow \infty$. In the measurements of the critical exponents, therefore, it is extremely important to take into account the logarithmic corrections to scaling.

In the lower sector of Table 4.3, the dynamic exponent z extracted from $A(t)$, $M^{(2)}(t)$ and $M(m_0, t)$ is given. The values are bigger than 2 by 2 or 3 percent. This probably indicates that the logarithmic correction is still not perfect in the time intervals we simulate.

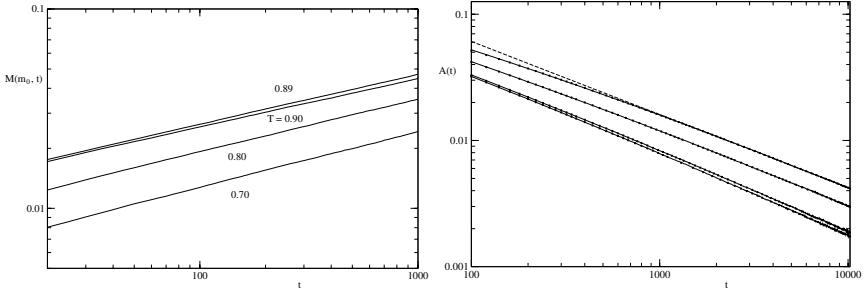


Fig. 4.3. The 2D XY model with the heat-bath algorithm: (*Left*) the magnetization $M(t)$ with disordered initial states and a non-zero m_0 . (*Right*) The auto-correlation $A(t)$ with disordered initial state and a zero m_0 . *Solid lines* are for temperatures $T = 0.90, 0.89, 0.80$ and 0.70 (from below). The *dashed line* shows a power-law fit to $T = 0.70$. *Dots fitted to the solid lines* are with logarithmic corrections to scaling. But the slope of the *dashed line* is 0.579, far from 0.695 with a logarithmic correction to scaling

Table 4.3. The exponents for the 2D XY model after taking into account the corrections to scaling. $\eta/2z$ in the third row for $M(t)$ is obtained with a fixed correction exponent $b = 1$. The value z_1 of the dynamic exponent z is estimated from d/z ; η is calculated from $\eta/2z$ by taking z_1 as input; z_2 is from $(d - \eta)/z$ with η as input; z_3 is calculated from $d/z - \theta$ and θ

		$T = 0.90$	0.89	0.80	0.70
$U(t)$	d/z	1.000(10)	0.995(5)	0.999(4)	0.995(5)
	z_1	2.00(2)	2.01(1)	2.00(1)	2.01(1)
$M(t)$	$\eta/2z$	0.0614(4)	0.0581(2)	0.0441(3)	0.0358(2)
	b	1.13	1.03	0.95	1.07
$M^{(2)}(t)$	$\eta/2z$	0.0611	0.0580	0.0442	0.0357
	fixed b	1	1	1	1
	η	0.246(3)	0.234(2)	0.176(2)	0.144(1)
[57]	η	0.239	0.229	0.179	0.146
$A(t)$	$(d - \eta)/z$	0.860(12)	0.877(9)	0.897(10)	0.920(8)
	z_2	2.04(3)	2.01(2)	2.03(2)	2.02(2)
$M(m_0, t)$	$d/z - \theta$	0.756(5)	0.738(4)	0.711(5)	0.695(6)
	θ	0.241(2)	0.249(2)	0.263(4)	0.280(4)
	z_3	2.01(2)	2.02(2)	2.05(2)	2.05(2)

Relaxation with Ordered Start

We have performed numerical simulations of the dynamic relaxation with an ordered initial state for both the 2D XY and $FFXY$ models [21]. For the 2D $FFXY$ model, the magnetization is defined as the projection of the

Table 4.4. The exponents for the 2D *FFXY* model after taking into account the corrections to scaling. The correction exponent is fixed to be $b = 1$. The dynamic exponent z is estimated from d/z ; η is calculated from $\eta/2z$ by taking z as input

		$T = 0.446$	0.440	0.40	0.30
$U(t)$	d/z	1.019(6)	0.994(10)	1.010(3)	1.007(3)
	z	1.96(1)	2.01(2)	1.98(1)	1.99(1)
$M(t)$	$\eta/2z$	0.0581(18)	0.0506(4)	0.0334(3)	0.0207(2)
	η	0.228(7)	0.203(3)	0.132(2)	0.0824(9)
[51]	η		0.196	0.122	0.064

spin configuration onto the ground state, i.e., the magnetization and its second moment should be calculated for each sublattice separately, and then summed up.

Our careful analysis shows that corrections to scaling here, if any, can be described by a power law in (4.14), and the correction exponent b takes a “universal” value $b = 1$.

The values of d/z and $\eta/2z$ are given in Tables 4.3 and 4.4. The estimated dynamic exponent z is very close to the theoretical value $z = 2$. Our η for the XY model is somewhat bigger than that in [57] at temperatures around T_{KT} , but smaller at lower temperatures. Our value η for the *FFXY* model is bigger than that in [51].

4.3.3 Weak First-Order Phase Transitions

Due to the success of the short-time dynamic approach in second order and KT phase transitions, naturally it is attractive to explore its possible applications to *first* order phase transitions. Especially, due to large correlation lengths and small discontinuities, a *weak* first order transition presents quite similar behavior as a second order one. It has long been challenging how to distinguish one from the other.

To distinguish a first order transition from a second order one, refined methods are typically based on the *finite size scaling* of the specific heat, susceptibility, order parameter, Binder cumulant of energy, or the transition point, e.g. see [58–63]. However, when a first order transition is very weak, it becomes subtle. The double peak structure of the energy distribution together with the finite size scaling shows its merit in this respect [64, 65], but further efficient methods are still desired.

In this subsection, we suggest a short-time dynamic approach to weak first order transitions [33]. The idea is inspired by the existence of two *pseudo* critical points K^* and K^{**} near the weak first order transition point K_c with $K^{**} < K_c < K^*$ [66, 67]. In equilibrium, numerical measurements of K^* and K^{**} are not easy since they are induced by *metastable* states. However, in

Table 4.5. Pseudo critical points K^{**} and K^* measured from short-time dynamics for the 2D 5-state and 7-state Potts models, in comparison with the transition point K_c

	K^{**}	K_c	K^*
$q = 5$	1.174322(2)	1.174359	1.174404(7)
$q = 7$	1.293008(7)	1.293562	1.293854(29)

short-time dynamics K^* and K^{**} can be determined rather accurately from two dynamic processes starting from high temperature and zero temperature states. In second order transitions, K^* and K^{**} overlap with the transition point K_c . Therefore, difference of K^* and K^{**} gives a criterion for a weak first order transition.

As examples, we investigate the 2D q -state Potts models. The transition point is exactly known at $K_c = \ln(1 + \sqrt{q})$. The phase transition is second order for $q \leq 4$ and becomes first order for $q \geq 5$. For small q , the first order transitions are weak.

In Table 4.5, all results for K^* and K^{**} have been collected [33]. For both the 7-state and the 5-state Potts model, K^* and K^{**} are clearly above and below the transition point K_c respectively. The short-time dynamic approach indeed provides a safe criterion for a weak first order transition.

4.4 Numerical Solutions of Deterministic Dynamics

It is usually believed that statistical mechanics is originated from the fundamental deterministic equations of motion for many body systems, even though up to now there exists no general proof. For example, the $O(N)$ vector model and XY model have been numerically solved [68–70], phase transition points and critical exponents in equilibrium have been estimated. The results are in agreement with those obtained from a canonical ensemble in statistical mechanics.

In principle, the deterministic equations of motion should describe not only equilibrium properties but also dynamic properties of the systems. In statistical mechanics, dynamics is approximately given by some effective *stochastic* equations of motion, e.g., Langevin equations or Monte Carlo algorithms. It has long been challenging whether stochastic dynamics is equivalent to the fundamental deterministic dynamics, and vice versa.

The purpose here is to study whether there exists short-time universal scaling behavior in deterministic dynamics, and meanwhile to determine the critical point and all the static and dynamic exponents [10].

The model we choose is the 2D ϕ^4 theory. The Hamiltonian of the model on a square lattice is

$$H = \sum_i \left[\frac{1}{2} \pi_i^2 + \frac{1}{2} \sum_{\mu} (\phi_{i+\mu} - \phi_i)^2 - \frac{1}{2} m^2 \phi_i^2 + \frac{1}{4!} \lambda \phi_i^4 \right], \quad (4.21)$$

with $\pi_i = \dot{\phi}_i$ and it leads to the equations of motion

$$\ddot{\phi}_i = \sum_{\mu} (\phi_{i+\mu} + \phi_{i-\mu} - 2\phi_i) + m^2 \phi_i - \frac{1}{3!} \lambda \phi_i^3. \quad (4.22)$$

In the dynamic evolution governed by (4.22), energy is conserved. The solutions are assumed to generate a microcanonical ensemble. In our short-time dynamic approach, the total energy is a convenient controlling parameter of the system, since it can be input from the initial state. Therefore, from now τ will be understood as a reduced energy density $(\epsilon - \epsilon_c)/\epsilon_c$. Here ϵ_c is the critical energy density corresponding to a second order phase transition.

The order parameter of the ϕ^4 theory is the magnetization. The time-dependent magnetization $M \equiv M^{(1)}(t)$ and its second moment $M^{(2)}$ are defined as

$$M^{(k)}(t) = \frac{1}{L^{2k}} \left\langle \left(\sum_i \phi_i(t) \right)^k \right\rangle, \quad k = 1, 2. \quad (4.23)$$

The average for observables is *only over initial configurations*. We consider a *random* initial state with a zero or small magnetization. For simplicity, we set initial kinetic energy to be zero, i.e. $\dot{\phi}_i(0) = 0$. Following [68], we take parameters $m^2 = 2$. and $\lambda = 0.6$.

To numerically solve these equations of motion, we discretize $\ddot{\phi}_i$ by $(\phi_i(t + \Delta t) + \phi_i(t - \Delta t) - 2\phi_i(t))/(\Delta t)^2$. In our calculations, we use a fairly large lattice $L = 256$ and samples of initial configurations for average range from 4 500 to 10 000. Errors are simply estimated by dividing total samples into two or three groups. Compared with numerical solutions in the long-time regime, the finite Δt effect in the short-time dynamic approach is not severe, since our updating time is limited.

In order to determine the critical point, we perform simulations with non-zero initial magnetization m_0 for a couple of energy densities in the critical regime. In Fig. 4.4 (left), the magnetization has been plotted. Careful analysis leads to $\epsilon_c = 21.11(3)$. This agrees well with $\epsilon_c = 21.1$ given by the Binder cumulant in equilibrium in [68].

At ϵ_c , one measures the exponent $\theta = 0.146(3)$. Accurately speaking, however, the exponent θ is defined in the limit $m_0 \rightarrow 0$. In general, for finite m_0 the exponent θ may show some m_0 -dependence [3]. Therefore, we have performed another simulation with $m_0 = 0.009$ at ϵ_c . The magnetization is also displayed in Fig. 4.4 left with a dashed line. The corresponding exponent is $\theta = 0.158(2)$. If we linearly extrapolate the results to $m_0 = 0$, we obtain the final value $\theta = 0.176(7)$.

With the critical energy in hand, we set $m_0 = 0$ and proceed to measure the auto-correlation $A(t)$ and second moment $M^{(2)}(t)$. In Fig. 4.4 right, $A(t)$

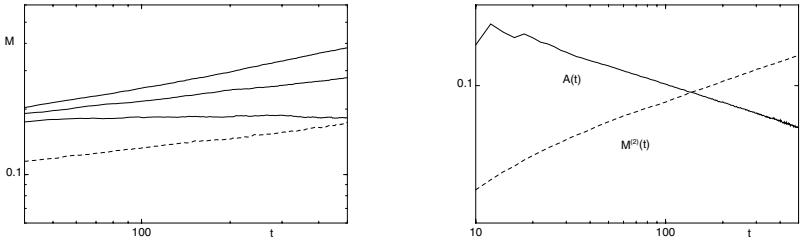


Fig. 4.4. Numerical solutions of the deterministic equations of motion of the 2D ϕ^4 theory: (*Left*) the magnetization with a disordered start in log–log scale. *Solid lines* are for $m_0 = 0.015$ with energy densities $\epsilon = 20.7, 21.1$ and 21.5 (from above), while the *dashed line* is for $m_0 = 0.009$ with $\epsilon_c = 21.11$. (*Right*) The auto-correlation $A(t)$ (*solid line*) and second moment $M^{(2)}(t)$ (*dashed line*) with $m_0 = 0$ at the critical point in log–log scale

Table 4.6. The critical exponents measured for the 2D ϕ^4 theory with the deterministic equations of motion in comparison with those of the Ising model with Monte Carlo dynamics. The values of $2\beta/\nu$ and ν for the Ising model are exact, while others are taken from Table 2 in [3]

	θ	$d/z - \theta$	$(d - 2\beta/\nu)/z$	$1/\nu z$	z	$2\beta/\nu$	ν
ϕ^4	0.176(7)	0.755(5)	0.819(12)	0.492(26)	2.148(20)	0.24(3)	0.95(5)
Ising	0.191(1)	0.737(1)	0.817(7)		2.155(3)	1/4	1

and $M^{(2)}(t)$ are displayed. We clearly see that after $t > t_{\text{mic}} \approx 50$, both curves nicely stabilize to a power law behavior. From the data for $t > 50$, we measure the exponents $d/z - \theta = 0.755(5)$ and $(d - 2\beta/\nu)/z = 0.819(12)$.

In Table 4.6, we summarize all the values of the critical exponents we obtain. For comparison, the exact values of $2\beta/\nu$ and ν for the Ising model and available results of other exponents measured with Monte Carlo dynamics are also given in Table 4.6. Remarkably, not only the static exponents, but also the dynamic exponents of the ϕ^4 theory are very close to those of the Ising model with Monte Carlo dynamics. Therefore, we conclude that the ϕ^4 theory with the deterministic equations of motion are in a same static as well as dynamic universality class of the Ising model with Monte Carlo dynamics.

4.5 Conclusions

We have reviewed recent progress in numerical simulations of critical dynamics far from equilibrium. The short-time dynamic scaling form is numerically verified. Based on the scaling form, a dynamic approach is developed for numerical measurements of both dynamic and static critical exponents as well

as the critical temperature. The method does not suffer from critical slowing down, and carries the feature of predicting the future. Applications of the dynamic approach to frustrated systems, Kosterlitz-Thouless phase transitions, first order phase transitions and the fundamental deterministic dynamics are demonstrated.

Compared with those methods developed in equilibrium, e.g., the non-local cluster algorithms, the dynamic approach does study the original local dynamics, and in principle can be applied to dynamic systems with metastable states, frustration and disorder. In general, however, corrections to scaling should be taken into account to obtain accurate results. For disordered systems, probably new concepts and techniques, e.g., pseudo-magnetization, should be developed.

Acknowledgements

Work supported in part by NNSF (China) under Grant No. 10325520 and 10275054, by the Project-sponsored by SRF for ROCS, SEM, and by DFG (Germany) under Grant No. TR 300/3-3.

References

1. H.K. Janssen, B. Schaub, B. Schmittmann: Z. Phys. B **73**, 539 (1989)
2. D.A. Huse: Phys. Rev. B **40**, 304 (1989)
3. B. Zheng: Int. J. Mod. Phys. B **12**, 1419 (1998), review article
4. D. Stauffer: Physica A **186**, 197 (1992)
5. N. Ito: Physica A **196**, 591 (1993)
6. Z.B. Li, U. Ritschel, B. Zheng: J. Phys. A **27**, L837 (1994)
7. R.E. Blundell, K. Humayun, A.J. Bray: J. Phys. A **25**, L733 (1992)
8. L. Schülke, B. Zheng: Phys. Lett. A **204**, 295 (1995)
9. P. Grassberger: Physica A **214**, 547 (1995)
10. B. Zheng, M. Schulz, S. Trimper: Phys. Rev. Lett. **82**, 1891 (1999)
11. A. Jaster, J. Mainville, L. Schülke, B. Zheng: J. Phys. A **32**, 1395 (1999)
12. H.J. Luo, L. Schülke, B. Zheng: Phys. Rev. Lett. **81**, 180 (1998)
13. H.J. Luo, L. Schülke, B. Zheng: Mod. Phys. Lett. B **13**, 417 (1999)
14. H.J. Luo, L. Schülke, B. Zheng: Phys. Rev. E **64**, 36123 (2001)
15. N. Ito, Y. Ozeki: Physica A **321**, 262 (2003)
16. B. Zheng, M. Schulz, S. Trimper: Phys. Rev. E **59**, R1351 (1999)
17. A.J. Bray, A.J. Briant, D.K. Jervis: Phys. Rev. Lett. **84**, 1503 (2000)
18. L.M. Jensen, B.J. Kim, P. Minnhagen: Phys. Rev. B **61**, 15412 (2000)
19. H.P. Ying, B. Zheng, Y. Yu, S. Trimper: Phys. Rev. E **63**, R35101 (2001)
20. Q.H. Chen, M.B. Luo, Z.K. Jiao: Phys. Rev. B **64**, 212403 (2001)
21. B. Zheng, F. Ren, H. Ren: Phys. Rev. E **68**, 046120 (2003)
22. Y. Ozeki, K. Ogawa, N. Ito: Phys. Rev. E **67**, 026702 (2003)
23. K. Medvedyeva, P. Holme, P. Minnhagen, B.J. Kim: Phys. Rev. E **66**, 026130 (2002)
24. H.P. Ying, H.J. Luo, L. Schülke, B. Zheng: Mod. Phys. Lett. B **12**, 1237 (1998)

25. K. Okano, L. Schülke, B. Zheng: Phys. Rev. D **57**, 1411 (1998)
26. M. Santos: Phys. Rev. E **61**, 7204 (2000)
27. J.F.F. Mendes, M.A. Santos: Phys. Rev. E **57**, 108 (1998)
28. A. Brunstein, T. Tomé: Phys. Rev. E **60**, 3666 (1999)
29. A. Jaster: Phys. Lett. A **258**, 59 (1999)
30. A. Jaster: Physica A **277**, 106 (2000)
31. H. Watanabe, S. Yukawa, Y. Ozeki, N. Ito: Phys. Rev. E **66**, 041110 (2002)
32. W.B. Zhang, X.W. Zou, H.Y. Liu, et al: Phys. Lett. A **272**, 408 (2000)
33. L. Schülke, B. Zheng: Phys. Rev. E **62**, 7482 (2000)
34. E.V. Albano: Phys. Lett. A **288**, 73 (2001)
35. R.A. Monetti, E.V. Albano: Europhys. Lett. **56**, 400 (2001)
36. G.P. Zheng: J. Phys. A **35**, 10549 (2002)
37. Y. Ozeki, K. Kasono, N. Ito, S. Miyashita: Physica A **321**, 271 (2003)
38. G.P. Saracco, E.V. Albano: J. Chem. Phys. **118**, 4157 (2003)
39. K. Zahn, G. Maret: Phys. Rev. Lett. **85**, 3656 (2000)
40. K. Zahn, J.M. Méndez-Alcaraz, G. Maret: Phys. Rev. Lett. **79**, 175 (1997)
41. A.J. Bray: Adv. Phys. **43**, 357 (1994), and references therein
42. K. Fischer, J. Hertz: *Spin Glasses* (Cambridge University Press, Cambridge 1991)
43. J. Kisker, L. Santen, M. Schreckenberg, H. Rieger: Europhys. Lett. **20**, 159 (1992)
44. Z.B. Li, L. Schülke, B. Zheng: Phys. Rev. Lett. **74**, 3396 (1995)
45. H.K. Janssen, in *From Phase Transition to Chaos*. Ed. by G. Györgyi, I. Kondor, L. Sasvári, T. Tél (Topics in Modern Statistical Physics) (World Scientific, Singapore 1992)
46. A.J. Bray: Phys. Rev. E **62**, 103 (2000)
47. P. Olsson: Phys. Rev. Lett. **77**, 4850 (1996)
48. J.V. José, G. Ramirez-Santiago: Phys. Rev. Lett. **77**, 4849 (1996)
49. P. Olsson: Phys. Rev. Lett. **75**, 2758 (1995)
50. P. Olsson: Phys. Rev. B **55**, 3585 (1997)
51. G. Ramirez-Santiago, J.V. José: Phys. Rev. B **49**, 9567 (1994)
52. E. Granato, M.P. Nightingal: Phys. Rev. B **48**, 7438 (1993)
53. S. Lee, K. Lee: Phys. Rev. B **49**, 15184 (1994)
54. D. Stauffer, R. Knecht: Int. J. Mod. Phys. C **7**, 893 (1996)
55. L. Schülke, B. Zheng: Phys. Lett. A **215**, 81 (1996)
56. R. Gupta, J. Delapp, G.G. Batrouni, G.C. Fox, C.F. Bailie, J. Apostlakis: Phys. Rev. Lett. **61**, 1996 (1988)
57. R. Gupta, C.F. Baillie: Phys. Rev. B **45**, 2883 (1992)
58. K. Binder, D.W. Heermann: *Monte Carlo Simulation in Statistical Physics* (Springer, Berlin Heidelberg New York 1992)
59. K. Binder: Rep. Prog. Phys. **60**, 487 (1997)
60. A.M. Ferrenberg, R.H. Swendsen: Phys. Rev. Lett. **63**, 1195 (1989)
61. P.M.C. de Oliveira, S.M.M. de Oliveira, C.E. Cordeiro, D. Stauffer: J. Stat. Phys. **80**, 1433 (1995)
62. W. Janke, R. Villanova: Phys. Lett. A **209**, 179 (1995)
63. A.M. Ferrenberg, D.P. Landau, K. Binder: Phys. Rev. E **58**, 3353 (1998)
64. J. Lee, J.M. Kosterlitz: Phys. Rev. Lett. **65**, 137 (1990)
65. J. Lee, J.M. Kosterlitz, E. Granato: Phys. Rev. B **43**, 11531 (1991)
66. P. de Gennes: *The Physics of Liquid Crystals* (Clarendon, Oxford 1975)

67. L.A. Fernández, J.J. Ruiz-Lorenzo, M.P. Lombardo, A. Tarancón: Phys. Lett. B **277**, 485 (1992)
68. L. Caiani, L. Casetti, M. Pettini: J. Phys. A **31**, 3357 (1998)
69. L. Caiani, L. Casetti, C. Clementi, G. Pettini, M. Pettini, R. Gatto: Phys. Rev. E **57**, 3886 (1998)
70. X. Leoncini, A.D. Verga: Phys. Rev. E **57**, 6377 (1998)
71. K. Okano, L. Schülke, K. Yamagishi, B. Zheng: Nucl. Phys. B **485**, 727 (1997)

Part II

Soft and Disordered Materials

5 Entropy Driven Phase Separation

R. Vink

Institut für Physik, Johannes-Gutenberg-Universität, Mainz, Germany

Abstract. A grand canonical Monte Carlo method for the simulation of a simple colloid-polymer mixture called the AO model will be described. The phase separation known to occur in this model is driven by entropy. The phase diagram of the unmixing transition, the surface tension and the critical point will be determined.

5.1 Introduction

Mixtures of particles are all around us. If you pour oil and water together you create a mixture. Another example is a solution containing colloids and polymers. Mixtures will sometimes phase-separate or unmix. When this happens particles of the same kind cluster together. A famous example is a mixture of oil and water: after unmixing, the lower part of the container contains mostly water and the upper part mostly oil with an interface in between.

There are different reasons for a mixture to unmix. For instance, the unmixed state might carry a lower energy. As an example consider a mixture of A and B particles interacting with the following pair potentials:

$$u_{AA}(r) = u_{BB}(r) = 0, \quad u_{AB}(r) = \frac{\epsilon}{r}, \quad (5.1)$$

where r is the distance between two particles and ϵ some positive constant. In this mixture, A particles do not feel each other, and neither do B particles, but when A and B particles are close together there is an energy penalty. At low temperature, the system tries to minimize the energy and can only do so by moving the A and B particles as far apart as possible. The system thus unmixes.

There is another mechanism that can induce unmixing. This mechanism has its origin in entropy and not in energy. The unmixing of colloid-polymer mixtures for example is driven by entropy. To show that this must be the case, it is instructive to briefly consider the nature of the interactions in a typical colloid-polymer mixture. To a crude approximation the colloids behave as hard spheres. The polymer interactions are more complicated. A real polymer consists of a chain of bonded monomers. In principle all monomers should be taken into account explicitly. However, under certain conditions it is reasonable to describe the polymer chain with only the coordinate of

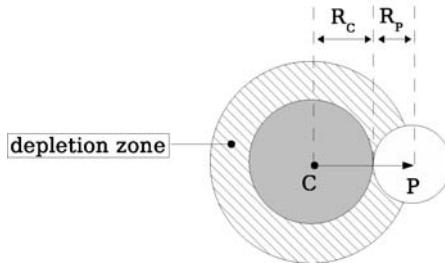


Fig. 5.1. Colloid and polymer at minimum separation (here and throughout this text C=colloid and P=polymer). The distance between the centers of mass equals $R_c + R_p$. This is the minimum allowed separation: any smaller separation is punished with infinite energy. The colloid thus carries a spherical depletion zone with volume V_δ around its center of mass which cannot contain any polymer centers

its center of mass and some effective radius called the radius of gyration. In 1954 Asakura and Oosawa [1] proposed a simple model for colloid-polymer mixtures based on precisely these approximations. In this model colloids and polymers are treated as spheres with respective radii R_c and R_p . Hard sphere interactions are assumed between colloid-colloid (cc) and colloid-polymer (cp) pairs while polymer-polymer (pp) pairs can interpenetrate freely. This leads to the following pair potentials:

$$\begin{aligned} u_{cc}(r) &= \begin{cases} \infty & \text{for } r < 2R_c \\ 0 & \text{otherwise,} \end{cases} & u_{cp}(r) &= \begin{cases} \infty & \text{for } r < R_c + R_p \\ 0 & \text{otherwise,} \end{cases} \\ u_{pp}(r) &= 0, \end{aligned} \quad (5.2)$$

with r again the distance between two particles. The above equations define what is nowadays called the AO model. Note that in 1976 the same model was proposed again and independently by Vrij [2].

According to (5.2) the energy of a valid AO mixture is always zero. Therefore, if unmixing occurs (and it does) it cannot be explained in terms of the energy argument used before. The reason for unmixing is a little more subtle, see Figs. 5.1 and 5.2. Shown in Fig. 5.1 is one colloidal particle just touching one polymer. When the particles touch, the distance between the centers of mass equals $R_c + R_p$. This is the minimum allowed separation between the particles because any smaller separation carries an infinite energy penalty. In the AO model every colloid is thus surrounded by a region which cannot contain any polymer centers of mass. This region is called the depletion zone. In three dimensions the volume of the depletion zone equals

$$V_\delta = \frac{4\pi}{3} (R_c + R_p)^3. \quad (5.3)$$

Consider now Fig. 5.2 which shows a container of volume V holding zero, one and two colloidal particles. Shown at the top is a polymer which we

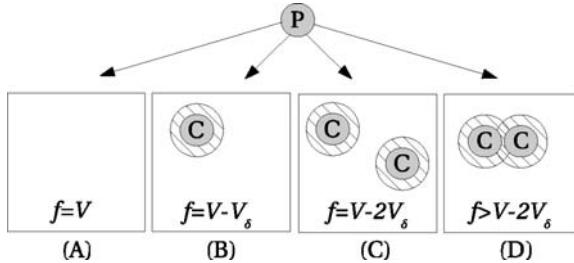


Fig. 5.2. Free volume f available to a single polymer in a container of volume V holding zero, one and two colloidal particles. See text for details

want to insert into the container. If the container is empty we can place the polymer anywhere so the free volume f available to the polymer is simply $f = V$ (Fig. 5.2A). If the container already holds one colloid the free volume decreases to $f = V - V_\delta$ (Fig. 5.2B). The physics becomes interesting when the container holds two or more colloids. Two cases can now be distinguished. In the first case both colloids are well separated and the free volume equals $f = V - 2V_\delta$ (Fig. 5.2C). In the second case the colloids are so close together that their depletion zones overlap and the free volume increases $f > V - 2V_\delta$ (Fig. 5.2D). This immediately has physical consequences: by clustering together the colloids can increase the free volume available to the polymers and hence the entropy of the polymers. Under certain conditions the gain in entropy is sufficient to drive unmixing.

The AO model thus provides a convenient framework in which to study entropy driven unmixing. It has sparked much theoretical work and many simulations [3–7]. Despite its apparent simplicity, predicting the phase behavior of the AO model is no easy task. In this paper an efficient grand canonical Monte Carlo (MC) method that can be used to simulate the AO model will be described. The use of the grand canonical ensemble allows one to bypass certain problems encountered for example in the canonical and Gibbs ensembles. In particular, one can accurately obtain the surface tension and also perform simulations close to the critical point. The method will be used to calculate the phase diagram and the surface tension of the interface. We will also present an accurate determination of the critical point using finite-size scaling.

5.2 Grand Canonical Monte Carlo

Imagine an AO mixture contained in some volume V at inverse temperature $\beta = 1/(k_B T)$. In the grand canonical ensemble V and β are fixed but the number of particles inside V is allowed to fluctuate. The probability of observing a volume containing N_c colloids and N_p polymers is given by the

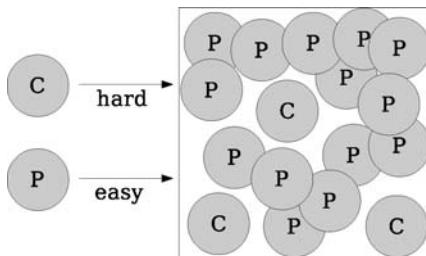


Fig. 5.3. Problem encountered in a standard grand canonical MC simulation of the AO model. When the number of polymers in the volume is substantial it becomes difficult to insert additional colloids. Inserting more polymers remains easy though

grand canonical distribution:

$$P = C z_c^{N_c} z_p^{N_p} e^{-\beta E}, \quad (5.4)$$

where C is a normalization constant, E the energy given by (5.2) and $\{z_c, z_p\}$ the fugacities of the colloids and polymers, respectively.

In a grand canonical MC simulation of the AO model one would like to generate configurations of colloids and polymers that sample (5.4). In the standard approach this is done by attempting to insert a single particle into V at a random location, or remove a single (randomly selected) particle from V . Insertion and removal are usually attempted with equal probability and accepted with probabilities that depend on the energy change of the attempted move and on the fugacities. The standard approach is however not efficient for the AO model as Fig. 5.3 illustrates. The figure shows a volume containing a substantial number of polymers and some colloids. According to (5.2) polymers do not interact with each other so they are happy to overlap. This is what generally will happen as the figure shows. Unfortunately, it is nearly impossible to insert an additional colloid into this system: no matter where the colloid is placed it will likely overlap with at least some of the polymers, and colloid-polymer overlaps are forbidden by (5.2). A standard grand canonical MC simulation of the AO model is thus characterized by a very low acceptance rate of colloid insertions.

The problem illustrated in Fig. 5.3 is typical of asymmetric binary mixtures of which the AO model is an example. The standard grand canonical MC algorithm does not deal well with such mixtures, essentially because it moves only one particle at a time. A MC move capable of removing entire clusters of polymers would be much more efficient. By using such a cluster move the formation of “holes” in the “sea” of polymers is enhanced. If the holes are large enough to contain a colloid, the acceptance rate of colloid insertions will increase. The MC move used in this work to simulate the AO model is aimed at doing precisely that.

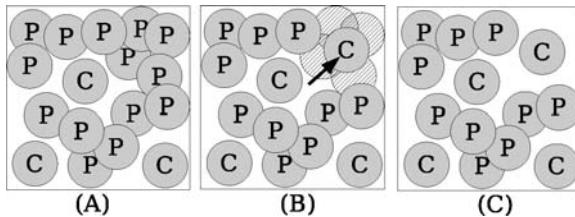


Fig. 5.4. Inserting an additional colloid into an AO mixture. **(A)** The starting configuration. **(B)** The colloid is inserted at a random location inside the volume. **(C)** All polymers that overlap with the colloid are removed

5.3 Cluster Moves

Consider now Fig. 5.4A which again shows an AO mixture of colloids and polymers. This time we are less shy and simply insert an additional colloid at some randomly selected location inside V . The resulting configuration is shown in Fig. 5.4B, with the colloid inserted in the upper right corner of the box. Note that the configuration in Fig. 5.4B is not a valid AO configuration because the colloid overlaps with four polymers. To fix this the overlapping polymers are simply removed, leading to the configuration shown in Fig. 5.4C.

The MC move of Fig. 5.4 is of course not sufficient to carry out a simulation. We also need a reverse MC move capable of bringing us back from the configuration of Fig. 5.4C to the starting configuration of Fig. 5.4A. In constructing the reverse move, the key idea is that for every colloid you take out, a number of polymers must be inserted into the empty depletion zone left behind by the colloid. But exactly how many polymers?

To answer this question it is instructive to consider an AO model containing only polymers. In this case the AO model reduces to a very simple system, namely that of an ideal gas (remember that the polymers do not interact with each other). According to basic statistical mechanics, the average density of an ideal gas is equal to its fugacity. Likewise, for an AO model free of colloids, the average polymer density equals z_p . If we attempt to insert a colloid into the pure polymer system it will on average overlap with $z_p V_\delta$ polymers with fluctuations in the average that are Poisson like, i.e. of order $\sqrt{z_p V_\delta}$. Because of these fluctuations, the number of polymers n_p that we must insert for every colloid that we take out cannot simply be a constant. Instead, n_p must be a random variable drawn from some probability distribution. One choice that works well (see Sect. 5.6) is to draw it uniformly from the interval $n_p \in [0, m]$ (so including 0 but excluding m) and m an integer given by

$$m = 1 + \max \left[1, \text{int} \left(z_p V_\delta + \alpha \sqrt{z_p V_\delta} \right) \right], \quad (5.5)$$

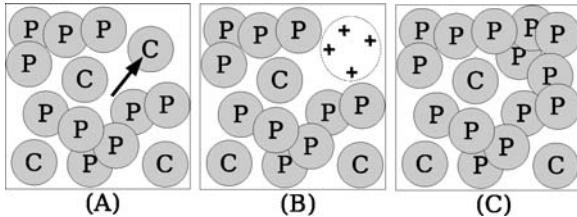


Fig. 5.5. MC move used to remove a colloid from an AO mixture. **(A)** The colloid to be removed is chosen randomly from among those present. **(B)** The colloid is removed and n_p random locations are selected inside the empty depletion zone left behind. The number n_p is uniformly drawn from the interval $[0, m]$. In this example $n_p = 4$. **(C)** Polymers are placed onto the random locations

with α a positive constant of order unity you are free to choose ($\alpha \approx 2.0$ usually gives good results). Recall that V_δ is the volume of the depletion zone given by (5.3).

The reverse move can now be constructed and is illustrated in Fig. 5.5. First, we randomly select a colloid (Fig. 5.5A). The colloid is removed and n_p random locations inside the depletion zone of this colloid are selected, with n_p uniformly drawn from the interval $[0, m]$ (Fig. 5.5B). Finally, polymers are placed onto these random locations resulting in the configuration shown in Fig. 5.5C.

5.4 Detailed Balance

We will now derive the acceptance rates of the MC moves illustrated in Figs. 5.4 and 5.5. The acceptance rates must be constructed such that detailed balance is obeyed. This is vital because the algorithm must sample the grand canonical distribution of (5.4). Recall that the condition of detailed balance demands that [8–10]:

$$P(\sigma)g(\sigma \rightarrow \tau)A(\sigma \rightarrow \tau) = P(\tau)g(\tau \rightarrow \sigma)A(\tau \rightarrow \sigma), \quad (5.6)$$

with $P(\sigma)$ the Boltzmann probability of state σ , $g(\sigma \rightarrow \tau)$ the probability that the MC scheme generates state τ from state σ , and $A(\sigma \rightarrow \tau)$ the probability of accepting the new state τ .

The derivation that follows is based on two assumptions:

1. The integer m of (5.5) is assumed a constant parameter of the algorithm. It must be set once at the start of the simulation and it may not be changed during the course of the simulation.
2. It is assumed that the insertion of a colloid and the removal of a colloid are attempted with equal probability.

In an implementation it is important that the above conditions are met. If they are not the acceptance rates to be derived next may yield wrong results!

5.4.1 Colloid Removal

The acceptance rate for the removal of a colloidal particle is derived first. This is the MC move shown in Fig. 5.5. Assume that we start in a state σ containing N_c colloids, N_p polymers and energy E_σ . It is convenient to label the state as $\sigma(N_c, N_p, E_\sigma)$. The energy E_σ will of course be zero but for the derivation it is convenient to write it down explicitly. After removing the colloid, n_p polymers are inserted into the depletion zone left behind. The state τ that we end up in can thus be labeled as $\tau(N_c - 1, N_p + n_p, E_\tau)$. Note that E_τ need not be zero: if a second colloid happens to be very close to the colloid that was removed, it could now overlap with one or more of the n_p polymers that were just inserted.

To enforce detailed balance the probabilities that appear in (5.6) must be written down. We begin with the easy ones $P(\sigma)$ and $P(\tau)$. These are simply given by the grand canonical distribution of (5.4):

$$P(\sigma) = C z_c^{N_c} z_p^{N_p} e^{-\beta E_\sigma}, \quad P(\tau) = C z_c^{N_c-1} z_p^{N_p+n_p} e^{-\beta E_\tau}. \quad (5.7)$$

The probabilities $g(\sigma \rightarrow \tau)$ and $g(\tau \rightarrow \sigma)$ are more complicated. If we are in state σ the probability of ending up in state τ is given by the product of the probabilities of the individual steps that were taken in the MC move. After close inspection of Fig. 5.5 the following steps can be identified:

1. Selecting a colloid at random. Since state σ contains N_c colloids the probability of choosing one particular colloid is $1/N_c$.
2. Selecting n_p . Since n_p is drawn uniformly from the interval $[0, m]$ the probability of this step is $1/m$.
3. Selecting n_p random locations inside a volume V_δ . The probability of choosing one particular location equals $1/V_\delta$ (see Appendix 5.9). Since n_p locations are selected we must raise this probability to the power n_p . Additionally, we pick up a factorial counting the number of ways in which the locations can be selected. The total probability of this step thus becomes $(n_p)!/V_\delta^{n_p}$.

The probability $g(\sigma \rightarrow \tau)$ is thus the product of the above three terms.

Finally, we derive $g(\tau \rightarrow \sigma)$. This is the probability that the reverse MC move brings us back from state τ to state σ . This move is illustrated in Fig. 5.4 and upon inspecting it we observe that state σ is regenerated if a colloid is placed inside V at precisely the same location it was just removed from. This carries a probability $1/V$ and so we find $g(\tau \rightarrow \sigma) = 1/V$.

Substitution of the above terms into (5.6) and using the Metropolis choice [8] we find that the removal of a colloid must be accepted with probability:

$$A(N_c \rightarrow N_c - 1) = \min \left[1, \frac{m N_c}{z_c V} \frac{(z_p V_\delta)^{n_p}}{(n_p)!} e^{-\beta(E_\tau - E_\sigma)} \right]. \quad (5.8)$$

Note that moves leading to configurations where E_τ is not zero are automatically rejected by the above equation.

5.4.2 Colloid Insertion

To make the algorithm complete the acceptance rate for the insertion of a colloid must still be derived, i.e. the MC move of Fig. 5.4. We can follow the same reasoning as before with perhaps one subtlety. We start again in a state σ containing N_c colloids, N_p polymers and energy E_σ , so with label $\sigma(N_c, N_p, E_\sigma)$. All polymers (say n_p of them) that overlap with the inserted colloid are removed, so the state τ that we end up in can be labeled as $\tau(N_c + 1, N_p - n_p, E_\tau)$. Again, the energy E_τ need not be zero: if we are unlucky the inserted colloid overlaps with one or more of the N_c colloids already present. According to (5.2) such overlaps also carry infinite energy.

To construct detailed balance we first write down $P(\sigma)$ and $P(\tau)$:

$$P(\sigma) = C z_c^{N_c} z_p^{N_p} e^{-\beta E_\sigma}, \quad P(\tau) = C z_c^{N_c+1} z_p^{N_p-n_p} e^{-\beta E_\tau}. \quad (5.9)$$

Next, we derive $g(\sigma \rightarrow \tau)$ which is the probability that the MC move of Fig. 5.4 generates state τ starting in state σ . This involves only the random selection of a location inside V so we obtain: $g(\sigma \rightarrow \tau) = 1/V$.

Finally, we consider $g(\tau \rightarrow \sigma)$. This is the probability that the reverse MC move of Fig. 5.5 transforms state τ back into state σ . This will happen only if the following conditions are met: 5

1. The newly inserted colloid is removed again. Since state τ contains $N_c + 1$ colloids the probability of this step equals $1/(N_c + 1)$.
2. Precisely the same number n_p of polymers that were removed are inserted again. This step requires some care. Assume that we chose m rather small in (5.5). In that case the forward move of Fig. 5.4 might remove more polymers than the reverse move of Fig. 5.5 can possibly put back. This will happen if $n_p \geq m$. Therefore, the probability p that exactly the same number of polymers are inserted must be written as

$$p = \begin{cases} 0 & \text{if } n_p \geq m \\ 1/m & \text{otherwise.} \end{cases} \quad (5.10)$$

3. Selecting the same n_p coordinates inside V_δ . As was explained before this probability equals $(n_p)!/V_\delta^{n_p}$.

The probability $g(\tau \rightarrow \sigma)$ is thus the product of the above three terms.

Substitution into (5.6) and using the Metropolis choice [8] we find that the insertion of a colloid must be accepted with probability

$$A(N_c \rightarrow N_c + 1) = \begin{cases} 0 & \text{if } n_p \geq m \\ \min \left[1, \frac{z_c V}{m(N_c + 1)} \frac{(n_p)!}{(z_p V_\delta)^{n_p}} e^{-\beta(E_\tau - E_\sigma)} \right] & \text{otherwise.} \end{cases} \quad (5.11)$$

Again, moves leading to configurations where E_τ is not zero are automatically rejected by the above equation.

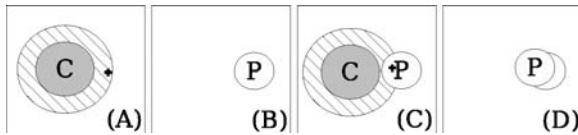


Fig. 5.6. Even low values of m may yield high polymer densities. (A) The starting configuration showing a colloid with its depletion zone. The colloid is removed and one polymer is inserted at a random location inside the depletion zone (*cross*) resulting in configuration (B). (C) Another colloid is inserted close to the polymer. (D) The colloid is removed again with one polymer inserted in its depletion zone at the cross resulting in two overlapping polymers

5.5 Ergodicity

Now that the MC scheme obeys detailed balance we need to check if it is ergodic. An algorithm is ergodic if every point in phase space can be reached with finite probability. It is easy to see that the algorithm is ergodic as far as the colloids are concerned. For every inserted colloid, a random location inside V was selected so the entire volume is correctly sampled. Similarly, polymers also sample the entire volume but do so indirectly, namely via the removal of colloids. Whenever a colloid is removed, a number of polymers are inserted in the depletion zone. Since the colloids sample the entire volume, so then do the polymers.

One might argue that the choice of m in (5.5) violates ergodicity. If at most $m - 1$ polymers are inserted for every colloid, the polymer density will always be less than $(m - 1)/V_\delta$ as a result. While this objection sounds reasonable it is in fact unjustified as Fig. 5.6 shows.

Figure 5.6 shows an example simulation of the AO model using the cluster algorithm just described with a low value of m , namely $m = 2$. This means that for every removed colloid, at most one polymer is inserted. The starting configuration is a volume containing one colloid, see Fig. 5.6A. The edge of the depletion zone of the colloid is also drawn. In Fig. 5.6B the colloid is removed and one polymer is placed very close to the edge of the depletion zone (but still inside of it, of course). The simulation is continued in Fig. 5.6C where another colloid is placed into the system. This colloid is placed close to the polymer already present, but not too close so the polymer survives. Finally, in Fig. 5.6D the colloid is removed again and one polymer is placed inside the depletion zone of this colloid, close to the other polymer. The configuration in Fig. 5.6D now shows two polymers practically stacked on top of each other, even though the algorithm was run with $m = 2$. This removes the above objection.

Running the algorithm with such a low value of m is not recommended though. The whole point of this discussion is to show that m does not influence the correctness of the algorithm, only its efficiency.

5.6 Early Rejection Scheme

In Sect. 5.3 it was argued that the number of polymers n_p that one needs to insert for every colloid that is removed cannot simply be a constant. In fact, it was shown that n_p is a random variable described by a Poisson distribution. Yet, in the subsequent description of the algorithm it was decided to draw n_p uniformly, and not from a Poisson distribution. One might wonder if the efficiency of the algorithm is not seriously impeded by this choice. The answer is it is not, provided one implements the so-called early rejection scheme.

In many MC simulations, the usual approach is to make a change to the system, calculate the involved energy change, select a random number between zero and one, compare this random number to the acceptance rate and finally accept or reject the move. In some cases, particularly in systems where the interactions are hard sphere like, one can do much better.

Consider for example the removal of a colloidal particle displayed in Fig. 5.5 and the associated acceptance rate (5.8). Most of the quantities appearing in (5.8) are already known at the start of the move. For example N_c , V_δ , E_σ , m , β and the fugacities. In fact, the only unknowns are n_p and E_τ . We also know that if the move is ever going to be accepted E_τ will be zero at the end. So the only remaining unknown is n_p which, if you recall, is a random number drawn from the interval $n_p \in [0, m]$. The early rejection scheme proceeds as follows:

1. Select a random number r between zero and one.
2. Select n_p uniformly from the interval $n_p \in [0, m]$.
3. Calculate the acceptance rate $A(N_c \rightarrow N_c - 1)$ using (5.8) with the above value for n_p and assuming that E_τ is zero.
4. If $r > A(N_c \rightarrow N_c - 1)$ reject the move immediately otherwise proceed to the next step.
5. Remove the colloid and insert the n_p polymers. If any of the inserted polymers produce overlap with other colloids reject the move, otherwise the move is accepted.

The most CPU time consuming step in the above scheme is step five. However, this step is only performed for those values of n_p that are reasonable. Unreasonable values for n_p were already filtered out in step four at the cost of only a few multiplications and selecting two random numbers. In this scheme it therefore does not matter so much what distribution n_p is drawn from. Needless to say, the early rejection scheme is highly recommended. To speed up the determination of overlap the link-cell method should also be used.

5.7 Application

In this section the grand canonical cluster scheme will be used to study bulk phase separation in the AO model. Since the AO energy is either zero or

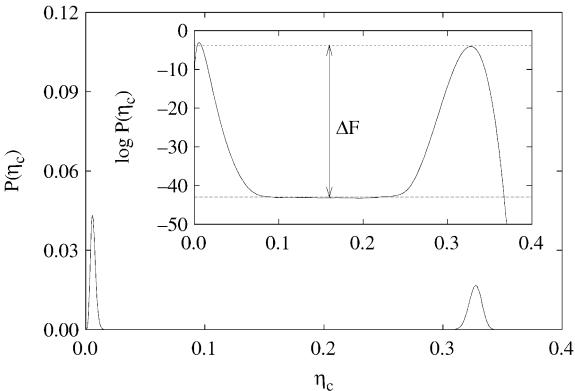


Fig. 5.7. Probability $P(\eta_c)$ of observing a colloid packing fraction η_c for an AO mixture with $\{q = 0.8; z_c = 87.2; \eta_p^r = 1.0\}$ in a simulation of dimensions $L_x = L_y = 16.7$ and $L_z = 33.4$. The probability is not normalized. The inset shows the logarithm of the probability distribution

infinity the temperature plays no role so we simply put $\beta = 1$ in the acceptance rates. The phase behavior of the AO model is thus fixed by the colloid to polymer size ratio $q = R_p/R_c$ and the fugacities $\{z_c, z_p\}$. We consider here a size ratio $q = 0.8$ and put $R_c = 1$ to set the length scale. The simulations are performed in a box with edges $L_x \times L_y \times L_z$ and using periodic boundary conditions. Following convention we define $\eta_p^r \equiv z_p(4\pi/3)R_p^3$ known as the polymer reservoir packing fraction. Note that η_p^r has no relation to the number of polymers actually in the system. It is just a different way of expressing the polymer fugacity z_p .

In a naive implementation of the scheme one sets the fugacities $\{z_c, \eta_p^r\}$ and starts the simulation. As the simulation proceeds colloids and polymers will enter and leave the box. The crucial quantity to measure is $P(\eta_c)$ defined as the probability of observing a box with colloid packing fraction $\eta_c \equiv (4\pi/3)R_c^3 N_c/V$. During the simulation one thus maintains a histogram counting how often a certain colloid packing fraction has occurred.

If phase separation occurs $P(\eta_c)$ is bimodal. An example distribution is shown in Fig. 5.7. The peak at low η_c corresponds to the colloid vapor phase (V), the peak at high η_c to the colloid liquid phase (L) and the region in between is the phase-separated regime. The distribution in Fig. 5.7 is at coexistence which means that the area under both peaks is equal. At coexistence, the simulation spends equal time in both phases on average. Also shown is the logarithm of $P(\eta_c)$. The physical significance of this curve is its relation to the free energy. The height of the barrier marked ΔF corresponds to the free energy barrier separating the phases. As was shown by Binder [11] this barrier is related to the surface tension via $\gamma = \Delta F/(2A)$ with A the area of the interface.

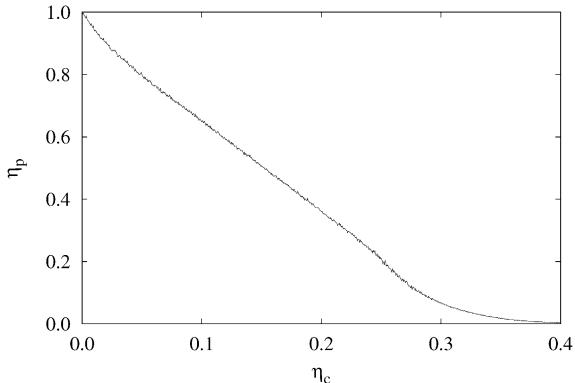


Fig. 5.8. Average polymer packing fraction η_p as a function of the colloid packing fraction η_c for an AO mixture at coexistence with $q = 0.8$ and $\eta_p^r = 1.0$

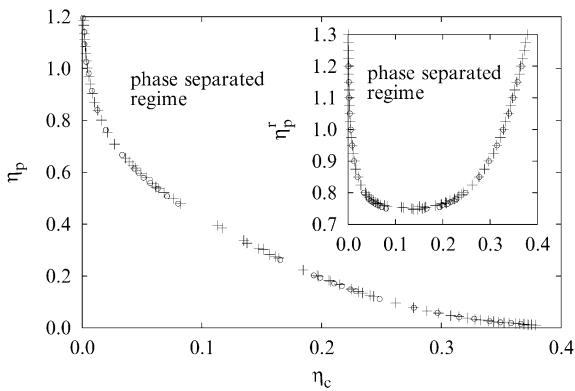


Fig. 5.9. Phase diagram of the AO model with $q = 0.8$ in system representation. Crosses were obtained using box dimensions $L_x = L_y = 16.7$ and $L_z = 33.4$; open circles were obtained in a smaller box with dimensions $L_x = L_y = 13.3$ and $L_z = 26.5$. The inset shows the phase diagram in reservoir representation

It is also interesting to measure the average polymer packing fraction $\eta_p \equiv (4\pi/3)R_p^3N_p/V$ as a function of η_c . This result is shown in Fig. 5.8. In the pure polymer phase ($\eta_c = 0$) we expect $\eta_p = \eta_p^r$ because the polymers then mimic the ideal gas. This is precisely what Fig. 5.8 shows. As the colloid packing fraction increases the polymer packing fraction in the system decreases to zero.

In the AO model η_p^r is the control parameter, much like temperature is for fluid-vapor transitions. To obtain the phase diagram one simply has to measure $P(\eta_c)$ at coexistence for a number of different η_p^r . The problem in a simulation is finding the value of z_c that yields coexistence for the chosen η_p^r of interest. Additionally, if the barrier ΔF is high, it will be difficult to sample

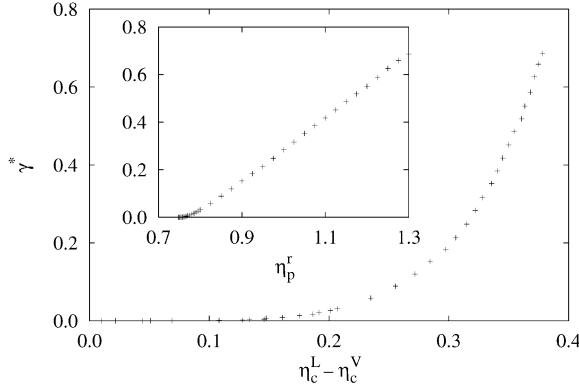


Fig. 5.10. Reduced surface tension $\gamma^* \equiv 4R_c^2\gamma$ for an AO mixture with $q = 0.8$ as a function of the difference in the colloid packing fractions in the coexisting liquid and vapor phases. The box dimensions were $L_x = L_y = 16.7$ and $L_z = 33.4$. The inset shows γ^* as a function of η_p^r

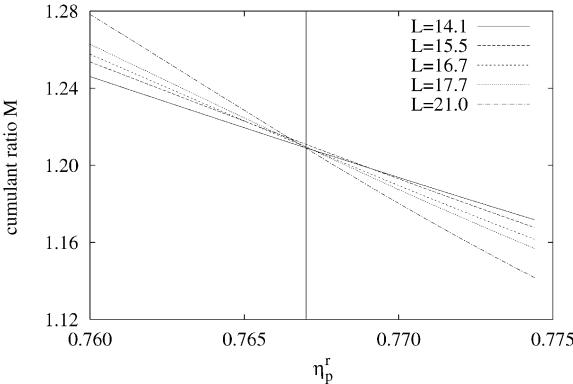


Fig. 5.11. Cumulant ratio M given by (5.12) as a function of η_p^r for an AO mixture with $q = 0.8$ for various system sizes. The simulations were performed in a cubic box with edge L as indicated. From the intercept (vertical line) we obtain for the critical polymer fugacity $\eta_{p,cr}^r = 0.766 \pm 0.002$

$P(\eta_c)$ in the region between the peaks. Fortunately, an array of techniques is available to overcome these problems [12]. The results in this work were obtained using a new technique called *Successive Umbrella Sampling* [13].

For each $P(\eta_c)$ at coexistence one reads off the colloid packing fraction of the vapor phase η_c^V and of the liquid phase η_c^L and plots the two points (η_c^V, η_p^r) and (η_c^L, η_p^r) in a graph. This yields the phase diagram in reservoir representation shown in the inset of Fig. 5.9. By using Fig. 5.8 we can convert the reservoir representation into the experimentally more relevant $\{\eta_c, \eta_p\}$ or system representation also shown in Fig. 5.9. For every $P(\eta_c)$ one also obtains

a value for the surface tension using the method of Binder. The results of this procedure are shown in Fig. 5.10, in two different representations.

Finally, we determine the critical polymer fugacity defined as the value of η_p^r above which phase separation begins to take place. From the inset of Fig. 5.9 we see that the critical fugacity is around $\eta_p^r \approx 0.75$. We have performed a finite size scaling analysis [14] by measuring the cumulant ratio

$$M = \frac{\langle (\eta_c - \langle \eta_c \rangle)^2 \rangle}{\langle |\eta_c - \langle \eta_c \rangle| \rangle^2}, \quad (5.12)$$

as a function of η_p^r close to the critical point for different system sizes. The results are shown in Fig. 5.11. The critical fugacity is at the intersection of the lines from which we obtain $\eta_{p,cr}^r = 0.766 \pm 0.002$.

5.8 Conclusions

The grand canonical cluster scheme described in this chapter is very successful at modeling phase separation in AO mixtures. In fact, at the time of writing, it is unsurpassed in speed and accuracy by other simulation methods [15]. However, the method does have its limitations. If the packing fraction of the colloids is high (say 0.40 and above) the algorithm is no longer efficient. In that case the insertion of colloidal particles fails, not because of overlap with polymers, but because of overlap with other colloids. This problem is well known in hard sphere simulations.

The algorithm could be improved for systems where the polymer-polymer interaction is not zero. In these cases the random insertion of n_p polymers into the depletion zone of a colloid may no longer be efficient. Fortunately, the algorithm is easily adapted to include smarter insertion moves such as configurational bias [16], recoil growth [17], wormhole MC [18] and perhaps PERM [19]. Work along these lines is in progress.

Note also that the presented method is general. The derivation of detailed balance for example can easily be modified to an AO model confined between walls or to systems interacting with smooth potentials. This would be the subject of further work.

Acknowledgments

I am grateful to the Deutsche Forschungsgemeinschaft for support (TR6/A5) and to J. Horbach, K. Binder, M. Müller, M. Schmidt and P. Virnau for collaboration and discussion.

5.9 Appendix: Random Points

Here we argue that the selection of a random point inside a volume V carries a probability $1/V$. The question is most easily answered by assuming that

the volume is made up of many tiny cells, each with a volume Γ . The total number of cells in the volume is thus equal to $n = V/\Gamma$ and the probability of picking one of them is $1/n = \Gamma/V$.

The difficulty lies in choosing Γ . In statistical mechanics it is assumed that the smallest element of phase space has a volume determined by the thermal wavelength Λ leading to $\Gamma = \Lambda^3$. This explains why the acceptance rates of grand canonical MC schemes often contain the thermal wavelength [9, 10]. Unfortunately, in a MC simulation this is not very useful. Since the thermal wavelength depends on temperature, particle mass and even the Planck constant, this choice suggests that a simulation of for example hard spheres is temperature and mass dependent.

The key observation is that the physics of the system is indifferent to the choice of Γ . To see this consider (5.8) which is the acceptance rate for the removal of a colloidal particle. Strictly speaking, the volumes V and V_δ appearing in (5.8) must be replaced by V/Γ and V_δ/Γ , respectively. These additional factors of Γ are, however, readily absorbed into the fugacities z_c and z_p . The fugacity is given by $z = \exp(\beta\mu)$ with μ the chemical potential, so a rescaling of the fugacity merely shifts $\beta\mu$ by a constant. This constant has no physical consequence because it in turn shifts the Hamiltonian by a constant. Thus, Γ has no physical consequence and may therefore be set to unity which was done throughout this text.

References

1. S. Asakura, F. Oosawa: J. Chem. Phys. **22**, 1255 (1954)
2. A. Vrij: Pure Appl. Chem. **48**, 471 (1976)
3. D.G.A.L. Aarts, R. Tuiner, H.N.W. Lekkerkerker: J. Phys. Condens. Matter **14**, 7551 (2002)
4. R. Roth, R. Evans, S. Dietrich: Phys. Rev. E **62**, 5360 (2000)
5. M. Schmidt, H. Löwen, J.M. Brader, R. Evans: Phys. Rev. Lett. **85**, 1934 (2000)
6. M. Dijkstra, R. van Roij: Phys. Rev. Lett. **89**, 208303 (2002)
7. P. Bolhuis, A. Louis, J. Hansen: Phys. Rev. Lett. **89**, 128302 (2002)
8. M.E.J. Newman, G.T. Barkema: Monte Carlo Methods in Statistical Physics (Clarendon, Oxford 1999)
9. D.P. Landau, K. Binder: A Guide to Monte Carlo Simulations in Statistical Physics (Cambridge University Press, Cambridge 2000)
10. D. Frenkel, B. Smit: Understanding Molecular Simulation (Academic Press, New York 2000)
11. K. Binder: Phys. Rev. A **25**, 1699 (1982)
12. N. Wilding: Am. J. Phys. **69**, 1147 (2001)
13. P. Virnau, M. Müller, submitted (2003), preprint cond-mat/0306678
14. See, for example, N. Wilding: Annual Reviews of Computational Physics IV, p. 37 (1996)
15. R.L.C. Vink, J. Horbach, submitted (2003), preprint cond-mat/0310404
16. J. Siepmann, D. Frenkel, Mol. Phys. **75**, 59 (1992)

17. S. Consta, N.B. Wilding, D. Frenkel, Z. Alexandrowicz: *J. Chem. Phys.* **110**, 3220 (1999)
18. J. Houdayer: *J. Chem. Phys.* **116**, 1983 (2002)
19. H. Hsu, V. Mehra, W. Nadler, P. Grassberger: *J. Chem. Phys.* **188**, 444 (2003)

6 Supercooled Liquids under Shear: Computational Approach

R. Yamamoto, R. Temam², J. Dean², D. Grove¹, C. Chambers²,
K.B. Bruce², and E. Bertino¹

¹ Department of Physics, Kyoto University, Kyoto 606-8502, Japan,

² PRESTO, Japan Science and Technology Agency
4-1-8 Honcho, Kawaguchi, Saitama, Japan

Abstract. We examine the microscopic dynamics of supercooled liquids under shear by performing extensive molecular dynamics simulations of a two-dimensional binary liquid with soft-core interactions near, but above, the glass transition temperature. Our simulations show that a drastic reduction of the structural relaxation time and the shear viscosity occurs due to shear. The structural relaxation time decreases as $\dot{\gamma}^{-\nu}$ with an exponent $\nu \leq 1$, where $\dot{\gamma}$ is the shear rate. The microscopic dynamics were confirmed to be surprisingly isotropic regardless of the strength of the anisotropic shear flow.

6.1 Introduction

As liquids are cooled toward the glass transition, the dynamics are drastically slowed down, while only small changes can be detected in the static properties. One of the main targets of theoretical investigations of the glass transition is to identify the mechanism of this drastic slowing-down. To this end, a number of molecular dynamics (MD) simulations have been carried out for supercooled liquids and revealed that the dynamics in supercooled liquids are spatially *heterogeneous* [1–10]. In our previous studies, we have examined bond breakage processes among adjacent particle pairs by MD simulations of two (2D) and three dimensional (3D) model fluids. We found that the broken bonds determined with an appropriate time interval ($\simeq 0.05\tau_b$, where τ_b is the average bond breakage time), are almost equivalent to the critical fluctuations in Ising spin systems. To support this picture, the structure factor of the broken bonds can be excellently fit to the Ornstein-Zernike form [4, 5]. The correlation length ξ determined from this analysis increases with decreasing temperature T and is related to τ_b or the structural α relaxation time τ_α via the dynamic scaling law, $\tau_\alpha \simeq 0.1\tau_b \sim \xi^z$, with $z = 4$ in 2D and $z = 2$ in 3D. The heterogeneous structure of the bond breakage is essentially the same as that in local diffusivity [6].

Another striking example occurs when one brings supercooled liquids away from equilibrium. By rapidly changing the temperature or applying shear flow, supercooled liquids undergo unique phenomena known as aging or shear thinning [11–15]. These phenomena are not only conceptually new

but also practically important. However, physical properties of glassy materials have not yet been well-understood in nonequilibrium conditions. In previous work, we performed extensive MD simulations of binary soft-core mixtures in two and three dimensions with and without shear flow. We found that the dynamical properties of the supercooled liquids under shear can be mapped onto those at quiescent states at higher temperatures [5]. In the present study, we calculate intermediate scattering functions in shear flow by using a method proposed by Onuki [17, 18] to examine the microscopic dynamics of sheared supercooled liquids. Simulations have been done in 2D to compare the present computational results directly with a theory developed recently for sheared supercooled liquids in 2D [19, 20].

6.2 Simulation Method

To prevent crystallization and obtain stable amorphous states via MD simulations, we choose a model system composed of two different particle species, 1 and 2, which interact via the soft-core potential

$$v_{ab}(r) = \epsilon(\sigma_{ab}/r)^{12}, \quad (6.1)$$

with $\sigma_{ab} = (\sigma_a + \sigma_b)/2$, where r is the distance between the two particles, and a, b denote particle species ($\in \{1, 2\}$). We take the mass ratio to be $m_2/m_1 = 2$, the size ratio to be $\sigma_2/\sigma_1 = 1.4$, and the number of particles $N = N_1 + N_2$, where $N_1 = N_2 = 5000$. Simulations are performed in the presence and absence of shear flow keeping the particle density and the temperature fixed at $n = n_1 + n_2 = 0.8/\sigma_1^2$ ($n_1 = N_1/V$, $n_2 = N_2/V$) and $k_B T = 0.526\epsilon$. Space and time are measured in units of σ_1 and $\tau_0 = (m_1\sigma_1^2/\epsilon)^{1/2}$, respectively. The size of the unit cell is $L = 118$. In the absence of shear, we impose microcanonical conditions and integrate Newton's equations of motion

$$\frac{d\mathbf{r}_i^a}{dt} = \frac{\mathbf{p}_i^a}{m_a}, \quad \frac{d\mathbf{p}_i^a}{dt} = \mathbf{f}_i^a. \quad (6.2)$$

Very long equilibration periods are used so that no appreciable aging (slow equilibration) effect is detected in various thermodynamic quantities, such as the pressure, or in time correlation functions. Here, $\mathbf{r}_i^a = (r_{xi}^a, r_{yi}^a)$ and $\mathbf{p}_i^a = (p_{xi}^a, p_{yi}^a)$ denote the position and the momentum of the i -th particle of species a , and \mathbf{f}_i^a is the force acting on the i -th particle of species a . In the presence of shear, the momentum $\mathbf{p}'_i^a = \mathbf{p}_i^a - m_a \dot{\gamma} r_{yi}^a \hat{\mathbf{e}}_x$ is defined as the momentum deviations relative to mean Couette flow. Using the Lee-Edwards boundary condition, we integrate the so-called SLLOD equations of motion keeping the temperature $k_B T (\equiv N^{-1} \sum_a \sum_i (\mathbf{p}'_i^a)^2 / m_a)$ at a desired value using a Gaussian-constraint thermostat to eliminate viscous heating effects [21]. We impose shear for $t \geq 0$ after a long equilibration time. Data

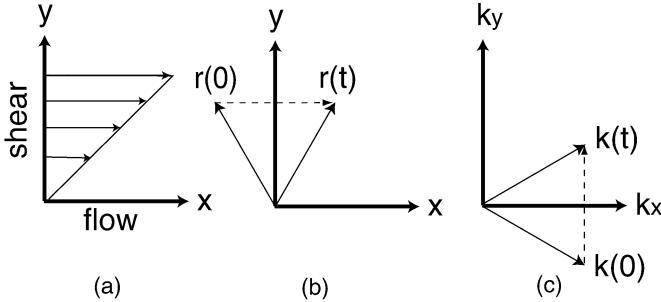


Fig. 6.1. (a) Geometry of shear flow. (b) Shear advection in real space. (c) Shear advection in Fourier space

for the analysis has been taken and accumulated in steady states which can be realized after transient waiting periods.

Figure 6.1a shows the geometry of shear flow in the present simulation. As shown in Fig. 6.1b, shear flow with the rate $\dot{\gamma}$ advect a positional vector \mathbf{r} as

$$\mathbf{r}(t) = \mathbf{r} + \dot{\gamma} t \mathbf{r}_y \mathbf{e}_x, \quad (6.3)$$

for time t , where \mathbf{e}_α is a unit vector in $\alpha \in \{x, y\}$ axis. The corresponding time-dependent wave vector

$$\mathbf{k}(t) = \mathbf{k} + \dot{\gamma} t k_x \mathbf{e}_y, \quad (6.4)$$

is shown in Fig. 6.1c. The above definition enable us to calculate the Fourier component \mathbf{k} of the time correlation function

$$C(\mathbf{k}, t) \equiv \langle A_{-\mathbf{k}(t)}(t) A_{\mathbf{k}}(0) \rangle \quad (6.5)$$

in shear flow [17,18]. We thus calculate the incoherent and the coherent parts of the scattering function for the binary mixture using the definitions [22]

$$F_{sa}(\mathbf{k}, t) = \frac{1}{N_a} \left\langle \sum_{i=1}^{N_a} e^{-i\{\mathbf{k}(-t) \cdot \mathbf{r}_i^a(t) - \mathbf{k} \cdot \mathbf{r}_i^a(0)\}} \right\rangle \quad (6.6)$$

and

$$F_{ab}(\mathbf{k}, t) = \frac{1}{N} \left\langle \sum_{i=1}^{N_a} e^{-i\mathbf{k}(-t) \cdot \mathbf{r}_i^a(t)} \sum_{j=1}^{N_b} e^{i\mathbf{k} \cdot \mathbf{r}_j^b(0)} \right\rangle, \quad (6.7)$$

with $a, b \in \{1, 2\}$. The α relaxation time τ_α of the present mixture, defined by

$$F_{11}(\mathbf{k}_0, \tau_\alpha) \simeq F_{s1}(\mathbf{k}_0, \tau_\alpha) = e^{-1}, \quad (6.8)$$

is equal to $\tau_\alpha \simeq 1800$ time units in the quiescent state for $|\mathbf{k}_0| = 2\pi/\sigma_1$.

6.3 Simulation Results

6.3.1 Microscopic Structure

The partial static structure factors $S_{ab}(\mathbf{k})$ are defined as

$$S_{ab}(\mathbf{k}) = \int d\mathbf{r} e^{i\mathbf{k}\cdot\mathbf{r}} \langle \hat{n}_a(\mathbf{r}) \hat{n}_b(\mathbf{0}) \rangle, \quad (6.9)$$

where

$$\hat{n}_a(\mathbf{r}) = \sum_j^{N_a} \delta(\mathbf{r} - \mathbf{r}_j^a) \quad (a \in \{1, 2\}), \quad (6.10)$$

is the local number density of the species a . Note, the dimensionless wave vector \mathbf{k} is measured in units of σ_1^{-1} . For a binary mixture, there are three combinations of partial structure factors, $S_{11}(\mathbf{k})$, $S_{22}(\mathbf{k})$, and $S_{12}(\mathbf{k})$. These are plotted in Fig. 6.2a in the quiescent state after taking an angular average over \mathbf{k} . A density variable representing the degree of particle packing, corresponding to the density of an effective one component system, can be defined for the present binary system as

$$\hat{\rho}_{\text{eff}}(\mathbf{r}) = \sigma_1^2 \hat{n}_1(\mathbf{r}) + \sigma_2^2 \hat{n}_2(\mathbf{r}). \quad (6.11)$$

The corresponding dimensionless structure factor is given by

$$S_{\rho\rho}(\mathbf{k}) = \sigma_1^{-4} \int d\mathbf{r} e^{i\mathbf{k}\cdot\mathbf{r}} \langle \delta \hat{\rho}_{\text{eff}}(\mathbf{r}) \delta \hat{\rho}_{\text{eff}}(\mathbf{0}) \rangle \quad (6.12)$$

$$= n_1^2 S_{11}(\mathbf{k}) + n_2^2 (\sigma_2/\sigma_1)^4 S_{22}(\mathbf{k}) + 2n_1 n_2 (\sigma_2/\sigma_1)^2 S_{12}(\mathbf{k}), \quad (6.13)$$

where $\delta \hat{\rho}_{\text{eff}} = \hat{\rho}_{\text{eff}} - \langle \hat{\rho}_{\text{eff}} \rangle$. One can see from Fig. 6.2b that $S_{\rho\rho}(k)$ has a pronounced peak at $k \simeq 5.8$ and becomes very small (~ 0.01) for small k , demonstrating that our system is highly incompressible at long wavelengths. Because $S_{\rho\rho}(k)$ behaves quite similarly to $S(k)$ of one component systems, we examine space-time correlations in $\hat{\rho}_{\text{eff}}(\mathbf{r})$ rather than those in the partial number density $\hat{n}_a(\mathbf{r})$ for the present binary system. The use of $\hat{\rho}_{\text{eff}}(\mathbf{r})$ makes comparison of our simulation data with the mode-coupling theory developed for a one component system more meaningful.

We next examine the anisotropy in the static structure factor $S_{\rho\rho}(\mathbf{k})$ in the presence of shear flow. Figures 6.3 and 6.4 show $S_{\rho\rho}(\mathbf{k})$ plotted on a two-dimensional k_x - k_y -plane (upper part) and the angular averaged curves (lower part) within the regions a-d obtained at $\dot{\gamma} = 10^{-3}$ and 10^{-2} , respectively. One sees that at a lower shear rate ($\dot{\gamma} = 10^{-3}$), the shear distortion is negligible at all the regions except for region d where the peak heights of $S_{\rho\rho}(\mathbf{k})$ start decreasing. But, at higher shear rate ($\dot{\gamma} = 10^{-2}$) the distortion becomes prominent. This seems to indicate that the nonlinear effects due to shear become important in this region. Ronis has explored the higher shear regime

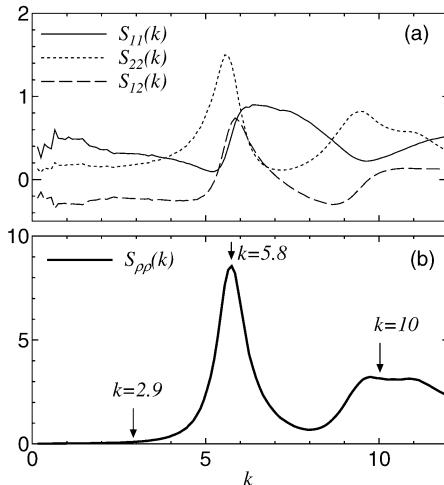


Fig. 6.2. Partial structure factors $S_{ab}(k)$ in (a) and $S_{\rho\rho}(k)$ in (b) defined by (6.13) for the present binary mixture

for the structure factor of hard sphere colloidal suspensions and concluded that, at higher shear, the peak should always be lower than the equilibrium value plus a shift that depends on the direction [23]. At even higher shear rate ($\dot{\gamma} = 10^{-1}$, though the figure is not shown here), peaks in all directions have been lowered. The peak with maximal distortion (region b) is shifted to lower wave vectors while the opposite is true for the shift of the region with minimal distortion (region d). The qualitative agreement with Ronis' theory is good, but it is not clear that our results can be explained by a simple two-body theory such as that of Ronis. Recently, Szamel has analyzed $S(k)$ for hard-sphere colloidal suspensions up to linear order in $\dot{\gamma}$ [24]. He took three-body correlations into account and found quantitative agreement with the shear viscosity evaluated using $S(k)$.

It should be noted that because our system is an atomic liquid without solvent, we cannot directly refer to the Péclet number since the bare diffusion coefficient D_0 does not exist. Therefore, a direct and quantitative comparison with the theories discussed above is not possible. However, we can estimate the relaxation times from the self-diffusion coefficients measured in [5]. This allows us to estimate the Péclet number in the range between 10^{-1} and 10^2 , which corresponds to the highest shear rates explored in the theoretical analysis of our paper [20].

6.3.2 Tagged Particle Motions

We generalize the displacement vector of the tagged (j th) particle as

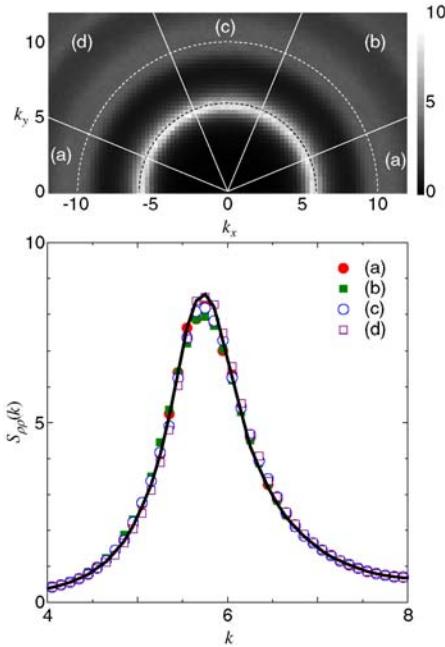


Fig. 6.3. *Upper:* $S_{\rho\rho}(\mathbf{k})$ at $\dot{\gamma} = 10^{-3}$ plotted in a two-dimensional (k_x, k_y) plane. *Lower:* The solid line is $S_{\rho\rho}(k)$ at equilibrium. Dots represent those observed in the region indicated in the (k_x, k_y) plane above

$$\Delta\mathbf{r}_j(t) = \mathbf{r}_j(t) - \dot{\gamma} \int_0^t dt' y_j(t') \mathbf{e}_x - \mathbf{r}_j(0), \quad (6.14)$$

where \mathbf{e}_x is the unit vector in the x (flow) direction. In this displacement, the contribution from convective transport by the average flow has been subtracted. We then analyze the mean square displacement,

$$r_2(t) \equiv \langle [\Delta\mathbf{r}(t)]^2 \rangle = \frac{1}{N_1} \sum_{j=1}^{N_1} \langle [\Delta\mathbf{r}_j(t)]^2 \rangle, \quad (6.15)$$

of tagged particles for the smaller component (species 1). MSDs obtained at $T = 0.53$ with shear rates $10^{-4} \leq \dot{\gamma} \leq 10^{-1}$ are plotted in Fig. 6.5, where the mean square displacements of the x and y components of the vector $\Delta\mathbf{r}_j(t)$ are separately displayed. Analogous to the effect of increasing temperature, the shear flow enhances the mobility of particles. It is also demonstrated that the statistical distribution of $\Delta\mathbf{r}_i(t)$ is surprisingly isotropic even at the highest shear rate.

Figure 6.6 shows the non-Gaussian parameter for the smaller component

$$\alpha_2(t) \equiv \frac{3\langle \Delta\mathbf{r}(t)^4 \rangle}{5\langle \Delta\mathbf{r}(t)^2 \rangle^2 - 1}, \quad (6.16)$$

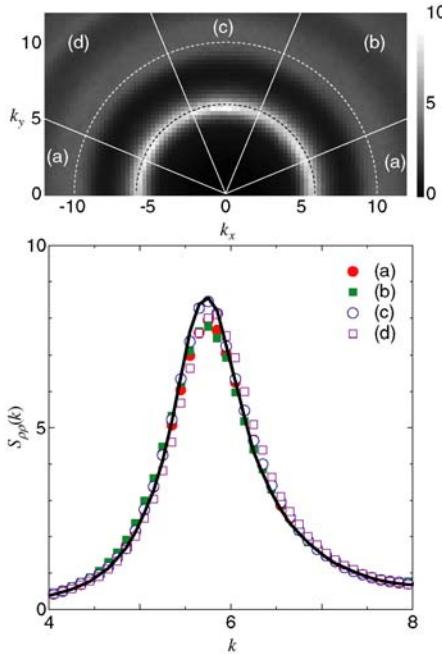


Fig. 6.4. The same as Fig. 6.3 but shear is increased to $\dot{\gamma} = 10^{-2}$

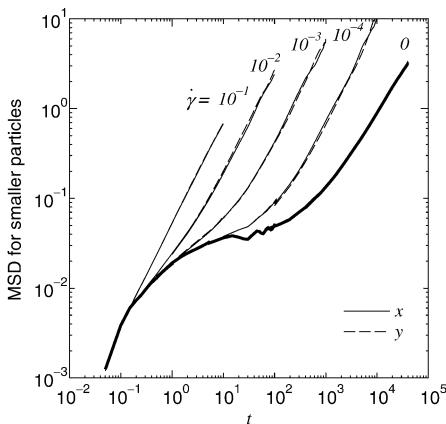


Fig. 6.5. The mean square displacements (MSD) obtained at $T = 0.53$ under shear flow

which assumes a maximum value at $t = t^*$. One finds that both t^* and $\alpha_2^* \equiv \alpha_2(t^*)$ tend to decrease with increasing shear rate. This is again analogous to the behavior observed when the temperature is increased without shear flow. A quantitative comparison of these two effects (increasing temperature

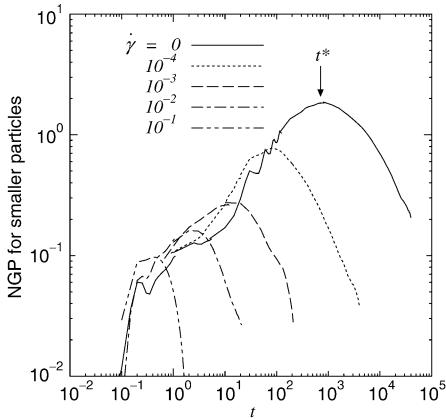


Fig. 6.6. The non-Gaussian parameter (NGP) obtained at $T = 0.53$ under shear flow

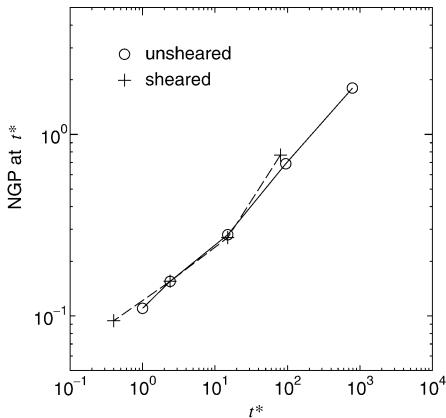


Fig. 6.7. Peak values of NGP, α_2^* , are plotted against t^* at which NGP shows a peak. Data are taken by both changing temperature without shear and changing shear rate at a fixed temperature $T = 0.53$ (+)

and increasing shear rare) is done in Fig. 6.7, which supports the analogy between the two effects.

We may also conclude that the decrease in NGP^* with increasing temperature or shear rate represents the suppressed heterogeneity, which becomes significant in glassy states.

6.3.3 Intermediate Scattering Functions

Here we examine the dynamics of the local density variable $\hat{\rho}_{\text{eff}}(\mathbf{r}, t)$. To this end, we define the intermediate scattering function,

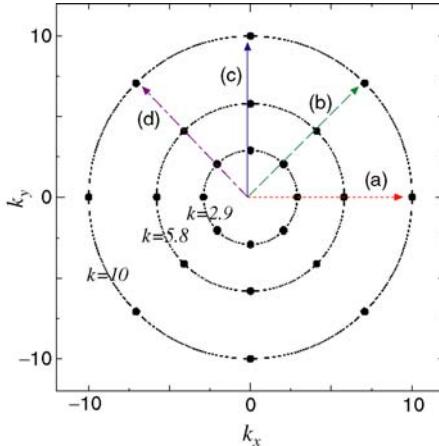


Fig. 6.8. Sampled wavevectors

$$F_{\rho\rho}(\mathbf{k}, t) = n_1^2 F_{11}(\mathbf{k}, t) + n_2^2 \left(\frac{\sigma_2}{\sigma_1} \right)^4 F_{22}(\mathbf{k}, t) + 2n_1 n_2 \left(\frac{\sigma_2}{\sigma_1} \right)^2 F_{12}(\mathbf{k}, t), \quad (6.17)$$

by taking a linear combination of the partial scattering functions defined in (6.7). Note that $F_{\rho\rho}(\mathbf{k}, 0) = S_{\rho\rho}(\mathbf{k})$ by definition. To investigate anisotropy in the scattering function $F_{\rho\rho}(\mathbf{k}, t)$, the wave vector \mathbf{k} is taken in four different directions \mathbf{k}_{10} , \mathbf{k}_{11} , \mathbf{k}_{01} , and \mathbf{k}_{-11} , where

$$\mathbf{k}_{\mu\nu} = \frac{k}{\sqrt{\mu^2 + \nu^2}} (\mu \hat{\mathbf{e}}_x + \nu \hat{\mathbf{e}}_y), \quad (6.18)$$

and $\mu, \nu \in \{0, 1\}$ as shown in Fig. 6.8. The wave vector k (in reduced units) is taken to be 5.8 (see also Fig. 6.2b). Because we use the Lee-Edwards periodic boundary condition, the available wave vectors in our simulations should be given by

$$\mathbf{k} = \frac{2\pi}{L} (n \hat{\mathbf{e}}_x, (m - n D_x) \hat{\mathbf{e}}_y), \quad (6.19)$$

where n and m are integers and $D_x = L \dot{\gamma} t$ is the difference in x -coordinate between the top and bottom cells as depicted in Fig. 6.5 of [21]. To suppress statistical errors, we sample about 80 available wave vectors around $\mathbf{k}_{\mu\nu}$ and calculate $F_{\rho\rho}(\mathbf{k}, t)$ using (6.17) and (6.7). Then we average $F_{\rho\rho}(\mathbf{k}, t)$ over the sampled wave vectors. The sampled wave vectors are indicated by dots in Fig. 6.8.

Figure 6.9 displays $F_{\rho\rho}(\mathbf{k}, t)/S_{\rho\rho}(\mathbf{k})$ for $k = 5.8$. Several features are noticeable. First, the quantitative trends as a function of $\dot{\gamma}$ are similar for different values of $\dot{\gamma}$. Secondly, shear drastically accelerates microscopic structural relaxation in the supercooled state. The structural relaxation time τ_α decreases strongly with increasing shear rate as $\tau_\alpha \sim \dot{\gamma}^{-\nu}$ with $\nu \simeq 1$ as shown

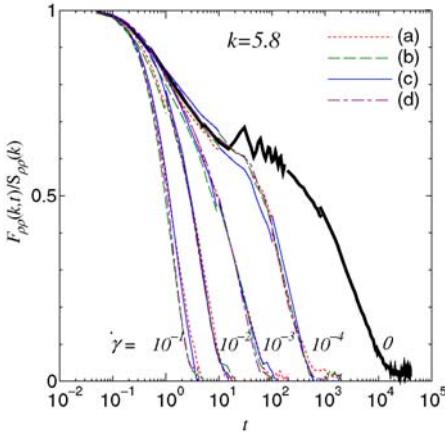


Fig. 6.9. $F(\mathbf{k}, t)/S(\mathbf{k})$ at $k\sigma_1 = 5.8$ for various shear rates and at the different observing points (a) \mathbf{k}_{10} , (b) \mathbf{k}_{11} , (c) \mathbf{k}_{01} , and (d) \mathbf{k}_{-11} as explained in Fig. 6.8

in Fig. 6.10. Thirdly, the acceleration in the dynamics due to shear occurs almost isotropically. Finally, almost the same trend is found for $k = 2.9$ and $k = 10$ as shown in Fig. 6.11. We observe surprisingly small anisotropy in the scattering functions, even under extremely strong shear, $\dot{\gamma}\tau_\alpha \simeq 10^3$. A similar isotropy in the tagged particle motions has already been reported in Fig. 6.5 [5]. The observed isotropy is more surprising than that observed in single particle quantities. In particular, the fact that different particle labels are correlated in the collective quantity defined in (6.7) means that a simple transformation to a frame moving with the shear flow cannot completely remove the directional character of the shear. Our results provide *post facto* justification for the isotropic approximation of [25]. The observed simplicity in the dynamics is quite different from the behavior of other complex fluids such as critical fluids or polymers, where the dynamics become noticeably anisotropic in the presence of shear flow.

6.4 Conclusions

We have performed extensive MD simulations for two- and three-dimensional binary liquids with soft-core interactions near, but above, the glass transition temperature. In our previous studies, we identified *weakly-bonded* or *relatively-active* regions from breakage of appropriately defined bonds. [4, 5] We also found that the spatial distributions of such regions resemble the critical fluctuations in Ising spin systems, so the correlation length ξ can be determined. The correlation length goes up to the system size as T is lowered, but no divergence seems to exist at nonzero temperatures. We have also demonstrated that the diffusivity in supercooled liquids is spatially hetero-

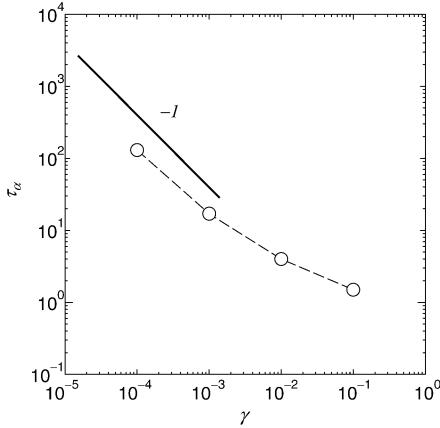


Fig. 6.10. The alpha relaxation time τ_α at $T = 0.53$ under shear flow. The solid line indicates $\nu = 1$

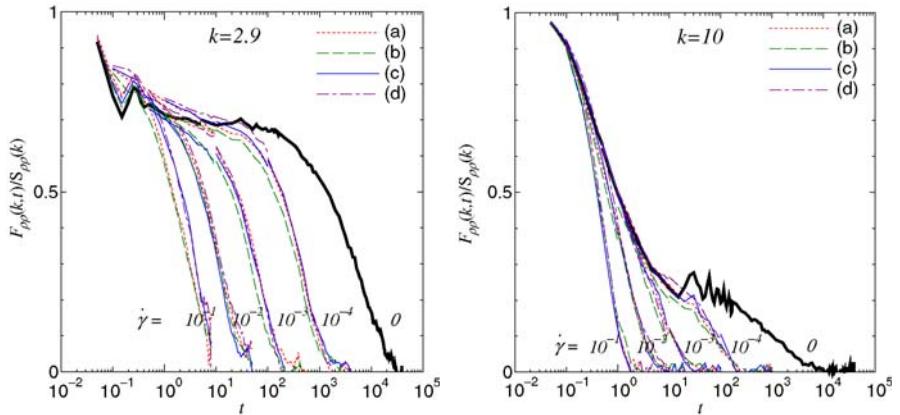


Fig. 6.11. Left: $F(\mathbf{k}, t)/S(\mathbf{k})$ at (a) $k\sigma_1 = 2.9$. Right: $F(\mathbf{k}, t)/S(\mathbf{k})$ at $k\sigma_1 = 10$. All symbols are as in Fig. 6.9

geneous on time scales shorter than $3\tau_\alpha$, which leads to the breakdown of the Stokes-Einstein relation [6]. The heterogeneity detected is essentially the same as that of the bond breakage in our former works [4, 5].

In the present study, we examined the microscopic dynamics of a two-dimensional supercooled liquid under shear. The numerical analysis illustrated several interesting features such as (i) drastic reduction of relaxation times and the viscosity (ii) almost isotropic relaxation irrespective of the direction of the flow. The fact that the dynamics are almost isotropic supports the simplified view of sheared supercooled liquids proposed so far. In this view, the effect of shear is transformed into scalar parameters, such as the ef-

fective temperature, and the anisotropic nature of the nonequilibrium states is not explicitly considered [15, 25].

Acknowledgments

The author would like to thank A. Onuki, D.R. Reichman, and K. Miyazaki for their collaborations and helpful discussions. Numerical calculations have been carried out at the Human Genome Center, Institute of Medical Science, University of Tokyo.

References

1. T. Muranaka, Y. Hiwatari: Phys. Rev. E **51**, R2735 (1995); Prog. Theor. Phys. Suppl. **126**, 403 (1997)
2. M.M. Hurley, P. Harrowell: Phys. Rev. E **52**, 1694 (1995)
3. D.N. Perera, P. Harrowell: Phys. Rev. E **54**, 1652 (1996)
4. R. Yamamoto, A. Onuki: J. Phys. Soc. Jpn. **66**, 2545 (1997)
5. R. Yamamoto, A. Onuki: Phys. Rev. E **58**, 3515 (1998)
6. R. Yamamoto, A. Onuki: Phys. Rev. Lett. **81**, 4915 (1997)
7. W. Kob, C. Donati, S.J. Plimton, P.H. Poole, S.C. Glotzer: Phys. Rev. Lett. **79**, 2827 (1997)
8. C. Donati, J.F. Douglas, W. Kob, S.J. Plimton, P.H. Poole, S.C. Glotzer: Phys. Rev. Lett. **80**, 2338 (1998)
9. P. Ray, K. Binder: Europhys. Lett. **27** 53 (1994)
10. C. Dasgupta, A.V. Indrani, S. Ramaswamy, M.K. Phani: Europhys. Lett. **15**, 307 (1991)
11. J.H. Simmons, R.K. Mohr, C.J. Montrose: J. Appl. Phys. **53**, 4075 (1982); J.H. Simmons, R. Ochoa, K.D. Simmons: J. Non-Cryst. Solids **105**, 313, (1988). Y. Yue, R. Brückner: J. Non-Cryst. Solids **180**, 66 (1994)
12. R. Yamamoto, A. Onuki: Europhys. Lett. **40**, 61 (1997)
13. R. Yamamoto, A. Onuki: J. Chem. Phys. **117**, 2359 (2002)
14. A.J. Liu, S.R. Nagel: Nature **396**, 21 (1998)
15. J.-L. Barrat, L. Berthier: Phys. Rev. E **63**, 012503 (2001); L. Berthier, J.-L. Barrat: Phys. Rev. Lett. **89**, 095702 (2002); L. Berthier, J.-L. Barrat: J. Chem. Phys. **116**, 6228 (2002)
16. I.K. Ono, C.S. O'Hern, D.J. Durian, S.A. Langer, A.J. Liu, S.R. Nagel: Phys. Rev. Lett. **89**, 095703 (2002)
17. A. Onuki, K. Kawasaki: Ann. Phys. (N.Y.) **121**, 456 (1979); A. Onuki: J. Phys.: Condens. Matter **9**, 6119 (1997)
18. A. Onuki: Phase Transition Dynamics (Cambridge University Press, Cambridge 2002)
19. K. Miyazaki, D.R. Reichman: Phys. Rev. E **66**, 050501 (2002)
20. K. Miyazaki, D.R. Reichman, R. Yamamoto: cond-mat/0401528, Phys. Rev. E (submitted)
21. D.J. Evans, G.P. Morriss: Statistical Mechanics of Nonequilibrium Liquids (Academic Press, New York 1990)

22. A similar but slightly different definition for $Fs(\mathbf{k}, t)$ was given in [5]. We however confirmed that the both definitions give almost identical results for the present case
23. D. Ronis: Phys. Rev. A **29**, 1453 (1984)
24. G. Szamel: J. Chem. Phys. **114**, 8708 (2001)
25. M. Fuchs, M.E. Cates: Phys. Rev. Lett. **89**, 248304 (2002); F. Matthias, M.E. Cates: J. Phys. Condens. Matter **15**, S401 (2003); F. Matthias, M.E. Cates: Faraday Discuss. **123**, 267 (2003)

7 Optimizing Glasses with Extremal Dynamics

S. Boettcher¹ and A.G. Percus²

¹ Physics Department, Emory University, Atlanta, Georgia 30322, USA

² Computer & Computational Sciences Division, Los Alamos National Laboratory, Los Alamos, NM 87545, USA

Abstract. A local-search heuristic for finding high-quality solutions for many hard optimization problems is explored. The method is inspired by recent progress in understanding far-from-equilibrium phenomena in terms of *self-organized criticality*. This method, called *extremal optimization*, successively replaces the value of extremely undesirable variables in a sub-optimal solution with new, random ones. Large, avalanche-like fluctuations in the cost function emerge dynamically. These enable the search to effectively scaling barriers to explore local optima in distant neighborhoods of the configuration space while eliminating the need to tune parameters. Extremal optimization complements approximation methods inspired by equilibrium statistical physics, such as *simulated annealing*. Although the method is quite general, we focus here on applications to spin glass ground states.

Natural systems that exhibit self-organizing qualities [1] often possess a common feature: a large number of strongly coupled entities with similar properties. Hence, at some coarse level they permit a statistical description. An external resource drives the system which then takes its direction purely by chance. Like descending water breaking through the weakest of all barriers in its wake, biological species are coupled in a global comparative process that persistently washes away the least fit. In this process, unlikely but highly adapted structures surface inadvertently. Optimal adaptation thus emerges naturally, from the dynamics, simply through a selection *against* the extremely “bad”. In fact, this process prevents the inflexibility inevitable in a controlled breeding of the “good”. Based on this notion, we introduce a new optimization heuristic, Extremal Optimization (EO) [2].

To introduce EO, let us consider a spin glass as a specific example of a hard optimization problem. We wish to minimize the cost function,

$$H(\mathbf{x}) = -\frac{1}{2} \sum_{\langle i,j \rangle} J_{ij} x_i x_j , \quad (7.1)$$

where the sum extends over all nearest-neighbor pairs of spins. Due to frustration, ground state configurations S_{\min} are hard to find, and it has been shown that it is among the hardest optimization problems [3] for $d > 2$.

To find near-optimal solutions for a particular optimization problem, EO performs a neighborhood search on a single configuration $S \in \Omega$. As in the

spin problem in (7.1), S consists of a large number n of variables x_i . We assume that we can define for each S a neighborhood $N(S)$ that rearranges the state of merely a small number of the variables. Those neighborhoods are a characteristic of a local search, in contrast to a genetic algorithm, say, where cross-overs may effect $O(n)$ variables on each update. The cost $C(S)$ is assumed to consist of the individual cost contributions, or “fitnesses”, λ_i for each variable x_i (analogous to the fitness values in the Bak-Sneppen model [4]). The fitness of each variable assesses its contribution to the total cost:

$$C(S) = - \sum_{i=1}^n \lambda_i . \quad (7.2)$$

Typically, each λ_i depends on the state of x_i in relation to connected variables. Here, for the Hamiltonian in (7.1), we assign to each spin x_i the fitness

$$\lambda_i = \frac{1}{2} x_i \sum_j J_{ij} x_j , \quad (7.3)$$

so that (7.2) is satisfied. Each spin’s fitness thus corresponds to (the negative of) its local energy contribution to the overall energy of the system. In similarity to the Bak-Sneppen model, EO then proceeds through a neighborhood search of Ω by sequentially changing variables with “bad” fitness on each update, for instance, via single spin-flips. After each update, the fitnesses of the changed variable and of all its connected neighbors are reevaluated according to (7.3). The basic EO algorithm proceeds as follows:

1. Initialize configuration S at will; set $S_{\text{best}} := S$.
2. For the “current” configuration S ,
 - a) evaluate λ_i for each variable x_i ,
 - b) find j satisfying $\lambda_j \leq \lambda_i$ for all i , i.e., x_j has the “worst fitness,”
 - c) choose $S' \in N(S)$ such that j *must* change its state,
 - d) accept $S := S'$ *unconditionally*,
 - e) if $C(S) < C(S_{\text{best}})$ then set $S_{\text{best}} := S$.
3. Repeat at step (2) as long as desired.
4. Return S_{best} and $C(S_{\text{best}})$.

The algorithm operates on a single configuration S at each step. Each variable x_i in S has a fitness, of which the “worst” is identified. This *ranking* of the variables provides the only measure of quality on S , implying that all other variables are “better” in the current S . In the move to a neighboring configuration, typically only a small number of variables change state, so only a few connected variables need to be re-evaluated [step (2a)] and re-ranked [step (2b)]. Note that there is not a single parameter to adjust for the selection of better solutions. It is merely the memory encapsulated in the ranking that directs EO into the neighborhood of increasingly better

solutions. Similar to the Bak-Sneppen model, those “better” variables possess punctuated equilibrium: their memory only get erased when they happen to be connected to one of the variables forced to change. On the other hand, in the choice of move to S' , there is no consideration given to the outcome of such a move, and not even the worst variable x_j itself is guaranteed to improve its fitness. Accordingly, large fluctuations in the cost can accumulate in a sequence of updates. Merely the bias *against* extremely “bad” fitnesses produces improved solutions.

Tests have shown that this basic algorithm is very competitive for optimization problems where EO can choose randomly among many $S' \in N(S)$ that satisfy step (2c), as is the situation for graph bipartitioning [2, 5]. But in cases such as the single spin-flip neighborhood for the spin Hamiltonian, focusing on only the worst fitness [step (2b)] leads to a deterministic process, leaving no choice in step (2c): If the “worst” spin x_j has to flip and any neighbor S' differs by only one flipped spin from S , it must be $S' = (S/x_j) \cup \{-x_j\}$. This deterministic process inevitably will get stuck near some poor local minimum. To avoid these “dead ends” and to improve results [2], we introduce a single parameter into the algorithm. To this end, we rank all x_i according to fitness λ_i , *i.e.*, we find a permutation Π of the variable labels i with

$$\lambda_{\Pi(1)} \leq \lambda_{\Pi(2)} \leq \dots \leq \lambda_{\Pi(n)}. \quad (7.4)$$

The worst variable x_j [step (2b)] is of rank 1, $j = \Pi(1)$, and the best variable is of rank n . Consider a scale-free probability distribution over the *ranks* k ,

$$P_k \propto k^{-\tau}, \quad 1 \leq k \leq n, \quad (7.5)$$

for a fixed value of τ . At each update, select a rank k according to P_k . Then, modify step (2c) so that the variable x_j with $j = \Pi(k)$ changes its state.

For $\tau = 0$, this “ τ -EO” algorithm is simply a random walk through Ω . Conversely, for $\tau \rightarrow \infty$, the process approaches a deterministic local search, only updating the lowest-ranked variable, and is bound to reach a dead end (see Fig. 7.1). However, for finite values of τ the choice of a *scale-free* distribution for P_k in (7.5) ensures that no rank gets excluded from further evolution, while still maintaining a bias against variables with bad fitness. In all problems studied so far, a value of

$$\tau - 1 \sim 1/\ln n \quad (n \rightarrow \infty) \quad (7.6)$$

seems to work best [6, 7]. We have studied a simple model problem for which the asymptotic behavior of τ -EO can be solved exactly [8]. The model reproduces (7.6) exactly in cases where the model develops a “jam” amongst its variables, which is quite a generic feature of frustrated systems: After many update steps most variables freeze into a near-perfect local arrangement and resist further change, while a finite fraction remains frustrated in a poor local arrangement. More and more of the frozen (slow) variables have to be

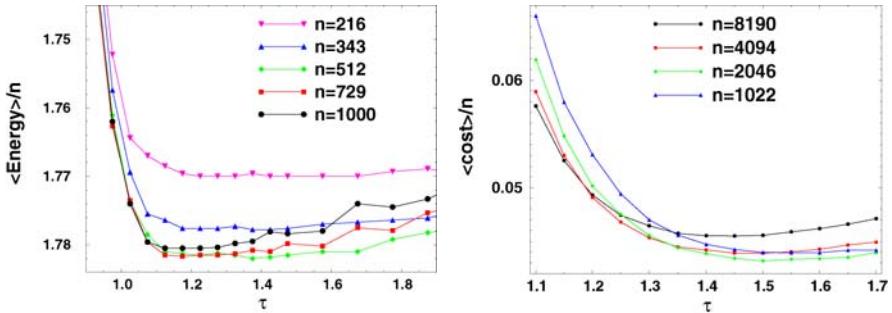


Fig. 7.1. Plot of the average costs obtained by EO for a $\pm J$ spin glass (*left*) and for graph bipartitioning (*right*) as a function of τ . A number of instances were generated at each n . For each instance, 10 different EO runs were performed at each τ . The results were averaged over runs and over instances. Although both problems are quite distinct, in either case the best results are obtained at a value of τ that behaves according to (7.6)

dislocated collectively to accommodate the frustrated (fast) variables before the system as a whole can improve its state. In this highly correlated state, slow variables block the progression of fast variables, and a “jam” emerges. Asymptotic analysis of the flow equations for a jamming system gives (7.6).

With this implementation, we have studied the ground states of spin glasses in $d = 3$ and 4. Our data is discussed in detail in [6]. A fit of our data with $e_d(n) \sim e_d(\infty) + A/n$ for $n \rightarrow \infty$ predicts $e_3(\infty) = 1.7865(3)$ for $d = 3$ and $e_4(\infty) = 2.093(1)$ for $d = 4$. Both values are consistent with the findings of [10–12], providing independent confirmation of those results.

More recently, we have used EO to test theoretical predictions for the $T = 0$ properties of spin glasses on Bethe lattices [13]. Mezard and Parisi [13] proposed a replica symmetry breaking (1RSB) solution for the ground state energy which differs much from replica symmetric (RS) results. As Fig. 7.2 shows, the EO results are consistent with the 1RSB results and rule out the RS solution.

Since EO never freezes into a local minimum, it is also well suited to enumerate the distribution of near-optimal states of a system, which is of great importance for the low-temperature dynamics in glassy materials [14, 15]. Thus, we have also counted all lowest-energy states that EO could find for the Bethe lattice up to $n = 256$ (when memory became exhausted). The data extrapolates to a rather low entropy per spin of $S/n = 0.009(2)$; there are no “free spins” [15] in an odd-connected lattice.

To gauge EO’s performance for larger n , we have run our implementation also on two 3d lattice instances, *toruspm3-8-50* and *toruspm3-15-50*, with $n = 8^3$ and $n = 15^3$, considered in the 7th DIMACS challenge for semi-

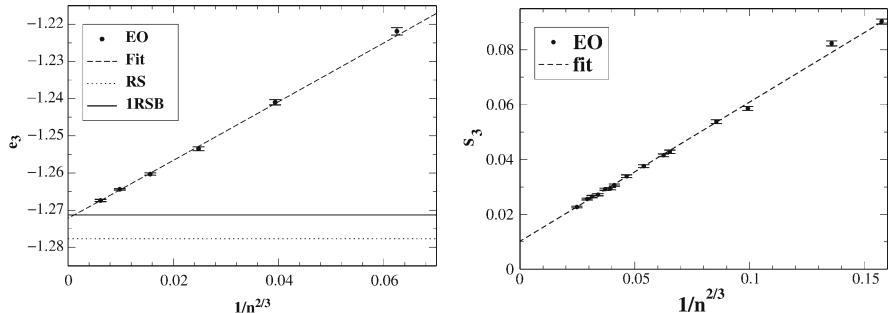


Fig. 7.2. Plot of the extrapolation in n for the ground state energy e_3 and entropy s_3 on a 3-connected Bethe lattice. Assuming a fit of the energies with $e_3(n) \sim e_3(\infty) + a_1/n^{a_2}$ we obtain $e_3(\infty) = -1.2722(10)$, $a_1 = 0.80(5)$, and $a_2 = 0.67(1)$, in good agreement with 1RSB predictions and well above the RS result. Similarly, we obtain $s_3 = 0.009(2)$

definite problems.³ Bounds [16] on the ground-state cost established for the larger instance are $C_{\text{lower}} = -6138.02$ (from semi-definite programming) and $C_{\text{upper}} = -5831$ (from branch-and-cut). EO found $C(S_{\text{best}}) = -6049$ (or $C/n = -1.7923$), a significant improvement on the upper bound and even lower than $e(\infty)$ from above. Furthermore, in that single run EO found 116 different states of that energy with Hamming-distances as far separated as 1500 mutually distinct spins! For the smaller instance the bounds given are -922 and -912 , resp., while EO finds -916 (or $C/n = -1.7891$). Here we found 10^5 such states before we terminated the run. While this run (including sampling degenerate states!) took only a few minutes of CPU (at 800 MHz), the results for the larger instance require about 16 h.

We like to thank the organizers of the annual workshop at the Center for Simulation Physics for their kind invitation, and the NSF for financial support for this work under grant DMR-0312510.

References

1. P. Bak: *How Nature Works* (Springer, New York 1996)
2. S. Boettcher and A.G. Percus: *Artificial Intelligence* **119**, 275 (2000)
3. F. Barahona: *J. Phys. A: Math. Gen.* **15**, 3241 (1982)
4. S. Boettcher and M. Paczuski: *Phys. Rev. Lett.* **76**, 348 (1996)
5. S. Boettcher: *J. Math. Phys. A: Math. Gen.* **32**, 5201 (1999)
6. S. Boettcher and A.G. Percus: *Phys. Rev. Lett.* **86**, 5211 (2001)
7. S. Boettcher and A.G. Percus: *Phys. Rev. E* **64**, 026114 (2001)
8. S. Boettcher and M. Grigni: *J. Math. Phys. A: Math. Gen.* **35**, 1109 (2002)

³ <http://dimacs.rutgers.edu/Challenges/Seventh/>

9. See *Frontiers in problem solving: Phase transitions and complexity*, Special issue of Artificial Intelligence **81**:1–2 (1996)
10. K.F. Pal: Physica A **223**, 283 (1996)
11. A.K. Hartmann: Europhys. Lett. **40**, 429 (1997)
12. A.K. Hartmann: Phys. Rev. E **60**, 5135 (1999)
13. M. Mezard and G. Parisi: J. Stat. Phys **111**, 1 (2003)
14. A.K. Hartmann: Phys. Rev. B **59**, 3617 (1999)
15. T. Klotz and S. Kobe: J. Phys. A: Math. Gen. **27**, L95 (1994)
16. M. Jünger and F. Liers, private communication

8 Stochastic Collision Molecular Dynamics Simulations for Ion Transfer Across Liquid–Liquid Interfaces

S. Frank^{1,2} and W. Schmickler¹

¹ Abteilung Elektrochemie, Universität Ulm, 89069 Ulm, Germany

² Current address: Center for Materials Research and Technology and School of Computational Science and Information Technology, Florida State University, Tallahassee, FL 32306-4350, USA

wolfgang.schmickler@chemie.uni-ulm.de, sfrank@csit.fsu.edu

Abstract. We compute the potential-energy surface for ion transfer across liquid–liquid interfaces from a lattice gas model and simulate the transfer as a random walk of the ion coupled to a heat bath. The kinetics obey Tafel behavior. The reaction rate is slowed down due to friction, and the friction effect is stronger than for a free particle.

8.1 Introduction

Ion transfer across liquid–liquid interfaces, though of considerable experimental interest, still lacks an established theoretical description. It is not clear whether this process should be viewed as a chemical reaction requiring an activation energy, or simply as a mass transport across a viscous boundary. Molecular dynamics simulations have shown a continuous increase of the chemical part of the free energy of ion transfer and no barrier (see, e.g., [1]). However, the simulations were performed in the presence of a high field driving the ion across the interface, and in the absence of space charge regions. Thus, an essential part of the interaction energy of the transferring ion has been missing. In a model proposed by Schmickler [2], it is the combination of several interactions that constitutes a barrier at the interface.

Here, we follow Schmickler’s ideas and treat ion transfer as a chemical reaction. The reaction coordinate – simply the distance from the average interface position – is singled out, and all the other degrees of freedom are represented as a heat bath; the same approach as in Kramers’ theory [3]. With this simplification, we can observe the reaction directly in a simulation.

8.2 Potential-energy Surface

We calculate the potential-energy surface of a transferring ion as the configurational energy of a positively charged test particle with fixed position in a simple cubic lattice gas, as a function of the distance z from the average

interface position. Our model contains two solvents S_1 and S_2 and a different base electrolyte in each phase, and each lattice site is occupied by one particle. The configurational energy is given by the sum over all nearest-neighbor interactions, plus for ions the energy in the instantaneous electrostatic potential caused by all ions in the system. We calculate the equilibrium properties of this model using the Metropolis Monte Carlo algorithm. Details are given elsewhere [4]. The system is polarizable in a certain potential window, and the absolute value of the free energy of ion transfer, which is governed by a single interaction parameter $\pm r$ for the interaction with the two solvents, must be low enough for the ion to be transferable within this window. At the potential of zero charge, the potential-energy surface is calculated exactly as described above, and for finite Galvani potential differences between the two phases, the electrostatic contribution obtained from independent simulations without the test particle is added. For the highest simulated Galvani potential difference $\Delta\phi$, we additionally compute the potential-energy surface exactly, in order to test the underlying approximation that chemical and electrostatic contributions add up independently. We note that, strictly, the free-energy surface, or potential of mean force, should be calculated. Figure 8.1 shows results for various values of the potential difference. Thermodynamically, a higher potential difference favors transfer from left to right.

8.3 Simulations of Ion Transfer

We simulate the ion transfer as a random walk of the ion, coupled to a heat bath, on the potential-energy surface. Each run starts with the ion in the well at negative z , and it ends when the ion reaches a point far enough across the saddle at positive z not to recross. We discuss the ion-transfer kinetics and dynamics and compare the results with transition-state theory.

8.3.1 Stochastic Molecular Dynamics Simulations

For the simulations of the ion dynamics on the potential-energy surface, we adopt the *stochastic collision model* introduced by Kast et al. [5, 6]. In this model, a particle with mass $m_1 = 1$ is subject to an elastic collision with a heat-bath particle of mass m_2 at every integration time step of the molecular dynamics. The velocities of the heat-bath particles are randomly drawn from a Gaussian distribution. For sufficiently small time steps Δt and $\alpha = m_2/(m_1 + m_2)$, the free particle follows a Langevin dynamic with a friction coefficient $\gamma_{\text{free}} = 2\alpha/\Delta t$. For motion in a harmonic potential, the kinetic temperature equals the bath temperature under the same conditions. Details, and an implementation within the Verlet algorithm, can be found in the original literature [5, 6].

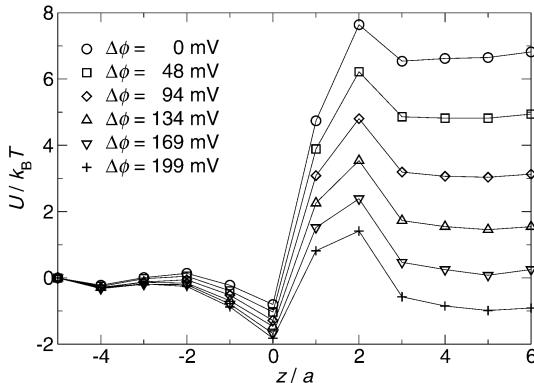


Fig. 8.1. Potential-energy surface for the ion transfer at several $\Delta\phi$. Distance z from the average interface position in units of the lattice constant a

8.3.2 Kinetics

From simulations with $m_2 = 0.01$ and $\Delta t = 0.01$, we obtain the escape rate $\Gamma = 1/\langle t_{\text{run}} \rangle$ to cross the barrier far enough not to recross. According to [2], correcting for the concentration in the well, which is in equilibrium with the bulk, one obtains for the rate constant

$$k_{f,\text{MD}} = \sqrt{\frac{2\pi k_B T}{m_1 \omega_{\min}^2}} \exp\left(-\frac{U_{\min}}{k_B T}\right) \Gamma, \quad (8.1)$$

with the frequency in the well ω_{\min} from a fit of the potential around its minimum U_{\min} to a parabola. Independently, we calculate the rate constants from transition-state theory [2]

$$k_{f,\text{TST}} = \sqrt{\frac{k_B T}{2\pi m_1}} \exp\left(-\frac{U_{\max}}{k_B T}\right). \quad (8.2)$$

Note that the formal activation energy is the maximum of the potential energy U_{\max} relative to the bulk, due to the correction for the concentration in the well, and can be negative. In a logarithmic plot vs. the overpotential, both sets of rate constants follow straight lines (Tafel behavior) with the same transfer coefficient, but the rates from the MD simulations are lower. We ascribe this difference to a friction effect of the heat bath on the ion. This can be quantified as an almost constant transmission factor $\kappa \approx 0.4$. We note that deviations from Tafel behavior at higher absolute overpotentials cannot be ruled out.

8.3.3 Dynamics

According to Kramers' Langevin-type theory of the effect of friction on reaction rates [3], the transmission factor is

$$\kappa = \frac{(\gamma^2/4 + \omega_b^2)^{1/2} - \gamma/2}{\omega_b}, \quad (8.3)$$

where γ is a friction coefficient. With the barrier frequency ω_b from parabolic fits to the potential maximum, we can obtain an effective friction coefficient γ_{eff} that reproduces the correct κ and compare it with γ_{free} . Results for a series of different m_2 at fixed potential difference are shown in Table 8.1. As for the free particle, greater m_2 increases friction, but there are strong quantitative differences. The effect of friction on the reaction rate is always greater than on diffusion.

A comparison of typical trajectories close to the transition state for different m_2 (Fig. 8.2) clearly shows that a higher friction increases the probability of barrier recrossing, since the movement of the ion is much less straight. Thus, higher friction makes it both less likely that the ion reaches the saddle point, and that it crosses the saddle point.

8.4 Conclusions

We have calculated the potential-energy surface for ion transfer across a liquid–liquid interface from a lattice-gas model and simulated the transfer as a random walk of the ion, coupled to a heat bath, for various potential differences across the interface, i.e., depending on the thermodynamic driving force. The model, as introduced by Kast et al. [5, 6], subjects the ion to elastic collisions with a bath particle at every time step. With our simulation parameters, the system achieves the bath temperature.

Tafel plots of the rate constants give straight lines, indicating a linear relationship between the free energies of the reaction and of activation (Brønsted relationship). Comparison with transition-state theory results shows a slowing-down of the reaction due to friction in the bath. As for a free particle, the friction increases with the mass of the bath particles, but for ion transfer, the effect of friction is always greater than in the former case. Moreover, higher friction makes barrier recrossing more likely.

Table 8.1. Effect of friction on the transfer rate at $\Delta\phi = 94$ mV. Transmission factor κ , (effective) friction coefficients for the particle in the potential field (γ_{eff}) and for the free particle γ_{free}

m_2	κ	γ_{eff}	γ_{free}
0.005	0.57	2.5	1.0
0.01	0.43	4.0	2.0
0.02	0.27	7.3	3.9
0.03	0.19	11.0	5.8
0.04	0.15	14.0	7.7
0.05	0.12	17.3	9.5

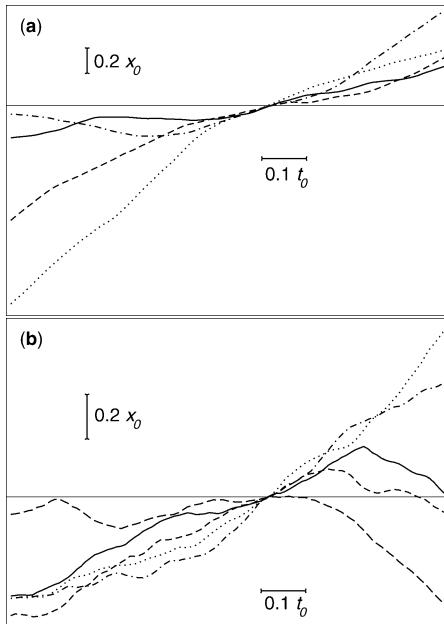


Fig. 8.2. Trajectories of MD simulations near the transition state, position of which is shown by the horizontal line. Bath particle mass $m_2 = 0.01$ (a), $m_2 = 0.05$ (b). $\Delta\phi = 94$ mV

Our treatment in the framework of reaction-rate theory, as opposed to a transport process, is conclusive and in line with experimental results (e.g. [7,8]). The barrier is a result of the superposition of chemical and electrostatic contributions, which drop over different length scales.

Acknowledgments

Support by the Deutsche Forschungsgemeinschaft is gratefully acknowledged. S. Frank's work at Florida State University was supported by U.S. National Science Foundation Grant No. DMR-0240078.

References

1. P. Fernandes, M. Cordeiro, J. Gomes: *J. Phys. Chem. B* **103**, 8930 (1999)
2. W. Schmickler: *J. Electroanal. Chem.* **426**, 5 (1997)
3. H. Kramers: *Physica* **7**, 284 (1940)
4. T. Huber, O. Pecina, W. Schmickler: *J. Electroanal. Chem.* **467**, 203 (1999)
5. S. Kast, K. Nicklas, H. Bär, J. Brickmann: *J. Chem. Phys.* **100**, 566 (1994)
6. S. Kast, J. Brickmann: *J. Chem. Phys.* **104**, 3732 (1996)
7. T. Kakiuchi, J. Noguchi, M. Senda: *J. Electroanal. Chem.* **327**, 63 (1992)
8. Y. Shao, J. Campbell, H. Girault: *J. Electroanal. Chem.* **300**, 415 (1991)

Part III

Biological Systems

9 Generalized-Ensemble Simulations of Small Proteins

U.H.E. Hansmann¹

Dept. of Physics, Michigan Technological University
Houghton, MI 49931-1295, USA

Abstract. We review the generalized-ensemble approach to protein studies. Presenting results from simulations of small proteins (of ≈ 35 amino acids) we show that these techniques allow efficient simulations of all-atom protein models and may lead to a deeper understanding of the folding mechanism in proteins.

Because a protein is only functional if it folds into its characteristic shape, it is important to understand how the structure of proteins emerge from their sequence of amino acids. Computer experiments offer one way to gain such knowledge but are extremely difficult for realistic protein models. This is because all-atom models lead to a rough energy landscape with a huge number of local minima separated by high energy barriers; and sampling of low-energy conformations becomes a hard computational task. *Generalized-ensemble* simulations are one way to overcome this multiple-minima problem and to optimize sampling of low-energy protein configurations [1].

In generalized ensembles the canonical weights, that suppress the crossing of an energy barrier of height ΔE by a factor $\propto \exp(-\Delta E/k_B T)$, are replaced with such that allow the system to escape out of local minima. For instance, in multicanonical sampling [2] the weight $w(E)$ leads to a distribution $P(E) \propto n(E)w(E) = \text{const}$, where $n(E)$ is the spectral density. A free random walk in the energy space is performed that allows the simulation to escape from any local minimum. From this simulation one can calculate the thermodynamic average of any physical quantity A by re-weighting [3]

$$\langle A \rangle_T = \frac{\int dx \mathcal{A}(x) w^{-1}(E(x)) e^{-E(x)/k_B T}}{\int dx w^{-1}(E(x)) e^{-E(x)/k_B T}}, \quad (9.1)$$

where x labels the configurations of the system.

Another realization of this idea is *energy landscape paving* (ELP) where one performs low-temperature Monte Carlo simulations with a modified energy expression: $w(\tilde{E}) = e^{-\tilde{E}/k_B T}$ with $\tilde{E} = E + f(H(q, t))$. Here, T is a (low) temperature, \tilde{E} serves as a replacement of the energy E and $f(H(q, t))$ is a function of the histogram $H(q, t)$ in a pre-chosen “order parameter” q . It follows that the weight of a local minimum state decreases with the time the system stays in that minimum till the local minimum is no longer favored and the system starts again to explore higher energies. Obviously, for

$f(H(q,t)) = f(H(q))$ the method reduces to the various generalized-ensemble methods (for instance for $f(H(q,t)) = \ln H(E)$ to multicanonical sampling).

Another way of enhancing the sampling of low-energy configurations in protein simulations is parallel tempering (also known as replica exchange or Multiple Markov Chain method) [4], a technique that was first introduced to protein folding in [5]. In its most common form, one considers in this technique an artificial system built up of N *non-interacting* copies of the molecule, each at a different temperature T_i . In addition to standard Monte Carlo or molecular dynamics moves that affect only one copy, parallel tempering introduces now a new *global* update [4]: the exchange of conformations between two copies i and $j = i + 1$ with probability

$$w(\mathbf{C}^{\text{old}} \rightarrow \mathbf{C}^{\text{new}}) = \min(1, \exp(-\beta_i E(C_j) - \beta_j E(C_i) + \beta_i E(C_i) + \beta_j E(C_j))). \quad (9.2)$$

This exchange of conformations leads to a faster convergence of the Markov chain than is observed in regular canonical simulations.

Our investigation of the two small proteins PTH(1-34) and HP-36 relies on an all-atom representation of these molecules. The interactions between the atoms are described by a standard force field, ECEPP/3 [6] (as implemented in the program package SMMP [7]). The protein-water interactions are approximated by a solvent-accessible surface term [8].

We find in multicanonical simulations of PTH(1-34) [9] as lowest-energy state an elongated helix with 31 residues part of the helix. This structure has not only the lowest *total* energy ($E_{\text{TOT}} = -277.8$ kcal/mol), but also the lowest *intramolecular* energy: $E_{\text{ECEPP/3}} = -136.5$ kcal/mol. It is very similar to the crystal structure of PTH(1-34) whose energy is $E_{\text{TOT}} = -277.9$ kcal/mol ($E_{\text{ECEPP/3}} = -187.0$ kcal/mol). The root-mean-square deviation (rmsd) between the two structures is only 0.8 Å for backbone atoms. At $T = 300$ K, our lowest energy structure appears with a frequency of $(99 \pm 0.5)\%$, i.e. almost all observed configurations resemble the crystal structure. However, in near-physiological solution, one observes instead two helices separated by a disordered and flexible region. Hence, our simulation of PTH(1-34) does not reproduce the experimental results for that peptide in solution.

A similar situation is observed in our simulation of HP-36 [10]. Here we observe in parallel tempering simulations two groups of low-energy configurations at room temperature. The predominant structure (group A) appears with 80% frequency and is elongated helix ($r_{\text{rgy}} = 12.5$ Å) with 85% of its residues are part of an α -helix. This structure has little similarity with the native one and differs from it by a RMSD of at least $r_{\text{RMSD}} = 9$ Å. However, a second group (B) of low-energy configurations is also observed and appears with 15% frequency at $T = 250$ K. Its RMSD is with $r_{\text{RMSD}} = 5.7$ Å comparable to the structures found in [11] (an RMSD of $r_{\text{RMSD}} = 5.8$ Å) and [12] (an RMSD of $r_{\text{RMSD}} = 5.7$ Å). The structure has 79% of the native helical content, that is 79% of all residues which are part of a helix in the experimental structure are also part of a helix in our structure, and 60% of

the native contacts are formed. Both values are only slightly smaller than the results of [12] (the optimal structure of a 1 μ s molecular dynamic folding simulation had 80% of native helical content and 62% of native contacts) and the lowest energy configuration of [11] that had 90% of helical content and 65% of native contacts. However, these native-like configurations cannot be identified by their energies. The average energy of the predominant configurations of group A is $\langle E_{\text{tot}} \rangle = -273(17)$ kcal/mol and differs little from the corresponding value of $\langle E_{\text{tot}} \rangle = -266(10)$ kcal/mol for the native-like configurations of group B. Since the configurations of group A have a much larger entropy, native-like configurations are also not the global minimum in free energy. This is in contradiction to the experimental results of [13]. Hence, our energy function is again not accurate enough for prediction of the native state.

In summary, we have performed all-atom simulations of two small proteins, PTH(1-34) and HP-36, in generalized ensembles. While previous applications of these techniques were restricted to small peptides with less than 20 residues, our results establish that the generalized-ensemble approach is also a useful tool for investigation of much larger polypeptides with 30–40 residues. Applications to even larger molecules seem to be restricted less by the sampling technique but by the accuracy of the energy function.

Acknowledgments

Support by a research grant from the National Science Foundation (CHE-0313618) is gratefully acknowledged.

References

1. U.H.E. Hansmann, Y. Okamoto: *J. Comp. Chem.* **14**, 1333 (1993)
2. B.A. Berg, T. Neuhaus: *Phys. Lett. B* **267**, 249 (1991)
3. A.M. Ferrenberg, R.H. Swendsen: *Phys. Rev. Lett.* **61**, 2635 (1988); *Phys. Rev. Lett.* **63**, 1658(E) (1989), and references given in the erratum
4. K. Hukushima, K. Nemoto: *J. Phys. Soc. (Jpn.)* **65**, 1604 (1996); G.J. Geyer: *Stat. Sci.* **7**, 437 (1992)
5. U.H.E. Hansmann: *Chem. Phys. Lett.* **281**, 140 (1997)
6. G. Némethy, K.D. Gibson, K.A. Palmer, C.-N. Yoon, G. Paterlini, A. Zagari, S. Rumsey, H.A. Scheraga: *J. Phys. Chem.* **96**, 6472 (1992)
7. F. Eisenmenger, U.H.E. Hansmann, S. Hayryan, C.-K. Hu: *Comp. Phys. Comm.* **138**, 192 (2001)
8. T. Ooi, M. Obatake, G. Nemethy, H.A. Scheraga: *Proc. Natl. Acad. Sci. USA* **8**, 3086 (1987)
9. U.H.E. Hansmann: *J. Chem. Phys.* **120**, 417 (2004)
10. C.-Y. Lin, C.-K. Hu, U.H.E. Hansmann: *Proteins* **52**, 436 (2003)
11. U.H.E. Hansmann, L.T. Wille: *Phys. Rev. Lett.* **88**, 068105 (2002)
12. Y. Duan, P.A. Kollman: *Science* **282**, 740 (1998)
13. C.J. McKnight, D.S. Doehring, P.T. Matsudaria, P.S. Kim: *J. Mol. Biol.* **260**, 126 (1996)

10 A Biological Coevolution Model with Correlated Individual-Based Dynamics

V. Sevim and P.A. Rikvold

School of Computational Science and Information Technology,
Center for Materials Research and Technology and Department of Physics,
Florida State University, Tallahassee, FL 32306-4120, USA
sevim@csit.fsu.edu, rikvold@csit.fsu.edu

Abstract. We study the effects of interspecific correlations in a biological coevolution model in which organisms are represented by genomes of bitstrings. We present preliminary results for this model, indicating that these correlations do not significantly affect the statistical behavior of the system.

10.1 Introduction

The dynamics of biological coevolution poses many problems of interest to the statistical-physics and complex-systems communities [1, 2]. Recently, we studied a coevolution model that is a simplified version of the one introduced by Hall et al. [3–6]. Our model, in which individuals give birth, mutate and die, displays punctuated equilibria-like, quiet periods interrupted by bursts of mass extinctions [7, 8]. Interactions in this model were given by a random interaction matrix. Here we report on a modified version of the model, in which we have added correlations to the interaction matrix in order to increase the biological realism. The modified model is compared with the original one to assess the effects of correlations.

10.2 Model

We used a bitstring genome of length L to model organisms in the Monte Carlo (MC) simulations [9, 10]. In this sense, each genotype, which is just an L -bit number, is considered a different haploid species. Therefore, the terms “genotype” and “species” are used in the same sense in this paper. We denote the population of species I at a discrete time t as $n_I(t)$, and several of the 2^L possible species can be present at the same time in our “ecosystem.” All species reproduce asexually (by cloning) in discrete, non-overlapping generations. In each generation t , every individual of species I is allowed to give birth to F offspring with a probability P_I . Whether it reproduces or not, the individual dies at the end of the generation, so that only offspring can survive to the next generation. The reproduction probability for an individual of species I is given by [3–6]

$$P_I(\{n_J(t)\}) = \frac{1}{1 + \exp[-\sum_J M_{IJ} n_J(t)/N_{\text{tot}}(t) + N_{\text{tot}}(t)/N_0]} . \quad (10.1)$$

The Verhulst factor N_0 represents the carrying capacity of the “ecosystem” and prevents the total population $N_{\text{tot}}(t) = \sum_I n_I(t)$ from diverging to infinity [11]. M is the interaction matrix, in which a matrix element M_{IJ} represents the effect of the population density of species J on species I . A positive M_{IJ} means that species I benefits from species J , while a negative M_{IJ} corresponds to a situation in which species I is harmed or inhibited by the presence of J . The form of the interaction matrix M is discussed in the next section.

In each generation, all individuals undergo mutation with a probability μ . If an individual is chosen to mutate, one bit in its genome is picked randomly and flipped. Since we consider different genotypes as different species, mutations lead to speciation, i.e., creation of another species.

10.3 The Interaction Matrix

In this study, the interaction matrix M was set up in four different ways. In the first case, all off-diagonal elements were randomly and uniformly distributed on $[-1, 1]$, while all diagonal elements were set to zero. This corresponds to the case in [5, 6]. We shall call this the uniform-uncorrelated model. In the second case, we added correlations between interaction constants of similar species to make this model more realistic. In a real ecosystem, two different but closely related species X and Y interact with another species Z in a similar way. Therefore, the interaction constant M_{XZ} should be positively correlated with M_{YZ} . To implement these correlations, we modified the interaction matrix by averaging all terms over their nearest neighbors (nn) in Hamming space. Thus,

$$M_{IJ} = \frac{\left[M_{IJ}^0 + \sum_{(K,L) \in nn(I,J)} M_{KL}^0 \right]}{(2L+1)^{1/2}}, \quad (10.2)$$

where M_{IJ}^0 are independent variables uniformly distributed on $[-1, 1]$, and $nn(I, J)$ are those bitstring pairs that differ from the bitstring pair (I, J) by one bit (a Hamming distance of one). A square-root appears in the denominator because we multiply the average with the square-root of the normalization constant in order not to change the standard deviation of the matrix elements.

To investigate the results of longer-range correlations we also modified the random interaction matrix by averaging all terms over their nearest and next-nearest neighbors (nnn) in Hamming space:

$$M_{IJ} = \frac{\left[M_{IJ}^0 + \sum_{(K,L) \in nn(I,J)} M_{KL}^0 + \sum_{(K,L) \in nnn(I,J)} M_{KL}^0 \right]}{(2L^2 + L + 1)^{1/2}}, \quad (10.3)$$

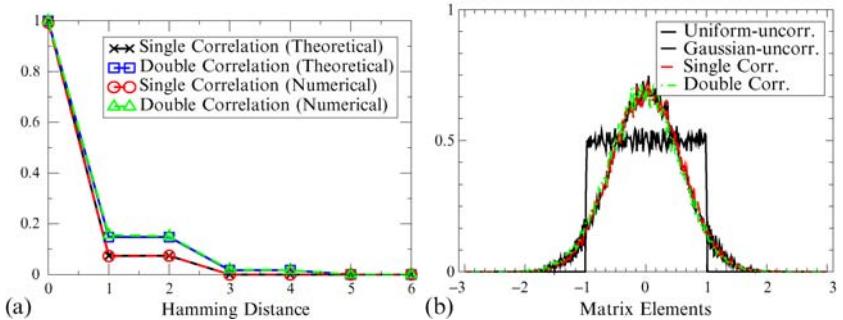


Fig. 10.1. Correlation functions and matrix-element distributions after averaging. (a) Theoretical (solid) and numerical (dashed) results for the correlation functions for $2^{13} \times 2^{13}$ interaction matrices of the single-correlated (*lower*) and double-correlated (*upper*) models. The theoretical and numerical results agree. (b) Distributions of matrix elements: uniform and Gaussian uncorrelated (*solid curves* with square and Gaussian distributions, respectively), single-correlated (*dashed curve*) and double-correlated (*dot-dashed curve*) models with $L = 8$. The approximately Gaussian distributions for the two correlated models and the distribution for the Gaussian-uncorrelated model practically overlap

where $\text{nnn}(I, J)$ is the set of bitstring pairs that differ from (I, J) by two bits. We shall call the nn-averaged and nnn-averaged models single-correlated and double-correlated, respectively.

As a result of the central limit theorem, after averaging, the distributions of the matrix elements in both the single and double-correlated models take an approximately Gaussian form with the same standard deviation as the uniform distribution of the elements of the initial, random matrix (see Fig. 10.1b). In order to see whether a possible difference in the results for different models is due to the correlations or the matrix-element distributions, we also set up another uncorrelated interaction matrix. In this fourth model, the uncorrelated, random matrix elements are distributed with a Gaussian distribution with the same standard deviation, $\sqrt{1/3}$, as in the correlated models. We shall call this the Gaussian-uncorrelated model.

The correlation functions and the distributions of interaction-matrix elements for all the models are shown in Fig. 10.1. The steps in the correlation functions are a result of the chosen metric. We use a city-block metric in which the Hamming distance between two matrix elements M_{IJ} and M_{KL} is given by $H(I, K) + H(J, L)$ where $H(I, K)$ is the Hamming distance between two L -bit bitstrings, I and K .

10.4 Simulation Results

The interaction matrix is set up at the beginning and is not modified during the course of the simulation. For all models, we performed eight sets of

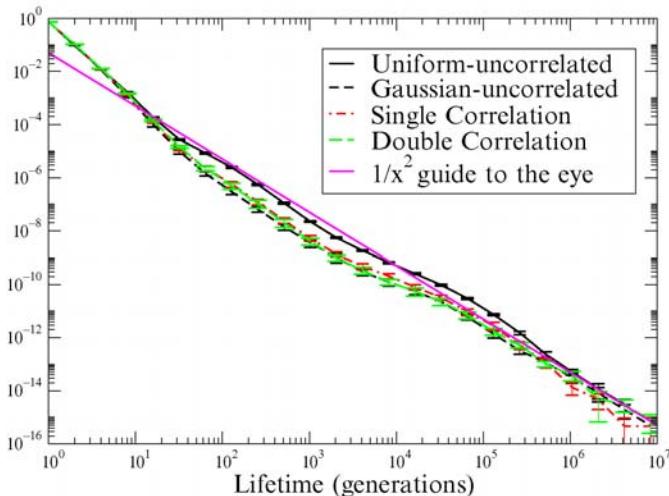


Fig. 10.2. Normalized histograms of species-lifetimes based on simulations of 2^{25} generations each: uniform-uncorrelated (*solid curve*), Gaussian-uncorrelated (*short-dashed curve*), single-correlated (*dot-dashed curve*) and double correlated (*long-dashed curve*). The distributions for the models with a Gaussian matrix-element distribution overlap (within the margin of error), and they differ significantly from the lifetime distribution for the uniform-uncorrelated model. The histograms exhibit a power-law like decay with an exponent near -2 . Results are averaged over eight runs each

simulations for $2^{25} = 33\,554\,432$ generations with the same parameters as in [5, 6]: genome length $L = 13$, mutation rate $\mu = 10^{-3}$ per individual per generation, carrying capacity $N_0 = 2000$, and fecundity $F = 4$. We began the simulations with 200 individuals of genotype 0. In order to compare the models, we constructed histograms corresponding to the species-lifetime distributions. The lifetime of a species is defined as the number of generations between its creation and extinction. As seen in Fig. 10.2, the lifetime distributions of the correlated and Gaussian-uncorrelated models overlap within the margin of error, and they differ significantly from the lifetime distribution of the uniform-uncorrelated model. They all exhibit a power-law like decay with an exponent near -2 . The correlations between the matrix elements do not seem to affect the behavior of the lifetime distribution to a statistically significant degree, at least not for the relatively weak correlations that were introduced here. On the other hand, changes in the marginal probability density of the individual matrix elements do have statistically significant effects, even though gross features, such as the approximate $1/x^2$ behavior of the species-lifetime distribution are not changed. Similar conclusions are reached also for other quantities that we studied. In particular, the power-spectral density of the Shannon–Wiener species diversity index shows $1/f$ noise [5, 6]

with an overall intensity that depends more on the marginal matrix-element distribution than on correlations in M .

Although we have tested only weak correlations, and so our conclusion is only preliminary, there appears to be no disadvantage in using a random interaction matrix to model such an “ecosystem.” This has obvious computational advantages as it makes it possible to simulate systems with larger numbers of completely different species.

Acknowledgments

This research was supported by U.S. National Science Foundation Grant No. DMR-0240078, and by Florida State University through the School of Computational Science and Information Technology and the Center for Materials Research and Technology.

References

1. B. Drossel, A.J. McKane, in: *Handbook of Graphs and Networks: From the Genome to the Internet*, ed. by S. Bornholdt, H.G. Schuster (Wiley-VCH, Berlin 2002)
2. B. Drossel: *Adv. Phys.* **50**, 209 (2001)
3. M. Hall, K. Christensen, S.A. di Collobiano, H.J. Jensen: *Phys. Rev. E* **66**, 011904 (2002)
4. K. Christensen, S.A. di Collobiano, M. Hall, H.J. Jensen: *J. Theor. Biol.* **216**, 73 (2002)
5. P.A. Rikvold, R.K.P. Zia: *Phys. Rev. E* **68**, 031913 (2003), and references therein
6. P.A. Rikvold, R.K.P. Zia, in: *Computer Simulation Studies in Condensed-Matter Physics XVI*, ed. by D.P. Landau, S.P. Lewis, H.-B. Schüttler (Springer, Berlin Heidelberg New York 2004)
7. S.J. Gould, N. Eldredge: *Paleobiology* **3**, 115 (1977)
8. S.J. Gould, N. Eldredge: *Nature* **366**, 223 (1993)
9. M. Eigen: *Naturwissenschaften* **58**, 465 (1971)
10. M. Eigen, J. McCaskill, P. Schuster: *J. Phys. Chem.* **92**, 6881 (1988)
11. P.F. Verhulst: *Corres. Math. Phys.* **10**, 113 (1838)

11 An Image Recognition Algorithm for Automatic Counting of Brain Cells of Fruit Fly

T. Shimada^{1,2}, K. Kato², and K. Ito^{1,2}

¹ Institute for Bioinformatics Research and Development
Japan Science and Technology Agency

² Institute of Molecular and Cellular Biosciences, The University of Tokyo

11.1 Introduction

Today's computing power enables us to simulate systems consist of tens or hundreds of thousands of realistic model-neurons such as the Hodgkin-Huxley neuron. What is absorbing is the number ($10,000 \sim 1,000,000$) here is roughly same to the number of brains of little but higher animals. An excellent example is the fruit fly (*Drosophila melanogaster*), whose brain consists of no more than 200,000 neurons while it has abilities of flight and the sense of sound, visual, and smell [1]. Furthermore, observations on behavioral disturbances have shown that it has even higher level of intelligence such as learning and memory. Therefore it is expected that the simulations on coupled tens of thousands of model neuron system will reveal some secrets of the mechanisms of such intelligence. However, this "brain reproduction" is actually impossible since we do not know how the neurons in the fly's brain are connected. At present *C. elegans* is the unique example that the connectivity among all of its 300 neurons' is known [2]. About the fruit fly, in fact, we do not have even the accurate number of the neurons.

Though the circuit information of its brain is not available, *Drosophila* has the following advantages in comparison with other higher animals. First, its all genome is known [3]. Second, the molecular genetics methods to label specific cells of the brain is well established [4, 5]. Third, its brain is small enough to observe the whole system virtually at the same time. The second and third facts together permit us to obtain the three-dimensional image data of the labeled cells [4, 5]. Thus, what should be tried is to extract the topological information of the neural circuit from the obtained image data.

In the following we introduce a new algorithm of counting cells automatically from source images. Counting cells is essential not only because its biological importance of just giving an accurate value of the total number of neurons against the current rough estimation ($40,000 \sim 200,000$), but because getting the detailed positions of cells is necessary to start the extractions of brain circuit information.

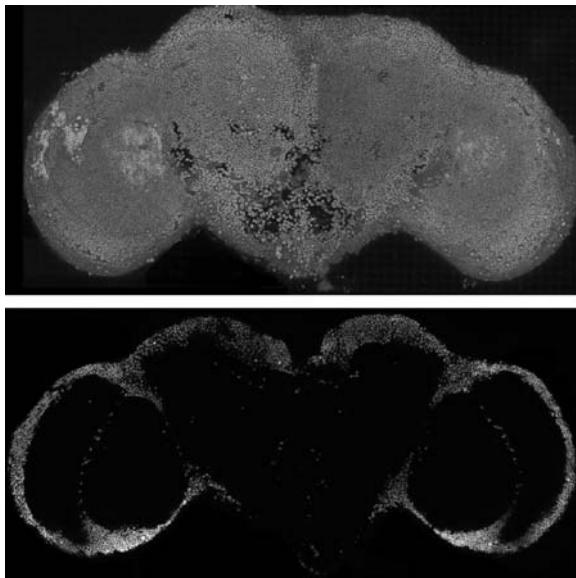


Fig. 11.1. A three-dimensional image of fruit fly's brain (*top*) and one of its slice images (*bottom*). The typical diameter of the cell nucleus is only $5 \sim 7$ voxels

11.2 Data

Instead of counting the neural cells, we count its nucleus. It is because nucleus have relatively uniform ball-like shapes while the shapes of neural cells are very complicated and diverse. The three-dimensional images of nucleus, obtained as the stack of two-dimensional slice images as shown in Fig. 11.1, consists of about $(x, y, z) = (2,000 \times 1,000 \times 300)$ voxels and the size of each voxel is $(0.32 \times 0.32 \times 0.64 \mu\text{m})$.

The typical size of the whole brain image is $(x, y, z) = (2,000 \times 1,000 \times 300)$ with the finest voxel size of $(0.32 \times 0.32 \times 0.64 \mu\text{m})$. Since the average size of nucleus is about $2 \mu\text{m}$, most of those images are only $5 \sim 7$ voxels in diameter. In addition, the source image contains the noise. Main causes of the noise are the diffusion of the labeling dye, defocusing due to the limit of resolution, and the noise on the photo-multiplier. The littleness of the nucleus's image and the inevitable noise make it quite difficult to count the cells by image-restoration-based approaches.

11.3 Counting Algorithm

In our algorithm, counting is realized by putting objects, each of which has an excluded volume whose shape approximates the shape of nucleus. We here apply the ball with approximate diameter, $2 \mu\text{m}$ for example, for the putting

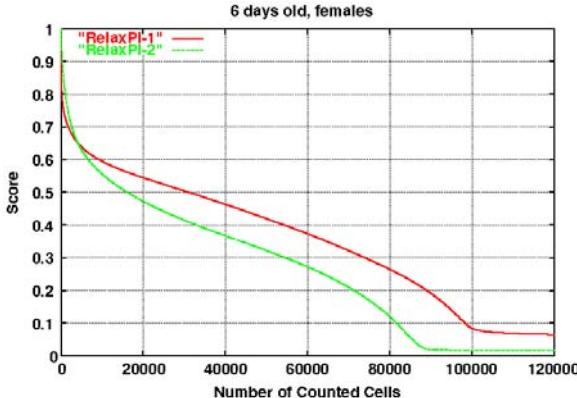


Fig. 11.2. Profiles of the normalized score obtained from the counting over the whole brains. Two lines correspond to different samples. The sharp inflexion points (about 100,000 for the upper line and 90,000 for the lower line) indicates the end of the correct counting stage

objects. More precisely, the ball-shaped volume is approximately expressed as the set of voxels with rugged shape.

To decide where the objects should be placed we first calculate the *score* for each voxel. The *score* S of a voxel, which represents the likelihood of regarding the voxel as a center of a nuclei, is calculated from the intensities of voxels around the voxel as

$$S(x, y, z) = \frac{\sum_{(x', y', z') \in V(x, y, z)} \rho(x', y', z') I(x', y', z')}{\sum_{(x', y', z') \in V(x, y, z)} \rho(x', y', z')}, \quad (11.1)$$

where V, ρ, I are the volume of the embedding object, weight of the convolution, and the data intensity, respectively. For the case of $\rho = \text{constant}$ this score becomes average of intensity over the volume. In the following we use this case.

Getting a source image, we first calculate the *score* for all voxels. Then we place excluded volumes on voxels in descending order of those *scores*. Voxels in the excluded volume of the newly placed objects are excluded from the counting candidates of the center of the nuclei.

It turns out that this simple procedure counts many “fake cells” though it surely can count the correct cell nucleus. The main reason of this miscount is the nonuniformity of voxel intensities. The local averages of the voxel intensity varies in the wide range depending on their positions therefore there is no clear “critical *score*”, which separates the correct counting from miscounting. To overcome this problem, we further require the candidate voxels to satisfy a locally maximum condition of the *score*

$$S(x, y, z) = \max\{S(x', y', z'); (x', y', z') \in v(x, y, z)\}, \quad (11.2)$$

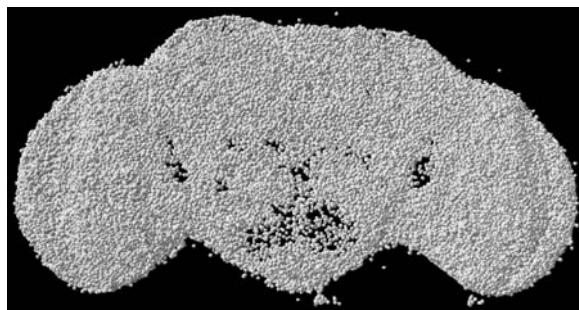


Fig. 11.3. Distribution of the “truly” counted cells, whose scores are higher than the cut-off score

where v is the neighbor region of (x, y, z) . In the following, we adopt $v(x, y, z)$ composed by the sum of the nearest-neighbors and the part of next-nearest-neighbors: $(x \pm 1, y \pm 1, z)$, $(x \pm 1, y, z \pm 1)$, and $(x, y \pm 1, z \pm 1)$. This modification corresponds to assessing the validity of counting cell from the intensity landscape nearby the voxel. Since the major case of miscounts is counting the voxel which is not on a cell but near the bright cells, the modified algorithm efficiently exclude miscounts. As a result, the accuracy of counting improves drastically.

11.4 Results

The sample of our world’s first whole-brain cell counting is the brain of female flies 6 days after eclosion. Applying the algorithm, we obtain the list of scores of the counted cells in the descending order as shown in Fig. 11.2. This profile shows a cross-over from the decreasing stage to the nearly constant stage. Since the constant behavior in this plot corresponds the divergence of the density of cells against the score, the nearly flat region is considered as the miscounting stage. Thus we decide the cut-off score from the cross-over point and regard the cells having score higher than the cut-off score as true cells. It is confirmed that the positions of the obtained “true” cells show quite good agreement with the partial visual counting by professionals (Fig. 11.3). The sharpness of the cross-over allows us to obtain the cut-off score, which decides the total number of cells, in good accuracy. The estimated total numbers of the cells of the two sample brains are $97,000 \pm 1,000$ and $88,500 \pm 1,000$, respectively. We expect that our algorithm will enable us to answer essential questions such as “How many cells, and in which position, are there in the fly’s brain?”, “How large is its individual/sexual/age/gene difference?”, and so on.

Acknowledgement

This work was supported by BIRD/JST grants and a JSPS Research Fellowship for Young Scientists to K.K..

References

1. N. Tanaka, K. Awasaki, T. Shimada, K. Ito: Integration of chemosensory pathways in the second-order olfactory centers of *Drosophila*. To be published in Curr. Biol.
2. J.G. White *et al.*: Phil. Trans. R. Soc. Lond. B **314**, 1 (1986)
3. Special Issue in Science **287**, 2181 (2000)
4. K. Ito: Saibou-Kougaku **15**, 1760 (1996) (in Japanese)
5. K. Ito: Seibutu-Buturi **40**, 179 (2000) (in Japanese)

12 Preferred Binding Sites of Gene-Regulatory Proteins Based on the Deterministic Dead-End Elimination Algorithm

R.G. Endres¹, T.C. Schulthess¹, and N.S. Wingreen²

¹ Center for Computational Sciences and Computer Science
and Mathematics Division, Oak Ridge National Laboratory
Oak Ridge, TN 37831-6164, USA

² NEC Laboratories America, Inc.
4 Independence Way, Princeton, NJ 08540, USA

Abstract. Identifying the specific DNA-binding sites of regulatory proteins is essential to understanding gene expression in the biological cell. Rather than employing statistical bioinformatics tools, we choose an atomistic approach based on the protein-DNA co-crystal structure and binding energies obtained from the AMBER force field. Our method uses the deterministic dead-end elimination and A* algorithms to search a library of discrete amino-acid sidechain rotamers and DNA basepairs for the optimal binding conformations. As an example, we search all possible 10-basepair DNA sequences (4^{10}) of Zif268, a well-studied mouse regulatory protein.

12.1 Introduction

Gene expression is regulated by proteins which bind to the DNA in the vicinity of the gene. However, these proteins can regulate many different genes by binding to different, albeit similar DNA sequences. Fast methods to determine binding sites are demanded in order to keep up with the sequencing of new genomes, and the desire to find new therapeutics for diseases.

While experimental methods based on DNase I footprinting [1] and DNA microarrays [2] are laborious and error-prone [3], standard statistical bioinformatics tools like the consensus sequence and the weight-matrix require prior knowledge of a lot of binding sites to succeed [4]. Both methods make the approximation of treating each DNA-sequence position as independent [5].

In contrast to these methods, we emphasize that protein-DNA binding is based on physical interactions at the atomic scale, i.e. electrostatic, van-der Waals, and hydrophobic interactions as well as (water mediated) hydrogen bonds. The deformability of DNA is also an important part of the binding specificity, since a protein-induced deformation of DNA upon binding will occur more readily for certain DNA sequences [6]. Previous atomistic methods have focused on this aspect of binding, but either neglected the protein completely [7] or modeled it as a rigid object [8]. We were therefore motivated

to develop a predictive theory of protein-DNA interactions which includes protein-sidechain flexibility in a fully atomistic description.

Our model consists of the fixed protein and DNA backbones obtained from a relevant co-crystal structure as shown in Fig. 12.1. For the unbound system we take a canonical B-DNA form, widely separated from the protein. No explicit water molecules are included in the current approach. To find preferred binding sites, we search the combined space of all DNA sequences and a library of discrete sidechain rotamers for the DNA-contacting amino acids [9].³ Since temperature-dependent statistical algorithms often fail to find the global minimum on rough potential energy surfaces [10], we resort to the deterministic dead-end elimination (DEE) [11] and A* algorithms [12]. Binding energies are evaluated using the molecular-mechanics package AMBER 7 (parm99 force field) [13]. An initial fast search for preferred DNA sequences is made using vacuum energies, but the final ranking is done with the generalized-Born implicit-solvent model (GB/SA) [14]. The AMBER force field is able to capture the fine structure of DNA, including the A- to B-DNA transition, even with the implicit-solvent model [15]. The same force field and conformational-search algorithm have been used successfully for protein design [12].

12.2 Numerical Methods

To find the highest-affinity binding sites, we use the dead-end-elimination (DEE) [11]. If it converges, DEE deterministically finds the global energy minimum. As a necessary condition to apply DEE, the total binding energy is decomposed in a pairwise fashion,

$$\Delta E^{\text{vac}} = \sum_{i=1}^P \left[\Delta\epsilon(i_r) + \sum_{j=i+1}^P \Delta\epsilon(i_r, j_s) \right], \quad (12.1)$$

where $P = N_{aa} + N_{bp}$ is the sum of the total numbers of protein amino acids N_{aa} and DNA basepairs N_{bp} . The quantities $\Delta\epsilon(i_r)$ and $\Delta\epsilon(i_r, j_s)$ are the self-energies and interaction energies of rotamers/basepairs of the bound minus unbound structures, where r and s describe rotamers/basepairs, while i and j are protein/DNA sequence positions. The resulting number of all possible conformations is extremely large, $O(n^P)$, if n is the average number of rotamers/basepairs at each position. Since a pairwise expression for the binding energy is strictly only correct in vacuum [16], we *pre-calculated* the self-energies and interaction energies without the solvent.

³ All rotamers in the library *bbdep02.May.sortlib* (downloaded from <http://dunbrack.fccc.edu/bbdep>) were employed, except for Arg and Glu. For these long-sidechain amino acids, we used only the 20 statistically most significant rotamers (probability $\gtrsim 0.1\%$).

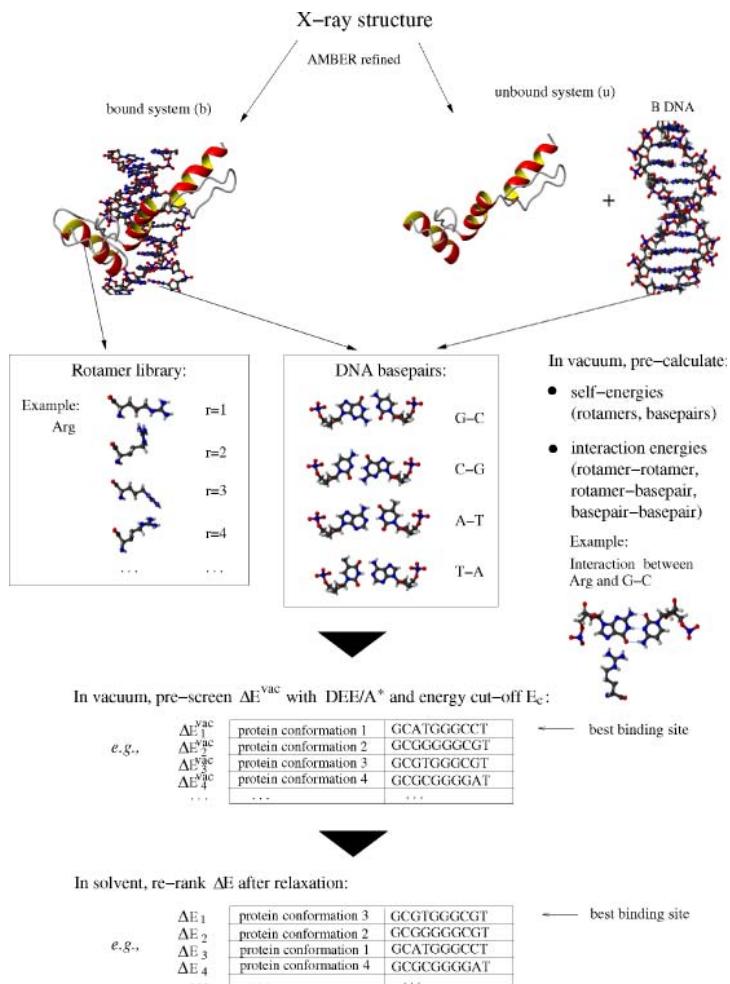


Fig. 12.1. Overview of atomistic method to identify DNA binding sites. (As an example, we choose the protein Zif268 whose “ribbon”-structure is shown)

We implemented several levels of DEE algorithms as recently described in [17, 18], including the so-called super-rotamer approach [19]. Just to mention two illustrative examples of DEEs, there is Desmet’s original dead-end elimination criterion [11] which is the least-efficient but best scaling DEE algorithm regarding CPU-time, $\approx O(P^2 n^2)$. It uses the fact that a potential rotamer/basepair i_r is inconsistent with binding energies within E_c of the global minimum and hence can be eliminated (i.e. is a “dead-ending single” or DES), if its best possible energy contribution to the total binding energy is still worse than the worst possible energy contribution of an alternative rotamer/basepair t at the same position i by more than E_c ,

$$\Delta\epsilon(i_r) + \sum_{j \neq i}^P \min \Delta\epsilon(i_r, j_s) > \Delta\epsilon(i_t) + \sum_{j \neq i}^P \max \Delta\epsilon(i_t, j_s) + E_c. \quad (12.2)$$

Here, \min (\max) is the minimum (maximum) interaction energy between rotamer/basepair i_r (i_t) and all possible rotamers/basepairs at position j . Our second example is Goldstein's [19] sharpest but worst scaling, $\approx O(P^3 n^5)$, criterion for identifying so-called "dead-ending pairs" (DEPs). Pair $\Delta\epsilon(i_r, j_s)$ is dead-ending, if

$$\begin{aligned} & \Delta\epsilon(i_r, j_s) - \Delta\epsilon(i_u, j_v) \\ & + \sum_{k \neq i, j}^P \min \{\Delta\epsilon[(i_r, j_s), k_t] - \Delta\epsilon[(i_u, j_v), k_t]\} > E_c, \end{aligned} \quad (12.3)$$

where $\Delta\epsilon[(i_{r/u}, j_{s/v}), k_t] = \Delta\epsilon(i_{r/u}, k_t) + \Delta\epsilon(i_{s/v}, k_t)$. These DEPs cannot be eliminated directly, but do not need to be included in subsequent calculations of the minimum in (12.2). As illustrated in Fig. 12.2, the DEE/ E_c algorithms for DESs and DEPs are applied iteratively until no further DESs can be eliminated. Note, however, that according to (12.2) the energy cut-off E_c is applied to each position i separately. Consequently, even with perfect convergence of the DEEs, the resulting total binding energies can still be many times E_c above the global minimum. The number of unwanted, high-energy rotamer and basepair combinations can be further reduced by combining rotamers/basepairs from different positions to super-rotamers [19]. For instance, pairing rotamers/basepairs from two different positions eliminates all of their previously flagged DEPs, and effectively applies the energy cut-off to the two positions simultaneously, leading to a stricter enforcement of the cut-off criterion. Unfortunately, the memory requirement for storing the increased number of interaction energies is a limiting factor. After several rounds of DEE algorithms and of forming super-rotamers, the remaining combinatorial space can be searched using the graph-search A* algorithm [12]. In principle, this *pre-screening* procedure returns an exhaustive list of rotamer-DNA sequence combinations with vacuum binding energies within E_c of the best.

Finally, we *re-rank* the best vacuum protein-DNA conformations using the solvated binding energy $\Delta E = E_b - E_u$, where E_b and E_u are the total energies of the bound and unbound protein-DNA systems in implicit solvent. Since the re-ranking in solvent is likely to change the order of the binding energies compared to vacuum, the cut-off E_c has to be chosen large enough to include the top solvated configurations.

12.3 Results and Discussion

As an example, we apply our method to the mouse protein, Zif268 (PDB file 1AAY, Fig. 12.1). Zif268 has a tandem repeat of three specific binding motifs,

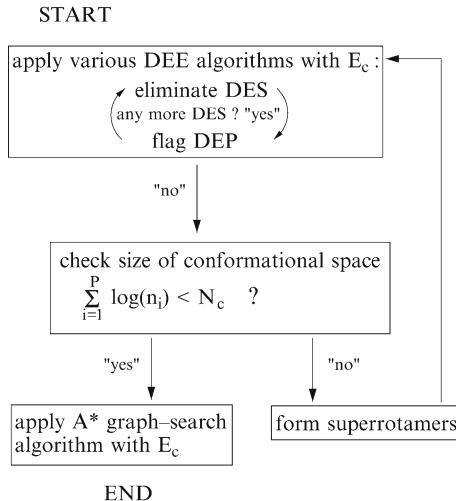


Fig. 12.2. Flowchart of the dead-end elimination (DEE) and A* graph-search algorithms. DES and DEP are defined in (12.2), (12.3). The parameters E_c and N_c are the energy cut-off and the upper limit of conformational size manageable by the A* algorithm, respectively. Furthermore, $P = N_{aa} + N_{na}$ (see (12.1)), and n_i is the number of rotamers/basepairs at position i . Due to convergence problems of the DEE/A*, we restricted ourselves to only a single (best) set of rotamers for each vacuum DNA sequence that was not eliminated

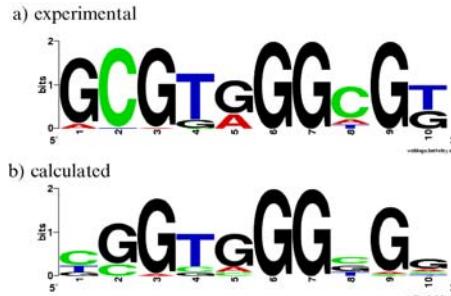


Fig. 12.3. (a) Zif268 10-base sequence logo from experiment [21], and (b) from our calculation. Figure prepared with the WebLogo [23] from <http://weblogo.berkeley.edu/>

so-called zinc fingers, each stabilized by a zinc ion (for a review see [3]). A high resolution co-crystal structure is available [20] with the protein bound to its consensus sequence, GCGTGGGCGT [21].

The best calculated binding sequence in solvent was TGGTGGGCGG, the 10th was GCGTGGGCGA, while the 17th matches the experimental consensus GCGTGGGCGT, as indicated by the underlining. Note that

the 10th (last) base position of the binding sequence is experimentally only weakly conserved, and was not considered part of a previously derived consensus sequence [22].

Figure 12.3 compares the experimental logo (a graphical presentation of the weight-matrix [23]) (a) with our calculations using the 30 best binding sequences (b). The logos clearly show strong similarities, but the lack of explicit water molecules in our model, end-effects caused by the short DNA segment, and an additional approximation (see caption of Fig. 12.2) lead to some disagreement – particularly at the DNA-positions 1 and 2 contacted by the zinc finger which has the most water-mediated contacts [24].

In summary, we have introduced a novel methodology for identifying potential protein-DNA binding sites through a global search of DNA sequences and DNA-contacting-sidechain conformations. We tested performance by calculating the DNA-sequence logo for the mouse transcription factor Zif268, and found good agreement with experiment.

Acknowledgement

Research sponsored by the Division of Mathematical Information and Computational Sciences, U.S. Department of Energy, under Contract DE-AC05-00OR22725 with UT-Battelle LLC.

References

1. M.L. Bulyk, X. Huang, Y. Choo, G.M. Church: Proc. Natl. Acad. Sci. USA **98**, 7158 (2001)
2. T.I. Lee, N.J. Rinaldi, F. Robert, D.T. Odom, Z. Bar-Joseph, G.K. Gerber, N.M. Hannett, C.T. Harbison, C.M. Thompson, I. Simon I. *et al.*: Nature **298**, 799 (2002)
3. C.O. Pabo, E. Peisach, R.A. Grant: Annu. Rev. Biochem **70**, 313 (2001)
4. G.D. Stormo: Bioinformatics **16**, 16 (2000)
5. P.V. Benos, M.L. Bulyk, G.D. Stormo: Nucleic Acids Res. **30**, 4442 (2002)
6. W.K. Olson, A.A. Gorin, X.J. Lu, L.M. Hock, V.B. Zhurkin: Proc. Natl. Acad. Sci. USA **95**, 11163 (1998)
7. N.R. Steffen, S.D. Murphy, L. Tolleri, G.W. Hatfield, R.H. Lathrop: Bioinformatics **18**, S22 (2002)
8. I. Lafontaine, R. Lavery: Biopolymers (Nucleic Acid Sciences) **56**, 292 (2001)
9. R.L. Dunbrack, Jr.: Curr. Opin. Struct. Biol. **12**, 431 (2002)
10. C.A. Voigt, D.B. Gordon, S.L. Mayo: J. Mol. Biol. **299**, 789 (2000)
11. J. Desmet, M. de Maeyer, B. Hazes, I. Lasters: Nature **356**, 539 (1992)
12. A.R. Leach, A.P. Lemon: Proteins **33**, 227 (1998)
13. D.A. Pearlman, D.A. Case, J.W. Caldwell, W.S. Ross, T.E. Cheatham III, S. Debolt, D. Ferguson, G.L. Seibel, P.A. Kollman: Comp. Phys. Commun. **91**, 1 (1995)
14. A. Onufriev, D. Bashford, D.A. Case: J. Phys. Chem. B **104**, 3712 (2000)

15. V. Tsui, D.A. Case: *J. Am. Chem. Soc.* **122**, 2489 (2000)
16. A.G. Street, S.L. Mayo: *Fold. Des.* **3**, 253 (1998)
17. D.B. Gordon, S.L. Mayo: *J. Comp. Chem.* **19**, 1505 (1998)
18. N.A. Pierce, J.A. Spriet, J. Desmet, S.L. Mayo: *J. Comp. Chem.* **21**, 999 (2000)
19. R. Goldstein: *Biophys. J.* **66**, 1335 (1994)
20. M. Elrod-Erickson, M.A. Rould, L. Nekludova, C.O. Pabo: *Structure* **4**, 1171 (1996)
21. A. Swirnoff, J. Milbrandt: *Mol. Cell. Biol.* **15**, 2275 (1995)
22. B. Christy, D. Nathans: *Proc. Natl. Acad. Sci. USA* **86**, 8737 (1989)
23. T.D. Schneider, R.M. Stephens: *Nucleic Acids Res.* **18**, 6097 (1999)
24. G. Roxström, I. Velázquez, M. Paulino, O. Tapia: *J. Biomol. Struct. Dyn.* **16**, 301 (1998)

Part IV

Algorithms and Methods

13 Geometric Cluster Algorithm for Interacting Fluids

E. Luijten and J. Liu

University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA

Abstract. We discuss a new Monte Carlo algorithm for the simulation of complex fluids. This algorithm employs geometric operations to identify clusters of particles that can be moved in a rejection-free way. It is demonstrated that this *geometric cluster algorithm* (GCA) constitutes the continuum generalization of the Swendsen–Wang and Wolff cluster algorithms for spin systems. Because of its nonlocal nature, it is particularly well suited for the simulation of fluid systems containing particles of widely varying sizes. The efficiency improvement with respect to conventional simulation algorithms is a rapidly growing function of the size asymmetry between the constituents of the system. We study the cluster-size distribution for a Lennard–Jones fluid as a function of density and temperature and provide a comparison between the generalized GCA and the hard-core GCA for a size-asymmetric mixture with Yukawa-type couplings.

13.1 Introduction and Motivation

The presence of multiple time and length scales constitutes one of the major problems in computer simulations of matter. In simulations that faithfully capture the dynamic evolution of a system, the fastest particles in the system dictate the required time resolution. If different types of particles with widely varying diffusion rates are present, then the slower particles may be unable to explore the entire configuration space during the course of the simulation, leading to ergodicity problems.

In the computational study of complex fluids, such as colloidal suspensions, this problem frequently occurs, since such systems typically contain particles of different sizes. A concrete example is the “nanoparticle haloing” phenomenon discovered by Lewis and coworkers [1]. This experimental work deals with a new approach to the stabilization of suspensions of micron-sized spherical particles, which tend to aggregate under the influence of their mutual van der Waals attraction. Conventional approaches to prevent this gelation, such as charge stabilization (variation of the pH to alter the surface charge of the particles), lead to complications in certain applications, e.g., in the formation of colloidal crystals, where the electrostatic repulsion prevents the close packing of the particles. It was found that these complications can be avoided by generating an effective colloidal repulsion through the addition of highly-charged nanometer-sized particles to the suspension of the

(near-neutral) colloids. A concrete explanation for the underlying mechanism leading to the effective repulsions is currently lacking. Evidently, a computational approach to this problem must involve both the microspheres and the nanoparticles, which typically differ by a factor 100 in diameter. Traditionally, the effective pair interaction between colloids (potential of mean force) is then calculated by “integrating out” the smaller species, e.g., by simulating a system containing two spheres at fixed separation embedded in a sea of smaller particles [2, 3]. Here, we discuss a new simulation algorithm [4] that is capable of explicitly incorporating both species and computing equilibrium properties of the suspension, without suffering from the disparity in time scales. This algorithm can be viewed as a continuum version of the widely-used cluster algorithms for lattice spin models [5, 6].

13.2 Cluster Monte Carlo Algorithms

Equilibrium Monte Carlo methods are aimed at obtaining static thermodynamic and structural properties by generating system configurations according to the Boltzmann distribution. Nonphysical moves may be employed in the underlying Markov process, which – in principle – offers an exquisite way to overcome the presence of multiple time scales. Thus, Monte Carlo algorithms are the method of choice for, e.g., simulations of critical phenomena and phase transitions. In the context of size-asymmetric fluids, collective moves have been devised in which groups of particles are moved simultaneously. Unless these groups are identified in a careful way, which typically involves knowledge about the physical properties of the system, the Monte Carlo step will entail a large energy change and consequently have only a very small acceptance rate. Thus, while this approach has been used quite successfully in specific situations [7–10], it typically involves a tunable parameter and cannot be generalized in a straightforward fashion.

The situation is rather different for spin models, for which Swendsen and Wang (SW) [5] introduced a *cluster algorithm* in which groups of parallel spins are identified in a probabilistic manner, based upon the Fortuin–Kasteleyn mapping of the Potts model onto the random-cluster model [11]. These groups can subsequently be flipped independently. This rejection-free algorithm (every completed cluster can be flipped without an additional evaluation of the resulting energy difference) suppresses critical slowing down. Accordingly, its generalization to off-lattice fluids has been a widely-pursued goal. However, the SW algorithm – as well as the even more efficient single-cluster variant introduced by Wolff [6] – relies on invariance of the Hamiltonian under a spin-inversion operation. For a fluid, this translates into particle-hole symmetry, which is only obeyed for lattice gases. Consequently, efficient off-lattice cluster algorithms have only been designed for a small number of specific fluid models [12, 13], and it is not clear how these methods can be generalized. Hard-sphere fluids constitute another special case. Here

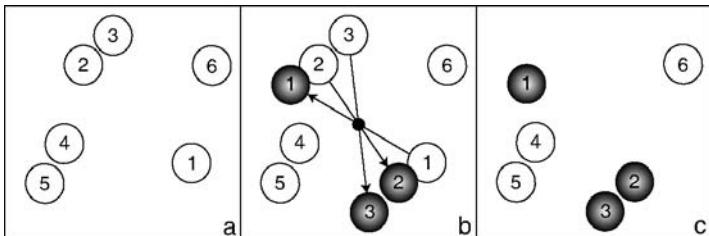


Fig. 13.1. Two-dimensional illustration of the interacting geometric cluster algorithm. Open and shaded circles label the particles before and after the geometrical operation, respectively. The small filled disk denotes the pivot. (a) Initial configuration; (b) construction of a new cluster via point reflection of particles 1–3 with respect to the pivot; (c) final configuration

every configuration without particle overlaps has the same energy. Dress and Krauth [14] devised a strategy to create such configurations by means of geometric transformations. Their approach is particularly advantageous in size-asymmetric mixtures, but cannot be applied to systems with other interactions without supplementing it with a costly acceptance criterion. Nevertheless, a remarkable feature of this geometric cluster algorithm is that it relies on the invariance of the Hamiltonian under geometric operations. Here, we exploit this property to formulate a cluster Monte Carlo algorithm that is applicable to arbitrary pair potentials without the imposition of an acceptance criterion.

13.3 Generalized Geometric Cluster Algorithm

13.3.1 Single-Cluster Variant

The geometric cluster algorithm as formulated by Dress and Krauth starts from a configuration of particles, with periodic boundary conditions. This configuration is rotated around an arbitrarily chosen pivot. Groups of particles that overlap between the original and the rotated configuration are exchanged between these configurations. In practice, it is more convenient to construct only a single cluster and to carry out a point reflection on the fly [15]. For each cluster a new pivot is chosen. This constitutes the counterpart of the Wolff algorithm for spin models [6]. In the presence of a general, isotropic pair potential $V(r)$ the cluster construction proceeds as follows (cf. Fig. 13.1):

1. Choose a random pivot point.
2. Choose the first particle i at random and move it from its original position \mathbf{r}_i to its new position \mathbf{r}'_i , via a point reflection with respect to the pivot.

3. Identify all particles that interact with i in its original position or in its new position. These particles j are considered for a point reflection with respect to the pivot. This reflection is carried out with a probability $p_{ij} = \max[1 - \exp(-\beta\Delta_{ij}), 0]$, where $\Delta_{ij} = V(|\mathbf{r}'_i - \mathbf{r}_j|) - V(|\mathbf{r}_i - \mathbf{r}_j|)$ and $\beta = 1/k_B T$. Note that p_{ij} solely depends on the interaction strength between i and j .
4. Repeat step 3 in an iterative fashion for each particle j that is added to the cluster. If j is moved with respect to the pivot, then all its interacting neighbors that have not yet been added to the cluster are considered for inclusion as well. The cluster construction is completed once all interacting neighbors have been considered.

It is instructive to compare this prescription to the Wolff cluster algorithm, in which a cluster of parallel spins is grown from a randomly chosen initial spin. Parallel spins with a ferromagnetic coupling constant K are added to the cluster with a probability $1 - \exp(-2K)$. The energy difference between the parallel and the antiparallel configuration indeed equals $2K$. An antiparallel spin is never added to the cluster, in accordance with the fact that such a pair will be in a state of lower energy if only the first spin is flipped.

This algorithm is ergodic, since each particle can be moved over an arbitrarily small distance. Namely, there is a non-vanishing probability that a cluster consists of only a single particle and the pivot can be located arbitrarily close to the center of this particle. Detailed balance is proven by considering a configuration X which is transformed into a new configuration Y by means of a cluster move. The energy change $E_Y - E_X$ results from every particle that is *not* included in the cluster but interacts with one or more particles that are part of the cluster. Each “broken bond” has a probability $1 - p_k$, so that the total probability of forming a cluster is given by

$$T(X \rightarrow Y) = C \prod_k (1 - p_k). \quad (13.1)$$

The prefactor C accounts for the probability of choosing a specific pivot and for the probability of creating a specific arrangement of bonds inside the cluster. The total set $\{k\}$ of broken bonds can be divided into two subsets. The broken bonds l that lead to an increase Δ_l in pair energy have a probability $\exp(-\beta\Delta_l)$, whereas the broken bonds m that lead to a decrease in pair energy have a probability equal to unity. Accordingly, the transition probability can be written as

$$T(X \rightarrow Y) = C \exp \left[-\beta \sum_l \Delta_l \right]. \quad (13.2)$$

The reverse move, in which the configuration Y is transformed into the configuration X by moving a cluster that is constructed in the same way, has a probability

$$T(Y \rightarrow X) = C \prod_k (1 - \bar{p}_k), \quad (13.3)$$

where the sum runs over the same set $\{k\}$ of broken bonds as in (13.1), but the sign of all energy differences Δ_k has been reversed (indicated by \bar{p}_k). Accordingly, the sum over Δ_l in (13.2) is replaced by the negative sum over the complementary set $\{m\}$,

$$T(Y \rightarrow X) = C \exp \left[+\beta \sum_m \Delta_m \right]. \quad (13.4)$$

The point reflection is a self-inverse operation, so that detailed balance is obeyed without the need to impose an acceptance criterion:

$$\frac{T(X \rightarrow Y)}{T(Y \rightarrow X)} = \exp \left[-\beta \sum_k \Delta_k \right] = \frac{\exp(-\beta E_Y)}{\exp(-\beta E_X)}. \quad (13.5)$$

Since energy differences are taken into account on a pair-wise basis during the construction of the cluster, rather than as a total energy difference after the cluster has been completed, the cluster is constructed in such a way that large energy differences are avoided. The clusters are representative of the actual structure of the system.

13.3.2 Multiple-Cluster Variant

In order to demonstrate that the generalized geometric cluster algorithm (GCA) indeed constitutes the off-lattice counterpart of the SW and Wolff cluster algorithms, it is instructive to formulate a multiple-cluster variant. This formulation yields a full decomposition of an off-lattice fluid configuration into *stochastically independent clusters*. In the following, we demonstrate how it can be phrased as a natural extension of the single-cluster version.

First, a cluster is constructed according to the Wolff version of the GCA, with the exception that the cluster is only *identified*; particles belonging to the cluster are marked but not actually moved. The chosen pivot will also be used for the construction of all subsequent clusters in this decomposition. These subsequent clusters are built just like the first cluster, except that particles that are already part of an earlier cluster will never be considered for a new cluster. Once each particle is part of a cluster the decomposition is completed and each cluster is moved with a probability f .

Although all clusters except the first are built in a *restricted* fashion, every individual cluster is constructed according to the rules of the Wolff formulation of the GCA. The exclusion of particles that are already part of another cluster simply reflects the fact that every bond should be considered only once. If a bond is broken during the construction of an earlier cluster it must not be re-established during the construction of a subsequent cluster.

In order to establish that this prescription is a true equivalent of the SW algorithm, we prove that each cluster can be moved (reflected) independently while preserving detailed balance. If only a single cluster is actually moved, this essentially corresponds to the Wolff version of the GCA, since each cluster is built according to the GCA prescription. The same holds true if several clusters are moved and no interactions are present between particles that belong to different clusters (the hard-sphere algorithm is a particular realization of this situation). If two or more clusters are moved and broken bonds exist between these clusters, i.e., a non-vanishing interaction exists between particles that belong to disparate (moving) clusters, then the shared broken bonds are actually preserved and the proof of detailed balance provided in the previous section no longer applies in its original form. However, since these bonds are identical in the forward and reverse move, the corresponding factors cancel out. This is illustrated for the situation of two clusters whose construction involves, respectively, two sets of broken bonds $\{k_1\}$ and $\{k_2\}$. Each set comprises bonds l that lead to an *increase* in pair energy and bonds m that lead to a *decrease* in pair energy. We further subdivide these sets into *external* bonds that connect cluster 1 or 2 with the remainder of the system and *joint* bonds that connect cluster 1 and 2. Accordingly, the probability of creating cluster 1 is given by

$$C_1 \prod_{i \in \{l_1\}} (1 - p_i) = C_1 \prod_{i \in \{l_1^{\text{ext}}\}} (1 - p_i) \prod_{j \in \{l_1^{\text{joint}}\}} (1 - p_j). \quad (13.6)$$

Upon construction of the first cluster, the creation of the second cluster has a probability

$$C_2 \prod_{i \in \{l_2^{\text{ext}}\}} (1 - p_i), \quad (13.7)$$

since all joint bonds in $\{l_1^{\text{joint}}\}$ already have been broken. The factors C_1 and C_2 account for the probability of realizing a particular arrangement of internal bonds in clusters 1 and 2, respectively. Hence, the total transition probability for moving both clusters (upon fixing the pivot) is given by

$$T_{12} = C_1 C_2 \exp \left[-\beta \sum_{i \in \{l_1^{\text{ext}}\}} \Delta_i - \beta \sum_{j \in \{l_2^{\text{ext}}\}} \Delta_j - \beta \sum_{n \in \{l_1^{\text{joint}}\}} \Delta_n \right]. \quad (13.8)$$

In the reverse move, the energy differences for all external broken bonds have changed sign, but the energy differences for the joint bonds connecting cluster 1 and 2 are the same as in the forward move. Thus, cluster 1 is created with probability

$$C_1 \prod_{i \in \{m_1^{\text{ext}}\}} (1 - \bar{p}_i) \prod_{j \in \{l_1^{\text{joint}}\}} (1 - p_j)$$

$$= C_1 \prod_{i \in \{m_1^{\text{ext}}\}} \exp[+\beta \Delta_i] \prod_{j \in \{l_1^{\text{joint}}\}} \exp[-\beta \Delta_j], \quad (13.9)$$

where the \bar{p} reflects the sign change compared to the forward move and the product over the external bonds involves the complement of the set $\{l_1^{\text{ext}}\}$. The creation probability for the second cluster is

$$C_2 \prod_{i \in \{m_2^{\text{ext}}\}} (1 - \bar{p}_i) = C_2 \prod_{i \in \{m_2^{\text{ext}}\}} \exp[+\beta \Delta_i] \quad (13.10)$$

and the total transition probability for the reverse move is

$$\tilde{T}_{12} = C_1 C_2 \exp \left[+\beta \sum_{i \in \{m_1^{\text{ext}}\}} \Delta_i + \beta \sum_{j \in \{m_2^{\text{ext}}\}} \Delta_j - \beta \sum_{n \in \{l_1^{\text{joint}}\}} \Delta_n \right]. \quad (13.11)$$

Accordingly, detailed balance is still fulfilled,

$$\frac{T_{12}}{\tilde{T}_{12}} = \exp \left[-\beta \sum_{i \in \{k_1^{\text{ext}}\}} \Delta_i - \beta \sum_{j \in \{k_2^{\text{ext}}\}} \Delta_j \right] = \exp[-\beta(E_Y - E_X)], \quad (13.12)$$

in which $\{k_1^{\text{ext}}\} = \{l_1^{\text{ext}}\} \cup \{m_1^{\text{ext}}\}$ and $\{k_2^{\text{ext}}\} = \{l_2^{\text{ext}}\} \cup \{m_2^{\text{ext}}\}$ and E_X and E_Y refer to the total internal energy before and after the move, respectively. This treatment applies to any simultaneous move of clusters, so that each cluster in the decomposition indeed can be moved independently without violating detailed balance. It is noteworthy that the probabilities for breaking joint bonds in the forward and reverse moves cancel only because the probability in the cluster construction factorizes into pairwise probabilities, as opposed to the probability for a multiple-particle move in a Metropolis-type algorithm.

As demonstrated in Fig. 13.2 for a binary mixture, the results obtained by means of the multiple-cluster geometric algorithm agree perfectly with those obtained using the single-cluster version.

13.4 Performance

The most striking feature of the generalized geometric cluster algorithm, apart from the fact that it creates clusters that can be moved in a rejection-free manner, is the speed at which it relaxes size-asymmetric mixtures. We illustrate this here for a binary fluid mixture consisting of N_s small and N_l large spherical particles with size ratio $\alpha \equiv \sigma_l/\sigma_s > 1$. The particles are contained in a fixed volume, at equal packing fractions $\eta_s = \eta_l = 0.1$. While $N_l = 150$ is kept fixed, N_s increases from 1 200 to 506 250 as α is varied

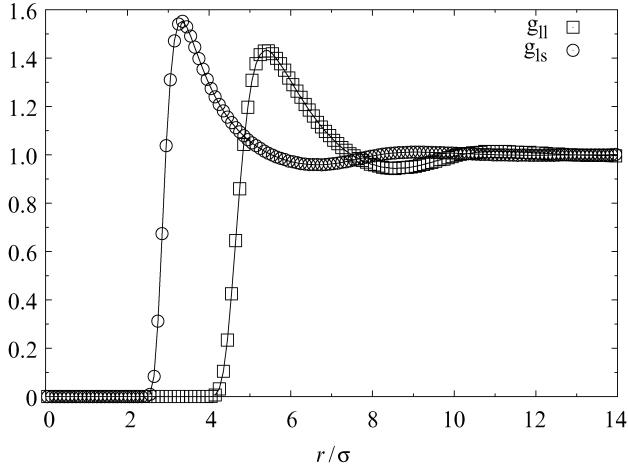


Fig. 13.2. Comparison between the single-cluster version (*solid lines*) and the multiple-cluster version (*symbols*) of the generalized geometric cluster algorithm. The figure shows pair correlation functions for the size-asymmetric Lennard–Jones mixture described in [4]. The system contains 800 small (diameter σ) and 400 large particles (diameter 5σ) at a total packing fraction $\eta \approx 0.213$. g_{ll} and g_{ls} represent the large–large and large–small correlation functions, respectively

from 2 to 15. Pairs of small particles and pairs involving a large and a small particle act like hard spheres. However, in order to prevent depletion-driven aggregation of the large particles, they have a Yukawa repulsion,

$$U_{22}(r) = \begin{cases} +\infty & r \leq \sigma_1 \\ J \exp[-\kappa(r - \sigma_1)]/(r/\sigma_1) & r > \sigma_1 \end{cases}, \quad (13.13)$$

where $\beta J = 3.0$ and the screening length $\kappa^{-1} = \sigma_s$. In the simulation, the exponential tail is cut off at $3\sigma_1$. As a measure of efficiency we consider the integrated autocorrelation time τ obtained from the energy autocorrelation function [16],

$$C(t) = \frac{\langle E(0)E(t) \rangle - \langle E(0) \rangle^2}{\langle E(0)^2 \rangle - \langle E(0) \rangle^2}. \quad (13.14)$$

For conventional MC calculations, τ rapidly increases with increasing α , because the large particles tend to get trapped by the small particles. Accordingly, an accurate estimate for τ could only be obtained for $\alpha \leq 7$. By contrast, the generalized GCA has an autocorrelation time that only weakly depends on the size ratio, as illustrated in Fig. 13.3. At $\alpha = 7$ the resulting efficiency gain already amounts to more than three orders of magnitude.

A crucial limitation of the generalized GCA is that each cluster must only occupy a fraction of the entire system. As observed by Dress and Krauth [14],

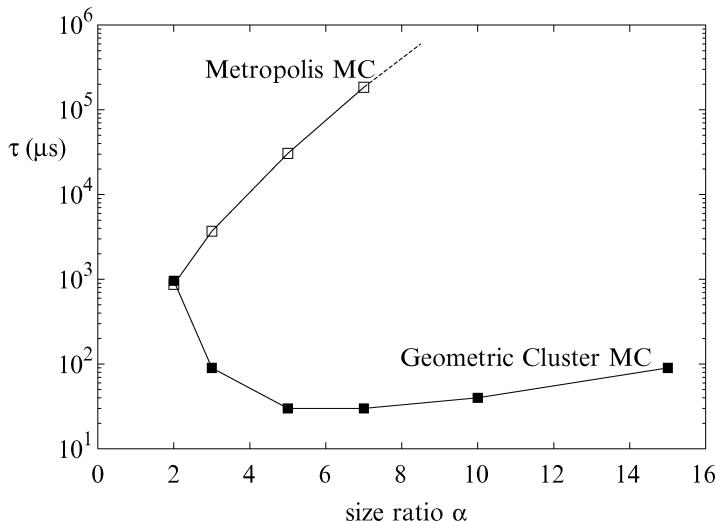


Fig. 13.3. Efficiency comparison between a conventional local update algorithm (*open symbols*) and the generalized geometric cluster algorithm (*closed symbols*), for a binary mixture (see text) with size ratio α . Whereas the autocorrelation time per particle (expressed in μs of CPU time per particle move) rapidly increases with size ratio, the GCA features only a weak dependence on α

the entire system typically will be occupied by a single cluster once the percolation threshold is reached in the combined system containing a configuration and its point-reflected counterpart. For the original system this leads, in three dimensions, to a practical upper limit in the packing fraction around 0.23–0.25. This number will vary as a function of size asymmetry and, if additional pair interactions are present, temperature. It is therefore instructive to study the cluster-size distribution as a function of reduced density ρ^* and temperature T . Figure 13.4 illustrates the cluster-size distributions as obtained by means of the multiple-cluster GCA for a regular (one-component) Lennard–Jones fluid. In the top panel, $\rho^* = 0.32 \approx \rho_c^*$. Already at temperatures that are far above the critical temperature $T_c \approx 1.19$, the cluster-size distribution starts to tend toward a bimodal form, indicative of the formation of large clusters. In the vicinity of the critical temperature, the average cluster size has become very large. For comparison, the bottom panel displays the cluster-size distributions at a twice smaller density, $\rho^* = 0.16$. In this case the bimodal shape does not appear until close to T_c .

The properties suggest that the generalized GCA, at least for the one-component Lennard–Jones fluid, will not suppress critical slowing – the primary advantage of lattice cluster algorithms. For the off-lattice GCA, this property is less essential, because of the speed-up it delivers for the simulation of size-asymmetric fluids over a wide range of temperatures and packing fractions. Nevertheless, we have investigated the integrated autocorrelation

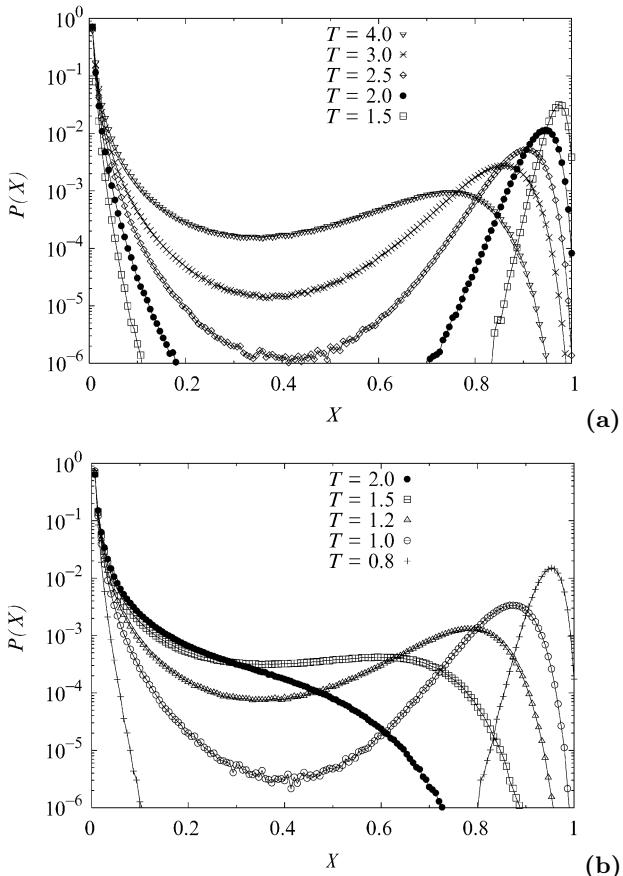


Fig. 13.4. Cluster size distributions as a function of relative cluster size X , for a monodisperse Lennard–Jones fluid. (a) Reduced density $\rho^* \equiv \rho\sigma^3 = 0.32$; (b) reduced density $\rho^* = 0.16$. Identical symbols compare to identical temperatures in both panels. All temperatures are indicated in terms of ε/k_B . See text for discussion

time for the energy at the critical point, as a function of linear system size. In Fig. 13.5 these times are collected for three algorithms. (1) Conventional local-update Metropolis algorithm; (2) Wolff version of the GCA; (3a) SW version of the GCA, in which each cluster is reflected with a probability 0.50; (3b) SW version of the GCA, in which each cluster is reflected with a probability 0.75. Just as for spin models, the single-cluster version outperforms the Swendsen–Wang type cluster decompositions. However, all variants of the GCA exhibit the same power-law behavior, which outperforms the Metropolis algorithm by a factor $\sim L^{2.1}$. It is important to emphasize that this acceleration may be due to the suppression of the hydrodynamic slowing down [17] caused by the conservation of the density (which may couple to the energy

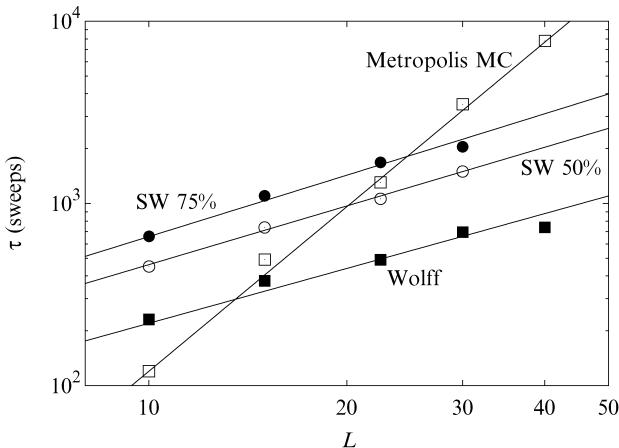


Fig. 13.5. Energy autocorrelation time τ as a function of linear system size for a critical Lennard–Jones fluid, in units of particle sweeps, for three different Monte Carlo algorithms: Local moves (“Metropolis MC”); GCA with Swendsen–Wang type cluster decomposition and probability 0.50 (“SW 50%”) and 0.75 (“SW 75%”) of moving each cluster; single-cluster GCA (“Wolff”)

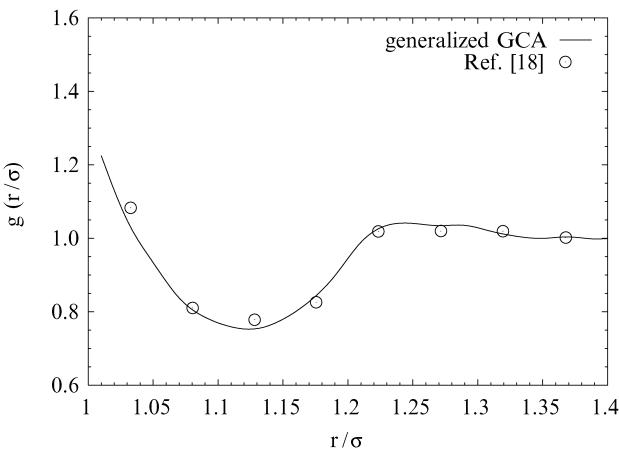


Fig. 13.6. Pair correlation function of dilute colloidal particles (diameter σ) in an environment of smaller particles (diameter $\sigma/5$, packing fraction $\eta = 0.116$) that experience a Yukawa-type attractive interaction with the colloids. The symbols represent data obtained by means of the hard-core GCA [18]; the solid line was obtained from the generalized GCA

correlations [4]). Another striking point is that already for moderate system sizes the generalized GCA outperforms the Metropolis algorithm, despite the time-consuming construction of large clusters (cf. Fig. 13.4a) which lead to only small configurational changes.

13.5 Illustration

In order to illustrate the capabilities of the generalized GCA, we have computed the pair correlation function of dilute colloidal particles (packing fraction $\eta_l = 0.001$) in an environment of particles with a five times smaller diameter (packing fraction $\eta_s = 0.116$). Large and small particles act as hard spheres, but unlike pairs (i.e., large–small) experience a Yukawa attraction which promotes the accumulation of small particles around the colloids. This system has been studied in [18] by means of the original hard-core GCA, in which clusters are moved according to an acceptance criterion, as proposed in [14]. This potentially greatly deteriorates performance, as entire clusters will be constructed that are subsequently rejected. Indeed, the authors report [18] that the colloid pair correlation function $g(r)$ had to be obtained via numerical differentiation of the integrated pair correlation function, rather than through direct sampling. The generalized GCA can handle this system without complication, as demonstrated in Fig. 13.6. In order to obtain data comparable to the solid curve in this figure, the authors of [18] utilized a polynomial fit to the integrated pair correlation function, which potentially leads to ambiguities.

13.6 Conclusion and Outlook

We have discussed the generalized geometric cluster algorithm, a Monte Carlo method for the simulation of fluids by means of geometric operations. This algorithm creates nonlocal multiple-particle moves that are capable of rapidly decorrelating fluid configurations that contain particles of widely different sizes. The multiple-particle moves are constructed in such a way that the typical decrease in acceptance rate is avoided; every proposed move is accepted without violating detailed balance. It is anticipated that this algorithm will find widespread application, in particular in the simulation of complex fluids and suspensions in which the solvent is modeled as an implicit background. Potential generalizations include the treatment of particles with internal degrees of freedom, such as nonspherical particles, and other geometries, such as layered systems.

Acknowledgements

This material is based upon work supported by the National Science Foundation under grants nos. DMR-0346914 and CTS-0120978, and by the U.S. Department of Energy, Division of Materials Sciences, under Award No. DEFG02-91ER45439, through the Frederick Seitz Materials Research Laboratory at the University of Illinois at Urbana-Champaign.

References

1. V. Tohver, J.E. Smay, A. Braem, P.V. Braun, J.A. Lewis: Proc. Natl. Acad. Sci. USA **98**, 8950 (2001)
2. R. Dickman, P. Attard, V. Simonian: J. Chem. Phys. **107**, 205 (1997)
3. E. Allahyarov, I. D'Amico, H. Löwen: Phys. Rev. Lett. **81**, 1334 (1998)
4. J. Liu, E. Luijten: Phys. Rev. Lett. **92**, 35504 (2003)
5. R.H. Swendsen, J.-S. Wang: Phys. Rev. Lett. **58**, 86 (1987)
6. U. Wolff: Phys. Rev. Lett. **62**, 361 (1989)
7. D. Wu, D. Chandler, B. Smit: J. Phys. Chem. **96**, 4077 (1992)
8. A. Jaster: Physica A **264**, 134 (1999)
9. L. Lue, L.V. Woodcock: Mol. Phys. **96**, 1435 (1999)
10. V. Lobaskin, P. Linse: J. Chem. Phys. **111**, 4300 (1999)
11. C.M. Fortuin, P.W. Kasteleyn: Physica **57**, 536 (1972)
12. G. Johnson, H. Gould, J. Machta, L.K. Chayes: Phys. Rev. Lett. **79**, 2612 (1997)
13. R. Sun, H. Gould, J. Machta, L.W. Chayes: Phys. Rev. E **62**, 2226 (2000)
14. C. Dress, W. Krauth: J. Phys. A **28**, L597 (1995)
15. J.R. Heringa, H.W.J. Blöte: Phys. Rev. E **57**, 4976 (1998)
16. K. Binder, E. Luijten: Phys. Rep. **344**, 179 (2001)
17. P.C. Hohenberg, B.I. Halperin: Rev. Mod. Phys. **49**, 435 (1977)
18. J.G. Malherbe, S. Amokrane: Mol. Phys. **97**, 677 (1999)

14 Polymer Simulations with a Flat Histogram Stochastic Growth Algorithm

T. Prellberg¹, J. Krawczyk¹, and A. Rechnitzer²

¹ Institut für Theoretische Physik, Technische Universität Clausthal
Arnold-Sommerfeld-Straße 6, 38678 Clausthal-Zellerfeld, Germany,

² Department of Mathematics and Statistics, The University of Melbourne
Parkville 3010, Australia

Abstract. We present Monte Carlo simulations of lattice models of polymers. These simulations are intended to demonstrate the strengths of a powerful new flat histogram algorithm which is obtained by adding microcanonical reweighting techniques to the pruned and enriched Rosenbluth method (PERM).

14.1 Introduction

Monte Carlo simulations of polymer models have played a significant role in the development of Monte Carlo methods for more than fifty years [1]. We present here results of simulations performed with a powerful new algorithm, flatPERM [2], which combines a stochastic growth algorithm, PERM [3], with umbrella sampling techniques [4]. This leads to a flat histogram in a chosen parameterization of configuration space.

The stochastic growth algorithm used is the pruned and enriched Rosenbluth method (PERM) [3], which is an enhancement of the Rosenbluth and Rosenbluth algorithm [5], an algorithm that dates back to the early days of Monte Carlo simulations. While PERM already is a powerful algorithm for simulating polymers, the addition of flat-histogram techniques [6] provides a significant enhancement, as has already been exploited in [7], where it has been combined with multicanonical sampling [8].

Before we describe the algorithm in detail and present results of the simulations, we give a brief motivating introduction to the lattice models considered here.

If one wants to understand the critical behavior of long linear polymers in solution, one is naturally led to a coarse-grained picture of polymers as beads of monomers on a chain. There are two main physical ingredients leading to this picture. First, one needs an *excluded volume* effect, which takes into account the fact that different monomers cannot simultaneously occupy the same region in space. Second, the quality of the solvent can be modeled by an effective monomer–monomer interaction. Monomers in a good solvent are surrounded by solvent molecules and hence experience an effective monomer–monomer repulsion. Similarly, a bad solvent leads to an effective monomer–monomer attraction.

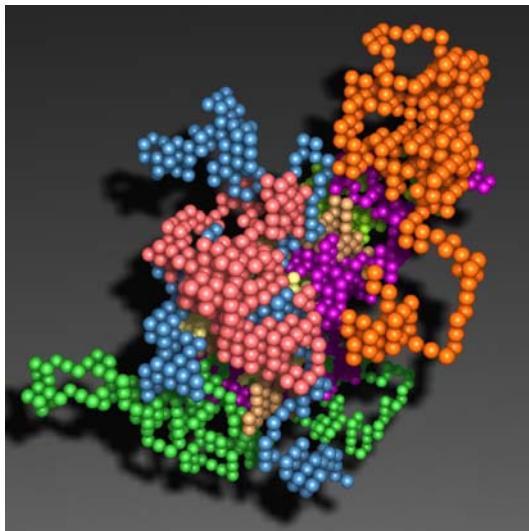


Fig. 14.1. Eight lattice polymers in a bad solvent (picture courtesy of H. Frauenkron, FZ Jülich)

Consequently, polymers in a good solvent form swollen *coils*, whereas polymers in a bad solvent form collapsed *globules* and also clump together with each other (see Fig. 14.1). In order to study the transition between these two states, it is advantageous to go to the limit of an infinitely dilute solution, in which one considers precisely one polymer in an infinitely extended solvent.

As we are interested in critical behavior, it is also possible to further simplify the model by discretizing it. Due to universality, the critical behavior is expected to be unchanged by doing so. We therefore consider random walks on a regular lattice, e.g. the simple cubic lattice for a three-dimensional model. One can think of each lattice site corresponding to a monomer and the steps as monomer-monomer bonds.

We model excluded volume effects by considering *self-avoiding* random walks which are not allowed to visit a lattice site more than once. The quality of the solvent is modeled by an attractive short-range interaction between non-consecutive monomers which occupy nearest-neighbor sites on the lattice. At this point we may add more structure to our polymer model by considering monomer-specific interactions. Specific properties of monomers i and j on the chain lead to an interaction $\epsilon_{i,j}$ depending on i and j .

In this paper, we will consider three examples in detail. First, we consider as pedagogical example, the problem of simulating polymers in a two-dimensional strip. The interaction energy is simply $\epsilon_{i,j} = 0$, however, the introduction of boundaries makes simulations difficult [9].

Our second example is the HP model which is a toy model of proteins [10]. It consists of a self-avoiding walk with two types of monomers along the sites

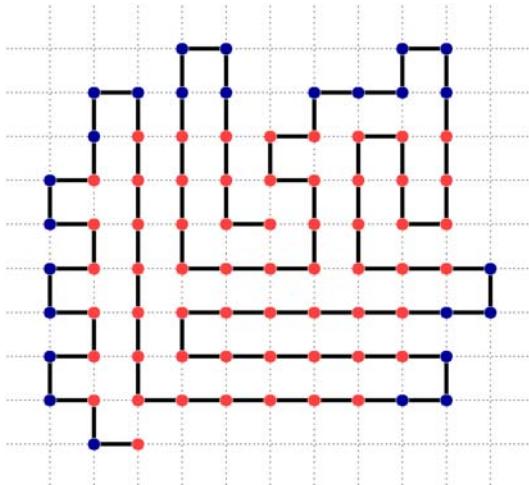


Fig. 14.2. HP model: shown is the conjectured groundstate of a sequence with 85 monomers on the square lattice. The monomers with a lighter shade are of type H (hydrophobic), the monomers with a darker shade are of type P (polar)

visited by the walk – hydrophobic (type H) and polar (type P). One considers monomer-specific interactions, mimicking the interaction with a polar solvent such as water. The interaction strengths are chosen so that HH-contacts are favored, e.g. $\epsilon_{HH} = -1$ and $\epsilon_{HP} = \epsilon_{PH} = \epsilon_{PP} = 0$. The central question is to determine the density of states (and to find the ground state with lowest energy) for a given sequence of monomers. An example of a conjectured ground state is given in Fig. 14.2 for a particular sequence of 85 monomers on the square lattice (the sequence is taken from [11]).

Our third example is the interacting self-avoiding walk (ISAW) model of (homo)-polymer collapse; it is obtained by setting $\epsilon_{i,j} = -1$ independent of the individual monomers. Here, one is interested in the critical behavior in the thermodynamic limit, *ie* the limit of large chain lengths. An example of a 26-step interacting self-avoiding walk with 7 interactions is shown in Fig. 14.3.

The partition function of n -step interacting self-avoiding walks can be written as

$$Z_n(\beta) = \sum_{\varphi} e^{-\beta E(\varphi)} = \sum_m C_{n,m} e^{\beta m}, \quad (14.1)$$

where $E(\varphi)$ is the energy of an n -step walk, φ . Note that the second sum is over the number m of interactions, and $C_{n,m}$ is the number of configurations of n -step self-avoiding walks with precisely m interactions.

While the motivation for simulations of the various models is different, the central problems turn out to be similar. For interacting self-avoiding walks, the collapse transition is in principle understood. One has a tri-critical phase

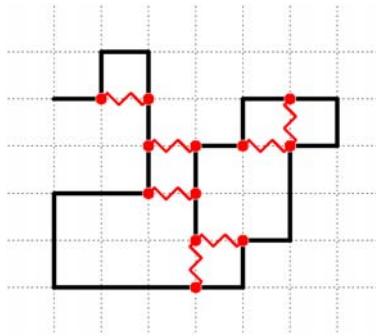


Fig. 14.3. A 26 step interacting self-avoiding walk on a square lattice with 7 interactions

transition with upper critical dimension $d_u = 3$, so that one can derive the critical behavior from mean-field theory for $d \geq 3$ [12], whereas for $d = 2$ one obtains results from conformal invariance [13]. However, even though this transition is in principle understood, there are surprising observations above the upper critical dimension [14]. Most importantly, there is no good understanding of the collapsed regime, which is also notoriously difficult to simulate.

Similarly, in the HP model one is interested in low-temperature problems, i.e. deep inside the collapsed phase. In particular, one wishes to understand the design problem, which deals with the mapping of sequences along the polymer chain to specific ground state structures. Again, the most important open question is in the collapsed regime.

It is therefore imperative, to find algorithms which work well at low temperatures. In the following section, we present just such an algorithm.

14.2 The Algorithm

This section describes our algorithm, as proposed in [2]. The basis of the algorithm is the Rosenbluth and Rosenbluth algorithm, a stochastic growth algorithm in which each configuration sampled is grown from scratch. The growth is kinetic, which is to say that each growth step is selected at random from all possible growth steps. Thus, if there are a possible ways to add a step then one selects one of them with probability $p = 1/a$.

For example, for a self-avoiding walk on the square lattice there may be between one and three possible ways of continuing, but it is also possible that there are no continuing steps, in which case we say that the walk is *trapped* (see Fig. 14.4).

As the number of possible continuations generally changes during the growth process, different configurations are generated with different probabilities and so one needs to reweight configurations to counter this. If one

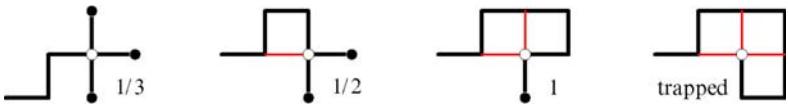


Fig. 14.4. For a self-avoiding walk on the square lattice, there can be between three and one ways of continuing, and the next step is chosen with equal probability from all possible continuations. In the right-most configuration, there is no way to continue, and the walk is trapped

views this algorithm as *approximate counting* then the required weights of the configurations serve as estimates of the total number of configurations.

To understand this point of view, imagine that we were to perform a *complete* enumeration of the configuration space. Doing this requires that at each growth step we would have to choose *all* the possible continuations and count them each with equal weight. If we now select *fewer* configurations, then we have to change the weight of these included configurations in order to correct for those that we have missed. Thus, if we choose one growth step out of a possible, then we must replace a configurations of equal weight by one “representative” configuration with a -fold weight. In this way, the weight of each grown configuration is a direct estimate of the total number of configurations.

Let the *atmosphere*, $a_n = a(\varphi_n)$, be the number of distinct ways in which a configuration φ_n of size n can be extended by one step. Then, the weight associated with a configuration of size n is the product of all the atmospheres a_k encountered during its growth, i.e.

$$W = \prod_{k=0}^{n-1} a_k . \quad (14.2)$$

After having started S growth chains, an estimator C_n^{est} for the total number of configurations C_n is given by the mean over all generated samples, $\varphi_n^{(i)}$, of size n with respective weights $W_n^{(i)}$, i.e.

$$C_n^{\text{est}} = \langle W \rangle_n = \frac{1}{S} \sum_i W_n^{(i)} . \quad (14.3)$$

Here, the mean is taken with respect to the total number of growth chains S , and *not* the number of configurations which actually reach size n . Configurations which get trapped before they reach size n appear in this sum with weight zero.

The Rosenbluth and Rosenbluth algorithm suffers from two problems. First, the weights can vary widely in magnitude, so that the mean may become dominated by very few samples with very large weight. Second, regularly occurring trapping events, i.e. generation of configurations with zero atmosphere can lead to exponential “attrition”, i.e. exponentially strong suppression of configurations of large sizes.

To overcome these problems, enrichment and pruning steps have been added to this algorithm, leading to the pruned and enriched Rosenbluth method (PERM) [3]. The basic idea is that one wishes to suppress large fluctuations in the weights $W_n^{(i)}$, as these should on average be equal to C_n .

If the weight of a configuration is too large one “enriches” by making copies of the configuration and reducing the weights by an appropriate factor. On the other hand, if the weight is too small, one throws away or “prunes” the configuration with a certain probability and otherwise continues growing the configuration with a weight increased by an appropriate factor. Note that neither S nor the expression (14.3) for the estimate, C_n^{est} , are changed by either enriching or pruning steps.

A simple but significant improvement of PERM was added in [15], where it was observed that it would be advantageous to force each of the copies of an enriched configuration to grow in *distinct* ways. This increases the diversity of the sample population and it is this version of PERM that we consider below.

We still need to specify enrichment and pruning criteria as well as the actual enrichment and pruning processes. While the idea of PERM itself is straightforward, there is now a lot of freedom in the precise choice of the pruning and the enrichment steps. The “art” of making PERM perform efficiently is based to a large extent on a suitable choice of these steps – this can be far from trivial! Distilling our own experience with PERM, we present here what we believe to be an efficient, and, most importantly, *parameter free* version.

In contrast to other expositions of PERM (e.g. [11]), we propose to prune and enrich constantly; this enables greater exploration of the configuration space. Define the *threshold ratio*, r , as the ratio of weight and estimated number of configurations, $r = W_n^{(i)} / C_n^{\text{est}}$. Ideally we want r to be close to 1 to keep weight fluctuations small. Hence if $r > 1$ the weight is too large and so we enrich. Similarly if $r < 1$ then the weight is too small and so we prune. Moreover, the details of the pruning and enrichment steps are chosen such that the new weights are as close as possible to C_n^{est} :

- $r > 1 \rightarrow \text{enrichment step}:$
make $c = \min(\lfloor r \rfloor, a_n)$ distinct copies of the configuration, each with weight $\frac{1}{c}W_n^{(i)}$ (as in nPERM [15]).
- $r < 1 \rightarrow \text{pruning step}:$
continue growing with probability r and weight $\frac{1}{r}W_n^{(i)} = C_n^{\text{est}}$ (i.e. prune with probability $1 - r$).

Note that we perform pruning and enrichment *after* the configuration has been included in the calculation of C_n^{est} . The new values are used during the *next* growth step.

Initially, the estimates C_n^{est} can of course be grossly wrong, as the algorithm knows nothing about the system it is simulating. However, even if

initially “wrong” estimates are used for pruning and enrichment the algorithm can be seen to converge to the true values in all applications we have considered. In a sense, it is self-tuning.

We also note here, that the number of samples generated for each size is roughly constant. Ideally, in order to effectively sample configuration space, the algorithm traces an unbiased random walk in configuration size. This means that PERM is, in some sense, already a flat histogram algorithm. We shall return to this central observation below.

It is now straight-forward to change PERM to a thermal ensemble with inverse temperature $\beta = 1/k_B T$ and energy E (defined by some property of the configuration, such as the number of contacts) by multiplying the weight with the appropriate Boltzmann factor $\exp(-\beta E)$, which leads to an estimate of the partition function, $Z_n(\beta)$, of the form

$$Z_n^{\text{est}}(\beta) = \langle W \exp(-\beta E) \rangle_n . \quad (14.4)$$

The pruning and enrichment procedures are changed accordingly, replacing W by $W \exp(-\beta E)$ and C_n^{est} by $Z_n^{\text{est}}(\beta)$. This gives threshold ratio $r = W_n^{(i)} \exp(-\beta E^{(i)}) / Z_n^{\text{est}}(\beta)$. This is the setting in which PERM is usually described.

Alternatively, however, it is also possible to consider microcanonical estimators for the total number $C_{n,m}$ of configurations of size n with energy m (i.e. the “density of states”). An appropriate estimator $C_{n,m}^{\text{est}}$ is then given by the mean over all generated samples $\varphi_{n,m}^{(i)}$ of size n and energy m with respective weights $W_{n,m}^{(i)}$, i.e.

$$C_{n,m}^{\text{est}} = \langle W \rangle_{n,m} = \frac{1}{S} \sum_i W_{n,m}^{(i)} . \quad (14.5)$$

Again, the mean is taken with respect to the total number of growth chains S , and *not* the number of configurations $S_{n,m}$ which actually reach a configuration of size n and energy m . The pruning and enrichment procedures are also changed accordingly, replacing C_n by $C_{n,m}$ and using $r = W_{n,m}^{(i)} / C_{n,m}^{\text{est}}$.

As observed above, the pruning and enrichment criterion for PERM leads to a flat histogram in length, i.e. a roughly constant number of samples being generated at each size n for PERM. In fact, one can motivate the given pruning and enrichment criteria by stipulating that one wishes to have a roughly constant number of samples, as this leads to the algorithm performing an unbiased random walk in the configuration size. Similarly, in the microcanonical version described above, the algorithm performs an unbiased random walk in both size and energy of the configurations, and we obtain a roughly constant number of samples at each size n and energy m .

It is because of the fact that the number of samples is roughly constant in each histogram entry, that this algorithm can be seen as a “flat histogram” algorithm, which we consequently call flat histogram PERM, or flatPERM. In

indsight in becomes clear that PERM itself can be seen as a flat histogram algorithm, at it creates a roughly flat histogram in size n . Recognizing this led us to the formulation of this algorithm in the first place.

We have seen that by casting PERM as an approximate counting method, the generalization from PERM to flat histogram PERM is straight-forward and (nearly) trivial. One can now add various refinements to this method if needed. For examples we refer the reader to [2]. We close this section with a summary of the central steps to convert simple PERM to flatPERM by comparing the respective estimators and threshold ratios, r :

1. *athermal PERM*: estimate the number of configurations C_n
 - $C_n^{\text{est}} = \langle W \rangle_n$
 - $r = W_n^{(i)} / C_n^{\text{est}}$
2. *thermal PERM*: estimate the partition function $Z_n(\beta)$
 - $Z_n^{\text{est}}(\beta) = \langle W \exp(-\beta E) \rangle_n$
 - $r = W_n^{(i)} \exp(-\beta E_m^{(i)}) / Z_n^{\text{est}}(\beta)$
3. *flat histogram PERM*: estimate the density of states $C_{n,m}$
 - $C_{n,m}^{\text{est}} = \langle W \rangle_{n,m}$
 - $r = W_{n,m}^{(i)} / C_{n,m}^{\text{est}}$

One can similarly generalize the above to several microcanonical parameters, m_1, m_2, \dots , to produce estimates of $C_{n,m_1, m_2, \dots}$.

Once the simulations have been performed the average of an observable, Q , defined on the set of configurations can be obtained from weighted sums:

$$Q_{n,m}^{\text{est}} = \frac{\langle QW \rangle_{n,m}}{\langle W \rangle_{n,m}} = \frac{\sum_i Q_{n,m}^{(i)} W_{n,m}^{(i)}}{\sum_i W_{n,m}^{(i)}}. \quad (14.6)$$

These can then be used for subsequent evaluations. For instance, the expectation value of Q in the canonical ensemble at a given temperature β can now be computed via

$$Q_n^{\text{est}}(\beta) = \frac{\sum_m Q_{n,m}^{\text{est}} C_{n,m}^{\text{est}} \exp(-\beta E_m)}{\sum_m C_{n,m}^{\text{est}} \exp(-\beta E_m)}, \quad (14.7)$$

i.e. only a single simulation is required to compute expectations at *any* temperature.

For many problems we are interested in their behavior at low temperatures where averages of observables are dominated by configurations with high energy. Such configurations are normally very difficult to obtain in simulations. The flatPERM algorithm is able to effectively sample such configurations because it obtains a roughly constant number of samples at all sizes and energies (due to the constant pruning and enrichment). This means that it is possible to study models even at very low temperatures. Examples of this are given in the next section.

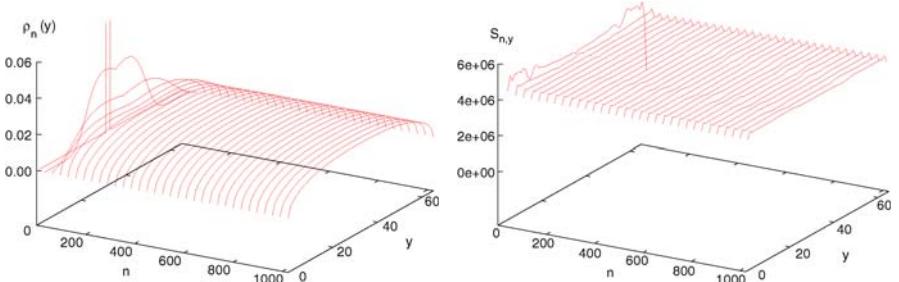


Fig. 14.5. Probability density $\rho_n(y)$ (left) and number of generated samples $S_{n,y}$ (right) versus length n and vertical endpoint coordinate y for self-avoiding walk on a strip of width 64 on the square lattice

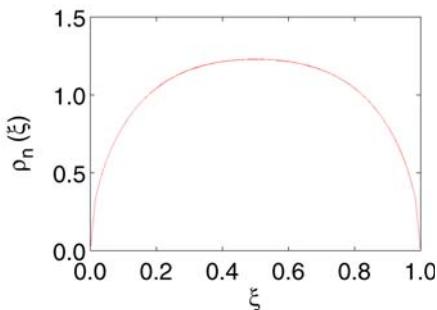


Fig. 14.6. Endpoint densities (scaled to the interval [0, 1]) for lengths 512, 768, and 1024. The different curves are indistinguishable

14.3 Simulations

A good way of showing how flatPERM works is to simulate two-dimensional polymers in a strip. This kind of simulation has previously been performed with PERM using Markovian anticipation techniques [9] which are quite complicated. With flatPERM one simply chooses the vertical position of the endpoint of the walk in the strip as an “energy” for the algorithm to flatten against. We have found that this produces very good results.

Figure 14.5 shows the results of our simulations of 1024-step self-avoiding walks in a strip of width 64. The left-hand figure is the probability density $\rho_n(y)$ of the endpoint coordinate y shown as a function of walk length n . The right-hand figure shows the actual number of samples generated for each length n and end point position y . One sees that the histogram of samples is indeed nearly completely flat. One can now extract several quantities from such simulations (see [9]), but we restrict ourselves here to the scaled endpoint density shown in Fig. 14.6.

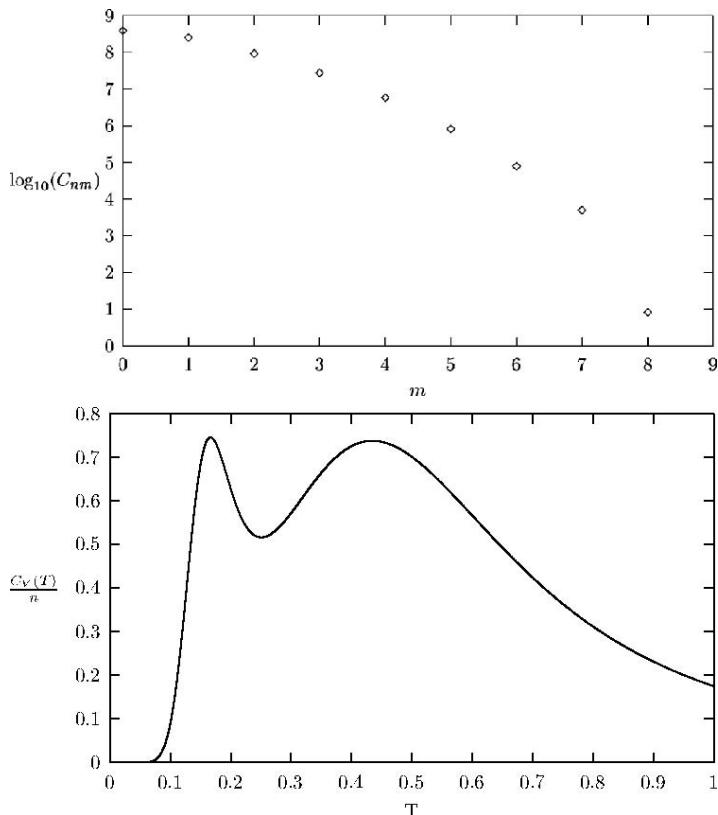


Fig. 14.7. Sequence I (14 Monomers, HPHPHHPHPHHPPH, $d = 3$): density of states versus energy (*above*) and specific heat C_V/n versus temperature T (*below*)

Next we show results from simulations of the HP-model. Here, we have obtained the whole density of states for small model proteins with fixed sequences. The first sequence considered (taken from [7]) is small enough to enable comparison with exact enumeration data. It has moreover been designed to possess a unique ground state (up to lattice symmetries).

Figure 14.7 shows our results. We find (near) perfect agreement with exact enumeration even though the density of states varies over a range of eight orders of magnitude! The derived specific heat data clearly shows a collapse transition around $T = 0.45$ and a sharper transition into the ground state around $T = 0.15$.

The next sequence (taken from [11]) is the one for which Fig. 14.2 shows a state with the lowest found energy. Figure 14.8 shows our results for the density of states and specific heat. We find the same lowest energy $E = 53$ as [11] (though this is not proof of it being the ground state). The density of states varies now over a range of 30 orders of magnitude! The derived specific

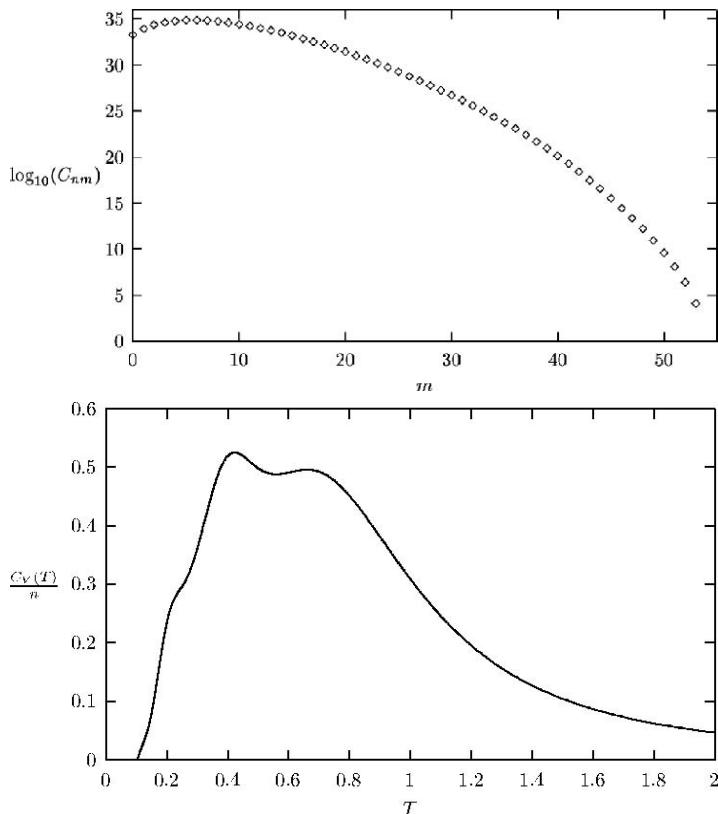


Fig. 14.8. Sequence II (85 Monomers, $d = 2$): density of states versus energy (above) and specific heat C_V/n versus temperature T (below)

heat data clearly shows a much more complicated structure than the previous example.

For several other sequences taken from the literature we have confirmed previous density of states calculations and obtained identical ground state energies. The sequences we considered had $n = 58$ steps (3 dimensions, $E_{\min} = -44$) and $n = 85$ steps (2 dimensions, $E_{\min} = -53$) from [11], and $n = 80$ steps (3 dimensions, $E_{\min} = -98$) from [16]. We studied also a particularly difficult sequence with $n = 88$ steps (3 dimensions, $E_{\min} = -72$) from [17], but the lowest energy we obtained was $E = -69$. While we have not been able to obtain the ground state, neither has any other PERM implementation (see [11]).

We now turn to the simulation of interacting self-avoiding walks (ISAW) on the square and simple cubic lattices. In both cases have we simulated walks up to length 1024. Here, we encounter a small additional difficulty; when PERM is initially started it is effectively blind and may produce poor

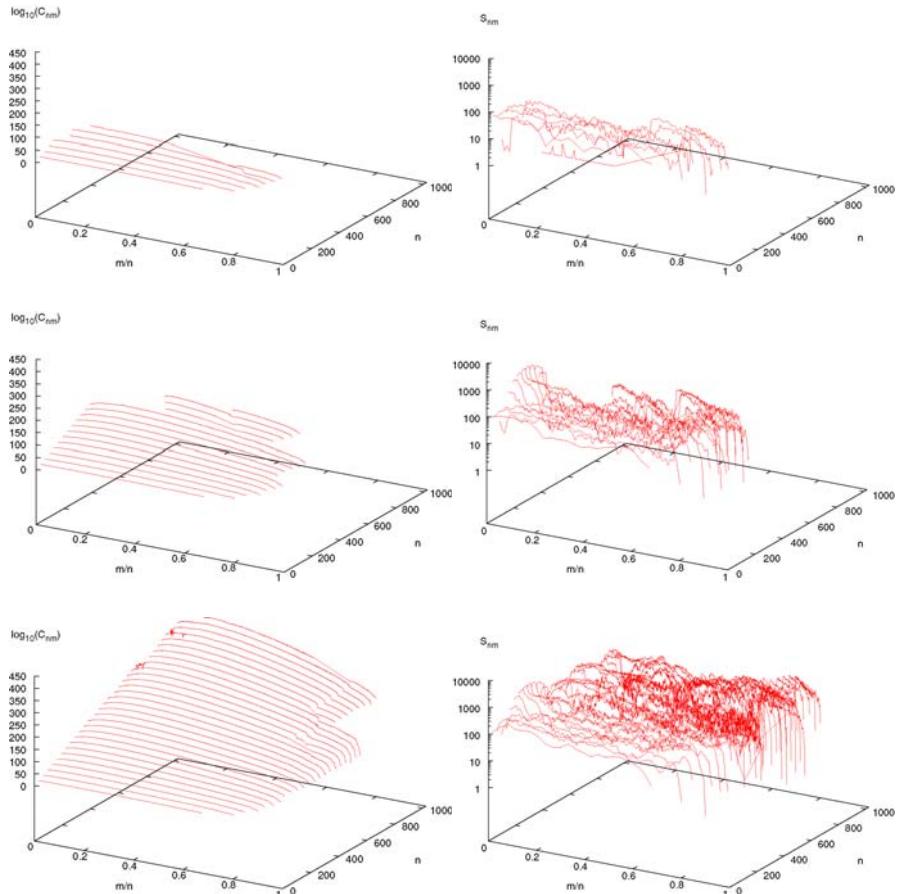


Fig. 14.9. Number of configurations $C_{n,m}$ (left) and number of generated samples (right) versus internal energy m/n and length n for ISAW on the square lattice, for various total sample sizes: 10^6 (top), 10^7 (middle), and 10^8 (bottom)

estimates of $C_{n,m}$ and this may in turn lead to overflow problems. It is therefore necessary to stabilize the algorithm by delaying the growth of large configurations. For this, it suffices to restrict the size of the walks by only allowing them to grow to size n once $t = cn$ tours (the number of times the algorithm returns to an object of zero size) has been reached. We found a value of $c \approx 0.1$ sufficient.

Figure 14.9 shows the equilibration of the algorithm due to the delay. Snapshots are taken after 10^6 , 10^7 , and 10^8 generated samples. While the sample histogram looks relatively rough (even on a logarithmic scale) the density of states is already rather well behaved. In the plots one clearly sees

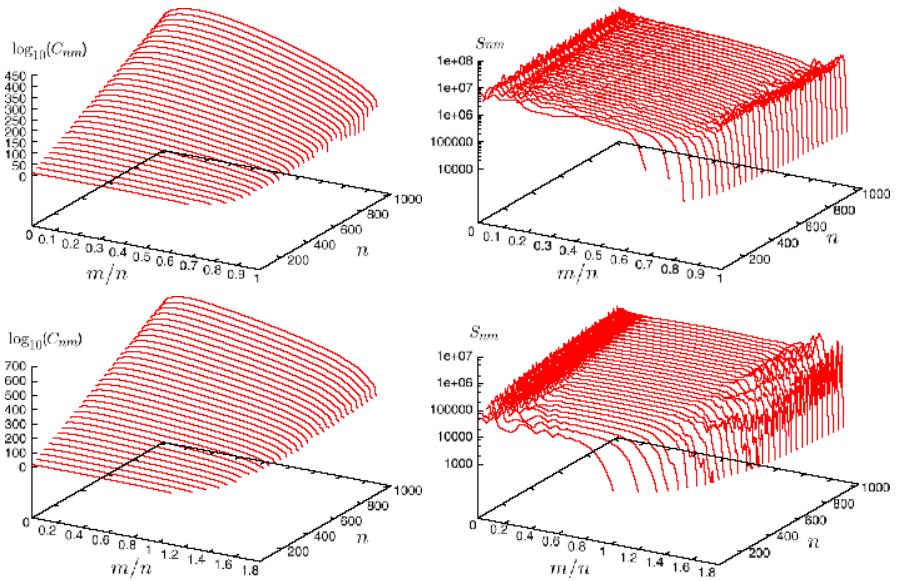


Fig. 14.10. Number of configurations $C_{n,m}$ (left) and number of generated samples (right) versus internal energy m/n and length n for ISAW on the square lattice (above) and simple cubic lattice (below)

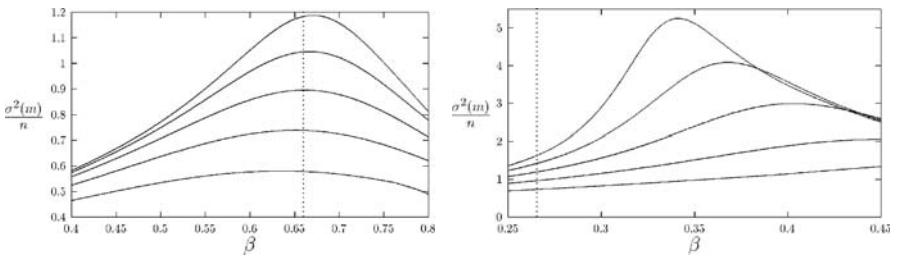


Fig. 14.11. Normalized fluctuations $\sigma^2(m)/n$ versus inverse temperature $\beta = 1/k_B T$ for ISAW on the square lattice (above) and the simple cubic lattice (below) at lengths 64, 128, 256, 512, and 1024. The curves for larger lengths are more highly peaked. The vertical lines denote the expected transition temperature at infinite length

the effect of large correlated tours in which large number of enrichments produce many samples with the same initial walk segment.

The final result of our simulations for interacting self-avoiding walks in two and three dimensions is shown in Fig. 14.10. It clearly shows the strength of flatPERM: with one single simulation can one obtain a density of states which ranges over more than 300 orders of magnitude!

From these data one can now, for example, compute the specific heat curves $C_n = k_B(\beta\epsilon)^2\sigma^2(m)/n$. The results for both systems are shown in Fig. 14.11. We see that the data is well behaved well into the collapsed low-temperature regime.

14.4 Conclusion and Outlook

We have reviewed stochastic growth algorithms for polymers. Describing the Rosenbluth and Rosenbluth method as an approximate counting method has enabled us to present a straight-forward extension of simple PERM to flat histogram PERM. Using this algorithm one can obtain the complete density of states (even over several hundred orders of magnitude) from one single simulation.

We demonstrated the strength of the algorithm by simulating self-avoiding walks in a strip, the HP-model of proteins, and interacting self-avoiding walks in two and three dimensions as a model of polymer collapse.

Further applications are in preparation, e.g. simulations of branched polymers, and simulations of higher-dimensional densities of states.

References

1. G.W. King, in: Monte Carlo Method, Vol. 12 (Applied Mathematics Series) (National Bureau of Standards, 1951)
2. T. Prellberg, J. Krawczyk: Phys. Rev. Lett., in print
3. P. Grassberger: Phys. Rev E **56**, 3682 (1997)
4. G.M. Torrie, J.P. Valleau: J. Comput. Phys. **23**, 187 (1977)
5. M.N. Rosenbluth, A.W. Rosenbluth: J. Chem. Phys. **23**, 356 (1955)
6. F. Wang, D.P. Landau: Phys. Rev. Lett. **86**, 2050 (2001)
7. M. Bachmann, W. Janke: Phys. Rev. Lett. **91**, 208105 (2003)
8. B.A. Berg, T. Neuhaus: Phys. Lett. B **267**, 249 (1991)
9. H.-P. Hsu, P. Grassberger: Eur. Phys. J. B **36**, 209 (2003)
10. K.A. Dill: Biochemistry **24**, 1501 (1985)
11. H.-P. Hsu, V. Mehra, W. Nadler, P. Grassberger: J. Chem. Phys. **118**, 444 (2003)
12. I.D. Lawrie, S. Sarbach, in: Phase Transitions and Critical Phenomena, ed. by C. Domb, J.L. Lebowitz, Vol. 9 (Academic Press, London 1984)
13. J.L. Cardy, in: Phase Transitions and Critical Phenomena, ed. by C. Domb, J.L. Lebowitz, Vol. 11 (Academic Press, New York 1987)
14. T. Prellberg, A.L. Owczarek: Phys. Rev. E **62**, 3780 (2000)
15. H.-P. Hsu, V. Mehra, W. Nadler, P. Grassberger: Phys. Rev. E **68**, 21113 (2003)
16. H. Frauenkron, U. Bastolla, E. Gerstner, P. Grassberger, W. Nadler: Phys. Rev. Lett. **80**, 3149 (1998)
17. T.C. Beutler, K.A. Dill: Protein Sci. **5**, 2037 (1996)

15 Convergence of the Wang–Landau Algorithm and Statistical Error

C. Zhou and R.N. Bhatt

Department of Electrical Engineering, Princeton University
Princeton, NJ 08544, USA

Abstract. We present a mathematical analysis of the Wang–Landau algorithm, show the density of states is well represented by the average histogram, and identify sources of error in simulation. In particular, we find that after an initial stage of data accumulation, the histogram increases uniformly with small fluctuation, and obtain the dependence of the statistical error on the modification factor.

The Wang–Landau (WL) algorithm [1] is a new type of Monte Carlo (MC) simulation method, which has received much attention lately. It is simple and versatile, and more importantly, circumvents many of the difficulties faced by conventional MC methods, such as long relaxation times due to critical slowing down near phase transitions, or frustration, or complex energy landscape in complex systems such as proteins. One especially attractive feature of the WL algorithm is that one performs a *single* simulation to obtain thermodynamic information at *all* temperatures. To date, the WL algorithm has been used to study classical discrete spin models [1,2], models with continuous degrees of freedom [3], quantum mechanical models [4] as well as glassy systems [5].

Though there has been substantial research in augmenting the original algorithm [6], to improve performance, correct errors, and expand the sphere of applicability, most has relied on innovative, empirical tricks, justified numerically for each specific implementation. Many fundamental issues remain unanswered. The WL algorithm, in its quest for the density of states, uses a parameter called the modification factor, and aims for a flat histogram. This raises several questions, e.g. (a) what is the relationship between flatness of the histogram and the accuracy of the method; (b) how does the modification factor affect the error; and (c) how does the simulation actually find the density of states. These are among the issues we address in our analytical study of the WL algorithm.

The WL algorithm aims to calculate the density of states $\varrho(E)$ of the system. It performs a random walk in state space by biasing the walk with a function $g(E)$, which is a function of energy. Using an acceptance ratio $\min\{1, g(E_f)/g(E_i)\}$, where E_i and E_f refer to energies before and after this transition, the algorithm produces a final histogram $h(E)$. $h(E)$ is related to the equilibrium distribution of the free random walk $\varrho(E)$ by $\varrho(E)g(E) =$

$h(E)$, provided that both sides of the identity are normalized. This identity is essentially a result of detailed balance. The aim of the WL algorithm is to learn $g(E) = 1/\varrho(E)$ from the simulation so that the histogram $h(E)$ is expected to be flat. Here $\varrho(E)$ is understood to be the density of states of the physical model, which is a priori unknown in the simulation.

We begin our analysis by clarifying the relevant parameters. Suppose the phase space of our physical model is divided into N classes with density (number of configurations) ϱ_i for class i ($i = 1, 2 \dots N$). These class labels could refer to energy, magnetization, or other macroscopic variables. Each configuration in the phase space uniquely belongs to one class. We use γ for the modification factor f defined in [1]. The histogram h_i with $1 \leq i \leq N$ is defined as the number of visits of each class in the simulation. A free random walk in the phase space generates a histogram proportional to the density ϱ_i . In the original implementation [1], one record is inserted into the histogram every sequential sweep of the system, hence the histogram is a record of the outcomes of a series of trial flips. Therefore, in general, we have a second tunable parameter of the algorithm, namely, the separation S between successive records in the histogram, defined as the number of trial flips (random steps) that precedes each increment of the histogram. S successive steps of random walk should be regarded as a single transition $i \rightarrow f$. In what follows, we assume that S is large enough to decorrelate the initial class i and final class f .

In general, we can start the algorithm with a guess of ϱ_i , which we call θ_i . Initially, the simulation uses an acceptance ratio $\min\{1, \theta_i/\theta_f\}$ to bias the random walk. The resultant distribution $\{\varrho_i/\theta_i\}$ is calculated in the following simulation. The first record k_1 in the histogram is generated with a probability distribution $Z_1^{-1} \varrho_{k_1}/\theta_{k_1}$, where $Z_1 = \sum \varrho_k/\theta_k$ is the normalization constant. At this point, the WL algorithm reduces the acceptance ratio of state k_1 by a factor γ . Next the second record is generated; its conditional probability is $P(k_2|k_1) = (\varrho_{k_2}/\theta_{k_2})\gamma^{-\delta_{k_1,k_2}}$, where the Kronecker δ ensures that if k_2 happens to be the same as k_1 then its probability is reduced by a factor of γ . Because of the assumption of large separation S between k_2 and k_1 , these two entries in the histogram are not correlated. In general, the conditional probability of n -th record is given by 5

$$P(k_n|k_{n-1}, \dots, k_1) = Z_n^{-1} \left(\frac{\varrho_{k_n}}{\theta_{k_n}} \right) \gamma^{-\sum_{i=1}^{n-1} \delta_{k_n, k_i}}. \quad (15.1)$$

The probability of a sequence $\{k_1, k_2, \dots, k_M\}$ with M elements, where $k_i, i = 1, 2, \dots, M$ take values from $\{1, 2, \dots, N\}$ is a product of conditional probabilities given above: $P(\{k_1, \dots, k_M\}) = \gamma^{-\sum_{i>j} \delta_{k_i, k_j}} \prod_{i=1}^M Z_i^{-1} (\varrho_{k_i}/\theta_{k_i})$. We rewrite this probability in terms of the histogram $\{h_1, h_2, \dots, h_N\}$:

$$P_M(h_1, \dots, h_N) = \frac{M!}{\prod Z_i} \prod_{i=1}^N \frac{(\varrho_i/\theta_i)^{h_i}}{h_i!} \gamma^{-h_i(h_i-1)/2}. \quad (15.2)$$

The exponent of γ is the result of counting pairs (k_i, k_j) with $i > j$ in the sequence $\{k_1, k_2, \dots, k_M\}$ whose elements k_i and k_j are equal. However, $\prod Z_i$ still depends on the whole sequence $\{k_1, k_2, \dots, k_M\}$. To proceed analytically, we construct an artificial ensemble that allows $M = \sum h_i$ to fluctuate around its expectation value, and keep the fluctuation small compared to $\sqrt{\langle M \rangle}$. In (15.2), the quadratic term $h_i(h_i - 1)/2$ is responsible for guiding the simulation, so we keep it in our construction. A simple artificial ensemble is written down in analogy to Poisson distribution, in which h_i 's are expected to be independent random variables, and a “chemical potential” μ controls the average length of the sequence:

$$P(h_1, \dots, h_N) = \frac{1}{Z} \prod_{i=1}^N \frac{e^{\mu h_i}}{h_i!} \left(\frac{\varrho_i}{\theta_i} \right)^{h_i} \gamma^{-h_i(h_i-1)/2}. \quad (15.3)$$

We will show the standard deviation of M is asymptotically a constant so that this ensemble captures the major properties of real ensemble of the computer simulations. Let $x = \frac{1}{2} \ln \gamma$ and $y_i = (\frac{1}{2} + \ln \varrho_i - \ln \theta_i + \mu)$. The partition function Z becomes $Z = \prod_{i=1}^N Z_i$, where $Z_i = e^{y_i^2/4x} \mathcal{Z}(x, y_i)$, and

$$\mathcal{Z}(x, y) = \sum_{i=0}^{\infty} \frac{1}{n!} \exp \left[-x \left(n - \frac{y}{2x} \right)^2 \right]. \quad (15.4)$$

we calculate $\mathcal{Z}(x, y)$ by a Fourier transform to the variable $t = y/2x$ and summing up the series. The inverse Fourier transformation is evaluated conveniently by asymptotic series expansion at the saddle point, the result is

$$\mathcal{Z}(x, y) = (1 + 2xe^\eta)^{-1/2} \exp \left(\frac{\eta^2}{4x} + e^\eta - \frac{\eta y}{2x} \right) e^{V(d/dh)} e^{h^2/2A} |_{h=0}, \quad (15.5)$$

where η is the solution of the saddle point condition: $y = \eta + 2xe^\eta$, and $A = 1/2x + e^\eta$. The asymptotic series given by $V(\epsilon) = e^\eta \sum_{n=3}^{\infty} (-i\epsilon)^n / n!$ is written formally. $\ln Z_i$ is conveniently expressed as a function of x and y_i

$$\ln Z_i(x, y_i) = xe^{2\eta_i} + e^{\eta_i} - \ln \sqrt{1 + 2xe^{\eta_i}} + O(e^{-\eta_i}). \quad (15.6)$$

Here η_i is an implicit function of x and y_i . The $O(e^{-\eta_i})$ correction comes from $\ln \{\exp[V(d/dh)] \exp(h^2/2A_i)\} |_{h=0}$. In the relevant regime $2xe^{\eta_i} \gg 1$ (justified below), the first order correction $e^{\eta_i} [(8A_i^2)^{-1} - 5e^{\eta_i}(24A_i^3)^{-1}]$ is of order $e^{-\eta_i}$. Figure 15.1 compares the results of the approximation, truncated at the lowest and first order, with the exact numerical value. The convergence is good at large values of $t = y/2x$.

The expectation values are calculated from: $\langle h_i^2 \rangle = -\frac{\partial}{\partial x} \ln Z_i$ and $\langle h_i \rangle = \frac{\partial}{\partial y_i} \ln Z_i$. After some algebra, we get:

$$\langle h_i^2 \rangle = e^{2\eta_i} + \frac{e^{\eta_i}}{(1 + 2xe^{\eta_i})^2} + O(e^{-2\eta_i}), \quad (15.7)$$

$$\langle h_i \rangle = e^{\eta_i} - \frac{xe^{\eta_i}}{(1 + 2xe^{\eta_i})^2} + O(e^{-2\eta_i}). \quad (15.8)$$

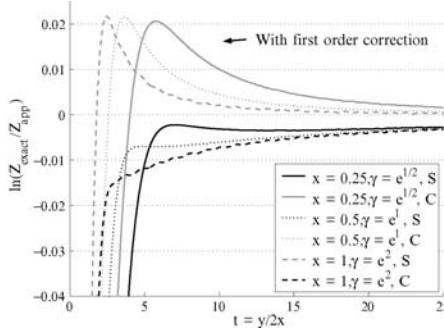


Fig. 15.1. The error of saddle point approximation, $\ln(\mathcal{Z}_{\text{exact}}/\mathcal{Z}_{\text{app}})$. $\mathcal{Z}_{\text{exact}}$ is calculated by adding the series in (15.4) up to machine accuracy. The lower three curves (S) are calculated by using the saddle point approximation (15.5) only. The upper three curves (C) take the first order correction into account. As can be seen, the latter converge faster at large t

The asymptotic behavior of $\langle h_i \rangle$ is clearly $\langle h_i \rangle = e^{\eta_i}$. We can control $\langle h_i \rangle$ in the simulation by allowing it to run for a certain time so that $2 \times e^{\eta_i} \gg 1$ is satisfied to ensure the self-consistency. Combining (15.8) and saddle point condition, we have $y_i(\langle h_i \rangle) = \ln \langle h_i \rangle + 2 \times \langle h_i \rangle$ asymptotically, which implies: $(\varrho_i \theta_j)/(\varrho_j \theta_i) = (\langle h_i \rangle / \langle h_j \rangle) \gamma^{\langle h_i \rangle - \langle h_j \rangle}$. Apart from a prefactor $\langle h_i \rangle / \langle h_j \rangle$, the ratio of density of states is reflected by the difference of histogram. This prefactor is a negligible weak correction, especially when the simulation runs for a long time, resulting in a “flat” histogram ($h_i/h_j \rightarrow 1$), although the difference $\langle h_i \rangle - \langle h_j \rangle$ is maintained.

The standard deviation $\sigma^2(h_i) = \langle h_i^2 \rangle - \langle h_i \rangle^2$ of the histograms is calculated directly from (15.7) and (15.8). Keeping only the $O(1)$ term, we find $\sigma(h_i) = 1/\sqrt{\ln \gamma}$ when $2xe^{\eta_i} \gg 1$. Because $\sigma(h_i)$ is finite, the length of the sequence $M = \sum_i h_i$ also has a *finite standard deviation independent of $\langle M \rangle$* . This property confirms the resemblance of our artificial ensemble with real computer simulations. The standard deviation of $\ln \varrho_i$ given by a single histogram is approximately $\sigma(\ln \varrho_i) = \sigma(h_i) \ln \gamma = \sqrt{\ln \gamma}$. By using $\gamma \rightarrow 1$, the resolution of the algorithm is improved: $\delta \varrho_i / \varrho_i \rightarrow 0$. However, the improvement in the asymptotic regime is somewhat slower than one could have hoped for, being dependent on $\sqrt{\ln \gamma}$. On the other hand, the relative insensitivity of the error to γ suggests that relatively large γ could be employed at early stages to speed up the convergence.

The histogram h_i enters the asymptotic regime of fluctuating linear growth $h_i(M) = \log_\gamma(\varrho_i/\theta_i) + M/N + r_i(M)$ when $2xe^{\eta_i} \gg 1$, where $r_i(M)$ is a random number which has zero mean and standard deviation $1/\sqrt{\ln \gamma}$. This behavior is illustrated in Fig. 15.2, which shows three snapshots of the histograms calculating the density of states of a two-dimensional Ising model. We have used $\gamma = e^4$ in this simulation to reveal the fluctuation of the his-

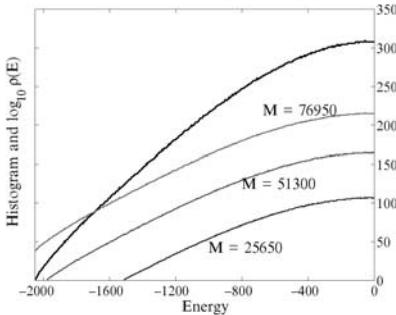


Fig. 15.2. Snapshots of the histogram of a simulation of two-dimensional Ising model on a 32 by 32 lattice with $\gamma = e^4$. Three parallel curves are histograms of three lengths labeled in the figure. The fourth curve is $\varrho'(E)$ calculated from the last histogram ($M = 76950$), which overlaps with the exact $\varrho(E)$ within the accuracy of the width of the line in this figure

togram. The simulation visits energies with high density of states first, then extends the histogram to the whole spectrum. Once the whole spectrum is visited, the histogram grows uniformly with small fluctuation.

Two important consequences follow. Firstly, a flat histogram is *not* required for convergence; the histogram is ready for evaluating ϱ_i in the asymptotic regime when the flat growth is seen. Secondly, the statistical error can be reduced by averaging over multiple results obtained with the same γ . Our strategy for a single iteration simulation is to run until a minimum number of visits have been accumulated ($h_i > 1/\ln \gamma$ for $i = 1, \dots, N$), followed by measurements separated by a short simulation which decorrelates $r_i(M)$. Usually $1/\sqrt{\ln \gamma}$ visits on each category is enough. With K measurements, the statistical error in $\ln \varrho_i$ is reduced to $c_i \sqrt{(\ln \gamma)/K}$, where c_i is an $O(1)$ prefactor that appears in a real simulation.

To summarize, we have analyzed the convergence of the WL algorithm based on an artificial ensemble which mimics the simulations very closely, and find: (i) the density of states is encoded in the average histogram; and (ii) the statistical error caused by fluctuations in the histogram is proportional to $\sqrt{\ln \gamma}$, which can be reduced by averaging over multiple measurements. Though the artificial ensemble (15.3) contains some implicit hypotheses, we have recently formulated a direct proof of flat growth of the histogram, which leads to the same conclusions, and will be presented elsewhere [7].

Acknowledgement

This research was supported by NSF DMR-0213706. We acknowledge discussions with D.A. Huse, M.S. Shell, and D. Guo.

References

1. F. Wang, D.P. Landau: Phys. Rev. Lett. **86**, 2050 (2001); Comput. Phys. Commun. **147**, 674 (2002)
2. C. Yamaguchi, Y. Okabe: J. Phys. A **34**, 8781 (2001)
3. M.S. Shell, P.G. Debenedetti, A.Z. Panagiotopoulos: Phys. Rev. E **66**, 056703 (2002)
4. M. Troyer, S. Wessel, F. Alet: Phys. Rev. Lett. **90**, 120201 (2003)
5. Y. Okabe, Y. Tomita, C. Yamaguchi: Comput. Phys. Commun. **146**, 63 (2002)
6. C. Yamaguchi, N. Kawashima: Phys. Rev. E, **65**, 056710 (2002); B.J. Schulz, K. Binder, M. Müller: Int. J. Mod. Phys. C **13**, 447 (2002); B.J. Schulz, K. Binder, M. Müller, D.P. Landau: cond-mat/0302123; P. Dayal, S. Trebst, S. Wessel, D. Wuertz, M. Troyer, S. Sabhapandit, S.N. Coppersmith: cond-mat/0306108
7. C. Zhou, R.N. Bhatt: submitted for publication

16 Wang–Landau Sampling with Cluster Updates

M. Körner and M. Troyer

Theoretische Physik, ETH Zürich, CH-8093 Zürich, Switzerland

Abstract. We present a new algorithm, combining cluster updates with Wang–Landau sampling and compare the performance to local updates.

16.1 Introduction

With the development of the Wang–Landau algorithm [1], a scalable and easy implementation of the multicanonical ensemble [2] has been achieved and has seen widespread use. A recent investigation [3] has however shown, that the multicanonical ensemble does not show the N^2 scaling of the autocorrelation times expected for an unbiased Markovian random walk over N energy levels, but performs slower due to the slowing down of the random walk around the critical energy. Two approaches can be taken to remedy this: changing the ensemble [4], or using cluster updates as proposed for the multibondic representation of the Ising model [5], or for quantum systems [6]. Here we will show that a combination of cluster updates with Wang–Landau sampling, as used for quantum systems [6] can also be used for classical systems and we will arrive at a generalization of the multibondic ensemble [5, 7].

We start the treatment as general as possible. We deal with the partition function of a classical model on a lattice with a Hamiltonian

$$H[C] = - \sum_{ij} h_{ij}[C_i, C_j]. \quad (16.1)$$

We write C for the configuration of the system and C_i for the configuration at site i . The energies of the Hamiltonian are shifted so that $h_{ij}[C_i, C_j] \geq 0$ for all configurations. The partition function of a classical system is usually expressed through the density of states $\rho(E)$ as

$$Z = \sum_C e^{-\beta H[C]} = \sum_E \rho(E) e^{-\beta E}. \quad (16.2)$$

We can also write the partition function in (16.2) as a Taylor series in the inverse temperature β

$$Z = \sum_{n \geq 0} g(n) \beta^n, \quad (16.3)$$

with $g(n) = \sum_C (-H[C])^n / n!$, which can be sampled using a flat histogram method in the order n with local configuration and order updates. In a practical computation, the sum over n will be truncated at some cutoff Λ restricting the largest inverse temperature that can be reached to $\beta_{\max} \sim \Lambda/N$ [6]. To implement cluster updates for the configurations we use

$$e^{-\beta H[C]} = \prod_{ij} e^{\beta h_{ij}[C_i, C_j]}. \quad (16.4)$$

Expanding the exponentials we obtain

$$Z = \sum_{\{n_{ij} \geq 0\}} \sum_C \prod_{ij} \frac{h_{ij}[C_i, C_j]^{n_{ij}}}{n_{ij}!} \beta^{\sum n_{ij}} \quad (16.5)$$

$$= \sum_{n \geq 0} \left[\sum_{\{n_{ij} | \sum n_{ij} = n\}} \sum_C W(C, \{n_{ij}\}) \right] \beta^n. \quad (16.6)$$

We have extended our state space from the configuration C to a configuration–order pair $(C, \{n_{ij}\})$ with a local n_{ij} for each bond. The weight of a state $(C, \{n_{ij}\})$ is given by

$$W(C, \{n_{ij}\}) = \prod_{ij} \frac{h_{ij}[C_i, C_j]^{n_{ij}}}{n_{ij}!} = \prod_{ij} w_{ij}(C_i, C_j, n_{ij}). \quad (16.7)$$

This representation can be viewed as the classical limit of the stochastic series expansion method for quantum systems [8]. Two types of updates are needed: updates in the $\{n_{ij}\}$ and updates in the configuration C . Applying a flat histogram method in the total order n , an order update $\{n_{ij}\} \rightarrow \{n'_{ij}\}$ (usually $n'_{ij} = n_{ij} \pm 1$ for one ij is a good choice) will be accepted with probability

$$P_{\text{accept}} = \min \left\{ 1, \prod_{ij} \left(h_{ij}[C_i, C_j]^{n'_{ij} - n_{ij}} \frac{n_{ij}!}{n'_{ij}!} \right) \frac{\hat{g}(\sum n_{ij})}{\hat{g}(\sum n'_{ij})} \right\}, \quad (16.8)$$

where $\hat{g}(n)$ is an estimate of $g(n)$. In a local update scheme, a configuration update $C \rightarrow C'$ (usually a change in the configuration of a single site $C_i \rightarrow C'_i$) at fixed $\{n_{ij}\}$ will be accepted with probability

$$P_{\text{accept}} = \min \left\{ 1, \prod_{ij} \left(\frac{h_{ij}[C'_i, C'_j]^{n_{ij}}}{h_{ij}[C_i, C_j]} \right) \right\}. \quad (16.9)$$

16.2 Cluster Updates

Because the weight of a state factors into bond weights, cluster updates can now be performed on the configurations using a cluster update scheme such

as the Swendsen–Wang [9] or Wolff algorithm [10]. The only difference to the canonical cluster algorithms is in the probability to insert a Swendsen–Wang bond connecting two sites. Denoting by w_{ij} the weight of a bond before a flip and by w'_{ij} the weight after a flip of one of the sites, the connection probability in the cluster update is given by

$$P_{\text{connect}} = 1 - \min \left\{ 1, \frac{w'_{ij}}{w_{ij}} \right\}, \quad (16.10)$$

which in the case of a canonical ensemble simulation of the Ising ferromagnet reduces to the well known $1 - \exp(-2\beta J)$ for parallel spins. In the series representation the connection probability becomes

$$P_{\text{connect}} = 1 - \min \left\{ 1, \left(\frac{h_{ij}[C'_i, C'_j]}{h_{ij}[C_i, C_j]} \right)^{n_{ij}} \right\}. \quad (16.11)$$

A direct consequence is that Swendsen–Wang bonds forming a cluster can be inserted only onto bonds with $n_{ij} > 0$. In contrast to the Wang–Landau sampling of the multibondic representation of the Ising model as introduced by Yamaguchi and Kawashima [5], we cannot update the n_{ij} globally. On the other hand, our method can be used to implement a cluster update for any $O(n)$ model using the Wolff embedding trick [10].

16.3 Performance

To compare the performance of single spin updates (16.9) and cluster updates (16.11) in the sampling of the partition function, we compare the tunneling times for both algorithms from a random state at $n = 0$ to a state with an energy within 5% of the ground state energy for a one-dimensional XY model with open boundary conditions. Because the n_{ij} updates (16.8) are the same for both algorithms, we measure the number of spin updates needed to reach the low energy state (τ_{Spin}), while performing many n_{ij} updates so that the n_{ij} are equilibrated for a particular spin configuration C (in the simulations for Fig. 16.1 only 0.1% of the updates were spin updates). The comparison of both updates clearly shows the improved scaling of the cluster updates. The fact that the cluster updates seem to follow the N^2 scaling of an unbiased random walk may be due to the fact that the final energy is still too high, but even for these energies the single spin updates have an extra $\log(L)$ factor.

16.4 Conclusions

We have demonstrated that cluster updates can be used in conjunction with flat histogram methods for classical systems by using a series representation of

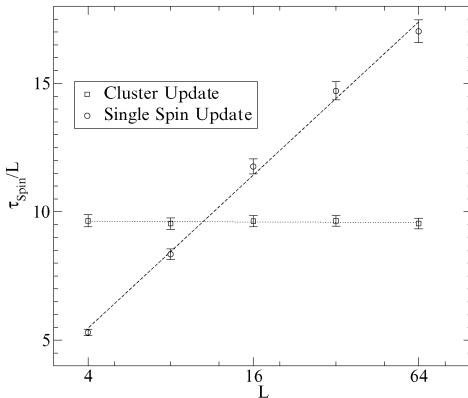


Fig. 16.1. Tunneling time τ_{Spin} to reach an energy 5% within the ground state energy for single spin and cluster updates as described in the text. The lines are drawn as a guide to the eye

the partition function and that the cluster updates show an improved scaling compared to single spin updates. The algorithm can be viewed as the classical limit of the quantum version of Wang–Landau sampling [6] and can be applied to any classical system for which cluster updates are known. In contrast to the multibondic cluster algorithm with Wang–Landau updates [5] however, the bond indices $\{n_{ij}\}$ cannot be updated globally leading to correlations between sequential $\{n_{ij}\}$. A detailed description and the relationship of this new algorithm to the multibondic cluster algorithm for Wang–Landau sampling will be presented in a forthcoming publication. We acknowledge financial support of the Swiss National Science Foundation.

References

1. F. Wang, D.P. Landau: Phys. Rev. Lett. **86**, 2050 (2001); Phys. Rev. E **64**, 056101 (2001)
2. B.A. Berg, T. Neuhaus: Phys. Rev. Lett. **68**, 9 (1992)
3. P. Dayal, S. Trebst, S. Wessel, D. Würtz, M. Troyer, S. Sabhapandit, S. Coppersmith: Phys. Rev. Lett. **92**, 097201 (2004)
4. D. Huse, S. Trebst, M. Troyer, cond-mat/0401195
5. C. Yamaguchi, N. Kawashima: Phys. Rev. E **65**, 056710 (2002)
6. M. Troyer, S. Wessel, F. Alet: Phys. Rev. Lett. **90**, 120201 (2003)
7. W. Janke, S. Kappler: Phys. Rev. Lett. **74**, 212 (1995)
8. A.W. Sandvik, J. Kurkijärvi: Phys. Rev. B **43**, 5950 (1991)
9. R.H. Swendsen, J.-S. Wang: Phys. Rev. Lett. **58**, 86 (1987)
10. U. Wolff: Phys. Rev. Lett. **62**, 361 (1989)

17 Multibaric-Multithermal Simulations for Lennard-Jones Fluids

H. Okumura^{1,2} and Y. Okamoto^{1,2}

¹ Institute for Molecular Science, Okazaki Aichi 444-8585, Japan

² The Graduate University for Advanced Studies, Okazaki, Aichi 444-8585, Japan

Abstract. We present a new Monte Carlo algorithm, which we refer to as the multibaric-multithermal algorithm. This algorithm allows the simulation to perform a random walk widely both in volume space and in potential-energy space. From only one simulation run, one can calculate average quantities in the isobaric-isothermal ensemble at any temperature and pressure values. We discuss the effectiveness of this algorithm by applying it to a Lennard-Jones potential system.

17.1 Introduction

Monte Carlo (MC) algorithm is one of the most widely used methods of computational physics. In order to realize desired statistical ensembles such as the canonical [1], isobaric-isothermal [2], and microcanonical [3] ensembles, corresponding MC techniques have been proposed. Besides the above physical ensembles, it is now almost a default to simulate in artificial, generalized ensembles so that the multiple-minima problem in complex systems can be overcome. The multicanonical algorithm [4, 5] is one of the most well known such methods in generalized ensemble. In the multicanonical ensemble, a non-Boltzmann weight factor is used so that a free one-dimensional random walk is realized in the potential-energy space. Thus, one can obtain various canonical-ensemble averages in a wide range of temperature T from one simulation run by the reweighting techniques [6]. However, it is impossible to change the volume V of the system and to specify pressure P as in experimental environments, since the simulation is performed in a fixed volume.

In order to overcome these difficulties, we recently proposed a new MC algorithm in which one can obtain various isobaric-isothermal ensembles from only one simulation run [7]. We refer to this method as the multibaric-multithermal algorithm. This MC simulation performs a random walk in two-dimensional space: volume space as well as potential-energy space.

17.2 Methods

In the multibaric-multithermal ensemble, every state is sampled by the multibaric-multithermal weight factor $W_{\text{mbt}}(E, V)$ so that a uniform distribution $P_{\text{mbt}}(E, V)$ of potential energy E and volume V may be obtained:

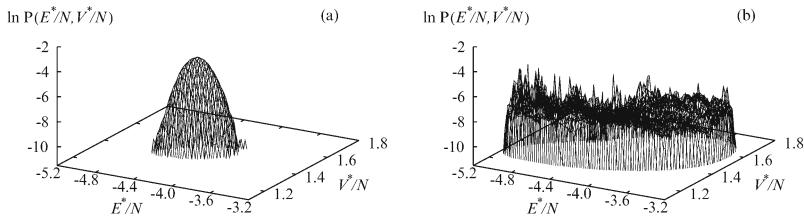


Fig. 17.1. (a) The probability distribution from the initial isobaric-isothermal simulation and (b) that from the production run after 12 iterations

$$P_{\text{mbt}}(E, V) = n(E, V)W_{\text{mbt}}(E, V) = \text{constant}, \quad (17.1)$$

where $n(E, V)$ means the density of states. By replacing $W_{\text{mbt}}(E, V)$ by $\exp[-\beta_0(E + P_0V)]$ ($\beta_0 = 1/k_B T_0$, T_0 and P_0 are temperature and pressure at which simulations are performed, respectively), (17.1) will return to the isobaric-isothermal distribution.

The multibaric-multithermal weight factor is, however, not a priori known and has to be determined by the usual iterations of short simulations [8–10]. After an optimal weight factor $W_{\text{mbt}}(E, V)$ is determined, a long production simulation is performed for data collection. We can apply the reweighting techniques [6] to the results of this production run in order to calculate the isobaric-isothermal-ensemble averages at the designated temperature T and pressure P . Namely, the probability distribution $P_{\text{NPT}}(E, V; T, P)$ in the isobaric-isothermal ensemble in wide ranges of T and P is given by

$$P_{\text{NPT}}(E, V; T, P) = \frac{P_{\text{mbt}}(E, V)W_{\text{mbt}}^{-1}(E, V)e^{-\beta(E+PV)}}{\int dV \int dE P_{\text{mbt}}(E, V)W_{\text{mbt}}^{-1}(E, V)e^{-\beta(E+PV)}}. \quad (17.2)$$

The expectation value of a physical quantity A at T and P is estimated from

$$\langle A \rangle_{\text{NPT}} = \int dV \int dE A(E, V)P_{\text{NPT}}(E, V; T, P). \quad (17.3)$$

Because of a random walk in the potential-energy space, we can calculate physical quantities in a wide range of T . Similarly, due to a random walk in the volume space, we can calculate physical quantities in a wide range of P/T . Here, we should note that the conjugate variable for E is $1/T$, that is, T , and the conjugate variable for V is P/T , rather than P only.

17.3 Computational Details

We now give the details of our simulations. We consider a Lennard-Jones 12-6 potential system. The length and the energy are scaled in units of the Lennard-Jones diameter σ and the depth of the potential ϵ , respectively. We use an asterisk (*) for the quantities reduced by σ and ϵ .

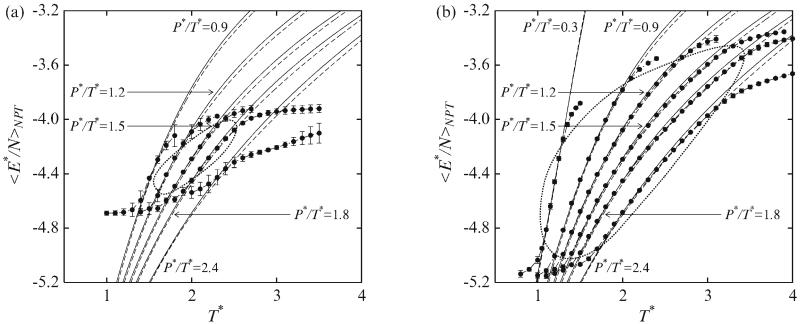


Fig. 17.2. (a) Average potential energies per particle $\langle E^*/N \rangle_{NPT}$ by the isobaric-isothermal MC simulation and (b) $\langle E^*/N \rangle_{NPT}$ by the multibaric-multithermal MC simulation. *Filled circles*: Combination of our MC simulations with the reweighting technique. *Solid curves*: Equation of states calculated by Johnson et al. [13]. *Broken curves*: Equation of states calculated by Sun and Teja [14]. The areas encircled by dotted curves indicate those in which the results by the reweighting techniques agree well with the equations of states

We used 500 particles in a cubic unit cell with periodic boundary conditions. We started the multibaric-multithermal weight factor determination from a regular isobaric-isothermal simulation at $T_0^* = 2.0$ and $P_0^* = 3.0$. These temperature and pressure values are respectively higher than the critical temperature T_c^* and the critical pressure P_c^* [11]. Recent reliable data are $T_c^* = 1.3207(4)$ and $P_c^* = 0.1288(5)$ [12]. The cutoff radius r_c^* was taken to be $L^*/2$. A cut-off correction was added for the pressure and the potential energy. In order to obtain flat probability distribution $P_{mbt}(E, V)$, we carried out relatively short MC simulations of 100,000 MC sweeps. In the present case, it was required to make 12 iterations to get an optimal multibaric-multithermal weight factor $W_{mbt}(E, V)$. We then performed a long production run of 400,000 MC sweeps.

For the purpose of comparisons of the new method with the conventional one, we also performed the isobaric-isothermal MC simulations of 400,000 MC sweeps with 500 Lennard-Jones 12-6 potential particles.

17.4 Results and Discussion

We now present the results of the multibaric-multithermal simulation of the Lennard-Jones system. Figure 17.1a shows the probability distribution $P_{NPT}(E^*/N, V^*/N)$ from the isobaric-isothermal simulation first carried out. It is a bell-shaped distribution. As the iteration proceeds, the probability distribution $P_{mbt}(E^*/N, V^*/N)$ will become flat and broad gradually. Figure 17.1b shows that the multibaric-multithermal MC simulation indeed sam-

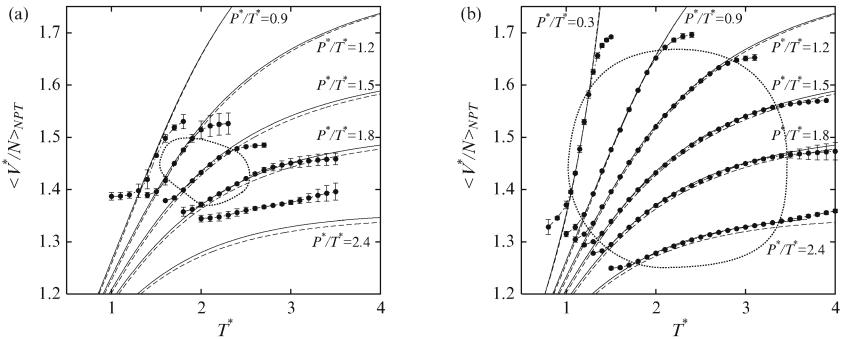


Fig. 17.3. (a) Average volumes per particle $\langle V^*/N \rangle_{\text{NPT}}$ by the isobaric-isothermal MC simulation and (b) $\langle V^*/N \rangle_{\text{NPT}}$ by the multibaric-multithermal MC simulation. See the caption of Fig. 17.2 for details

pled the configurational space in much wider ranges of E^* and V^* than the conventional isobaric-isothermal MC simulation.

In order to investigate the T^* and P^*/T^* ranges in which the isobaric-isothermal and the multibaric-multithermal method can determine accurately average physical quantities by the reweighting techniques, we show $\langle E^*/N^* \rangle_{\text{NPT}}$ and $\langle V^*/N^* \rangle_{\text{NPT}}$ as functions of T^* at several P^*/T^* values in Figs. 17.2 and 17.3. These figures also show the curves of two equations of states of the Lennard-Jones 12-6 potential fluid. One was estimated by Johnson et al. [13] and the other by Sun and Teja [14]. These curves are used as reference in the present work. The areas encircled by dotted curves roughly indicate those in which the results by the reweighting techniques agree well with the equations of states.

One cannot calculate the ensemble averages of physical quantities correctly by combining the isobaric-isothermal algorithm with the reweighting techniques except at T^* and P^*/T^* very close to T_0^* and P_0^*/T_0^* , respectively as shown in Figs. 17.2a and 17.3a. As T^* or P^*/T^* is going far from T_0^* or P_0^*/T_0^* , the error bars of any physical quantities and the deviations from its correct value will get large. The ranges in which $\langle E^*/N \rangle_{\text{NPT}}$ and $\langle V^*/N \rangle_{\text{NPT}}$ are correct are $1.6 \leq T^* \leq 2.5$ and $1.0 \leq P^*/T^* \leq 1.9$. On the other hand, the multibaric-multithermal simulation enables us to calculate physical quantities in wide ranges of both T^* and P^*/T^* . Figures 17.2b and 17.3b show that the multibaric-multithermal data agree with the equations of states well in $1.0 \leq T^* \leq 3.4$ and $0.1 \leq P^*/T^* \leq 2.8$. That is, the multibaric-multithermal simulation provides correct average quantities in the ranges of T^* and P^*/T^* about three times wider than the isobaric-isothermal simulation. These reliable ranges of T^* and P^*/T^* for the reweighting techniques depend on how much effort one is willing to spend for the weight factor determinations. In principle, the multibaric-multithermal simulation can provide

various isobaric-isothermal ensemble averages at any T^* and P^*/T^* values if the ideal weight factor is obtained.

17.5 Conclusions

In this article, we presented a new generalized-ensemble MC algorithm, the multibaric-multithermal algorithm. We can obtain various isobaric-isothermal ensemble averages from only one simulation run in wide ranges of T and P/T . Because of this advantage, we think that these new algorithms will be of use for investigating a large variety of more complex systems.

Acknowledgments

This work was supported, in part, by the Grants-in-Aid for the NAREGI Nanoscience Project and for Scientific Research in Priority Areas, “Water and Biomolecules”, from the Ministry of Education, Culture, Sports, Science and Technology, Japan.

References

1. N. Metropolis, A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller, E. Teller: *J. Chem. Phys.* **21**, 1087 (1953)
2. I.R. McDonald: *Mol. Phys.* **23**, 41 (1972)
3. M. Creutz: *Phys. Rev. Lett.* **50**, 1411 (1983)
4. B.A. Berg, T. Neuhaus: *Phys. Lett.* **B267**, 249 (1991)
5. B.A. Berg, T. Neuhaus: *Phys. Rev. Lett.* **68**, 9 (1992)
6. A.M. Ferrenberg, R.H. Swendsen: *Phys. Rev. Lett.* **61**, 2635 (1988); *ibid.* **63**, 1658(E) (1989)
7. H. Okumura, Y. Okamoto: *Chem. Phys. Lett.* **383**, 391 (2004)
8. B.A. Berg, T. Celik: *Phys. Rev. Lett.* **69**, 2292 (1992)
9. J. Lee: *Phys. Rev. Lett.* **71**, 211 (1993); *ibid.* **71**, 2353(E) (1993)
10. Y. Okamoto, U.H.E. Hansmann: *J. Phys. Chem.* **99**, 11276 (1995)
11. H. Okumura, F. Yonezawa: *J. Chem. Phys.* **113**, 9162 (2000)
12. H. Okumura, F. Yonezawa: *J. Phys. Soc. Jpn.* **70**, 1990 (2001)
13. J.K. Johnson, J.A. Zollweg, K.E. Gubbins: *Mol. Phys.* **78**, 591 (1993)
14. T. Sun, A.S. Teja: *J. Phys. Chem.* **100**, 17365 (1996)

18 A Successive Umbrella Sampling Algorithm to Sample and Overcome Free Energy Barriers

P. Virnau¹ and M. Müller²

¹ Massachusetts Institute of Technology, Department of Physics
Building 12-109, 77 Massachusetts Avenue, Cambridge, MA 02139-4307, USA
virnau@mit.edu

² Institut für Physik, Johannes Gutenberg-Universität Mainz
Staudinger Weg 7, 55099 Mainz, Germany

Abstract. We present an implementation of umbrella sampling which allows us to sample and overcome free energy barriers. To this end, we focus on a subdivision of relevant states into very small windows which are sampled consecutively. This allows us to simulate without a weight function or to generate it during simulation by extrapolating results of previously sampled windows to the neighboring window. We argue that the efficiency of our scheme is in principle comparable to a weighted simulation of all states without subdivision. In our case, however, we do not need to know the weight function before the simulations starts. As a first application we determine the phase behavior and interfacial properties of hexadecane with a coarse-grained model. In this context, we observe a very good agreement with experiments which is comparable to much more involved united-atom models.

18.1 Introduction

The efficient sampling of free energy barriers remains a considerable challenge in condensed matter physics, though substantial progress has been achieved in recent years [1,2]. In computer simulations microscopic states are sampled and linked to thermodynamical quantities in a probabilistic fashion. The Helmholtz free energy A of a system containing n molecules can for instance be derived from the probability distribution $P[n]$: $A[n] = -k_B T \ln P[n] + \text{const.}$, where k_B stands for Boltzmann's constant and T for temperature. Therefore, regions of low probability correspond to barriers in the free energy, which may eventually slow down the simulation. We address this problem with an implementation of umbrella sampling [3] and exemplify the scheme with a study of a coarse-grained model of hexadecane.

18.2 A Coarse-Grained Model for Hexadecane

Our bead-spring model [4,5] aims at reproducing the crude structure of alkanes. In this spirit we map roughly 3 CH₂-groups onto a coarse-grained bead such that hexadecane is represented by a chain of five monomers. Internal degrees of freedom are neglected, which results in a considerable reduction of

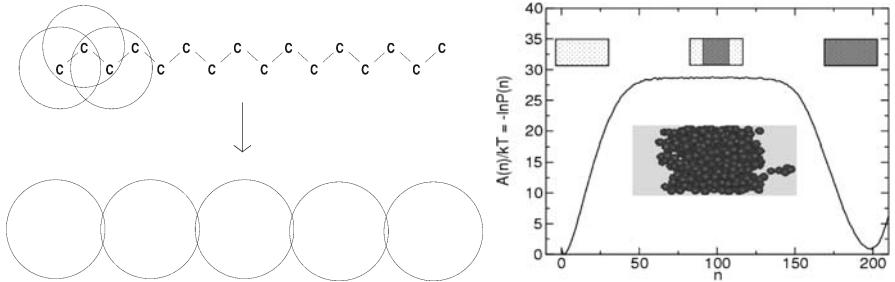


Fig. 18.1. (a) The upper drawing corresponds to a united-atom representation, the lower drawing depicts the coarse-grained model which was used to determine the phase diagrams in Sec.IV. (The diameter of the spheres is σ , respectively. For simplicity, only the stretched version of the molecules are drawn.) (b) Free energy as a function of pentamer number at $T = 1.38\epsilon/k_B$ and $V = 9 * 9 * 18 \sigma^3$. The coexistence densities at temperature T is given by $\rho = n_{g/l}/V$, the interfacial tension $\gamma = \Delta A/2L^2$. Inset: Representative configuration for $n = 100$ pentamers

computational effort in comparison to more involved united-atom representations (Fig. 18.1).

Coarse-grained beads interact via a truncated and shifted Lennard-Jones potential:

$$V_{\text{LJ}}(r) = \begin{cases} 4\epsilon \left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^6 + \frac{127}{16384} \right] & \text{if } r < r_c \\ 0 & \text{else.} \end{cases} \quad (18.1)$$

r denotes the distance between particles and $r_c = 2r_{\min} = 2\sqrt[6]{2} \sigma$ the cutoff. Adjacent beads along a chain are bonded together with an additional FENE potential:

$$V_{\text{FENE}}(r) = -33.75\epsilon \cdot \ln \left[1 - \left(\frac{r}{1.5\sigma_{\text{PP}}} \right)^2 \right]. \quad (18.2)$$

Parameters ϵ and σ in (18.1) and (18.2), which describe the depth and the radius of the interaction, are determined by equating the critical point of the simulation ($T = 1.725$, $\rho = 0.27/\sigma^3$) with the critical point of experiments ($T = 723$ K and $\rho = 0.219$ g/cm).

Simulations take place in the grand canonical (μ, V, T) ensemble. We employ periodic boundary conditions, particle insertion and deletion attempts are realized with a grand canonical configurational bias algorithm [6]. A typical free energy landscape at coexistence is shown in Fig. 18.1b. In an unweighted simulation, the system would either stay in the gas or the liquid phase because the free energy barrier between both states is too large. In such a situation our approach becomes advantageous.

18.3 Successive Umbrella Sampling

In successive umbrella sampling [7], the pertinent range of states is subdivided into small windows which are sampled consecutively. In the easiest implementation, we start with an empty box and allow the system to change only between 0 and 1 particle. A histogram $H(n)$ monitors how often each state is visited. After a predetermined number of insertion/deletion Monte Carlo steps, we determine the ratio $H(1)/H(0)$, move the window to the right (to allow 1 and 2 particles) and start all over again. After all relevant states have been sampled this way, the probability distribution is

$$\frac{P(n)}{P(0)} = \frac{H(1)}{H(0)} \cdot \frac{H(2)}{H(1)} \cdots \frac{H(n)}{H(n-1)}. \quad (18.3)$$

Free energy differences within a window are relatively small such that no weight function is required. However, the consecutive sampling also allows us to extrapolate $P(n)$ into the next window which is about to be sampled. This way we can generate an educated guess for the weight function $w(n+1) = \ln(P(n+1))$ and flatten even the small differences within the windows ($P_{\text{sim}}[n+1] = P[n+1] \exp(-w[n+1]) \approx 1$). The original distribution can be regained by multiplying the weighted distribution $P_{\text{sim}}(n)$ with $\exp(w[n])$. The first window is usually unweighted, $w(2)$ is obtained from a linear and $w(n > 2)$ from quadratic extrapolation.

In principle (i.e., if the restriction to small window sizes does not impart sampling difficulties) our scheme is as fast as a weighted simulation without subdivision. If the relative error of the ratio in a single window is Δ , Gaussian error propagation of (18.3) leads to an accumulated error of order $\sqrt{n}\Delta$. To have comparable overall errors, the error in each window needs to be smaller by a factor of \sqrt{n} when compared to the error of a weighted simulation without subdivision. In both cases, it takes $O(n^2) = n \cdot nO(1)$ Monte Carlo steps [7].

Our approach is easy to implement, does not violate detailed balance and we do not need to know a weight function before the simulation starts. Possible applications are manyfold and range from the study of quantum spin models [8] to off-lattice simulations of colloidal suspensions [9] (also compare with R.L.K Vink's contribution in this book). For a complete error analysis of the method and a discussion of possible limitations, the reader is referred to [7].

18.4 Phase Behavior and Interfacial Tension of Hexadecane

Both the simulated coexistence curve (Fig. 18.2a) and the interfacial tension (Fig. 18.2b) are in very good agreement with experiments, which affirms our

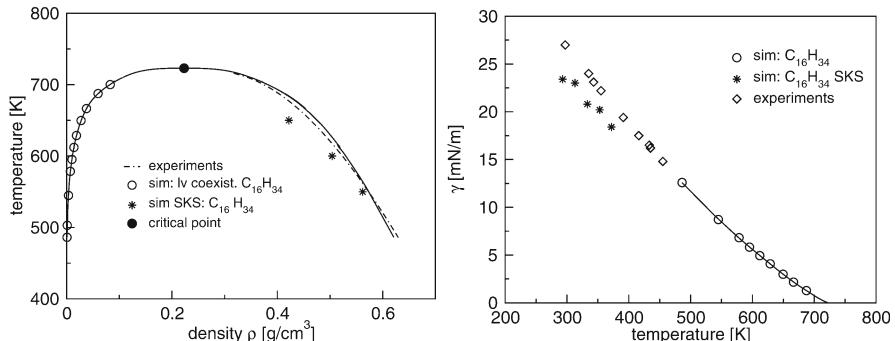


Fig. 18.2. (a) Coexistence curve, (b) interfacial tension of hexadecane as a function of temperature in comparison with experiments [10,11] and united-atom simulations [12]. The solid black lines are only “guides to the eye”. For the liquid phase of $\text{C}_{16}\text{H}_{34}$ we omitted symbols to emphasize the good agreement with experiments

initial assumption that the phase behavior of alkanes is mostly determined by the rough structure of the molecules. The mapping of the critical points only sets temperature and density scales, the similarity of the critical amplitudes, however, should not be taken for granted. An extension of this analysis to n-alkanes will be discussed in [5].

Acknowledgments

It is a great pleasure to thank K. Binder and L.G. MacDowell for stimulating discussions. Financial support by BASF AG, the BMBF (grant No. 03N6015) and DFG under grant No. Mu1674/1 and Vi237/1-1 is gratefully acknowledged. CPU time was provided by the NIC Jülich, the HLRZ Stuttgart and the computing center of the Johannes-Gutenberg-Universität Mainz.

References

1. B. Berg, T. Neuhaus: Phys. Rev. Lett. **68**, 9 (1992)
2. F. Wang, D.P. Landau: Phys. Rev. Lett. **86**, 2050 (2001)
3. G. Torrie, J. Valleau: J. Comput. Phys. **23**, 187 (1977)
4. P. Virnau, M. Müller, L.G. MacDowell, K. Binder: New J. Phys. **6**, 7 (2004)
5. P. Virnau, M. Müller, in preparation
6. J.I. Siepmann: Mol. Phys. **70**, 1145 (1990)
7. P. Virnau, M. Müller: cond-mat/0306678 (2003)
8. M. Troyer, S. Wessel, F. Alet: Phys. Rev. Lett. **90**, 120201 (2003)
9. R.L.C. Vink, J. Horbach: cond-mat/0310404 (2003)
10. B.D. Smith, R. Srivastava: Thermodynamic Data for Pure Compounds: Hydrocarbons and Ketones (Elsevier, Amsterdam 1986)
11. G.T. Dee, B.B. Sauer: Adv. Phys. **47**, 161 (1998)
12. J.P. Nicolas, B. Smit: Mol. Phys. **100**, 2471 (2002)

Part V

Computer Tools

19 C++ and Generic Programming for Rapid Development of Monte Carlo Simulations

G. Brown^{1,2}, H.K. Lee³, and T.C. Schulthess¹

¹ Center for Computational Sciences, Oak Ridge National Laboratory,
Oak Ridge, TN 37831-6164, USA

² School of Computational Science and Information Technology,
Florida State University, Tallahassee, FL 32306-4120, USA

³ Department of Physics, Tokyo Metropolitan University,
Tokyo 192-0397, Japan

Abstract. Flexible coding techniques are important for enabling the solution of many problems with one application or the rapid construction of new applications through code reuse. Generic programming offers the high-performance computing community an excellent method for code reuse without loss of efficiency. Here, solutions to both problems from the Ψ -Mag toolset are presented. The class `EHModel` and its associated `EHTerms` can be used to calculate energies in a wide variety of spin models. An allied generic programming solution for applying Metropolis Monte Carlo to estimation of thermodynamic quantities is also given.

19.1 Introduction

In a wide variety of science and engineering topics, rapid advances in computational resources make possible calculations that were impossible only a few years earlier. Frequently, the codes performing these calculations have been developed by one or two scientists interested in the calculation. While such codes may enable reuse of an entire application through flexibility in input, these codes rarely reuse existing code fragments or even produce fragments suitable for later reuse. Achieving better efficiency often requires the efforts of several programmers, some skilled in computational science but not the science investigated through the calculation. In this environment, production of efficient codes requires development of flexible code fragments that can be reused quickly by people who did not develop them.

Here we discuss a generic programming approach to high-performance computing. In the definition popularized by Austern [1], generic programming is an extension of object-oriented programming that focuses on “the set of requirements” which a data type must fulfill to qualify as an implementation of a particular abstraction. The set of requirements defines a *concept*, and an implemented data type which meets those requirements is called a *model* of the concept. A good example of a concept is `Container`, which is a collection of elements that can be iterated over. The addition of require-

ments leads to the *refinement* of one concept into another. `Container` can be refined, for example, into concepts such as `Sequence`, in which elements can only be accessed in some particular order, `Random-Access Container`, in which the elements can be indexed by an integer type, and `Associated Container`, in which elements can be indexed by an arbitrary type. Concepts such as these, along with algorithms such as `Count`, `Copy`, and `Sort` can be recognized as commonly recurring themes in computational science. The Standard Template Library (STL) defines many such concepts and algorithms. STL provides both the tools to simplify the implementation of high-performance codes and the inspiration to solve problems in a generic, problem-independent manner. The result is code that can easily be reused to create new applications.

To be useful for high-performance computing, the flexibility of generic programming must come without sacrificing efficiency or clarity. Using templates in C++ [2] is one programming paradigm in which this can be achieved. In fact, STL was first implemented as templated C++ classes [1]. Direct comparison has shown that, in general, modern C++ compilers generate executables with no performance penalty with respect to procedural languages such as FORTRAN (see, e.g., [3]). Templates, in particular, defer the specification of types associated with a function or class until compilation, but at no cost to execution speed. For example, a templated sorting function can use a template parameter for the type being sorted in the same manner as subroutines use variable arguments. Whether the type being sorted is a floating point number or an integer is specified at compilation, but the resulting machine code is not affected by the fact that a templated function was used to generate it.

The ψ -Mag project [4] is an effort to foster a community approach to code sharing and reuse in computational materials science. One goal of the project is developing specifications for important concepts for computational materials science in general and computational magnetism in particular. These concepts are then translated into implemented C++ models. Here we discuss the effort within the ψ -Mag project that is concerned with the generic implementation of Monte Carlo methods [5], which can be used in the context of computational magnetism to quickly develop new applications.

The natural flexibility in scientific applications based on making as few assumptions as possible is important. A thorough discussion of this point is necessary for a proper comparison to generic programming. Applications built with such flexibility can simulate many different problems, often with the differences specified only in the input files. There are often significant trade-offs between flexibility and computational efficiency, and making a few assumptions can often greatly simplify an algorithm. Much of the art of scientific programming involves optimizing among flexibility, speed, and coding complexity.

The flexible classes in ψ -Mag that use object-oriented programming to describe the energy and fields associated with spin models are described in Sect. 19.2. A set of generic-programming concepts for Metropolis-sampling Monte Carlo is outlined in Sect. 19.3. In that section, some implementation details particular to ψ -Mag are also discussed. Section 19.4 is a summary of this work.

19.2 Flexible Energy Calculation

Flexibility built on avoiding assumptions is not limited to a particular programming language and has been utilized throughout the history of computer programming. Object-oriented programming facilitates this flexibility through polymorphism [2], by which a set of types can be treated as equivalent even though the members of the set implement very different behaviors. An object-oriented approach, augmented by a trivial application of templates, is outlined here for the calculation of energies in spin models.

The collective behavior of spin systems, such as the Ising and Heisenberg models, has been studied extensively, often in the context of developing fundamental principles in statistical mechanics (see, e.g., [6]). In part, this interest stems from the relative simplicity of the models. For example, the Zeeman energy of a system of spins in a field is

$$E_z = - \sum_{i=1}^N M_i \mathbf{H} \cdot \mathbf{s}_i, \quad (19.1)$$

where the sum is over all spins, M_i is the moment of the i -th spin, \mathbf{s}_i is the orientation of that spin, and \mathbf{H} is the applied field. The difference between the two models lies principally in the nature of \mathbf{s}_i and \mathbf{H} . For the Ising model $\mathbf{s}_i \in \{-1, +1\}$, but for the Heisenberg model $\mathbf{s}_i \in \mathbb{R}^d$ with the constraint $|\mathbf{s}_i| = 1$. Likewise, $\mathbf{H} \in \mathbb{R}$ for the former and $\mathbf{H} \in \mathbb{R}^d$ for the latter. It is thus possible to create one flexible C++ class to handle the Ising model and the family of d -dimensional Heisenberg models if the class is templated on the `spin_type` and `field_type`. In fact, maximum flexibility is achieved if different containers are allowed. Therefore, templating on the type of container holding the orientations of the spins and the type of container holding fields provides maximum flexibility. The types of the individual spins and fields are then the `value_types` of the containers.

As an example of avoiding assumptions, consider exchange interactions. In most simple models, each spin interacts with a well-defined (often small) subset of the other spins. If J_{ij} is the interaction energy between the i -th and j -th spins, then the set of all interactions for a particular problem is $\{J_{ij}\}$, and the exchange energy is

$$E_x = - \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N J_{ij} \mathbf{s}_i \cdot \mathbf{s}_j. \quad (19.2)$$

Frequently $\{J_{ij}\}$ can be simply constructed, e.g. nearest-neighbor interactions on a cubic lattice. A function which uses such an assumption can calculate the energy using a geometry-dependent rule that calculates the neighbors $\{j\}$ on the fly quite easily. However, each particular geometry will require a separate function. A more flexible approach is to explicitly store $\{j\}$ and J_{ij} for each spin i , so that one function can calculate the energy for any geometry. All information about the dimension and symmetry of that lattice, as well as boundary conditions, are conveniently incorporated into the set $\{j\}$ for which J_{ij} are nonzero, and any other variations in exchange strength into an explicit list of $\{J_{ij}\}$. Such flexibility greatly enhances code reuse-ability and decreases development time for new codes.

With these two contributions to the energy of spin models in mind, it is possible to define a generic-programming concept of `Energy`. First define `orient_container` to be the type used to contain the orientations of an entire system of spins, `field_type` to be the type used to represent the direction and strength of a field, and `field_container` to be the type used to contain the field at the location of each spin. The required member functions are as follows.

1. `Real energy (Real time, const orient_container & spin)`
calculates the total energy for the configuration of the system specified in `spin`.
2. `Real energy (Real time, const orient_container & spin, size_type ispin)` calculates the energy associated with the i -th spin for the system configuration specified in `spin`.
3. `void field (Real time, const orient_container & spin, field_container & h)` calculates the field affecting each spin in the system. The calculated fields are *added* to `h`. The form of the field affecting the j -th spin should be the negative of the functional derivative of the energy with respect to s_j , which enforces $E(\mathbf{s}_i) = -\mathbf{h}_i \cdot \mathbf{s}_i$.
4. `void field (Real time, const orient_container & spin, size_type ispin, field_type & h)` calculate the field affecting only the i -th spin. Add the calculated field to `h(si)`.

Here `Real` is a floating-point type, and `size_type` is an integer type appropriate for indexing (it may be nonnegative). Note that the simulation time is made a parameter of the function call to allow for the possibility of, for example, time-dependent applied fields $\mathbf{H}(t)$. Two method functions are required for calculating energies. The first one calculates the total energy of the system, while the second calculates only the energy associated with the i -th spin. The latter method is required for efficiency in algorithms that consider only one spin at a time, such as the local update Monte Carlo discussed in the next section. Note that summing the $E(\mathbf{s}_i)$ of the latter function over all the spins in the system will not, in general, give the E of the former function. Double counting in E_x , which is related to the two-bodied nature of exchange, is a prime example of this fact. Methods for calculating the field at each spin

site $\mathbf{h}(\mathbf{s}_i)$ are also included in the concept definition. Some algorithms for highly efficient Monte Carlo sampling, such as overrelaxation [7], and numerical approaches like micromagnetics [8,9] use this field. The inclusion of both quantities in the `Energy` concept is also intended to enforce the fundamental definition $\mathbf{H}_i = -\delta E / \delta \mathbf{s}_i$. The flexibility outlined here is maximized when the inclusion or exclusion of particular contributions is not determined until the code is executed. Outside the object-oriented paradigm, such flexibility has been mimicked through logical flags or by setting all interactions, such as $\{J_{ij}\}$, to zero. Using object-oriented polymorphism is much more efficient. The `Ψ-Mag` toolset contains a class `EHModel` which is a model of `Energy`, but sums over contributions calculated by other models of the `Energy` concept.

Since the purpose of `EHModel` is to accumulate different energy contributions, it is derived from a container of pointers to those contributions. See Fig. 19.1. Since all of the pointers must be of the same type, a base class `EHTerm`, which is a model of `Energy`, is the target of those pointers. Each particular energy contribution is constructed as a derived class which inherits from `EHTerm`. For this polymorphism to work properly, the member functions of `EHTerm` must be virtual functions. The key property of a virtual function is that it actually calls the equivalent function in the derived class. For `EHModel`, this means that the appropriate function will be called for each term. The flexibility inherent in this approach is that different energy contributions can be incorporated into `EHModel`'s container as the program is executed.

Templating is also required for `EHModel`, since different types may be employed to represent spins and fields in different models. Since virtual functional calls must be employed in calculating the particular energy contributions, and virtual member functions cannot be templated, the classes `EHModel`, `EHTerm`, and the derived classes must all be templated. The templating of so many classes is undesirable as it increases code-complexity, but in this situation it is unavoidable.

`EHModel` and its associated classes obviously allow any number of different contributions to the energy. In addition to the Zeeman and exchange energies defined above, common contributions are the uniaxial anisotropy

$$E_u = - \sum_{i=1}^N K_i \left(1 - [\mathbf{s}_i \cdot \hat{\mathbf{n}}_i]^2 \right) \quad (19.3)$$

where $\hat{\mathbf{n}}_i$ is the axis of anisotropy for the i -th spin, and K_i is the strength of that anisotropy, and dipole-dipole interactions

$$E_m = -\frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N M_i M_j \mathbf{s}_j \cdot \left(\frac{3\hat{\mathbf{r}}_{ij}\hat{\mathbf{r}}_{ij} - 1}{r_{ij}^3} \right) \cdot \mathbf{s}_i \quad (19.4)$$

with \mathbf{r}_{ij} the displacement vector between the i -th and j -th spins. Extensibility is easily achieved, because new derived classes can be created without

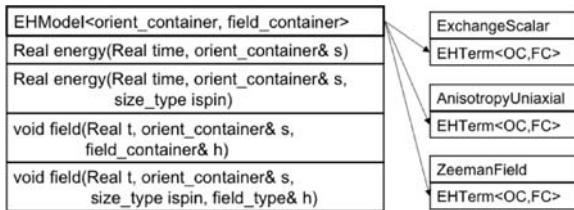


Fig. 19.1. Schematic representation of the polymorphism used in `EHModel`. `EHModel` is a container of pointers to `EHTerm`, and the various energy terms inherit from `EHTerm`. Virtual function calls, with the same interface as those in `EHModel`, enable `EHModel` to access the method functions of the derived classes

modifying the existing code. In addition to energy contributions not mentioned here, this can be exploited to increase the generality or the specificity of the energy contributions. As an example of the former, the toolset already contains an `EHTerm`-derived class for which J_{ij} can be a tensor. This is appropriate for anisotropic exchange. As an example of the latter, a derived class could be created for which J_{ij} is the same for all nearest neighbors. The advantage of this kind of specificity is that it can enhance the efficiency of the program. Such general or specific classes can be added to the toolset as the need for each is identified.

While the particular class system associated with `EHModel` and `EHTerm` outlined in the previous section is itself quite flexible, the real power of generic programming is associated more with the `Energy` concept itself. A complete class system for lattice-gas simulations (`LGModel`) or for Lennard-Jones fluids (`LJFluid`) are also possible. With generic programming, algorithms and classes that use `EHModel`, `LGModel`, or `LJFluid` interchangeably are then possible.

19.3 Monte Carlo Concepts

By defining a set of generic-programming concepts associated with Monte Carlo sampling, it is possible to create a large set of models which become interchangeable tools and can be assembled quite easily into new Monte Carlo simulations. This section provides one such set of concept definitions.

At its most fundamental, Monte Carlo is a technique for estimating the integral

$$I = \int g(x)f(x)dx \quad (19.5)$$

by randomly drawing N values of x with probability density function (pdf) $f(x)$ and evaluating the mean

$$E = (1/N) \sum_{i=1}^N g(x_i). \quad (19.6)$$

In the limit of large N , $E \approx I$ [5]. Monte Carlo techniques usually outperform deterministic quadrature techniques when the dimensionality of x is large, as is usually the case in statistical mechanics where one is interested in high-dimensional phase spaces.

Randomly drawing values of x from $f(x)$ is, in general, not easy. Although it is possible to find simple algorithms for particular forms of $f(x)$, one often has to resort to importance sampling techniques. One important example is a technique due to Metropolis et al. [10] which produces a Markov chain of x_i distributed proportionally to $f(x)$. The Metropolis algorithm is based on the idea of *detailed balance*, specifically that the probability of making a transition from state x to x' in the Markov chain must be equal to the probability of making the transition from x' to x . In the Metropolis approach [5]

$$A(x'|x)T(x'|x)f(x) = A(x|x')T(x|x')f(x'), \quad (19.7)$$

where $T(x'|x)$ is the probability of “proposing” the transition from x to x' , $A(x'|x)$ is the probability of accepting the transition, and $f(x)$ is the desired pdf. This can be rearranged to give a constraint on A ,

$$A(x'|x) = \frac{T(x|x')f(x')}{T(x'|x)f(x)}A(x|x') \equiv q(x'|x)A(x|x'). \quad (19.8)$$

It is obvious from (19.8) that $q(x'|x) = 1/q(x|x')$. This constraint does not completely specify A . Two commonly used forms that satisfy (19.8) are the Metropolis acceptance probability $A_M = \min(q, 1)$ and the Glauber acceptance probability $A_G = q/(1+q)$. In condensed-matter physics, the situation is often much simpler because the transitions in the Markov chain are chosen such that $T(x'|x) = T(x|x')$, then $q = f(x')/f(x)$.

For a generic approach to Metropolis sampling, several concepts need to be defined. The first concept covers Monte Carlo sampling in general, and Markov-chain importance sampling in particular. In the Ψ -Mag toolset, we define a class **MetropolisSampling** to implement a generic Monte Carlo concept, labeled **MCSample**, as follows.

1. **Real mc_time(Real val)**: This sets the “simulation time” in the sampling algorithm. The time is given in units of Monte Carlo steps (mcs). The simulation time is frequently needed. For instance, it is needed for measuring the correlation between the samples generated by the Metropolis method, and may be employed in algorithms such as simulated annealing.
2. **Real mc_time() const** returns the “simulation time” associated with the sampling object in units of mcs.
3. **int operator()(RNG & urng, TransitionOp & op, Weight & f, System & x, int ntransition=1)** performs **ntransition** Monte Carlo steps to generate a new configuration of the system to be used in Monte Carlo integration. The parameter **x** contains the initial configuration

when the function is called and the new configuration upon return. Here a function object $f(t, x)$ is used to determine the weight of configuration x at time t , with t the time (in mcs) supplied by the **MCSample** class. For example, for the canonical ensemble, the weight is $\exp[-\beta E(t, x)]$. The transitions associated with each mcs are accomplished via a model of **TransitionOp**, and random numbers are supplied by a model of **RNG**. The return value of the function is the number of accepted mcs, which can be used to calculate the acceptance rate.

Essentially **MCSample** is a function object that uses one function object as a uniform random number generator on $[0, 1]$) for making choices and another function object for determining the weight $f(t, x)$ of particular configurations. Note that the acceptance probability A is not passed as a function parameter. Instead it assumed to be an integral part of the sampling concept. In the C++ implementation of Ψ -Mag it is a template parameter. **MCSampling** also defines a nontrivial transition concept which can generate the sequences of the Markov chain. The **MTransition** concept, with **element_type** the type used to represent the individual components of the system, is as follows.

1. **Real step_size (Real val)**: Set the size of change associated with each transition. The meaning of this parameter is highly dependent on the individual transitions being implemented.
2. **Real step_size() const**: return the size of the change associated with each transition.
3. **size_type pick_dof (RNG & urng, Real time, const System & x)**: choose the index of a particular degree of freedom within x . This usually becomes the trial move.
4. **Real operator() (RNG & urng, Real time, System & x, size_type idof)**: perform on x a transition associated with the i -th degree of freedom. The return value is the ratio $T(x|x')/T(x'|x)$, i.e. the ratio of the probability of attempting a move from x' to x to the probability of attempting a move the other way.
5. **void undo(System & x)**: restores x to its configuration before the last move was made. This can only be done for the last transition, repeated calls are *not* guaranteed to move the system backwards through all configurations visited.

MTransition provides for selecting a degree of freedom (dof) within the system, performing a transition associated with that dof, and undoing that transition if it is not accepted.

Based on these concepts, the core of the generic single-site-update Metropolis sampling code looks like

```
int dof = op.pick_dof(urng,time,x);
Real fPrev = f(time,x,dof);
Real tRatio = op(urng,time,x,dof);
Real fNew = f(time,x,dof);
```

```

Real q = tRatio*fNew/fPrev;
if( urng() < A(q) )
    isuccess++;
else
    op.undo(x);

```

This algorithm should work as well for quadrature as it does for simulations of Ising models and compressible Lennard-Jones fluids. A cluster update algorithm may require a different implementation, but could be a model of `MCSample` and be able to easily use the all of the concepts employed here.

In condensed-matter physics the weight function is usually a Boltzmann factor. In spin systems frequently $f(t, x) = \exp[-\beta E(t, x)]$, where $E(t, x)$ is the energy of configuration x at time t , and $\beta = (k_B T)^{-1}$ is the inverse temperature. For compressible fluids, however, $f(x) = \exp(-\beta[E(x) + PV(t, x)])$, where the pressure P and volume $V(t, x)$ have to be introduced. Since the weight concept only calls for a function object that takes t and x as parameters, both possibilities are allowed for, but the weight concept is then refined into two cases. The incompressible refinement requires methods for setting and accessing the temperature $1/\beta$ and uses a model of the `Energy` concept to calculate the energy associated with x . The compressible refinement requires setting and accessing methods for both $1/\beta$ and P , and needs to supplement the `Energy` concept with a way to calculate $V(t, x)$. This can be very easily generalized to the grand canonical ensemble, and is a prime example of the power of generic programming.

The structure of the Boltzmann-weight functions is particularly simple. The various functions can be coded generically using the appropriate `Energy`, `Volume`, and `Number` concepts as template parameters which set the policies [11] appropriate to specific models. In fact, in this decomposition of the importance-sampling problem, all of the system-specific details have been pushed into the weight function. This is not a particularly large burden, care must simply be taken that the `System` passed to the Metropolis sampling object be compatible with the `Energy` model that is waiting for it at the other end of the function-call sequence. If the `System` and `Energy` objects are instantiated together in the application code, the potential for problems will be minimized.

To this point, the discussion has focused on the generic-programming concepts associated with Monte Carlo sampling. To have an actual application code, classes that are models of these concepts must be written, instantiated, and assembled. Schematic representations of two different applications are shown in Fig. 19.2. Figure 19.2a represents the simulation of a Heisenberg spin model. It indicates that R1279, a lagged-Fibonacci generator, is used for random numbers. The simulation is done in the canonical ensemble using a class `BoltzmannWeight` with `EHModel`, as defined in Sect. 19.2, used to actually calculate the energies. The transition model is a class that displaces a Heisenberg spin within some sphere around its current direction. The tran-

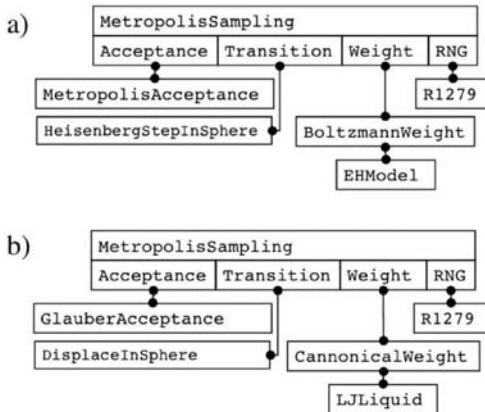


Fig. 19.2. Schematic representation of the particular classes used in to different Monte Carlo simulations assembled from the Ψ -Mag toolset. (a) Simulation of a Heisenberg model. (b) Simulation of a Lennard–Jones liquid

sition is a property of the simulation that can be expected to change most frequently. Finally, the Metropolis form of the acceptance probability A is used, in the Ψ -Mag toolset we have found it convenient to make the form of A a template parameter of the `MetropolisSampling` class.

Figure 19.2b represents canonical-ensemble simulations on a Lennard–Jones fluid at constant pressure. Compared to Fig. 19.2a, the same random number generator is used and a trivial change to the Glauber form of A has been made. However, this simulation requires a new class for calculating the weight $f(x)$, and `LJLiquid` must therefore be capable of calculating both E and V . Finally, since the rotation of a Heisenberg spin has no meaning in this context, a class which displaces particles within a sphere centered on the previous location has been used.

There is another indication of the power of generic programming techniques. The schematics shown in Fig. 19.2 were inspired by a graphical interface used for the Common Component Architecture (CCA) [12, 13]. CCA is based on the Scientific Interface Description Language (SIDL), and can be used to bind together code not originally written to interoperate; even code written in different languages. There is a very strong correlation between the SIDL and generic programming concepts. Using a simple wrapping method [14], it is possible to turn the Monte Carlo classes of the Ψ -Mag toolset into several components which can be combined easily in a graphical format similar to that shown in Fig. 19.2.

The concepts developed here have been verified with actual simulations. The results of simulations of a $6 \times 6 \times 6$ Heisenberg antiferromagnet with $K_i = 0.01$, $\hat{\mathbf{n}}_i = \hat{z}$, and $J_{ij} = -1$ for i and j nearest neighbors are presented in Fig. 19.3. The squares are the staggered magnetization estimated using the Monte Carlo scheme outlined here. The loss of order in-

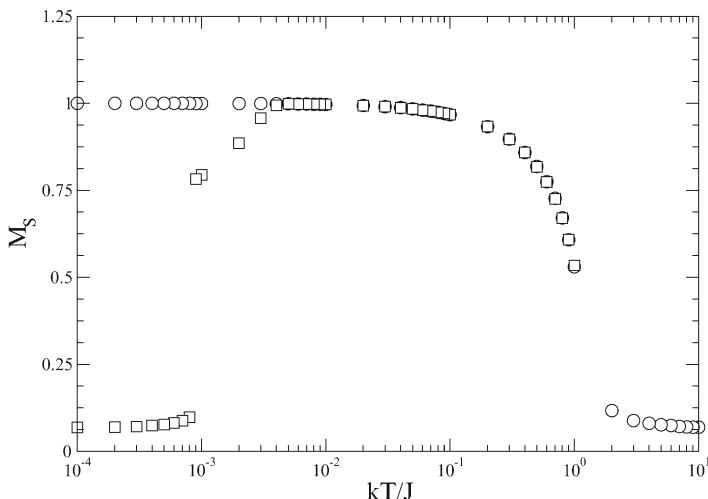


Fig. 19.3. Staggered magnetization, M_S , for a $6 \times 6 \times 6$ Heisenberg model with $K = 0.01$ calculated at various temperatures using the scheme described here. The implementation using `MetropolisSampling` and `BoltzmannWeight` as the model for the `Weight` concept (*squares*) suffers from underflow at low temperatures, while an implementation using `MetropolisSamplingBoltzmann` and `EnergyWeight` for the model (*circles*) does not

dicated by the decrease in the staggered magnetization at very low temperatures is caused by an underflow error in the exponentiation that occurs because in `MetropolisSampling` $q = \exp[-\beta E(x')]/\exp[-\beta E(x)]$. The creation of two new classes, one an `MCSampling` model that calculates using $q = \exp(-\beta[E(x') - E(x)])$ and one that makes $f(x) = E(x)$, solves this problem. The error-free results are the circles in Fig. 19.3.

The results shown here highlight some cautions that must be kept in mind. The most important is that generic programming will not enable non-experts to naively make calculations that would have been infeasible for them otherwise. In fact, the opposite may be true. Expertise and a thorough examination of the results may be required to validate an application assembled in this way. It is precisely the knowledge that calculations can go wrong in many subtle ways that makes the expertise so important.

19.4 Summary

Applications of object-oriented and generic programming to Monte Carlo simulation of spin models have been presented. Using object-oriented polymorphism and explicit descriptions of details such as $\{J_{ij}\}$, flexible calculations of the energy and fields associated with the spin models are possible. The

concepts needed for a generic-programming implementation of Metropolis-sampling Monte Carlo have been outlined. The Ψ -Mag toolset implements both solutions and has been used to successfully calculate thermodynamic properties of spin systems.

Acknowledgments

This work was supported by the DOE Office of Science through ASCR-MICS and the Computational Material Science Network of BES-DMSE as well as the Laboratory directed Research and Development program of ORNL under contract DE-AC05-00OR22725 with UT-Battelle LLC.

References

1. M.H. Austern: Generic Programming and the STL (Addison-Wesley, Reading, MA, 1999)
2. B. Stroustrup: The C⁺⁺ Programming Language (Addison-Wesley, Reading, MA, 2000)
3. J.R. Cary, S.G. Shasharina, J.C. Cummings, J.V.W. Reynders, P.J. Hinker: Comp. Phys. Comm. **105**, 20 (1997)
4. <http://www.ccs.ornl.gov/mri/psimag>
5. M.H. Kalos, P.A. Whitlock: Monte Carlo Methods (Wiley, New York 1986)
6. N. Goldenfeld: Lectures of Phase Transitions and the Renormalization Group (Addison-Wesley, Reading, MA, 1992)
7. F.R. Brown, T.J. Woch: Phys. Rev. Lett. **58**, 2394 (1987)
8. W. Brown: Micromagnetics (Wiley, New York 1963)
9. W.F. Brown: IEEE Trans. Magn. **15**, 1196 (1979)
10. N. Metropolis, A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller, E. Teller: J. Chem. Phys. **21**, 1087 (1953)
11. A. Alexandrescu: Modern C⁺⁺ Design (Addison-Wesley, Reading, MA, 2001)
12. R. Armstrong, D. Gannon, A. Geist, K. Keahey, S. Kohn, L. McInnes, S. Parker, B. Smolinski, in *Eighth IEEE International Symposium on High Performance Distributed Computing* (IEEE, 1998) [Lawrence Livermore National Laboratory technical report UCRL-JC-134475]
13. D.E. Bernholdt, W.R. Elwasif, J.A. Kohl, T.G.W. Epperly, in *Proceedings of the Workshop on Performance Optimization via High-Level Languages and Libraries*. Available at <http://www.ece.lsu.edu/jxr/ics02workshop.html>
14. W.R. Elwasif, G. Brown, D.E. Bernholdt, T.C. Schulthess: (unpublished)

20 Visualization of Vector Spin Configurations

R. Hihinashvili¹, J. Adler¹, S.-H. Tsai^{2,3}, and D.P. Landau³

¹ Physics Department, Technion, Haifa, 32000, Israel,

² The Center for Simulational Physics

³ Enterprise IT Services, University of Georgia Athens, GA 30602-2451, U.S.A.
`phr76ja@tx.technion.ac.il`

Abstract. This chapter describes the visualization of three-dimensional classical vector spins with the AViz atomistic visualization package.

20.1 Introduction

Ising models on a regular lattice with two different states are conceptually trivial to visualize – a simple 0,1 representation using 0 for one state and 1 for the other conveys the physics quite accurately, especially in two dimensions. Such a scheme could be (and often has been) implemented for two-dimensional systems even without a graphical screen. Three-dimensional Ising spin systems do require some minimal graphical capabilities, although any object that has two different shapes or can be presented in two different colors provides a suitable representation of the two states.

Simulations relating to structural aspects of atomistic systems are also conceptually easy to visualize – one uses colored balls for atoms and sticks to indicate bonds. Interesting cases of atomistic simulations are often off-lattice and three dimensional and therefore do require good graphical capability. The Computational Physics group at the Technion has simplified visualization of atomistic systems with their AViz package [1]. Different colors or sizes for the balls and or sticks represent different atoms or states of the atoms. AViz has built-in possibilities for easy implementation of such characterizations. It is free, as well as easy to install and apply.

AViz was actually designed from the beginning to visualize either atoms as balls or atomic spins as arrows. However, despite the success of AViz for the ball/stick representation, visualization of vector spins with AViz was not as straightforward. The visualization of classical spin systems described by three dimensional spin vectors in a three-dimensional space presents both a conceptual and a technical challenge, especially in three dimensions. The interest is to visualize not only static equilibrium spin configurations, but also their temporal evolution as governed by equations of motion [2]. We describe below some first steps towards meeting this challenge. The data visualized in this project originate from the Simulational Physics group at the University of Georgia, and details of the physical problems and simulation methods can be found in [3]. Technical details of the implementation can be found

on a website [4] where detailed instructions and color examples are given. In the present chapter we discuss conceptual issues and give some greyscale examples.

20.2 Models

A classical Heisenberg spin model has the Hamiltonian,

$$-\beta\mathcal{H} = \sum_{ij} J_{ij} \mathbf{s}_i \cdot \mathbf{s}_j ,$$

where spin \mathbf{s}_i on site i and \mathbf{s}_j on another site j interact with strength J_{ij} , ($\beta = 1/kT$ is the inverse temperature, where k is Boltzmann's constant and T is temperature). The spins are thought of as unit vectors pointing in the direction of the spin-current. The sum over sites i and j is usually taken over nearest neighbours which in the cases discussed here sit on translationally invariant lattices in one, two or three dimensional space. One vector is situated at each lattice site and, in general, can point in any direction in a three dimensional space. It is these vectors whose spatial structure and time-development we wish to observe. Possible representations include arrows, cones or pins, originating at the spatial position of the atom and pointing into the same three dimensional space where these sites are situated. It is this double role for the three-dimensional space that so complicates the visualization process.

20.3 AViz

The original technical development of AViz [5–7] was carried out by Adham Hashibon and Geri Wagner. The graphics use MESA/OpenGL and the Qt toolkit is used to prepare the GUI. The AViz homepage [1] contains examples, sample datafiles and downloadable installation files (including rpm format for Linux Redhat and SUSE and source tar files for general UNIX) and a manual (prepared by Geri Wagner and Anastasia Sorkin). Sources and instructions are freely available, under a GNU General Public License, Copyright [6].

The basic concept of AViz is that the location and nature of the objects simulated is recorded in an input file or list of files. These are visualized either concurrently or after the simulations are completed. The interactive interface enables selection of viewing angle, object type, color representation etc and once the desired values are selected they can be saved for future use. Animations of a series of states can be prepared and saved in multiple formats.

A recent introduction to the use of AViz for atomistic simulations is given in [7]. This article includes a discussion of the installation as well as general guidance for the implementation of AViz for visualization of atoms, as well as a survey of some recent implementations. The description below is a continuation of [7] with specific consideration and examples for the spin case.

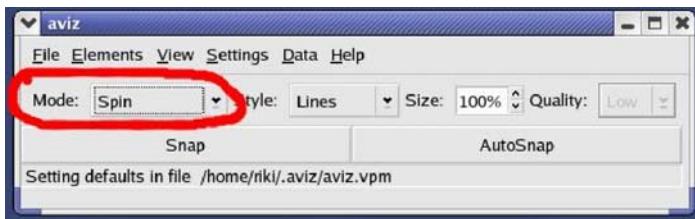


Fig. 20.1. The selection box that pops up when AViz is initialized, with the mode chosen as spin rather than atom

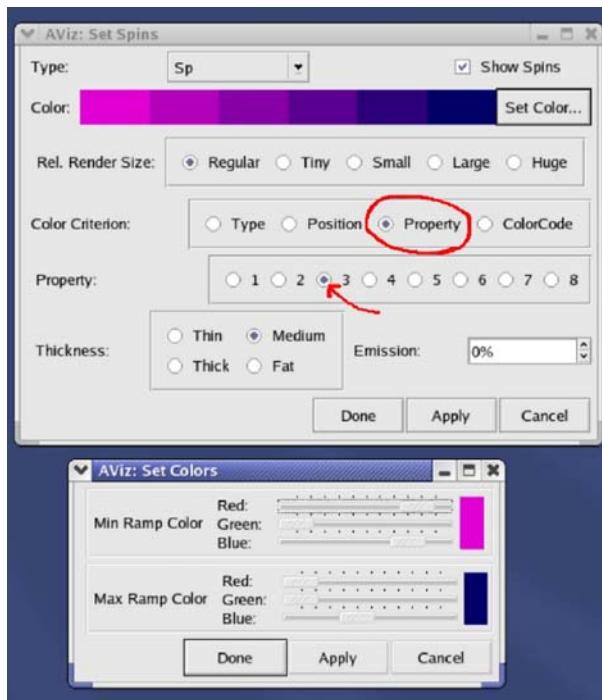


Fig. 20.2. This box is used to set the color coding for the z-coordinate of the atomic spin

20.4 Visualizing Vector Spins

In contrast to the case of visualization of an atom by a colored ball, where only four pieces of information (the atom type and three spatial coordinates) are needed, when an atomic spin is visualised in three dimensions, three additional numbers are required to indicate the direction of the spin vector. Full details of the process of spin visualization are given on the website [4], and we now give brief directions concerning the major points, using a two-

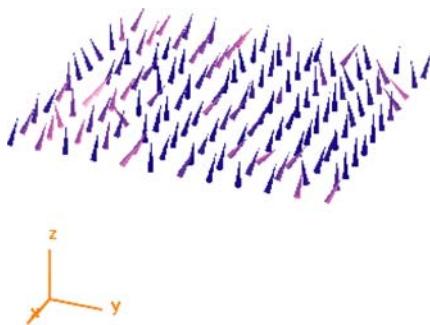


Fig. 20.3. A visualization of a two-dimensional vector spin array

dimensional array of three-dimensional vector spins. The first few lines of a typical spin datafile [8] have the format:

144

```
# 12x12 Heisenberg ferromagnet
Sp 1.0000 1.0000 1.0000 -0.5489 0.2497 0.7978
Sp 1.0000 2.0000 1.0000 -0.5951 0.2954 0.7474
Sp 1.0000 3.0000 1.0000 -0.4631 0.1305 0.8767
Sp 1.0000 4.0000 1.0000 -0.5206 0.0163 0.8536,
```

where the first row gives the total number of spins, the next row is a comment and the subsequent rows, (144 of them in the full file) give the type, three spatial and then three directional coordinates for a datafile of 12×12 Heisenberg spins in a two-dimensional space.

After the AViz interface is brought up, and the datafile loaded, one goes to the mode button on the pop-up selection box, presses atom and then selects spin from the pull-down options (see Fig. 20.1, where this interface is illustrated). Three style types then become available – line, cone and pin. Line is analogous to the quick first-draft dot for the atom case. We have found that cone, with a size larger than 100% and a quality of high or final give the best results.

Once the type is selected, there are several ways to proceed. Two sucessful directions are either to use the shading provided by lighting from a suitable viewpoint (see [4]) or using color gradation related to the z -component to guide the eye. For the latter, press elements, select spins and then edit the box as shown in Fig. 20.2.

Once these preliminaries are completed, the sample can be rotated, sliced, snapped etc as in other AViz applications. A final picture might look like Fig. 20.3, which has colors selected to reproduce well in greyscale and does not correspond exactly to the selection in Fig. 20.2 above or the color version on the website, [9]. Another useful trick to enhance visualization is to provide a vertical stick at each site around which the spin can rotate. This requires

“fooling” AViz by adding an additional set of line objects very close to the origins of the real spins.

20.5 Three Dimensions

Three dimensional samples require additional thought. One must visualize the location in three dimensions which can be confusing. Two useful ways (examples are given in [4]) to enhance perception are:

1. Build the sample up layer-by-layer
2. Zoom in and out of the sample to show its spatial structure.

During the workshop we were happy to hear about a different, powerful approach [10] to the visualization of three-dimensional spins using stereoscopic glasses and a power wall.

Acknowledgements

We acknowledge the support of NSF Grant #DMR-0341874 and of the US-Israel Binational Science Foundation for this project. We thank Arie Aharon and Yehudit Mond of the Technion Visualization Laboratory for their advice.

References

1. <http://phycomp.technion.ac.il/~aviz>
2. D.P. Landau, M. Krich: J. Phys. C **11**, R179 (1999)
3. S.-H. Tsai, A. Bunker, D.P. Landau: Phys. Rev. B **61**, 333 (2000)
4. <http://phycomp.technion.ac.il/~riki>
5. J. Adler, A. Hashibon, A. Kanigel, I. Rosenblum, D. Saada, in *Computer Simulation Studies in Condensed Matter Physics, XI*. D.P. Landau, H.-B Schüttler (Eds.) (Springer, Berlin Heidelberg New York Tokyo 1999) p. 186–190; J. Adler, A. Hashibon, G. Wagner, in *Recent Developments in Computer Simulation Studies in Condensed Matter Physics, XIV*. D.P. Landau, S.P. Lewis, H.-B. Schüttler (Eds.) (Springer, Berlin Heidelberg New York Tokyo 2002) p. 160-165; J. Adler, A. Hashibon, N. Schreiber, A. Sorkin, S. Sorkin, G. Wagner: Computer Physics Communications **147**, 665 (2002)
6. G. Wagner, A. Hashibon: AViz AtomicVizualization, 2001 Computational Physics Group, Technion
7. J. Adler: Computers in Science and Engineering **5**, 61 (2003)
8. <http://phycomp.technion.ac.il/~phr76ja/spin/sample.xyz>
9. <http://phycomp.technion.ac.il/~phr76ja/spin/sample.0001.png>
10. <http://www.csit.fsu.edu/~mitchell>

21 The BlueGene/L Project

D. Chen

IBM T.J. Watson Research Center, Yorktown Heights, NY 10598, USA

Abstract. BlueGene/L is a partnership between IBM, US ASIC Tri-Labs and Universities to develop and build a 360 Teraflops massively parallel supercomputer. In this article, we first discuss the background of this project. We then describe the detailed architecture of the BlueGene/L computer and give an update to the current project status.

21.1 Project Background

In December 1999, IBM Research announced a multi-year initiative to build a petaflops scale supercomputer for computational biology research [1]. The BlueGene/L project is a first step in this effort [2]. It is a jointly funded research partnership between IBM and the Lawrence Livermore National Laboratory (LLNL) as part of the US Department of Energy ASCI Advanced Architecture Research Program. The research partnership has since expanded to ASCI Tri-Labs and multiple Universities.

BlueGene/L is a massively parallel supercomputer. Its design philosophy has been influenced by special-purpose massively parallel machines dedicated for Lattice QCD physics simulations, for example, the QCDSF [3] and QCDOC [4] machines. BlueGene/L leverages IBM's system-on-a-chip silicon technology and builds a large parallel system of up to more than 65,000 nodes, yet at a significantly lower price/performance and power consumption/performance versus the conventional approach of clustering large, expensive commercial SMP machines.

A large BlueGene/L machine with 65536 compute nodes and a peak performance of 360 Teraflops is anticipated to be built at the LLNL in early 2005. A smaller, 100 Teraflops machine is expected to be built at the IBM T.J. Watson Research Center for computational biology studies.

21.2 BlueGene/L Architecture

A BlueGene/L compute node consists of a single custom designed ASIC chip and several standard DDR-SDRAM chips as local memory external to the ASIC. The architecture of the BlueGene/L node ASIC is shown in Fig. 21.1.

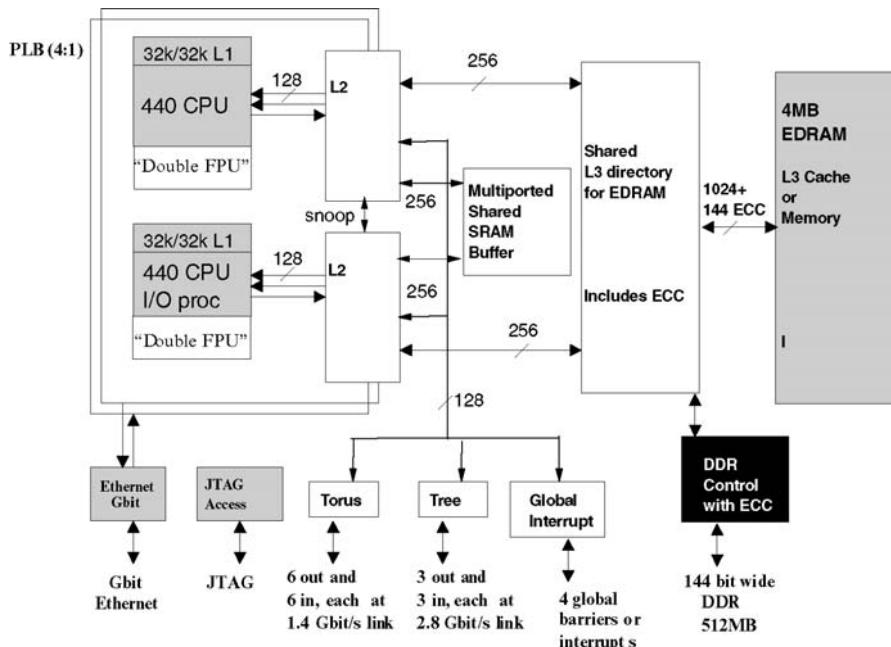


Fig. 21.1. BlueGene/L node ASIC. Each ASIC integrates two 32 bit PowerPC 440 integer CPU cores, 2 “double” floating point FP2 cores, L2 cache buffers, 4 MB EDRAM as L3 cache, DDR-SDRAM controller for external memory, high speed torus network logic, global tree network logic, Gigabit Ethernet and JTAG control interface

Each ASIC chip integrates two PowerPC 440 integer cores, 2 PowerPC 440 FP2 floating point cores, 2 small L2 buffers, 4 MB embedded DRAM configured as L3 cache, memory controller for external DDR-SDRAM memory, custom designed high speed torus and tree network logic, a Gigabit Ethernet and a JTAG control interface. The ASIC is manufactured on IBM CMOS CU-11 0.13 micron copper technology with an 11.1 mm² die size.

The PowerPC 440 core is a standard 32-bit integer PowerPC embedded core product from IBM Microelectronics. It has integrated 32 KB instruction and 32 KB data L1 caches. The 440 FP2 core is an enhanced “double” FPU consisting of a primary unit and a secondary unit, each of which is a complete FPU with their own register sets. The primary FPU supports standard PowerPC floating point instructions. With a SIMD like instruction extension, both FPU units can be utilized to execute 2 fused 64 bit multiply-adds per CPU cycle. The instruction extension is flexible to allow efficient double precision complex number operations. At a target CPU clock frequency of 700 MHz, each core can achieve a peak performance of 2.8 Gflops. The peak speed per node is either 2.8 Gflops when one CPU does computation

and the other CPU does communications, or 5.6 Gflops when both CPUs do computation and communications. A floating point load/store operation can be executed in parallel to the multiply-adds. The load/store unit supports 128 bit “quad-word” loads (stores) to (from) a pair of registers, one each from the two FPUs.

Each of the 440 processor core is directly connected to a small 2 KB L2 cache with data prefetch capabilities, then to a shared 4 MB on-chip embedded DRAM configured as the L3 cache. The L3 controller directly connects to a DDR-SDRAM controller for external memory. Currently, 512 MB memory per node is planned.

The BlueGene/L ASIC integrates several custom high speed networks on chip. The 3-d torus is the primary network for general-purpose, point-to-point message passing operations. Each ASIC has 6 torus links built in. On a compute node, these 6 links are connected to its 6 nearest neighbors, in $+x$, $-x$, $+y$, $-y$, $+z$ and $-z$ directions, respectively. Each link is a bi-directional serial connection with a target speed of 1.4 Gbit/s per direction. The torus logic supports routing through the 3-d torus from a packet’s source to its destination without any CPU intervention. A separate global tree network is implemented to reduce the latency of global reduction operations. The tree network also supports global broadcasts from a single node to all other nodes as well as point-to-point messages between a pair of nodes for I/O.

In a BlueGene/L system, in addition to compute nodes, there are also I/O nodes. An I/O node uses the same BlueGene/L ASIC. It is connected to compute nodes through the tree network. The integrated Gbit Ethernet on an I/O node is then connected through external Gbit Ethernet switches to the host computer and the I/O disk farm. The ratio of I/O nodes to compute nodes can vary. On the LLNL machine, 1 I/O node for every 64 compute nodes is planned.

Figure 21.2 shows the steps of building the BlueGene/L supercomputer. Each node consists of a compute ASIC and 512 MB local memory. Two such nodes are put onto a compute card. 16 compute cards are plugged into a node board. A BG/L cabinet contains 32 node boards, forming a $8 \times 8 \times 16$ torus with a peak performance of 2.9/5.7 Tflops. The large BlueGene/L system then consists of 64 cabinets connected as a $32 \times 32 \times 64$ torus. The total peak performance is 180/360 Tflops. The system will occupy an area of about 250 m^2 including space for the host computer and the disk farm. The total power is estimated at approximately 1.5 MW and the whole system is air cooled.

In terms of development tools, IBM’s XL family C, C++ and FORTRAN compilers are ported to support the PowerPC 440 instructions as well as the “double” FPU. The GNU compiler also supports the common set of PowerPC instructions, but it does not support the FP2 instruction extensions. A set of highly optimized math libraries will be provided to facilitate high perfor-

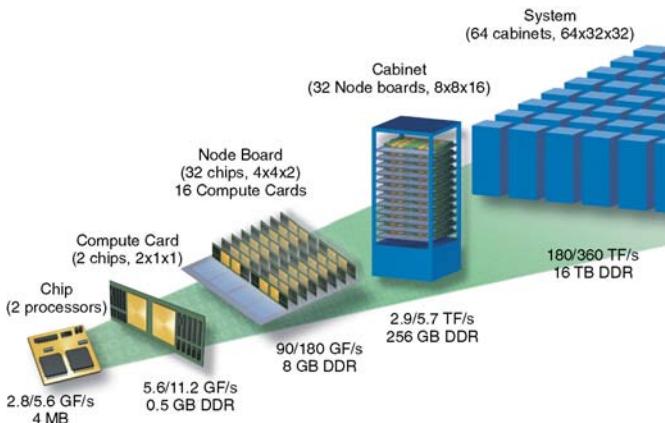


Fig. 21.2. Building BlueGene/L. A node ASIC contains 2 CPUs. 2 node ASICs along with their associated local external memory are built onto a compute card. A node board contains 16 compute cards. 32 node boards are then plugged into both sides of 2 mid-plane boards in a cabinet. A large BlueGene/L system contains 64 cabinets, forming a $32 \times 32 \times 64$ torus of compute nodes, with a peak performance of 180/360 Tflops

mance application development. MPI libraries are supported as the standard parallel programming model.

21.3 Project Status Update

During the design process, the BlueGene/L team has been doing extensive studies of a wide variety of applications [5]. The results of these studies show that a large class of applications should scale well on the BG/L architecture, even to the level of 65,536 nodes.

In November 2003, a half cabinet of 512 way BlueGene/L prototype from the 1st ASIC chip release (DD1) running at 500 MHz with a peak speed of 2 Teraflops was ranked as the 73rd fastest supercomputer in the world [7]. As of mid-2004, the number of nodes has been expanded. A 4096 node DD1 (500 MHz) 16.3 Teraflops peak prototype and a second 2048 node DD2 (700 MHz) 11.4 Teraflops peak prototype have been running in IBM labs. These two ‘small’ prototypes (compared to the final BlueGene/L machines) are already ranked among the top 10 fastest supercomputers in the world.

With the availability of the prototypes, a lot of application benchmarks from the BlueGene/L research partners has been run on the BlueGene/L hardware. As an example, the IBM Blue Matter team has studied the scaling of 3-d FFT code on the BlueGene/L prototype [8], and showed very good

speedup up to the 512 way system, the largest system available when the study was done.

21.4 Conclusion

BlueGene/L is a first step in IBM's commitment to petaflops scale computing by exploring a new architecture of building massively parallel machines. Utilizing system-on-a-chip technology and optimizing the design at the system level, great reduction in cost, power consumption and machine size over performance is achieved. With more computing power available, the performance and scalability of BlueGene/L will open up new possibilities for a large class of important scientific applications.

References

1. F. Allen et al.: Blue Gene. IBM System Journal **40**, 310 (2001)
2. IBM Research BlueGene project web page, www.research.ibm.com/bluegene"
3. D. Chen et al.: Nucl. Phys. B (Proc. Suppl.) **73**, 898 (1999). R.D. Mawhinney: Parallel Comput. **25**, 1281 (1999)
4. N.H. Christ et al.: QCDOC, A 10-teraflops scale computer for lattice QCD. Nucl. Phys. B (Proc. Suppl.) **94**, 825 (2001)
5. G. Almasi et al.: Cellular Supercomputing with System-On-A-Chip. Proc. IEEE International Solid-State Circuits Conf. (2002)
6. BG/L Team: An Overview of the BlueGene/L Supercomputer. SuperComputing 2002 Technical Papers, November (2002)
7. Top 500 supercomputer list: www.top500.org
8. M. Eleftheriou, J.E. Moreira, B.G. Fitch, R.S. Germain: A Volumetric FFT for BlueGene/L'. Lecture Notes in Computer Science, Vol. 2913 (Springer, Berlin Heidelberg New York 2003) pp. 194–203

Part VI

Molecules, Clusters and Nanoparticles

22 All-Electron Path Integral Monte Carlo Simulations of Small Atoms and Molecules

J. Shumway

Department of Physics and Astronomy, Arizona State University
Tempe AZ 85283, USA

Abstract. Imaginary-time Feynman path integrals describe quantum mechanics at finite temperatures. Monte Carlo simulations using path integrals have many applications to nanoscale physics. Fermions introduce negative terms in the partition function, which we remove with a fixed-node approximation. In this article, we give an introduction to path integral Monte Carlo simulations. We focus on a difficult problem: the direct application of fermion path integrals to light atoms and molecules. Work on this problem illustrates the technical challenges facing path integral simulations, and may lead to new *ab initio* methods.

22.1 Introduction: Path Integral Theory

In classical mechanics, the physical trajectory of a particle $\mathbf{r}_{\text{cl}}(t)$ is the unique function that minimizes the action,

$$S[\mathbf{r}(t)] = \int_{t_a}^{t_b} \frac{1}{2} m \dot{\mathbf{r}}^2 - V(\mathbf{r}) dt, \quad (22.1)$$

where the quantity in the integrand is the Lagrangian, $\mathcal{L} = T - V$. Motivated by the work of Dirac, Feynman showed that the classical principle of least action follows from the path integral representation of the quantum propagator,

$$P(\mathbf{r}, t; \mathbf{r}', t') \equiv \langle \mathbf{r} | e^{-i\hat{H}(t-t')} | \mathbf{r}' \rangle = \int \mathcal{D}\mathbf{r}(t) \exp \frac{i}{\hbar} S[\mathbf{r}(t)], \quad (22.2)$$

where the path integral sums *all* paths connecting the starting point (\mathbf{r}', t') to the ending point (\mathbf{r}, t) . The classical limit of (22.2) follows naturally from the stationary phase approximation: the action of classical paths are much greater than Plank's constant \hbar , so the phases of nearby paths fluctuate rapidly, leaving paths near the extremum of the action to dominate the integral.

Direct application of the real-time path integral, (22.2), is usually impossible because of the rapidly fluctuating phase. Switching to imaginary time solves this problem by turning all the phases into positive, real-valued weights. Remarkably, this is a sensible thing to do, because imaginary time path integrals give us thermal averages.

To see the connection to thermodynamics, look at the thermal density matrix operator, $\hat{\rho}_\beta = e^{-\beta \hat{H}}$ (where $\beta = 1/k_B T$ is the inverse temperature), which is central to quantum statistical mechanics. In the basis of energy eigenfunctions $\{\Psi_i\}$, it is easy to see that thermal averages involve traces over the thermal density matrix,

$$\langle \hat{O} \rangle_\beta = \frac{1}{Z} \sum_i e^{-\beta E_i} \langle \Psi_i | \hat{O} | \Psi_i \rangle = \frac{1}{Z} \text{tr} \hat{O} \hat{\rho}_\beta , \quad (22.3)$$

where $Z = \text{tr} \hat{\rho}$ is the partition function.

The substitution $i \rightarrow \beta$ turns the real time propagator $e^{-i\hat{H}t}$ into the thermal density matrix $\hat{\rho}_\beta = e^{-\beta \hat{H}}$. It follows from (22.2) that the thermal density matrix is equal to an imaginary time path integral

$$\rho(\mathbf{r}, \mathbf{r}'; \beta) \equiv \langle \mathbf{r} | e^{-\beta \hat{H}} | \mathbf{r}' \rangle = \int \mathcal{D}\mathbf{r}(t) \exp \left(-\frac{1}{\hbar} S_E^\beta[\mathbf{r}(t)] \right) . \quad (22.4)$$

$\mathbf{r}(0) = \mathbf{r}'$
 $\mathbf{r}(\beta) = \mathbf{r}$

The action $S_E^\beta[\mathbf{r}(t)]$ is the Euclidean, or imaginary-time, action,

$$S_E^\beta[\mathbf{r}(t)] = \int_0^\beta \frac{1}{2} m \dot{\mathbf{r}}^2 + V(\mathbf{r}) dt , \quad (22.5)$$

where the quantity in the integrand is the Euclidean Lagrangian, $\mathcal{L}_E = T + V$. The trace of the density matrix is taken by setting $\mathbf{r} = \mathbf{r}'$ (closing the path) and summing over all paths. Thus, statistical quantum mechanics may be formulated as a weighted sum over closed paths [1]. This is quantum mechanics without Schrödinger's equation: a 3-dimensional quantum problem has been mapped onto the classical partition function of a continuous, polymer-like path.

The real power of the imaginary-time path integral, (22.4), is that it is directly applicable to interacting quantum many-body systems. While the exact solution to N interacting quantum particles might seem to require determination of a $3N$ -dimensional wavefunction satisfying a $3N$ -dimensional Schrödinger equation, the path integral formula turns the problem into a computationally tractable $3N$ -dimensional path integral.

Bose and Fermi statistics enter naturally into the path integral formalism. Let $R = (\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N)$ represent the $3N$ coordinates of our particles. The partition function for distinguishable particles is simply a sum over closed paths, $R(0) = R(\beta)$. If the particles are indistinguishable, the path integral must be symmetrized or antisymmetrized for Bose and Fermi statistics,

$$\rho_{B/F}(R, R'; \beta) = \sum_P (\pm 1)^P \int \mathcal{D}R(t) \exp \left(-\frac{1}{\hbar} S_E^\beta[R(t)] \right) , \quad (22.6)$$

$R(0) = R'$
 $R(\beta) = PR$

where P is a permutation of particle coordinates. Thus, the partition function for identical quantum particles includes permuting paths, and these permutations are responsible for Bose and Fermi effects. The minus sign that occurs in the antisymmetrization of fermions is a distinctly quantum feature that generally prevents exact simulations of interacting fermions.

In this paper we discuss quantum Monte Carlo (QMC) calculations using path integrals, with an emphasis on fermion simulations. In Sects. 22.2–22.3 we summarize recent applications, and suggest reasons why all-electron calculations on atoms and molecules should be studied. In Sect. 22.4 we give a brief discussion of the simulation techniques, with an emphasis on new developments needed for atoms and molecules. The remaining sections of the article, Sects. 22.5–22.7, describe results of some recent tests and preliminary atomic and molecular calculations.

22.2 Recent Path Integral Simulations on Nanostructures

The imaginary-time path integral, (22.6), can be evaluated with Monte Carlo integration, as described in detail in the review by Ceperley [2]. For example, bosonic simulations have determined many properties of superfluid ^4He starting from the interatomic potential [2]. Fermionic simulations often require an additional “fixed-node” approximation [3–6] to remove the troublesome minus sign. For a recent review of QMC simulations in nanostructures (see [7]). To illustrate the utility of these simulations, we give a quick summary of some nanoscale applications.

He Nanodroplets

As a microscopic theory of ^4He superfluidity [2], the path integral QMC technique is a natural tool for studying nanoscale features in superfluid ^4He droplets. Sindzingre, Ceperley, and Klein simulated pure ^4He droplets in the late eighties [8], showing that droplets of 64 atoms are essentially superfluid. Later experiments have confirmed these theoretical predictions, and the study of ^4He droplets has become an exciting research area. Most interesting is the ability to suspend single molecules impurities in the superfluid drop, providing a novel environment for spectroscopy and other chemical studies. The simulations follow Feynman’s interpretation of the Bose–Einstein transition as a percolation of permuting bosonic paths. Permuting paths that encircle the droplet are a direct indicator of the superfluid response. At the surface of the molecular impurity, a layer structure is formed. Path integral simulations can predict the superfluid response of these layers at the temperature of the drops [9].

Self-assembled Dots

Path integrals and related QMC techniques have been used in effective-mass models of semiconductor quantum dots. The interest in these systems is to directly account for the few-body Coulomb interactions between electrons and holes in these dots. For example, the energy of an emitted photon from an electron-hole recombination is often shifted by several meV by correlation with other “spectator” electrons and holes in the dot [10, 11]. One of our recent applications is a low temperature (effectively ground state) calculation of recombination rates in dots [12]. The radiative recombination rates of interacting electron-hole pairs in a quantum dot are strongly affected by quantum correlations among electrons and holes in the dot. Recent measurements of recombination rates within biexcitons lie in the theoretically predicted range of one to two times the rate from single excitons, but the value depends on the degree of correlation. Theoretical approaches using effective mass wave functions have been unable to accurately determine biexciton recombination rates for a realistic, material-specific model. We have developed a path-integral formulation of the problem that allows the direct evaluation of the recombination rate, including important quantum-correlation effects. Using Monte Carlo simulations, we have studied recombination rates for realistic three-dimensional models of InGaAs/GaAs, CdSe/ZnSe, and InP/InGaP dots.

Electrons in Quantum Point Contacts

We are beginning to study electrons in quantum point contacts. The fixed-node fermion algorithm described in Sect. 22.4.3 scales well enough with temperature and system size that we can simulate hundreds of electrons in a gated structure at room temperature. We are in the process of adding Coulomb interactions and calculating correlation functions. Our goal is to study the effects of electron-electron interaction and Fermi statistics on quantum transport, especially at the onset of single-electron conductivity. The path integral QMC technique allows us to explicitly include the interacting two-dimensional electron gas that is present in the device contacts. This may be a general technique for other nanoscale systems that have open boundary conditions.

22.3 Motivation for Atomic and Molecular Calculations

There are several reasons to attempt room-temperature fermionic path integral calculations on atoms and molecules. If such calculations are feasible, this could be a high-accuracy finite-temperature ab initio method. The thermal and zero-point motion of the nuclei could be treated simultaneously with electron sampling. In fact, even the equilibrium atomic positions could

be thermally sampled without the explicit calculation of forces. Ultimately, the finite-temperature many-body formulation of electronic structure calculations may be a good framework for ab initio simulations of magnetism or superconductivity.

From a less-ambitious, skeptical point of view, atomic and molecular simulations provide a challenging test for the path integral QMC method. The well-known energies, bond-lengths, and other properties of simple atomic and molecular systems are good benchmarks for quantifying the accuracy and precision of current fermion path integral simulations.

Finally, it is fun and intriguing to see if well-known atomic and molecular properties emerge spontaneously from numerical evaluation of the ab initio partition function, (22.6).

22.4 Monte Carlo Simulation Technique

To numerically evaluate the thermal path integral, we discretize imaginary time into M slices of length $\tau = \beta/M$. This time step has units of inverse energy. For the atomic and molecular calculations presented here we take $\tau = 0.01 \text{ Ha}^{-1}$, giving $M = 10^5$ slices for room temperature ($T = 0.001 \text{ Ha}$) calculations. Thus, the path configuration for N particles can be stored as $N \times M$ points, plus N integers to represent the global permutation. Because of the very high-dimensionality of the discretized path integral (even for a single particle) we use Monte Carlo integration to sample the integral.

22.4.1 Monte Carlo Basics

The key equation for Metropolis Monte Carlo is the acceptance probability [13, 14],

$$A(S' \rightarrow S) = \min \left[1, \frac{\pi(S)T(S \rightarrow S')}{\pi(S')T(S' \rightarrow S)} \right], \quad (22.7)$$

where S is the state of the system, $\pi(S)$ is the probability distribution to be sampled, an $T(S' \rightarrow S)$ is the probability of attempting a transition from S' to S . To use this equation: (i) initialize the state of the system, (ii) make a trial move of the system, keeping track of the probabilities $T(S' \rightarrow S)$ and $T(S \rightarrow S')$ for the forward and reverse moves, (iii) calculate the old and new probabilities $\pi(S')$ and $\pi(S)$, (iv) change the system state to the trial move with a probability $A(S' \rightarrow S)$ given by (22.7), (v) collect data on the current system state, and (vi) repeat from step (ii). The collected data is then statistically analyzed.

In practice, one wishes to construct large moves, but avoid spending too much time on moves that will be rejected. In that case, a multilevel sampling scheme is preferable. A move, described in steps (ii)–(iv) above, can be done in multiple levels ($k = N_{\text{level}}$ down to 0), each with acceptance probability,

$$A_k(S' \rightarrow S) = \min \left[1, \frac{\pi_k(S)T_k(S \rightarrow S')\pi_{k+1}(S')}{\pi_k(S')T_k(S' \rightarrow S)\pi_{k+1}(S)} \right]. \quad (22.8)$$

The highest level crudely (and cheaply) rejects bad steps; and, as long as all the T_k are non-zero, the algorithm will sample π_0 . For path integrals, we can sample a section of $2^{N_{\text{level}}} - 1$ beads by first sampling a permutation for the level $k = N_{\text{level}}$, the midpoint of the section for level $k = N_{\text{level}} - 1$, then the midpoints of the subsections for lower k , and finally the finest resolution for $k = 0$. For a more detailed description of multilevel sampling of paths (see [2, 7]).

22.4.2 The Coulomb Action

The action for has contributions from the kinetic and Coulomb energies. To avoid problems with Coulomb singularities [15], one can define the action for a discretized slice to be the negative log of the density matrix satisfying the Bloch equation,

$$S = -\log \rho(R, R'; \tau); \quad \text{where} \quad H\rho(R, R'; \tau) = -\partial_\tau \rho(R, R'; \tau). \quad (22.9)$$

We make a further approximation, separating the action into free particle and pair interaction terms, $S = S_0 + \sum_{i < j} S_{ij}^{\text{coul}}$, which neglects three-body and higher order terms. The error can be controlled by decreasing the time step τ .

To determine the Coulomb action terms, we need to find the thermal density matrix for the Coulomb potential,

$$\left(-\frac{1}{2\mu_{ij}} \nabla^2 + \frac{q_i q_j}{r_{ij}} \right) \rho_{ij}^{\text{coul}}(\mathbf{r}_{ij}, \mathbf{r}'_{ij}; \tau) = -\partial_\tau \rho_{ij}^{\text{coul}}(\mathbf{r}_{ij}, \mathbf{r}'_{ij}; \tau), \quad (22.10)$$

where μ_{ij} is the reduced mass and q_i and q_j are the charges. Then

$$S_{ij}^{\text{coul}}(\mathbf{r}, \mathbf{r}'; \tau) = -\log \rho_{ij}^{\text{coul}}(\mathbf{r}, \mathbf{r}'; \tau) - S_0(\mathbf{r}_i, \mathbf{r}'_i; \tau) - S_0(\mathbf{r}_j, \mathbf{r}'_j; \tau). \quad (22.11)$$

Much like choosing a propagator in molecular dynamics, it is essential to have a very accurate expression for the short-time pair Coulomb action $S_{ij}^{\text{coul}}(\mathbf{r}, \mathbf{r}', \tau)$. There are several ways to numerically evaluate the Coulomb density matrix: (i) use a semiclassical expansion or other high temperature approximation [16], (ii) use density matrix squaring [17], (iii) expand in the bound and continuum states of the hydrogen atom [18], (iv) use an analytic power-series solution to the problem [19], or (v) repeatedly apply the kinetic and potential operators to a discrete radial grid, using fast Fourier transforms for efficiency [20]. In the present calculations, we use a careful combination of methods (i), (iv), and (v) to tabulate and store results on a logarithmic radial grid; this gives energies accurate to better than 0.01 Ha for one-electron ions with $Z \leq 10$ for a small timestep, $\tau = 0.01 \text{ Ha}^{-1}$.

22.4.3 Fixed-Node Fermion Approximation

The most crucial part of these fermionic calculations is the fixed-node approximation [3–6]. We desire the following features for our approximation

- It should eliminate the negative probability contributions to the fermionic integral.
- It should have a variational property, so that we can determine which approximation choices are more accurate.
- It should be improvable, in that the better we understand our system, the more accurate we can make the calculations.

Here we present an approximation that removes negative contributions, and may meet the other criteria. The remainder of the paper is devoted to this fixed node approximation.

In ground state QMC calculations, one typically uses a fixed node approximation [3, 4, 21] based on some trial function, Ψ_T , often constructed from a Slater determinant of Hartree–Fock or Kohn–Sham orbitals. During the simulation, random walks are not allowed to cross the nodes of the trial function. The approximation removes negative contributions to the random walk, is variational, and would be exact if we knew the nodes of the true ground-state wavefunction. For finite temperature, we no longer have a wavefunction, but rather have a thermal density matrix. While the ground state fixed-node approximation could be applied to a path integral calculation, it is a very poor approximation at finite temperatures [5].

With a density matrix $\rho_T(R, R')$, the restriction now must depend on two slices, R and R' . There are different ways to apply this restriction. Ceperley [5] introduced a *reference slice* R_* , a special slice at $t = 0$ that anchors the nodal restriction. At any other slice on the path, at imaginary time t ($-\beta/2 < t < \beta/2$), the quantity $\rho_T(R, R_*; |t|)$ is required to be non-zero. (Node crossing between slices is also forbidden, which can be difficult to enforce if τ gets to large). There are three complications with this approach: (i) The nodes must be time dependent and vary smoothly from infinite temperature down to twice the simulation temperature. (ii) The restriction breaks time translational symmetry along the path, a significant difference from the standard fermionic path integral (22.6). This has the practical consequence that estimators should be evaluated only at the references slice, rather than over the entire path. (iii) The reference is a global constraint, so all successful trial moves must refer back to the nodal slice to check for node crossing. Even worse, moves of the reference slice become very difficult at low temperatures, effectively preventing fermion path integral QMC applications to atomic and molecular problems at temperatures below 5000 K [6].

Here we introduce an alternative formulation of the nodal constraint. For any time t along the path, there is a time $t' = (t + \beta/2)\text{mod}\beta$ that lies opposite of t . We require that $\rho_T(R(t), R(t'); \beta/2) \neq 0$ for all times t along the path. For a discretized path with finite timestep τ , care must be taken

to ensure that the path did not cross the node between the discrete slices. In our calculations we prevent node crossing by choosing a small timestep $\tau = 0.01$ Ha and including the repulsive nodal action [6]. Note that the modulo β choice of t' could lead to an ambiguity of a permutation in $R(t')$. It turns out that only even permutations are allowed. Since the fermionic trial density matrix ρ_T is invariant under an even permutation, this ambiguity in the permutation is not a problem.

Now we consider the merits of our new nodal formulation. First it can be seen that it eliminates the negative contributions to the fermionic integral: any negative permutation of closed paths will have some times t_1 and t_2 for which $\rho_T(R(t_1), R(t'_1); \beta/2)$ and $\rho_T(R(t_2), R(t'_2); \beta/2)$ have opposite sign. If ρ_T is the exact density matrix for the inverse temperature $\beta/2$, then the only paths that are eliminated did not contribute to $\rho(R, R; \beta) = \int \rho^2(R, R'; \beta/2) dR'$ anyway. We suspect the restriction is variational in free energy, but we do not yet have a rigorous proof. In comparison to the formulation of Ceperley [5], we see that this new formulation is a great simplification and is bilocal in time. This enables new calculations at room temperatures and more efficient parallel implementations.

22.5 Examples and Tests for Non-Interacting Fermions

Since this is a new formulation of the fermion restriction, we have conducted several tests of the nodal constraint, which we present in this section.

22.5.1 Antisymmetric States of a Harmonic Oscillator

Before treating Fermi systems, we illustrate the fixed-node constraint on a much simpler system. A path integral simulation can easily sample the partition function of a quantum harmonic oscillator, $\text{tr} \rho(x, x'; \beta)$. Suppose you only wanted the antisymmetric contribution, $\rho_A(x, x'; \beta) = \frac{1}{2}[\rho_A(x, x'; \beta) - \rho_A(-x, x'; \beta)]$. You could explicitly antisymmetrize the path integral by allowing both $x = x'$ or $x = -x'$ when taking the trace and introducing a negative weight for the $x = -x'$ terms. Alternatively, you could recognize that all antisymmetric wavefunctions are zero at the origin, and throw away all paths that cross the origin. This restriction is equivalent to both the ground state and the finite temperature nodal restrictions.

Now consider a two-dimensional oscillator. The first excited state is doubly-degenerate. One could enforce a ground state restriction by placing a node at $x = 0$. Then you sample a density as illustrated in Fig. 22.1a. Alternatively, one could place the node at $y = 0$, resulting in the density illustrated in Fig. 22.1b. Ground state nodes assume a unique wavefunction, and a density matrix is needed to represent both degenerate nodal choices equally.

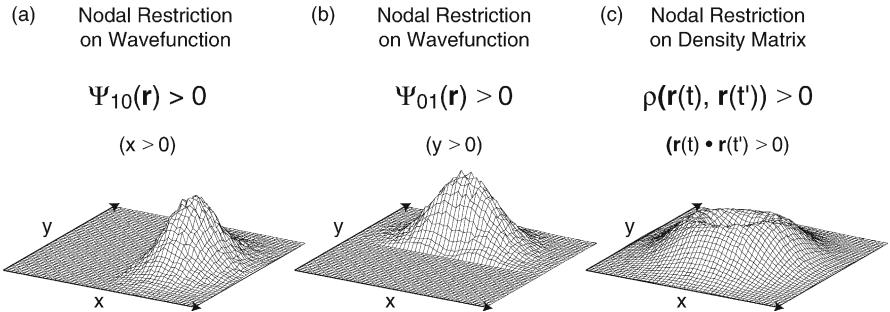


Fig. 22.1. Probability distribution for first excited state of a harmonic oscillator using different nodal restrictions: (a) wavefunction restriction $\psi_{1,0}(\mathbf{r}) > 0 \Rightarrow x > 0$, (b) wavefunction restriction $\psi_{0,1}(\mathbf{r}) > 0 \Rightarrow y > 0$, (c) density-matrix restriction $\rho(\mathbf{r}(t), \mathbf{r}(t')) > 0 \Rightarrow \mathbf{r}(t) \cdot \mathbf{r}(t') > 0$. The density-matrix restriction allows sampling of both degenerate states; for finite temperature the density-matrix restriction allows sampling of all low-energy excitations

For this two-dimensional oscillator, our density-matrix-based nodal approximation reduces to the requirement that $\mathbf{r}(t) \cdot \mathbf{r}(t') \neq 0$ for all t . The densities for this case are shown in Fig. 22.1c. This Fermi constraint allows both degenerate states to contribute equally to the simulation.

22.5.2 Free Fermi Gas

We have tested 27 spin-polarized (identical) fermions in a cubic box with side length $10a_0$. We choose this density to approximate the density of valence electrons in typical semiconductors. The results are shown in Fig. 22.2. We have compared this to an equivalent bosonic simulation, to clearly show the effect of the fermionic restriction. We have also indicated the classical energy per particle, $3/2 kT$, and the low temperature energy per particle of bulk fermions at this density. We see that the energy of the fermions with the density-matrix nodal restriction smoothly crosses over from the low temperature to high temperature limits.

22.5.3 Free Fermi Gas in a Sphere

Note that the free particle density matrix is exact for free fermions at any density. The free particle density matrix must also be appropriate for systems with slowly varying density. Next we tested the applicability of the free particle density matrix for an inhomogeneous system of free particles confined in a sphere. For 10 fermions (5 up and 5 down) confined in a sphere with radius $2a_0$, we should find a ground state energy of 25.915 Ha. Our path integral calculations with $\tau = 0.01$ gives 24.30 ± 0.36 Ha. The error is about what we expect from the finite time step and the hard spherical confinement, a

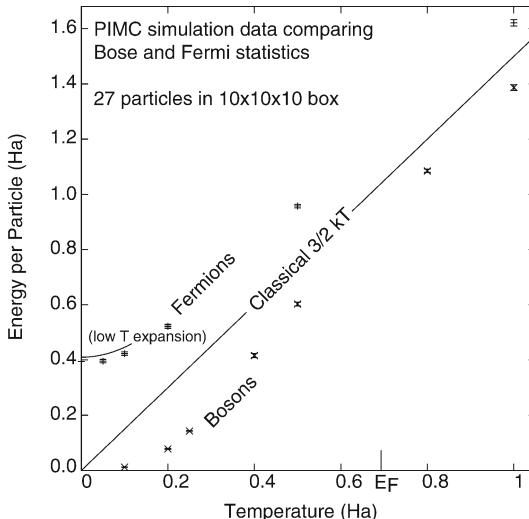


Fig. 22.2. Test of nodal restriction on 27 free particles. We plot the energy per particle for bosons and fermions and compare with the classical equipartition theorem energy, $3/2k_B T$. These data show that this fermion restriction smoothly crosses over from the high-temperature classical to low-temperature regime. Low temperature curve is the thermodynamic limit for free fermions at this density

trend supported with a second calculation with a smaller timestep. We will need to rerun the system with a more accurate action and conduct time-step analysis to quantify magnitude of the fixed node error for this system. For now we know that the error is smaller than the fermionic shell structure in this system.

22.6 Calculations on Atoms

We have performed calculations on atoms from H to Ne at room temperature. These test of fixed-node path integral electronic structure can be directly compared to well-known ground state results from Hartree-Fock, fixed-node diffusion QMC, density functional theory, and experiments.

22.6.1 Atomic Simulation Issues and Strategies

When testing calculations on atoms, several issues arise. First, the Coulomb action must be accurately tabulated, as we noted earlier. As a necessary condition for accuracy, we require that calculations of single electron ions reproduce the analytical energies $E = -0.5 Z^2 \text{ Ha}$ and $V = -Z^2 \text{ Ha}$ to within at least 5 mHa. We also tested the energy of He to similar accuracy. Future calculations may require even higher accuracy.

Table 22.1. Ionization potentials calculated by PIMC with comparison to experimental ionization energies from [22]. The $2s$ electrons in Li and Be (appearing in Li-I, Be-I, and Be-II) are significantly underbound in the free-particle nodal restriction used in the PIMC calculations

Element	Spectrum			
	I	II	III	IV
H	PIMC: 0.499(11) exp.: 0.4997			
He	PIMC: 0.9053(8)	2.000		
	exp.: 0.9038	1.9998		
Li	PIMC: 0.117(10)	2.781(12)	4.494(8)	
	exp.: 0.1982	2.7796	4.5000	
Be	PIMC: 0.111(40)	0.518(40)	5.657(17)	7.985(7)
	exp.: 0.3426	0.66924	5.6555	8.0008

Second, we must consider entropic ionization. To avoid the infinite entropy of the extended vacuum, we have put the atoms in small spherical hard-walled enclosures, with radii ranging from five to eight Bohr radii. While eight Bohr radii should be sufficient to prevent room temperature ionization for most atoms, fixed node errors can exacerbate the ionization problem. That is, the fixed node approximation may cause the ionization energies to be artificially low, leading to partial ionization and significant departure from ground-state behavior, even at room temperature.

Finally, we note that relaxation of the paths to thermal equilibrium may be accelerated by first relaxing the energy at high-temperature using a small confinement radius, then restarting the simulation at a lower temperature and larger confinement radius. Note that the form of our fixed node restriction is such that dropping the temperature by factors of two by duplicating relaxed paths in imaginary time will always satisfy the nodal constraint.

22.6.2 Results for Atoms

We quickly tested the method on atoms from H to Ne. We found a serious 10% error in total energy that underestimates the binding. Plots of path distributions revealed that the outer shell electrons were weakly bound and have a tendency to ionize. We believe the 10% errors are too large for most applications, since Hartree–Fock (HF) and the local density approximation to density function theory (LDA-DFT) have total energy errors that are closer to 3%.

To establish the source of these total energy errors, we have conducted careful room-temperature ($T = 0.001$ Ha) calculations on atoms and anions with $Z = 1\text{--}4$. Our calculated total energies are listed in Table 22.1, along with the experimental values. The columns indicate the electron being re-

Table 22.2. Total energies, in Hartrees. The PIMC calculations are restricted with free-particle and Hartree–Fock nodes. We find that the Hartree-Fock nodes give better results for Li and Be that are similar to FN-DMC with Hartree-Fock nodes

Element	E_{PIMC}					
	Free Nodes	HF Nodes	E_{LDA}	E_{HF}	$E_{\text{FN-DMC}}$	E_{exact}
H	−0.499(11)	−0.499(11)	−4.7867	−0.4997	−0.4997	−0.4997
He	−2.9053(8)	−2.9053(8)	−2.8348	−2.8617	−2.9040(13)	−2.9037
Li	−7.404(12)	−7.467(20)	−7.3440	−7.4327	−7.4780(11)	−7.4781
Be	−14.280(17)	−14.615(17)	−14.4472	−14.5730	−14.6579(14)	−14.6673

moved, i.e., column I is the energy to remove the first (outermost) electron. The diagonal of the table is the energy of the one electron ion. We see first of all that the energies of the ions with only 1s electrons are in excellent agreement with the experiments. (Note that the experimental data show some small relativistic effects that enhance bonding for Li and Be 1s states – these relativistic effects are *not* included in our simulations.) In contrast, the 2s electrons in Li and Be are bound too weakly, with binding energies about one-half the experimental value. We believe the source of this error is our choice of free particle nodes.

To test this hypothesis, we performed total energy calculations on Li and Be using Hartree–Fock nodes. The spherical symmetry of the Li and Be atoms reduces the Hartree–Fock nodal surface to the simple condition $|\mathbf{r}_1| = |\mathbf{r}_2|$, where \mathbf{r}_1 and \mathbf{r}_2 are the positions of two like-spin electrons relative to the nucleus. We have also performed fixed node calculations with a ground state method, fixed node diffusion Monte Carlo (FN-DMC). In Table 22.2 we see that the Hartree–Fock nodes improve the total energy to about the value found in single-determinant FN-DMC calculations, suggesting that it the free particle nodal surfaces were the source of error. For comparison, we have also listed the LDA-DFT, HF, and exact total energies in Table 22.2. Our conclusion is that free particle nodal restriction is too severe an approximation for most applications.

22.7 Calculations on Molecules

Molecules allow tests of an important question: can fixed-node path integral calculations determine geometric structure, such as optimal bond lengths? While failings of free-particle fixed nodes on atoms must be resolved before one expects accurate bonding calculations, we have run some preliminary tests on LiH and CH₄.

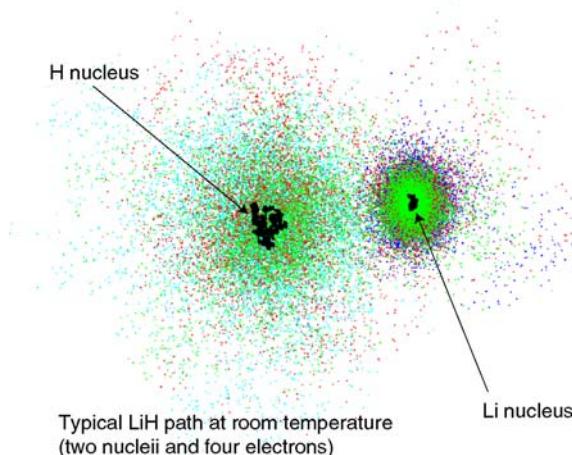


Fig. 22.3. Scatter plot of a path configuration from a simulation of an LiH molecule at room temperature. The relative sizes of the large H cation and smaller Li anion are apparent, as the quantum thermal wavelength of the Li and H nuclei

22.7.1 Molecular Simulation Issues and Strategies

The question of geometry and structure can be independent of total energy calculations. For example, LDA-DFT makes significant errors in total energy and band-gaps of bulk Si, yet calculates the equilibrium structure and phonon dispersion curves quite well. Thus tests of geometry calculation are important even if total energy calculations fall below chemical accuracy. Another reason for early tests on molecules is to get an idea about structure optimization ability: if we are able to improve our fixed-node approximation, will thermal simulations efficiently relax to equilibrium geometries?

One new strategy we tried was to speed up the initial equilibration process by starting from electrons already relaxed to a closed shell configuration. For example, we initialized our CH_4 simulations by starting from Ne atom paths, then changing the Ne nucleus to C and placing four bare protons at the edge of the electron cloud. Such strategies work well and may be useful in future calculations on larger molecular systems.

22.7.2 Results for Molecules

Our first calculations on LiH and CH_4 failed, but we think the method shows promise. We started our calculations away from equilibrium and allowed our nuclei to move. We found that the electrons and nuclei moved toward the expected structures, but eventually ionized and fell apart. In Fig. 22.3 we show a scatter plot of a path configuration of a room temperature LiH simulation.

We note the following correct features: (i) the molecule is highly ionic with a large H cation and smaller Li anion, (ii) the extended paths of the Li and H molecule (reflecting the thermal wavelength $\lambda_\beta = h/\sqrt{2\pi mk_B T}$) show the significance of zero-point motion in addition to thermal vibrations, (iii) quantitative measurement of bond lengths and dipole moments were within about 10% of the expected values, but they failed to converge, and eventually the molecule fell apart, and (iv) the paths can exchange between molecules, which is a feature of our fixed-node formalism not present in single-determinant ground state nodes. Calculations on CH₄ indicated that the simulations relax to the tetragonal symmetry, but again the system falls apart for long simulation times. The ionization and dissociation seems consistent with the weak binding of 2s electrons we found in our atomic test. We plan to study molecular calculations more once we have improved or nodes beyond the free-particle formalism. In summary, we think that converged room temperature calculations of molecules with geometry optimization are feasible, but free-particle fixed node errors prevent accurate and converged calculations on our test cases.

22.8 Conclusion and Future Work

We have developed a new formulation for fixed node constraints, and have used it for test calculations on atoms and molecules at room temperature. Initial tests on atoms from H to Ne and on LiH and CH₄ molecules indicate that the method does not yet have the accuracy for practical atomic and molecular calculations. We wish to stress that 1s states are essentially correct, and we believe that current H and He *high-temperature* simulations are reasonably accurate [23], especially when improved nodes are used [24]. Our careful tests on ionized states of Li and Be indicate that the $n = 2$ shell electrons are too weakly bound to the atom. The free particle nodes are the likely source of this error, and tests with Hartree-Fock nodes support this hypothesis.

The next step in this research will be to introduce effects of Coulomb interactions into the nodal structure. Since our free-particle nodes lead to bound 2s states, it may be possible that perturbative corrections to the nodes that account for Coulomb interactions are accurate enough to advance the method. Establishing practical constraints for nodal choices will require further systematic tests, similar to the Li and Be ion study presented here. We propose as a target to judge success of new nodal strategies for path integrals the determination of LiH, CH₄, and H₂O molecule geometries at or above LDA-DFT accuracy.

Acknowledgement

Work supported by DARPA/ONR N00014-01-1-1062 and NSF DMR-0239819.

References

1. R.P. Feynman: Statistical Mechanics (Addison-Wesley, Reading, MA, 1972)
2. D.M. Ceperley: Rev. Mod. Phys. **67**, 279 (1995)
3. J.B. Anderson: J. Chem. Phys. **63**, 1499 (1975)
4. J.B. Anderson: J. Chem. Phys. **65**, 4122 (1976)
5. D.M. Ceperley: Phys. Rev. Lett. **69**, 331 (1992)
6. D.Ceperley, in: Monte Carlo and Molecular Dynamics of Condensed Matter Systems, ed. by K. Binder, G. Ciccotti (Editrice Compositori, Bologna 1996)
7. J. Shumway, D.M. Ceperley, in: Handbook of Theoretical and Computational Nanotechnology, ed. by M. Rieth, W. Schommers (American Scientific Publishers, 2005) In press, article preprint online at <http://phy.asu.edu/shumway/papers/nanorev.pdf>
8. P. Sindzingre, M.L. Klein, D.M. Ceperley: Phys. Rev. Lett. **63**, 1601 (1989)
9. E.W. Draeger, D.M. Ceperley: Phys. Rev. Lett. **90**, 065301 (2003)
10. E. Dekel, D. Gershoni *et al.*: Phys. Rev. Lett. **80**, 4991 (1998)
11. J. Shumway, A. Franceschetti, A. Zunger: Phys. Rev. B **63**, 155316 (2001)
12. M. Wimmer, J. Shumway, S.V. Nair: (2004) Submitted to Phys. Rev. B
13. N. Metropolis, A.W. Rosenbluth *et al.*: J. Chem. Phys. **21**, 1087 (1953)
14. D.P. Landau, K. Binder: A Guide to Monte Carlo Simulations in Statistical Physics (Cambridge University Press, Cambridge 2000)
15. H. Kleinert: Path Integrals in Quantum Mechanics, Statistics, and Polymer Physics (World Scientific, Singapore 1990)
16. P. Vieillefosse: J. Stat. Phys. **74**, 1195 (1994)
17. R.G. Storer: J. Math. Phys. **9**, 964 (1968)
18. E.L. Pollock: Comp. Phys. Comm. **52**, 49 (1988)
19. P. Vieillefosse: J. Stat. Phys. **80**, 461 (1994)
20. K.E. Schmidt, M.A. Lee: Phys. Rev. E **51**, 5495 (1995)
21. P.J. Reynolds, D.M. Ceperley, B.J. Alder, W.A. Lester: J. Chem. Phys. **77**, 5593 (1982)
22. C.E. More: Ionization potentials and ionization limits derived from the analysis of optical spectra. Tech. Rep. 34, National Bureau of Standards (1970)
23. B. Militzer, D.M. Ceperley: Phys. Rev. E **63**, 066404 (2001)
24. B. Militzer, E.L. Pollock: Phys. Rev. E **61**, 3470 (2000)

23 Projective Dynamics in Realistic Models of Nanomagnets

S.H. Thompson^{1,2}, G. Brown^{1,3}, and P.A. Rikvold^{1,2}

¹ School of Computational Science and Information Technology,
Florida State University, Tallahassee, FL 32306-4120, USA

² Center for Materials Science and Technology and Department of Physics,
Florida State University, Tallahassee, FL 32306-4350, USA

³ Center for Computational Sciences, Oak Ridge National Laboratory,
Oak Ridge, TN 37831-6164 USA

sam@csit.fsu.edu, browngrg@csit.fsu.edu, rikvold@csit.fsu.edu

Abstract. The free-energy extrema governing the magnetization-reversal process for a model of an iron nanopillar are investigated using the projective dynamics method. Since the time evolution of the model is computationally intensive, one could question whether sufficient statistics can be generated, with current resources, to evaluate the position of the metastable configuration. Justification for the fine-grained discretization of the model that we use here is given, and it is shown that tractable results can be obtained for this system on realistic time scales.

23.1 Introduction

Magnetic nanopillars can be fabricated [1], which offer interesting opportunities for direct comparison with numerical models. In the application discussed here, the pillars are grown perpendicular to a surface. For an applied field parallel to the long axis of the pillar, there exists a stable free-energy minimum when the field is parallel to the average magnetization of the pillar, and a metastable minimum for the antiparallel orientation. Magnetization reversal involves a transition from the metastable minimum to the stable minimum across a saddle point in the free energy. It is the free-energy difference between the saddle point and the metastable minimum that determines the switching time of the pillar. While determining the magnetization of the metastable minimum is comparatively easy, the magnetization of the saddle point is more difficult to determine. Projective Dynamics (PD) is a technique that has proven capable of finding the saddle point in other magnetization-reversal systems [2, 3].

23.2 Model and Numerical Results

In order for the proper dynamics to emerge, the magnetization of the simulated pillar must be discretized on a sufficiently fine scale. Pillar models that have previously been studied by projective dynamics consisted only of

a one-dimensional chain of spins [3]. This model lacks richer dynamics, such as magnetization curling modes, that are seen in experiments and in more realistic models. In addition, the dependence of the switching field on the angle of misalignment between the pillar axis and the applied field observed in the one-dimensional model [4] does not correspond to that seen in the experimental pillars [5–7].

To achieve more realistic dynamics, a model pillar with physical dimensions $10 \times 10 \times 150$ nm was discretized onto a $6 \times 6 \times 90$ cubic lattice. The model parameters were chosen to be consistent with bulk iron: exchange length of 2.6 nm and magnetization density of 1700 emu/cm³. The temperature was fixed at 20 K.

The dynamics of the system are governed by the Landau–Lifshitz–Gilbert (LLG) equation [8],

$$\frac{d\mathbf{M}(\mathbf{r}_i)}{dt} = \frac{\gamma_0}{1+\alpha} \left(\mathbf{M}(\mathbf{r}_i) \times \left[\mathbf{H}(\mathbf{r}_i) - \frac{\alpha}{M_s} \mathbf{M}(\mathbf{r}_i) \times \mathbf{H}(\mathbf{r}_i) \right] \right), \quad (23.1)$$

which describes the time evolution of the magnetization in the presence of a local field. Here γ_0 is the electron gyromagnetic ratio with a value of 1.67×10^7 Hz/Oe, and α is a phenomenological damping parameter, chosen as 0.1 to give underdamped behavior. $\mathbf{M}(\mathbf{r}_i)$ is the magnetization density at position \mathbf{r}_i , with a constant magnitude M_s , and $\mathbf{H}(\mathbf{r}_i)$ is the total local field at \mathbf{r}_i . The latter contains contributions from the dipole-dipole interactions, exchange interactions, and the applied field.

A reversal simulation begins with the magnetization oriented along the long axis of the pillar in the $+z$ direction. It is allowed to relax in the presence of a field $+B_0\hat{z}$ before the field is quickly switched to $-B_0\hat{z}$, where $B_0 = 1125$ Oe has been chosen smaller than the coercive field.

Projective dynamics involves analyzing the probability of growing toward the stable equilibrium or shrinking away from it, as a function of a single variable that measures the progress of the reversal process. To collect growing and shrinking statistics here, the z -component of the total magnetization of the pillar, M_z , was recorded at each integration step. In addition, the change in this value from the previous integration step, ΔM_z , was also recorded. The magnetization axis was discretized into bins into which the collected data were sorted. The bin size was determined such that the discretized ΔM_z would only involve changes between adjacent bins. The growing and shrinking probabilities were calculated by counting the number of times that ΔM_z would cause the magnetization to change bins. Define G as the number of times the magnetization hopped to the next bin in the direction toward equilibrium, and S as the number of hops in the direction back toward the metastable state. If N is the total number of visits to the starting bin, then the growing and shrinking probabilities for each bin, P_G and P_S , can be written respectively as

$$P_G = \frac{G}{N}, \quad P_S = \frac{S}{N}. \quad (23.2)$$

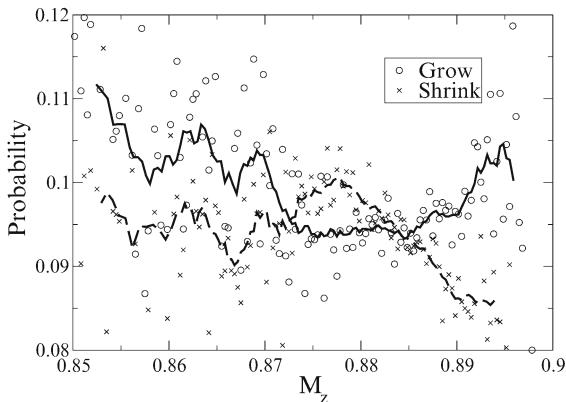


Fig. 23.1. Growing and shrinking probabilities versus the magnetization along the z -axis for a $6 \times 6 \times 90$ pillar. Grow/shrink data were collected into 6000 bins, which span the entire magnetization axis, $[-1, 1]$. The simulation was run at 20 K, 1125 Oe, and the applied field was directed along the long axis of the pillar, antiparallel to the initial magnetization. The continuous curves represent nine-point running averages

The crossings of the growing and shrinking probabilities give information about the locations of the metastable free-energy minimum and the saddle point. Where the two probabilities are equal, the system has no preferred direction of magnetization evolution, indicating that the free-energy landscape is flat in this region. This is the location of an extremum. The right-hand intersection in Fig. 23.1 corresponds to the metastable free-energy minimum, where the system spends most of its time. At the saddle point, the growing and shrinking probabilities are also equal, as indicated by the left-hand crossing. The region between these two crossings shows a higher probability for shrinkage than growth, corresponding a preference for moving toward the metastable minimum.

Previous simulations performed on a one-dimensional chain of seventeen spins exhibited Stoner–Wohlfarth behavior for the angular dependence of the switching field [4]. This is not representative of experimental data [5–7]. However, a full three-dimensional model qualitatively produced the proper angular dependence [4]. This is not surprising, given that the full three-dimensional model allows for magnetization curling modes which are simply not allowed in the one-dimensional model. PD has previously been applied to the one-dimensional spin chain [3], but several thousand switches had to be performed in order to accumulate sufficient statistics to reveal the free-energy extrema [3]. With typical switches taking around 1500 CPU-hours for the full three-dimensional model, it was not clear that sufficient statistics could be generated.

Figure 23.1 shows the results collected from only 20 switches for the full model. The estimated P_G is shown as circles, and P_S as crosses. Scatter in these data is primarily due to the counting statistics and the small number of samples in each bin. Nine-point running averages are used to reduce the scatter and improve estimates of the crossings of P_G and P_S . The running averages are represented by curves and are sufficient to determine the locations of the extrema with an uncertainty of about 1×10^{-3} in the magnetization. For the one-dimensional model the estimated curves representing P_G and P_S are almost parallel near the saddle point, making it difficult to locate crossings of the growing and shrinking probabilities (see Fig. 1 of [3]). Since they cross at a significant angle for the model studied here (see Fig. 23.1), it is possible to find an estimate for the magnetization of the free-energy saddle point from far fewer switches than were needed for the one-dimensional model.

23.3 Summary and Conclusions

The projective dynamics method was used to produce growing and shrinking probabilities during the magnetization reversal of a realistic model nanopillar. These probabilities can be used to locate the magnetization of the free-energy extrema. In particular, the magnetization of the saddle point is revealed. Compared to the one-dimensional model, far fewer switches are needed to find the location of the extrema in the three-dimensional model. In the future, a study of the dependence of the saddle point on temperature and misalignment between the pillar and the applied field are planned.

Acknowledgments

This work was supported by U.S. National Science Foundation Grant No. DMR-0120310, and by Florida State University through the School of Computational Science and Information Technology and the Center for Materials Science and Technology. This work was also supported by the DOE Office of Science through ASCR-MICS and the Computational Materials Science Network of BES-DMSE, as well as the Laboratory directed Research and Development program of ORNL under contract DE-AC05-00OOR22725 with UT-Battelle LLC.

References

1. M.A. McCord, D.D. Awschalom: *Appl. Phys. Lett.* **57**, 2153 (1990)
2. M. Kolesik, M.A. Novotny, P.A. Rikvold, D.M. Townsley, in: *Computer Simulation Studies in Condensed Matter Physics X*, ed. by D.P. Landau, K.K. Mon, H.-B. Schüttler (Springer, Berlin Heidelberg New York 1998) p. 246
3. G. Brown, M.A. Novotny, P.A. Rikvold: *Physica B* **343**, 195 (2003)

4. G. Brown, S.M. Stinnett, M.A. Novotny, P.A. Rikvold: *J. Appl. Phys.* (in press). E-print arXiv: cond-mat/0310497
5. S. Wirth, S. von Molnár: *Appl. Phys. Lett.* **76**, 3283 (2000); **77**, 762 (2000)
6. S. Wirth, M. Field, D.D. Awschalom, S. von Molnár: *Phys. Rev. B* **57**, R14028 (1998)
7. Y. Li, P. Xiong, S. von Molnár, Y. Ohno, H. Ohno: *Appl. Phys. Lett.* **80**, 4644 (2002)
8. W.F. Brown: *IEEE Trans. Magn.* **15**, 1196 (1979)

24 Cumulants for an Ising Model for Folded 1-d Small-World Materials

M.A. Novotny

Department of Physics and Astronomy and
ERC Center for Computational Sciences, Mississippi State University
P.O. Box 5167, Mississippi State, MS 39762-5167, USA

Abstract. An Ising ferromagnetic model on a one-dimensional lattice with small-world bond connections is studied by Monte Carlo simulations. The small-world bond connections are built with Ising spins with the same bond length as the starting one-dimensional lattice. A certain number of the longest of these small-world connections are, however, set to length one. This corresponds to a folded one-dimensional lattice. Results for the fourth-order cumulant are presented for folded and unfolded lattices. It is seen that the number of folds can dramatically change this cumulant, with ramifications for the critical behavior of the system.

24.1 Introduction

Our technological society is possible due to the unique characteristics of different classes of materials. Most materials are classified by the dimension d of their fixed point, $d = 3$ for bulk systems, $d = 2$ for thin films, and $d = 1$ for quasi-one-dimensional materials. Ferromagnetic Ising models can be formed for lattice systems in each of these dimensions and studied using Monte Carlo simulations [1].

Recently there has been a great interest in networks that are not regular lattices. One example is the small-world network [2], usually obtained from social or economic networks. Ising and other statistical mechanical models have been studied on such networks (see [3–5] and references therein).

Recently Novotny and Wheeler [6] have asked whether a small-world connected $d = 1$ Ising model with physical atoms making up the small-world links is a model with a finite-temperature phase transition. The result was not a finite-temperature phase transition, but a model where the ‘critical temperature’, T_c , approached zero very slowly as the system size L increased. Here we study using Monte Carlo simulations the same model, except we fold the $d = 1$ chain a finite number of times. We examine how the fourth-order Binder cumulant for the magnetization behaves for this case.

24.2 Models and Methods

The ferromagnetic Ising Hamiltonian for the model is

$$\mathcal{H}_{\text{sp}} = -J_1 \sum_{i=1}^L \sigma_i \sigma_{i+1} - J_2 \sum_i' \sigma_i \sigma_{\text{sw}(i)}, \quad (24.1)$$

where J_1 is the ferromagnetic nearest-neighbor interaction constant along the $d = 1$ chain and $\sigma = \pm 1$. Spin i is connected through the small-world connections described below to spin $\text{sw}(i)$ with ferromagnetic bond strength J_2 . The second sum includes all spins in the $d = 1$ lattice (single counted, as the ' denotes), as well as added spins to make up the small-world connections. The $d = 1$ system has L spins with periodic boundary conditions, and the entire system has N_{tot} spins.

The small-world connections are constructed randomly, with spins paired together at random so that every spin belongs to one and only one pair of spins. These pairs are constructed independent of the location of the spins. The lengths of these small-world connections are then calculated. The longest M_ℓ of these small-world connections are connected with a single bond of strength J_2 . This is equivalent to folding and connecting the $d = 1$ chain. For the rest of the small-world connections, a number of spins given by $|i - \text{sw}(i)| + 1$ are connected as a $d = 1$ lattice with all bonds of strength J_2 . In this way, one obtains a system with all bond lengths the same – a system which can be built, in theory, from atoms.

24.3 Results

We have calculated the fourth-order Binder cumulant of the magnetization for the case with a fixed number of short-cuts. This cumulant is given by [1]

$$U_4 = 1 - \frac{\langle m^4 \rangle}{3 \langle m^2 \rangle^2}. \quad (24.2)$$

Figure 24.1 shows the cumulant U_4 for two different systems sizes. The left-most curves are for no short-cuts in small-world connections, while the right curves are for all possible short-cuts in the small-world connections. Note that the left-most curves do not cross, indicating that there may be no finite-temperature phase transition. The right-most curves do cross, showing that there is a finite-temperature phase transition.

The effect of adding short-cuts is to change quantitatively the curves for U_4 for a given system size. This is shown in Fig. 24.2 for $L = 64$ for a number of short-cuts between the minimum and maximum.

The effects on crossings of U_4 is shown in Fig. 24.3 for $L = 64$ for a number of short-cuts between the minimum and maximum values.

24.4 Summary and Conclusions

As seen in the three figures, changing the number of short-cuts in the small-world lattice interpolates between a system with no phase transition and one

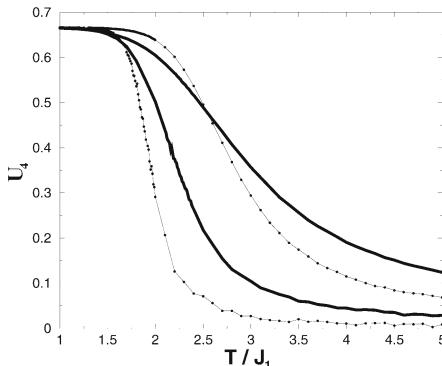


Fig. 24.1. The fourth-order Binder cumulant for the model is shown as a function of temperature for $L = 64$ (thin curves) and $L = 32$ (thick curves). The left-most curves correspond to no short-cuts in the small-world connections. The right-most curves correspond to all small-world bonds being short-cuts in the small-world connections

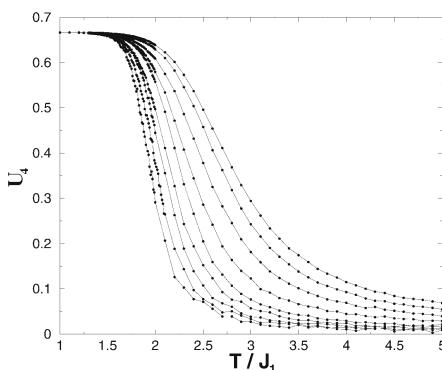


Fig. 24.2. The fourth-order Binder cumulant for the model is shown as a function of temperature for $L = 64$ for different numbers of short-cuts. The left-most curve corresponds to no short-cuts in the small-world connections. The right-most curve corresponds to all small-world bonds being short-cuts in the small-world connections. The number of short-cuts in the small-world connections increases by 4 for each curve, from zero for the left-most curve to the maximum value of 32 for the right-most curve

with a phase transition. In most models there is a unique way to take the thermodynamic limit $L \rightarrow \infty$. However, for these small-world connections with short-cuts, there is not a unique way to take this limit. Taking the limit in different ways should allow for different behavior in terms of whether or not there is a finite-temperature T_c . Keeping the number of short-cuts fixed as $L \rightarrow \infty$ will not give a finite-temperature T_c . Keeping the number of

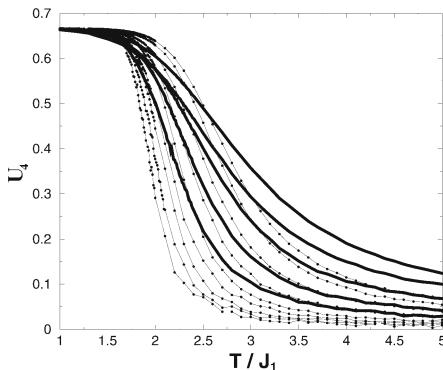


Fig. 24.3. The fourth-order Binder cumulant for the model is shown as a function of temperature for $L = 64$ (thin curves) and $L = 32$ (thick curves). The left-most curves correspond to no short-cuts in the small-world connections. The right-most curves correspond to all small-world bonds being short-cuts in the small-world connections. The number of short-cuts in the small-world connections increases by 4 for each curve for each L , from zero for the left-most curve to the maximum value of $L/2$ for the right-most curve

short-cuts a fixed ratio with L as $L \rightarrow \infty$ will give a finite-temperature T_C . It is possible to take the limit with a power-law relationship between the number of short-cuts and L . The effect this will have on T_c is a topic for future investigation.

Acknowledgements

Thanks to G. Korniss, P.A. Rikvold, and Z. Toroczkai for useful discussions – even if PAR won the bet. Partially funded by NSF DMR-0120310 and NSF DMR-0113049.

References

1. D.P. Landau, K. Binder: *A Guide to Monte Carlo Simulations in Statistical Physics*. (Cambridge University Press, Cambridge, UK 2000)
2. R. Albert, A.-L. Barabási: *Rev. Mod. Phys.* **74**, 47 (2002)
3. C.P. Herrero: *Phys. Rev. E* **65**, 066110 (2002)
4. D. Jeong, H. Hong, B.J. Kim, M.Y. Choi: *Phys. Rev. E* **68**, 027101 (2003)
5. M.B. Hastings: *Phys. Rev. Lett.* **91**, 098701 (2003)
6. M.A. Novotny, S.W. Wheeler: *Brazilian J. Physics*, in press; preprint cond-mat/0308602

25 Embryonic Forms of Nickel and Palladium: A Molecular Dynamics Computer Simulation

Z. El-bayyari

Department of Basic Sciences, Faculty of Science, Philadelphia University
P.O. Box 1 Philadelphia University 19392, Jordan, USA
E-mail: Zbayyari@rocketmail.com

Abstract. An empirical many Potential Energy Function (PEF) incorporating two-plus-three body atomistic potential derived by fitting experimental data pertaining to bulk nickel, and palladium has been applied to study the structural stability and energetics of the trimers, tetramers, and pentamers microclusters of Ni and Pd. A constant temperature Molecular Dynamics (MD) technique is employed in the simulation. It has been found that the energetically most stable structures for the trimers, tetramers and pentamers are all identical in the final optimized distorted geometry with a different bond length distances, binding energies, with palladium microclusters as the most stable.

25.1 Introduction

There has been a continuing interest in research on clusters in general and small metallic clusters in particular in the last decades from experimental and theoretical prospectives, because such species play a fundamental role in homogeneous nucleation, crystal growth, structure of amorphous materials, atmospheric chemistry, epitaxy and surface science in general [1–13].

Geometrical structure of transition metals clusters is an important factor determining their physical and chemical properties. Such atomic species containing two to a few hundred atoms are now known to exhibit novel electronic, magnetic, optical and chemical behaviors when they change size [1, 14].

Owing to the importance of nickel as a catalyst, and a ferromagnetic solid nickel have been widely studied by both theoreticians and experimentalists [1, 3–5, 14–23]. Clusters of palladium are especially of great interest since palladium is employed as catalysts in hydrogenation reactions due to the large hydrogenation energies of such element [24]. Such metal, which form hydrides, can be used for hydrogen storage and is interesting because properties such as conductivity or transparency change with the hydrogen concentration [25]. Palladium, particles supported on oxide surfaces play the role as active components in the control of automotive pollution [26].

In the present work, we focus our attention on studying the ground state and energetically low lying structures of Ni_n , and Pd_n ($n = 3–5$) clusters using MD simulations based on the use of an empirical many body Potential Energy Function (PEF) and Nordsieck–Gear algorithm, a field of current

Table 25.1. Calculated interatomic distances in (\AA) between the $i - j$ atom index for the clusters, shown in Fig. 25.1. The * indicates the value is not taken in the average bond length calculation, apex to apex atom in the trigonal bipyramidal

$i - j$	Ni ₃	Ni ₄	Ni ₅	Pd ₃	Pd ₄	Pd ₅
1–2	2.198688	2.189574	2.158147	2.569987	2.569528	2.578145
1–3	2.241405	2.214381	2.143275	2.581338	2.576214	2.569499
1–4		2.245574	2.236004		2.581518	2.584206
1–5			2.339133			2.575817
2–3	2.205809	2.204008	3.577597*	2.570044	2.569848	2.572945
2–4		2.267131	2.251268		2.583576	2.572929
2–5			2.272096			4.214275*
3–4		2.216046	2.234010		2.573801	2.571848
3–5			2.251630			2.600595
4–5			2.203067			2.573144
Average	2.215	2.223	2.232	2.574	2.576	2.578

Table 25.2. Calculated energies in (eV) for Ni _{n} , and Pd _{n} ($n = 3–5$) clusters. Φ represents the total energy of the cluster and Φ_n is the average energy per atom in the cluster

Energy	Ni ₃	Ni ₄	Ni ₅	Pd ₃	Pd ₄	Pd ₅
$-\Phi$	1.635200	2.996330	4.356740	2.052750	3.810940	5.524580
$-\Phi_n$	0.545067	0.749083	0.871348	0.684250	0.952735	1.104916

interest. The bond length distance between the atoms in each cluster, binding energies per atom will be calculated, and optimized geometries will be obtained. The present PEF has been used successfully to simulate the structural stability and energetics of Al clusters [27], bulk and surface properties of Al [28], Ag clusters [29], and for Au clusters [30]. We begin with a brief description of the method used. Finally, we discuss the results of our simulations and compare between them with the available literature data.

25.2 The Potential Energy Function

We employed a molecular-dynamics technique based on the Nordsieck–Gear algorithm [31], to study the energetics and the structural stability of Ni _{n} and Pd _{n} microclusters. The PEF has been discussed elsewhere [23, 32]. A set of different initial configurations were considered; energetically the most stable structure for each set is shown in Fig. 25.1, the corresponding interatomic distance, the average bond length, and the interaction energies are shown in Tables 25.1 and 25.2 respectively.

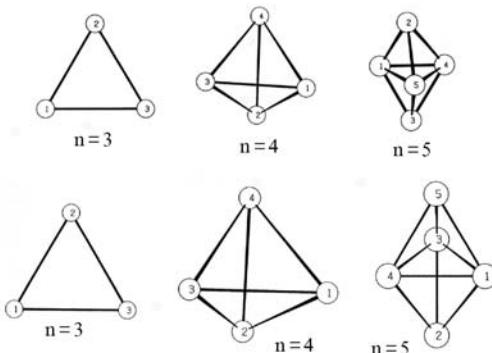


Fig. 25.1. The energetically most stable cluster geometries of Ni_n , and Pd_n , ($n = 3\text{--}5$) respectively

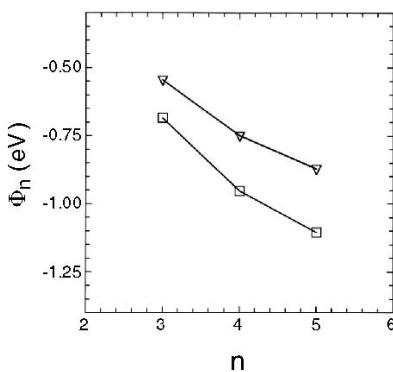


Fig. 25.2. Variation of the average interaction energy per atom vs. Cluster size for Ni_n (triangle down), and Pd_n (Square) for ($n = 3\text{--}5$) respectively

The cluster structures presented in this work are the snap-shots of the last molecular-dynamics step for each nickel and Palladium cluster. The variation of the average interaction energy per atom, namely $\Phi_n = \Phi/n$, as a function of the cluster size is plotted in Fig. 25.2. The clusters studied are discussed and compared with the available literature data in the next section.

25.3 Calculations and Discussion

In this section we will review and report the different methods and results of the experimental and theoretical work available in Tables 25.3 and 25.4 to keep the scope of the proceedings as short as possible.

It would be interesting to compare the relative stabilities of the Ni_n , and Pd_n clusters ($n = 3\text{--}5$) from the consideration of their binding energies per

Table 25.3. calculated properties of Ni_n clusters. The symbols BE/n is the binding energy per atom in (eV), BL is the bond length in Å, and the calculated results of this work using (MD), other symbols are explained below the table

Cluster	Symmetry	$-\text{BE}/n$	BL (Å)	Method	Ref.
Ni ₃	D _{∞h}	1.134	2.50	LSMOA	[1]
	D _{∞h}	0.50	2.38	ECP-SCF	[3]
	C _{2v}			UV visible Absorption	[4]
	D _{3h}	0.40		CNDO	[5]
	C _{2v}	1.960	2.148–2.185	LCAO-MO	[14]
	D _{3h}	1.436, 1.40, 1.547	2.122, 2.233, 2.202	CEM, MD/MC-CEM, DF	[15]
	D _{3h}	1.66	2.30	TBMD	[16, 19]
	D _{3h}			MD	[18]
	D _{3h}	1.750	2.280	EH	[20]
	D _{3h}	0.550	2.220	EM	[23]
Ni ₄	D _{3h} [*]	0.545	2.215 (average)	MD	*)
	D _{4h}	1.150	2.60	LSMOA	[1]
	D _{4h}	0.58		ECP-SCF	[3]
	D _{4h}	1.15		Semi-empirical	[5]
	D _{2d} , D _{4h}	2.34	2.10	LCAO-MO	[14]
	T _d	1.875, 2.333, 2.433	2.16, 2.323, 2.249	CEM, MD/MC-CEM, DF	[15]
	D _{4h}	1.870	2.260	TBMD	[16]
	T _d			MD, NH ₄ Experiments	[18, 21]
	D _{4h}	1.870	2.260	TBMD	[19]
	D _{2d}	2.250	2.310	EH	[20]
Ni ₅	T _d	0.750	2.250	EM	[23]
	T _d [*]	0.750	2.223 (average)	MD	*)
	D _{5h}	1.320	2.70	LSMOA	[1]
	C _{4v}	0.710		ECP-SCF	[3]
	D _{5h}	1.320		Semi-empirical	[5]
	D _{3h}	2.83	2.233–2.286	LCAO-MO	[14]
	D _{3h}	2.102, 1.948		CEM, MD/MC-CEM	[15]
	T _d	2.16	2.42	TBMD	[16, 19]
	D _{3h}			N ₂ -Uptake Experiments	[17]
	D _{3h}			MD	[18]
Ni ₆	D _{3h}	2.40	2.33	EH	[20]
	T _d or D _{3h}			NH ₄ Uptake Experiments	[21]
	D _{3h}			G, FS, LJ, M Potentials	[22]
	D _{3h}	0.88	2.26	EM	[23]
	D _{3h} [*]	0.87	2.232 (average)	MD	*)

Acronyms used refer to the methods LSMOA: Low Spin Molecular Orbital Approximation, ECP-SCF: Effective Core Potentail-Self Consistent Field, UV: Ultra Violet, CNDO: Complete Neglect of Differential Overlap, LCAO-MO: Linear Combination of Atomic Orbitals-Molecular Orbital, CEM: Corrected Effective Medium, MD/MC-CEM: Molecular Dynamics/Montecarlo-CEM, TBMD: Tight Binding Molecular Dynamics, MD: Molecular Dynamics, EH: Extended Hkel, EM: Energy Minimization, G: Gupta, FS: Finnis-Sinclair, LJ: Lennard-Jones, EAM: Embedded Atom Method, TBM: Tight Binding Method, HF: Hartree Fock

*) This work

Table 25.4. Same as in Table 25.3, but for Pd_n clusters

Cluster	Symmetry	$-\text{BE}/n$	BL (Å)	Method	Reference
Pd ₃	D_{3h}	0.685	2.575	EM	[23]
	D_{3h}	0.80	3.30	EH	[2]
		2.50		MD, EAM, TBM	[34]
Pd ₄	D_{3h}^*	0.684	2.574 (average)	MD	This Work
	T_d	1.463		MD/MC-CEM	[15]
	T_d	0.954	2.583	EM	[23]
	T_d			ECP	[5]
	T_d	0.74	2.10	EH	[2]
Pd ₅			3.30	HF	[33]
		2.70		MD, EAM, TBM	[34]
	T_d^*	0.953	2.576 (average)	MD	This work
	D_{3h}	1.678		MD/MC-CEM	[15]
	D_{3h}	1.107	2.586	EM	[23]
Pd ₆	D_{3h}	0.67	2.30	EH	[2]
		2.850		MD, EAM, TBM	[34]
	D_{3h}^*	1.105	2.578 (average)	MD	This work

Table 25.5. Experimental and theoretical binding energies (eV) per atom of Ni_n , and Pd_n clusters from [35], and the calculated results of this work using (MD). Experimental binding energies are within Parentheses

n	Ni_n	MD	Pd_n	MD
3	(0.9626)	0.5450	1.1834	0.6843
4	(1.0995)	0.7491	1.4050	0.9527
5	(1.4096)	0.8713	1.3960	1.1050

atom and the average bond length distance in each cluster. On the basis of the total binding energies of Pd_n ($n = 3\text{--}4$) clusters as presented in Table 25.5.

Its quite evident that the stability of the cluster increases with increasing size. The experimental and theoretical results as presented in Table 25.5. also supports this trend. The results of the Ni cluster differ from Pd clusters, which are the most stable, as shown from Table 25.5 and Fig. 25.2 in two respects. First, the Ni clusters have substantially smaller binding energies compared to Pd clusters. This trend could be attributed to the relativistic effects that are substantially larger for palladium. The other major difference between Ni and Pd clusters lie in the respective order of stability of the two metal clusters. The binding energy per atom of the Ni_n ($n = 3\text{--}5$) cluster is found to be 0.5450, 0.7491, and 0.8713 eV as shown in Table 25.5 respectively for Ni_3 , Ni_4 , and Ni_5 clusters, whereas in Pd_n ($n = 3\text{--}5$), the order of stability increases monotonically with the cluster size. Our results has a smaller binding energies compared to the experimental and theoretical work which could be due to

the thermodynamic state that describes the experiment or the different level of theory that is utilized to compute the binding energies by other authors.

Figure 25.2 contains the plots of the binding energies per atom of Ni_n and Pd_n , respectively with respect to cluster size (n). The results of Ni clusters indicate that they are less strongly bound than the Pd_n ($n = 3\text{--}4$) clusters. An interesting feature that emerges from this plot is that there is increased stability among the smaller clusters for $n = 4$. Thus this cluster size ($n = 4$) could be considered as a local magic number for group 8B metal clusters (Ni, Pd, and Pt). Another point worth mentioning concerning the lower binding energies in Ni_n ($n = 3\text{--}5$) clusters with respect to Pd_n ($n = 3\text{--}5$) clusters is the concept of less participation of the d orbitals in bond formation Majumdar et al. [35]. Concerning the bond length distance as for Pd_n clusters, the calculated average bond length distance is 2.574, 2.576, and 2.578 Å for Pd_3 , Pd_4 , and Pd_5 respectively. A comparison of these distances of the present work with the corresponding values of Ni reveals that they are considerably larger than Ni counterparts 2.215, 2.223, and 2.232 Å respectively for Ni_3 , Ni_4 , and Ni_5 . The Pd–Pd distances for Pd_3 , and Pd_4 reported by Majumdar et al. [35] is found to be 2.67, and 2.69 Å respectively. In the case of Pd_5 the ground state geometry is a tetragonal pyramid with a Pd–Pd distance of 2.611 Å for the tetragonal base. The Pd–Pd distance for the other bonds is 2.84 Å. Our results are qualitatively in agreement with Majumdar et al. [35] results regarding Ni and Pd clusters.

25.4 Conclusion

The main drawn conclusions from this study are as follows:

1. Most ab initio methods results obtained for Ni and Pd clusters, while satisfactorily describing trends in the variation of various cluster properties with the cluster size, fail to agree among themselves and with experiment. As a consequence of these difficulties the ab initio calculations even when considering optimization, limit themselves to symmetry restricted geometries with $n > 3$.
2. With the present method the most stable microclusters of Nickel and Palladium particles are in three dimensional distorted compact forms with average bond length distance around (2.215–2.232 Å) for Ni and (2.574–2.578 Å) for Pd.
3. The calculated binding energies of Ni_n ($n = 3\text{--}4$) clusters are less strongly bound than the Pd_n ($n = 3\text{--}4$) clusters as experiments shows.
4. In comparison between experiment and theory the role of temperature and kinetics should be remembered. The global minimum is only rigorously the equilibrium structure at zero temperature. At higher temperature other structures may become more stable due to the entropic effects. Furthermore, it is not always clear whether equilibrium has been

achieved under experimental conditions, especially for clusters formed at low temperature.

In conclusion semi-empirical methods can play a major role especially with MD technique to predict the most stable structures, bond distances and other properties for a molecular system. An analysis of various structures confirms the short range of the effective interatomic interactions in metals, and thus lends strong support to semiempirical methods [36], which describe the energetics of interatomic interactions in terms of only first and possibly second-neighbor interactions.

References

1. A.B. Anderson: *J. Chem. Phys.* **64**, 4046 (1976); **66**, 5108 (1977)
2. E.L. Muetterties, T.N. Rhodin, E. Band, C.F. Burcker, W.R. Pretzer: *Chem. Rev.* **79**, 91 (1979)
3. H. Basch, M.D. Newton, J.W. Moskowitz: *J. Chem. Phys.* **73**, 4492 (1980)
4. W. Weltner Jr., R.J. Van Zee: *Ann. Rev. Phys. Chem.* **35**, 291 (1984)
5. J. Koutecký, P. Fantucci: *Chem. Rev.* **86**, 539 (1986)
6. M.D. Morse: *Chem. Rev.* **86**, 1049 (1986)
7. J. Davenas, P.M. Rabette (Eds.): *Contribution of Cluster Physics to Material Science and Technology*. NATO ASI series E: NO. 104 (Martinus Nijhoff, Dordrecht 1986)
8. T. Halıcıoğlu, C.W. Bauschlicher Jr.: *Rep. Progr. Phys.* **51**, 883 (1988)
9. V.B. Koutecký, P. Fantucci, J. Koutecký: *Chem. Rev.* **91**, 1035 (1991)
10. G. Pacchioni, P.S. Bagus, F. Parmigiani (Eds.): *Cluster Models for Surface and Bulk Phenomena*. NATO ASI series B: Vol. 283 (Plenum Press, New York 1992)
11. P. Jena, S.N. Khanna, B. K. Rao (Eds.): *Physics and Chemistry of Finite Systems: From Clusters to Crystals*. Vols. I and II, NATO ASI series C: No. 374 (Kluwer, Dordrecht, The Netherlands 1992)
12. H. Haperland (Ed.): *Clusters of Atoms and Molecules* (Springer, Berlin Heidelberg New York 1994)
13. D.G. Vlachos, L.D. Schmidt, R. Aris: *J. Chem. Phys.* **96**, 6880 (1992)
14. F.A. Reuse, S.N. Khanna: *Chem. Phys. Lett.* **234**, 77 (1995); *Eur. Phys. J. D.* **6**, 77 (1999)
15. M.S. Stave, A.E. DePristo: *J. Chem. Phys.* **97**, 3386 (1992)
16. M. Menon, J. Connolly, N. Lathiotakis, A. Andriotis: *Phys. Rev. B* **50**, 8903 (1994)
17. E.K. Parks, L. Zhu, J. Ho, S.J. Riley: *J. Chem. Phys.* **100**, 7206 (1994); *J. Chem. Phys.* **102**, 7377 (1995)
18. A.P. Amarillas, I.L. Garzón: *Phys. Rev. B* **54**, 10362 (1996)
19. N.N. Lathiotakis, A.N. Andriotis, M. Menon, J. Connolly: *J. Chem. Phys.* **104**, 992 (1996)
20. E. Curotto, A. Matro, D.L. Freeman, J.D. Doll: *J. Chem. Phys.* **108**, 729 (1998)
21. E.K. Parks, B.J. Winter, T.D. Klots, S.J. Riley: *J. Chem. Phys.* **94**, 1882 (1991)

22. J.L. Rodriguez-López, F. Aguilera-Granja, A. Vega, J.A. Alonso: Eur. Phys. J. D. **6**, 235 (1999)
23. Ş. Erkoç: Phys. Stat. Sol. (b) **161**, 211 (1990)
24. K. Balasubramanian: J. Chem. Phys. **91**, 307 (1989)
25. C. Nützenadel, A. Züttel, D. Chartouni, G. Schmid, L. Schlapbach: Eur. Phys. J. D. **8**, 245 (2000)
26. U. Heiz, A. Sanchez, S. Abbet, W.D. Schneider: Eur. Phys. J. D. **9**, 35 (2002)
27. Z. El-Bayyari, Ş. Erkoç: Phys. Status Solidi b **170**, 103 (1992)
28. Z. El-Bayyari, Ş. Erkoç: Mat. Chem. Phys. **37**, 382 (1994)
29. Z. El-Bayyari: Phys. Status. Solidi b **208**, 339 (1998)
30. Z. El-Bayyari, in *Computer Simulation Studies of Condensed Matter Physics XIV*. D.P. Landau, S.P. Lewis, H.B. Schüttler (Eds.) Springer Proceedings in Physics, Vol. 89 (Springer, Berlin Heidelberg New York 2001) p. 47
31. A. Nordsieck: Math. Comp. **16**, 22 (1962). C.W. Gear: Numerical initial value problems in ordinary differential equations. (Prentice-Hall, Englewood Cliffs, NJ, 1971)
32. Ş. Erkoç: Phys. Stat. Sol. b **152**, 447 (1989)
33. E. Miyoshi, Y. Sakai, S. Mori: Chem. Phys. Lett. **113**, 457 (1985)
34. C. Rey, L.J. Gallego, J. García-Rodeja, J.A. Alonso, M.P. Iñiguez: Phys. Rev. B. **48**, 8253 (1993)
35. D. Majumdar, D. Dai, K. Balasubramanian: J. Chem. Phys. **113**, 7919 and 6928 (2000)
36. Y.J. Yel, J. Dirk, J. Bernholc: Phys. Rev. Lett. **67**, 1594 (1991)

Part VII

Surfaces and Alloys

26 Usage of Pattern Recognition Scheme in Kinetic Monte Carlo Simulations: Application to Cluster Diffusion on Cu(111)

C. Ghosh, A. Kara, and T.S. Rahman

Department of Physics, Cardwell Hall, Kansas State University
Manhattan, KS 66506, USA

Abstract. We have invoked a simple pattern recognition scheme in kinetic Monte Carlo simulations of post-deposition evolution of two dimensional islands on fcc(111) surfaces. On application of the technique to the diffusion of small Cu clusters (8–55 atoms) on Cu(111) we find that, at room temperature, clusters with certain magic numbers show stick-slip type of motion with striking patterns rather than the random paths followed by the others. At higher temperatures all clusters display random motion. The calculated diffusion coefficients show dependence on size and temperature with an effective barrier ranging between 0.64 eV and 0.71 eV. Small asymmetries in diffusion barriers lead to a large difference in the frequencies of adatom diffusion along the two types of microfacetted steps on Cu(111) leading to consequences in their shape evolution. The pattern recognition scheme revealed forty nine basic periphery diffusion processes whose activation energy barriers were calculated using the nudged elastic band technique and interatomic potentials from the embedded atom method.

26.1 Introduction

An important constituent of nanostructuring of material is the process of epitaxial growth which is intimately related to the diffusion, evaporation, condensation and related phenomena of atoms and atomic clusters on surfaces. Experimental studies of the diffusion of adatoms and small atomic clusters on metal surfaces using Field Ion Microscopy (FIM) have already provided a number of unexpected events such as the concerted motion of atoms [1] and the collective sliding motion of clusters [2]. These studies, together with the availability of reliable computational techniques, motivated a number of computational and theoretical scientists to examine in more detail the possible paths for the diffusion of adatom and vacancy clusters on metal surfaces and calculate their energetics and dynamics. Rapid developments in the technique of Scanning Tunneling Microscopy (STM) have further exposed interesting characteristics of the dynamics of large (a few hundred) adatom and vacancy islands. In fact, it was from the STM measurements of Ag/Ag(100) [3] and Ag/Ag(111) [4] that large two-dimensional islands were found to be significantly mobile. Furthermore, the mobility of large vacancy islands was found to be comparable to that of the adatom islands [5]. These and related observations led to a series of papers [6–9] with speculations about the microscopic

mechanisms that cause these large islands to move. Of particular interest here are the competing mechanisms of adatom periphery diffusion, evaporation/condensation, and terrace diffusion. Statistical mechanical calculations based on solid-on-solid (SOS) models predict specific scaling of the diffusion coefficient with the island diameter, depending on the preponderance of one of these three mechanisms. Since these dependencies are not unequivocally extracted from experimental data, because of the large error bars involved, the issue is not yet completely settled, although the bias is towards periphery diffusion. Molecular dynamics simulations of Ag vacancy island on Ag(111), on the other hand, have shown a preference for periphery diffusion, while in the case of the elastic-continuum models the three mechanisms lead to characteristic decays of the time and space correlation functions. Several questions about the elastic-continuum based models, however, remain as shown by Bogicevic et al. [10] who find that the exponents in the power law dependence of the diffusion coefficient on island size were themselves temperature dependent and material specific, unlike predictions of the simpler earlier model calculations. While the work of Bogicevic et al. points to the simplicity of the previous calculations, it also begs the question whether the atomistic model used by them truly brings out the inherent complexity of the system. Our reference here is the small number of diffusion processes that appear to be included in [10]. While we agree that the kinetic Monte Carlo technique (KMC) can provide a linkage between information at the atomic scale and simulated behavior at the mesoscopic one, the evolution of the system could be prejudiced by the usage of an insufficient set of atomic processes arising from a narrow local consideration.

The main point of departure in the present paper is the usage of a pattern recognition scheme in KMC simulations and the provision of a sufficiently large set of possible single atom processes that the system might undertake. The activation energy barriers for these processes are calculated using reliable interaction potentials and calculation schemes. While for the applications presented here, these interatomic potentials are based on the embedded atom method (EAM) [11], they could also be obtained from ab initio methods if sufficient computational resources are available. After presenting a framework for carrying out such KMC calculations we apply it to examine the diffusion of two-dimensional (2D) atomic clusters of several sizes and shapes on a fcc(111) surface. Our specific system here is Cu/Cu(111), as described by EAM potentials. In the following, we present summary of the theoretical technique in Sect. 26.2, followed by a description of the model system in Sect. 26.3, and a list of diffusion processes and activation energies in Sect. 26.4. The results for cluster diffusion are presented and discussed in section V, while some conclusions are offered in Sect. 26.4.

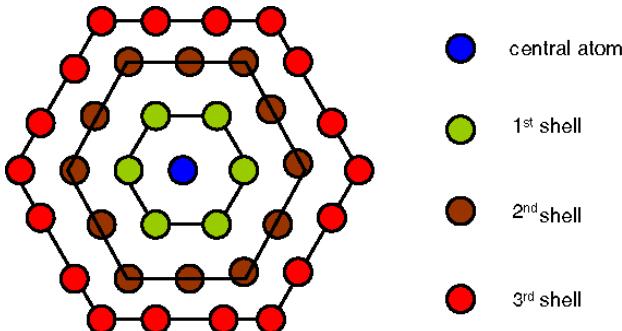


Fig. 26.1. Illustration of the 3-shell scheme for an (111)fcc 2D system

26.2 Theoretical Details

We use standard lattice gas model (solid on solid) with hexagonal symmetry which maps the fcc(111) surface. In this model only occupancy of fcc sites is allowed which is in general agreement with experimental observation on several metal surfaces. A typical sample consists of a 3-layer slab with periodic boundary conditions in the XY plane, parallel to the surface. The topmost layer is formed by the system that we are studying, be it an adatom island, a vacancy island or any other nanostructure. The two bottom layers of the sample are fixed. Thus only in plane (2D) motion of the surface layer is considered.

The goal of kinetic Monte Carlo (KMC) is to mimic real experiments through sophisticated simulations. For these simulations to be realistic, one has to implement increasingly complex scenarios requiring intensive use of state-of-the art software and hardware. At the heart of a KMC simulation of the time evolution of a given system lie the mechanisms that are responsible for determining the microscopic events to be performed at any given time. To illustrate the point, consider a system containing N particles at a given time with N_e possible types of events. Let us also associate with each event-type (i) n_i the number of particles in the system that are candidates for this event-type. We also associate with the event-type (i), the activation energy E_i and a pre-factor ν_i . The microscopic rate associated with event (i), within Transition State Theory (TST), is then, $r_i = \nu_i \exp(-E_i/k_b T)$, with k_b the Boltzmann factor and T the temperature. The macroscopic rate associated with event (i) is simply $R_i = n_i r_i$. The total rate is given by

$$R = \sum R_i .$$

In KMC simulations, the acceptance of a chosen event is always set to one. The choice of a given event is, however, dictated by the rates. First, an event-type is chosen according to its probability $P_i = R_i/R$ and then a particle is randomly chosen from the set n_i . For simplicity, we assume that any event in

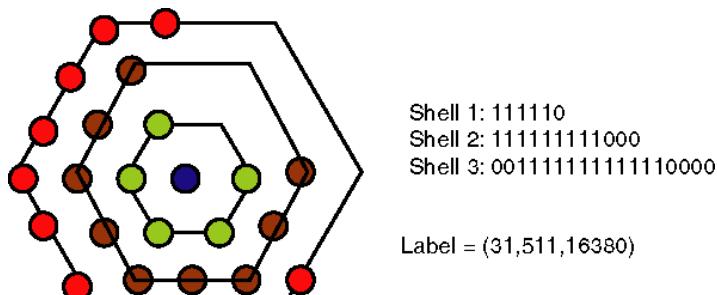


Fig. 26.2. Schematic representation of the labeling procedure in the 3-shell scheme, with 1 for occupied and 0 for unoccupied sites

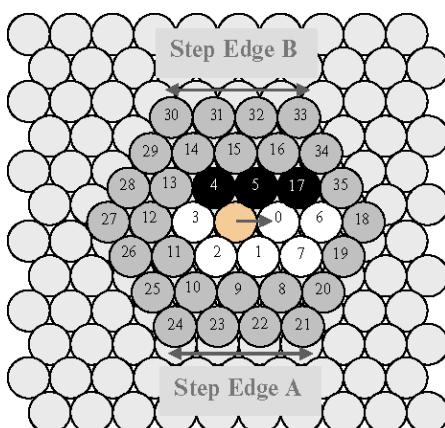


Fig. 26.3. Three shells of nearest neighbor atoms around the active atom (atom selected to move) and the pattern recognition scheme

this system can be described by a 3-shell scheme containing a central atom surrounded by three shells of neighbors, as illustrated in Fig. 26.1.

The labeling of the surrounding atoms is done in binary and a base ten number is subsequently associated with the first shell configuration. The same procedure is followed for the second and third shell. Hence, for an atom in the system to be active (i.e. central atom for a given event), it should have a vacancy in its first shell (or an occupancy number less than 63) as illustrated in Fig. 26.2.

In identifying the active atoms according to the above labeling scheme for the diffusion of adatom clusters on fcc(111) surface, forty-nine basic processes (and their equivalents obtained by applying the symmetry operations for a hexagonal lattice) are found for single atom peripheral diffusion. Furthermore at every MC step the active atoms are classified according to

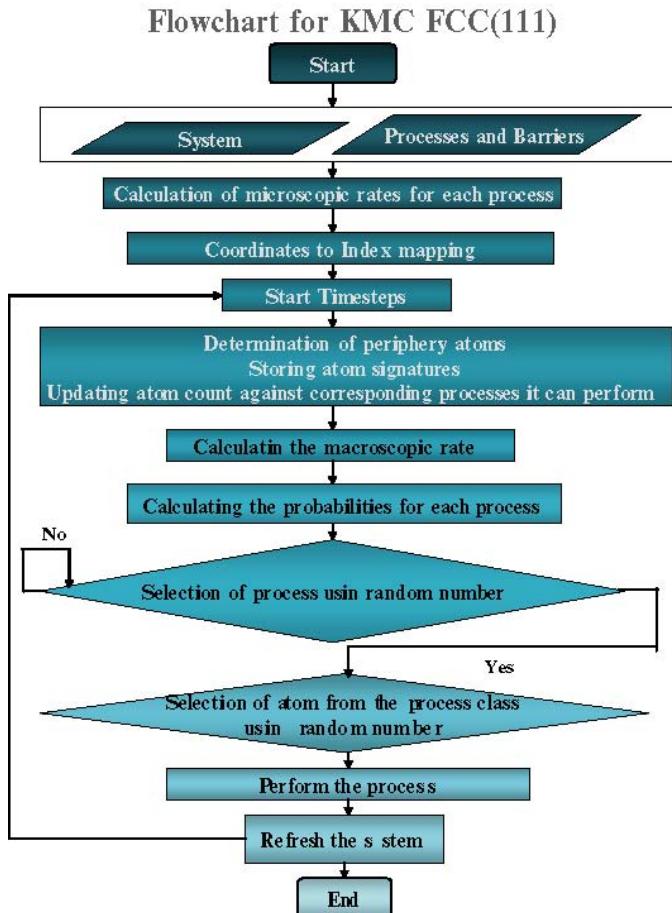


Fig. 26.4. Flowchart for a kinetic Monte Carlo simulation

the processes they may perform within the 3-shell scheme. The macroscopic rate for every process is thus computed and the corresponding probability determined. The interval between zero and one is then divided into intervals with widths corresponding to the probabilities of all the active processes. A first random number selects the process to be performed and a second selects the atom to be moved from the list belonging to the selected process. Not all the surrounding 36 atoms are used in the classification of an active atom. Instead a “minimum configuration” scheme is invoked in which we take into account (from the 36 atoms) only those sites (vacant or occupied) that uniquely determine the process that is associated with a given atom. In Fig. 26.3, as an example, we show the minimum configuration associated with diffusion along step A. Step A signifies a (100) microfaceted step while

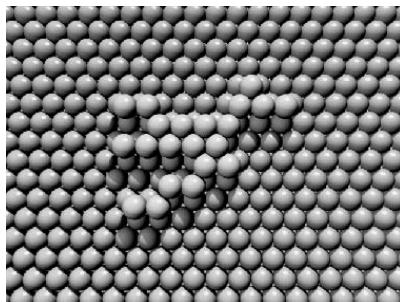


Fig. 26.5. Fractal-like starting configuration for the cluster

step B signifies a (111) microfacetted step. For this process, the conditions to be met are: (i) sites 4, 5, 17 (in black) *must* be filled and (ii) sites 0, 1, 2, 3, 6, 7 (in white) *must* be vacant. The rest of the sites could be either filled or vacant (grey). The simulation procedure described above is summarized in Fig. 26.4.

For the interactions between the atoms in the systems studied here, we use the embedded atom method (EAM). This is a semi-empirical, many-body interaction potential [11]. The EAM was very successful in reproducing many of the characteristics of the bulk and the surfaces of Cu, Ag, Pd, Pt, Au and Ni [11]. We have also found the EAM potentials to be reliable for examining the temperature dependent structure and dynamics of the low Miller index surfaces of Ag and Cu [12], and for describing the energetics and dynamics of Cu, Ag and Pd vicinals [13, 14], and self diffusion on the Ag, Cu, and Ni flat surfaces [15].

The activation energies associated with the different processes are calculated using the nudged elastic band method [16]. This method has been shown to be efficient for finding the (minimum energy path) MEP between given initial and final states of a process. The MEP is found by constructing a set of images of the system, thirty in our case, between the initial and final states. A spring type interaction potential between adjacent images is added to ensure continuity of the path. A minimization procedure for the forces acting on the images brings the images to the MEP. This method determines the activation energy with a typical error of about 10 meV for the processes involved in the studies presented here.

The diffusivity of the islands was studied by following the trace of the center of mass (CM) with time and quantitatively by calculating the mean square displacement (MSD) of the CM as function of time. While calculating the mean square displacement, we first divided our data into equidistant steps with respect to time having the corresponding CM coordinates at that time instant. From this set of data the mean square displacement and hence the diffusion coefficients were obtained by correlating these equidistant data points.

26.3 Model Systems

We have studied the dynamics of sixteen Cu islands containing between 8 and 55 atoms on Cu(111). We have classified these islands into four categories depending on the number of atoms they contain: (i) compact hexagon (19, 37); (ii) compact hexagon +1 (8, 20, 38); (iii) $6n + 3$ (9, 15, 21, 27, 33); and (iv) general (10, 12, 14, 18, 26, 55), where n is an integer. The last category contains all those islands not fitting in categories (i), (ii) and (iii). Each KMC simulations started with an arbitrary (fractal like) shaped cluster (Fig. 26.5) in which all atoms are connected.

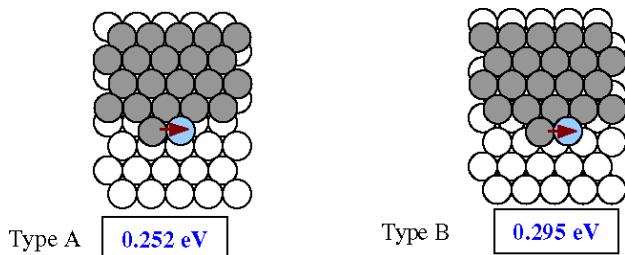
26.4 Diffusion Processes and Activation Energies

At the outset of the simulation for each active atom (shell label less than 63), all possible peripheral moves were recorded and the database checked for their activation energies. If the activation energy was not found, it was calculated as discussed above and stored in the database. In this manner 49 different peripheral diffusion processes were found. On the application of symmetry operations, these correspond to 294 possible processes.

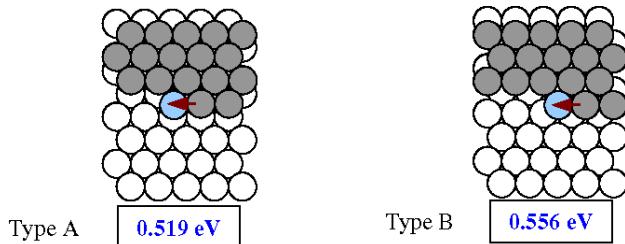
In this section we present a summary of the forty nine processes involved in 2D island diffusion on fcc(111) surface. These processes represent all possible events that could be encountered in the hopping of periphery atoms from one fcc site to another. The set of 49 figures (Fig. 26.6) illustrates these processes with the arrow showing the transition from the initial to the final state. The activation energy as determined by the NEB method is also recorded in the figure. As can be noted the EAM potential used here present a systematic difference in the energetics of processes along step A and step B. For example, the diffusion along step A has an activation energy 43 meV lower than that along the step B. A similar observation is made for the process involving kink detachment along a step. These asymmetries in the energetics will ultimately dictate the probabilities associated with the available processes during the simulation, as we will show later in this section. We would also like to point that some processes necessitate more than one step as our pattern labeling dictates motion from an fcc site to another fcc site. Hence, the rounding of an AA or a BB junction is done in three stages while that of an AB or a BA junction is performed in two.

26.5 Results

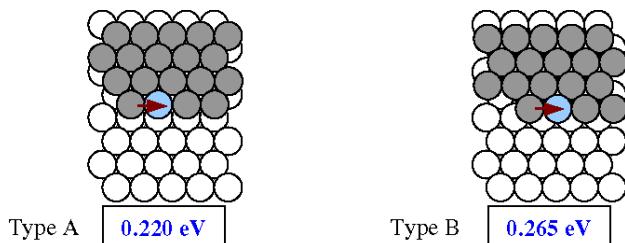
In the following we first present the results for the motion of the center of mass of each cluster. This is followed by a summary of the calculated mean square displacement and diffusion coefficients. Several tables showing the frequencies of the 49 events are included in the end of the section.



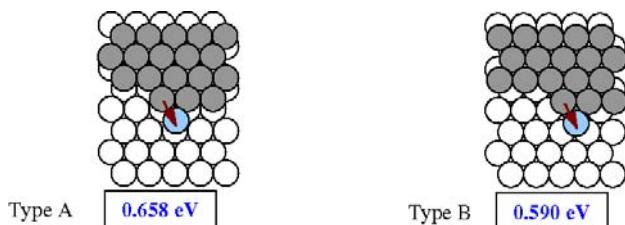
Step Edge Diffusion



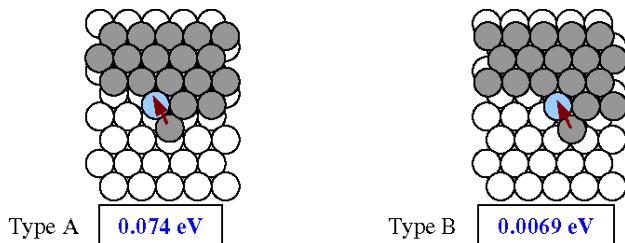
Kink detachment along step edge



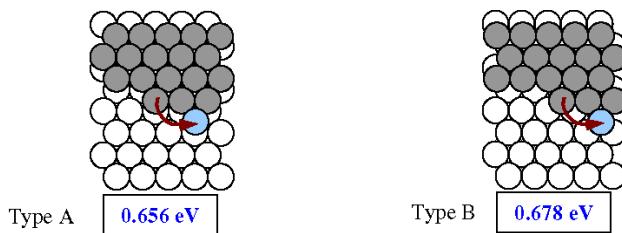
Kink Incorporation



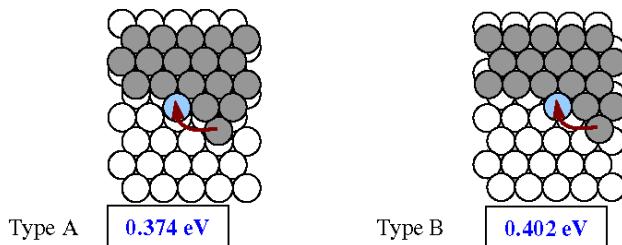
Kink detachment out of step



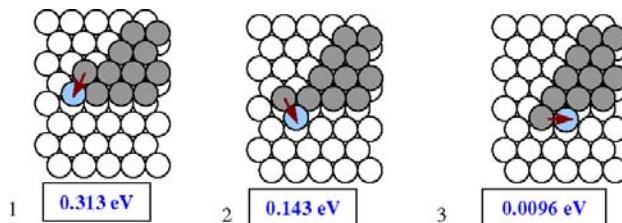
Falling into kink



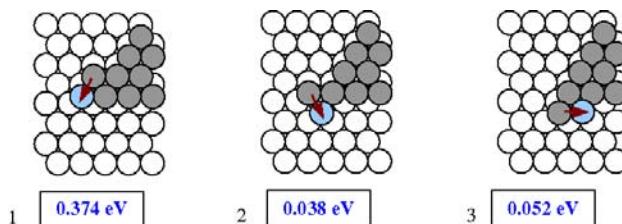
Kink Rounding



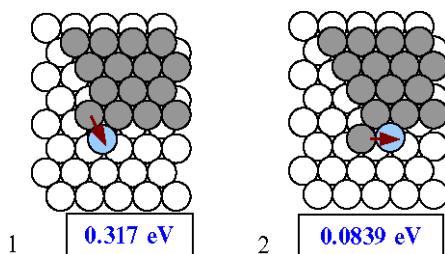
KESE



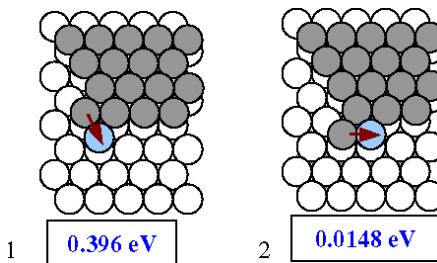
Corner rounding at AA junction – 3 stages



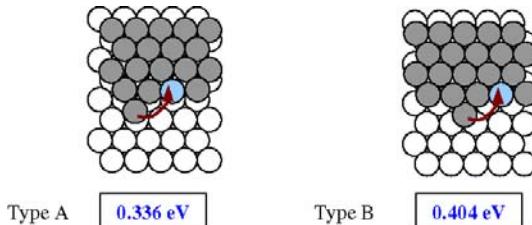
Corner rounding at BB junction – 3 stages



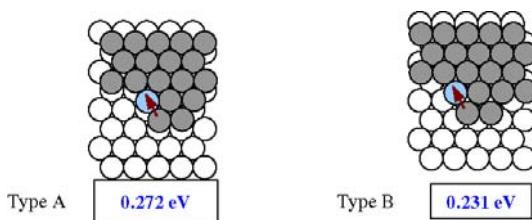
Corner rounding at AB junction – 2 stages



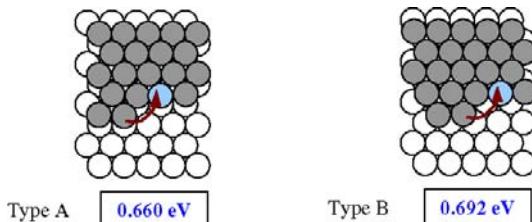
Corner rounding at BA junction – 2 stages



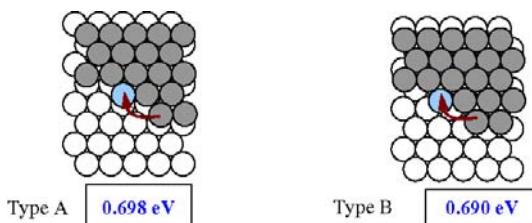
Step attachment



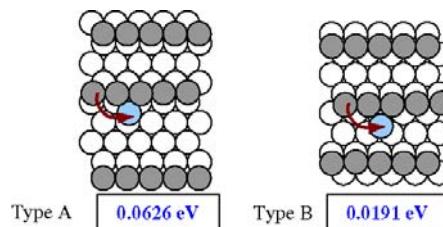
Detach from step and fall into kink



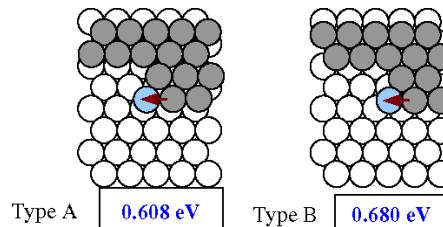
Detach from atom and attach to step



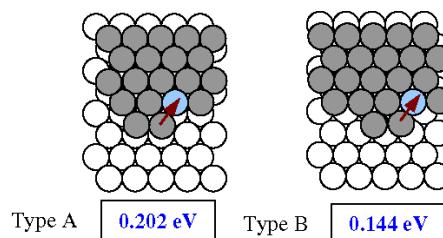
Detach from atom and fall into kink



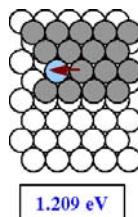
Rounding a chain



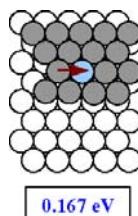
Kink detach along step (small)



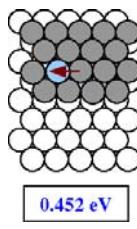
Detach from atom and attach to step (small)



Detach from 5 nn to vacancy at 6th nn

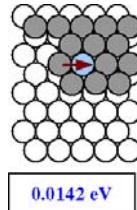


Reverse of above process-Detach from 6th nn to vacancy of 5 nn



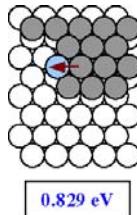
0.452 eV

Move to a vacancy within a closed shell



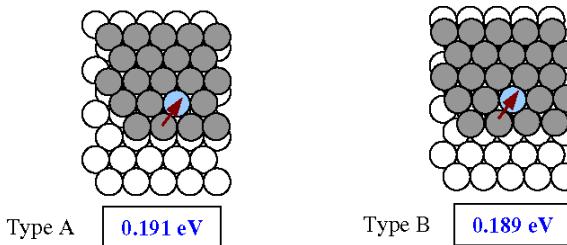
0.0142 eV

Detach from an AB corner into 5 nm vacancy



0.829 eV

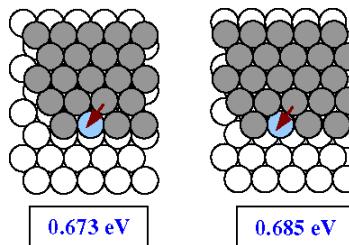
Reverse of above process – Detach from 5 nm to a vacancy at 6th nn AB corner



0.191 eV

0.189 eV

Detach from a step to a vacancy above



0.673 eV

0.685 eV

Reverse of the above process – Detach from 5 nm to a vacancy on the step

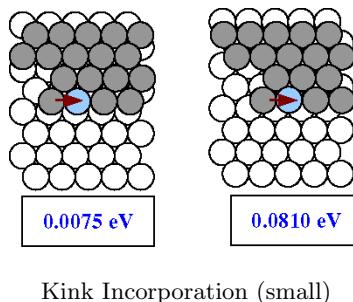


Fig. 26.6. Energy barriers for key diffusion processes for Cu/Cu(111) – NEB.

● Initial position, ● final position

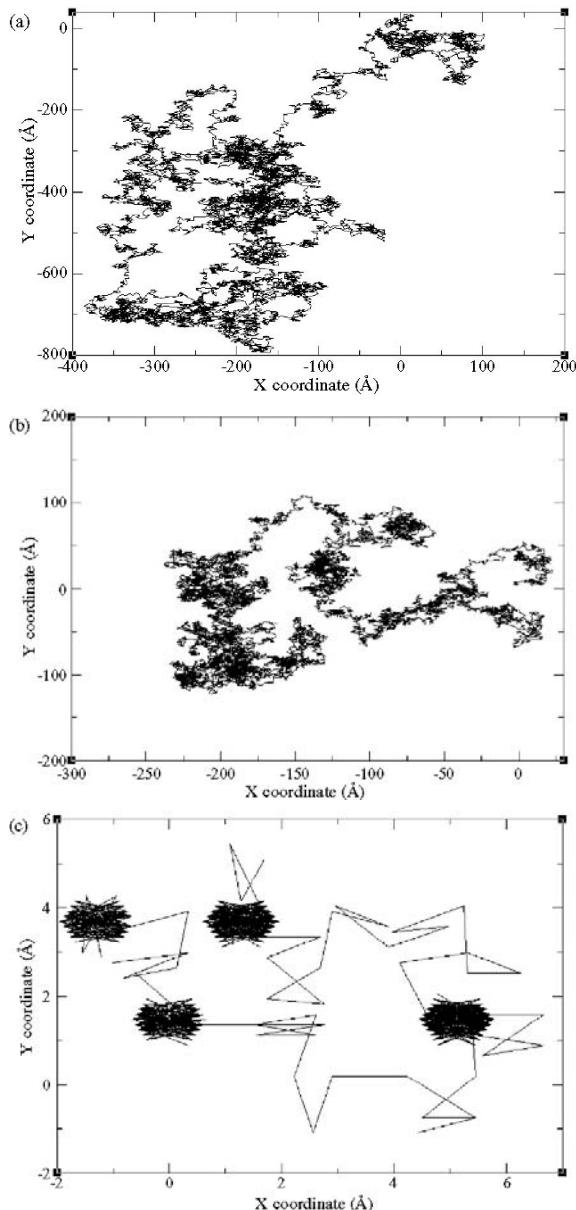
26.5.1 Trace of the Center of Mass

The diffusion of the above mentioned islands was monitored by recording the center of mass motion. We have used a substrate containing 480 atoms (20×24) with the island initially placed at the center of the template. In order to use our pattern recognition scheme without any complications, we shift the island back to the center of the substrate whenever an atom of the island reaches the border of the template. The shift is then recorded to recover the real diffusion path of the center of mass. In the following figures, we show the trace of the CM for 5 islands, containing 10 atoms (general), 18 atoms (general), 19 atoms (compact hexagon), 21 atoms ($6n + 3$) and 26 atoms (general). We choose these islands because the physical time elapsed during 100 million MC steps are comparable: 2686, 3634, 2291, 3753 and 3854 s respectively.

We focus here on a qualitative description of the diffusion of these five islands. For the 10 atom island (Fig. 26.7a), the trace of the CM covers an area of about 60×80 nm and appears to be random and resembles the diffusion path of a single atom.

The 18 atom island (Fig. 26.7b) also shows a large area of about 30×25 nm. Note that in this case one can see some dark areas separated by thin paths indicating a sort of stick slip motion, which is still random. Both these islands belong to the general category.

The 19 atoms island belongs to the compact hexagon category and one can see from Fig. 26.7c that the center of mass hardly moved during the 100 million MC steps corresponding to 2291 s. The center of mass clearly experiences a stick slip motion but with a very long residence time during the stick phase. The trace of the center of mass is regular and is close to a hexagon. The difference between the 18 and the 19 atoms islands is only one atom, which will make the detachment/attachment of atoms very important in the diffusion of these islands. One can imagine a scenario when a 19 atom island starts to diffuse very rapidly because of loss of a peripheral atom, and inversely, a rapidly diffusing 18 atom island may be brought to a halt by



incorporation of a diffusing single atom. Note that in these simulations the atoms are not allowed to detach from the cluster.

Another case of slowly diffusing island is the 21-atom cluster (category $6n+3$) as illustrated in Fig. 26.7d. Here again, the stick slip motion is obvious with a very long stick residence time. The trace of the CM during the stick

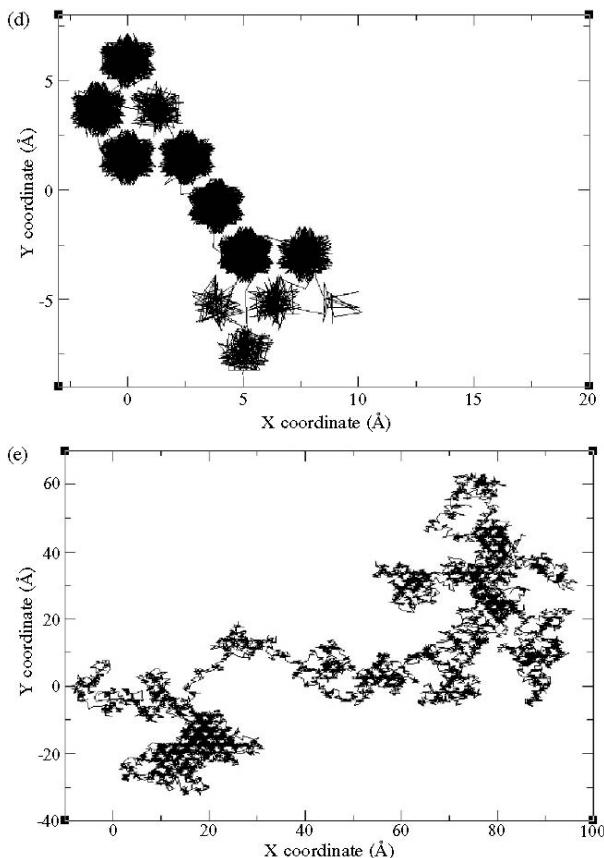


Fig. 26.7. (a) Trace of CM for a 10 atom cluster at 300 K for 100 million MC steps. (b) Trace of CM for a 18 atom cluster at 300 K for 100 million MC steps. (c) Trace of CM for a 19 atom cluster at 300 K for 100 million MC steps. (d) Trace of CM for a 21 atom cluster at 300 K for 100 million MC steps. (e) Trace of CM for a 26 atom cluster at 300 K for 100 million MC steps

phase present a 6-sided star and the area covered during these 3753 s (100 million MC steps) is only about $1 \times 1 \text{ nm}$.

The slow diffusion of these islands is not due to their large size as one would expect, but is rather due to the fact islands with magic numbers diffuse slowly. The 26 atoms (with category “general”, meaning no magic number) is larger than the 19 and the 21 atom ones but diffuses much more rapidly as one can see qualitatively in Fig. 26.7e. The area covered by this island during the 3854 s is about $10 \times 10 \text{ nm}$, considerably larger than the $1 \times 1 \text{ nm}$ covered by the 19 and 21 atom islands. It can also be noted that the trace is made of densely packed dark spots indicating stick-slip motion but with a very short stick residence time. But the overall trace is like a random walk.

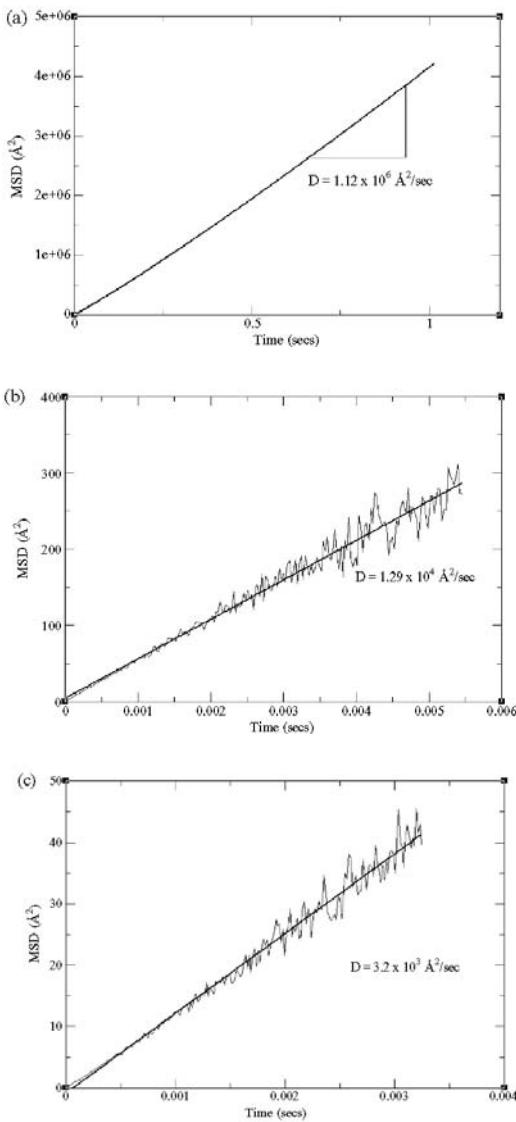


Fig. 26.8. (a) Plot of MSD vs. time for a 10 atom cluster at 500 K. (b) Plot of MSD vs. time for a 19 atom cluster at 500 K. (c) Plot of MSD vs. time for a 100 atom cluster at 500 K

26.5.2 Mean Square Displacement and Diffusion Coefficient

Since our KMC simulation scheme has a very high speed of performance and can run for several hundreds of millions of Monte Carlo steps in about one

day of CPU time on a 2 GHz chip, we have performed simulations in the range of 500 million MC-steps at different temperatures.

Since the physical time elapsed at each MC-step is governed by the rate of the process, they are unequal in length. Thus, to calculate the mean square displacement of the center of mass, we filter our data to a set of almost equidistant MC-step along with the corresponding center of mass coordinates. This filtration is easily possible, due to the large number of iterations (100–500 million) that we perform in the simulations, which enable us to extract enough data points that are almost equidistant in time. From this new filtered data set we calculate the mean square displacement using correlations between these time intervals. We find that if we consider an overlap between the time intervals in order to obtain better averaging, the plots of MSD vs. time appear as a soft parabola. This indicates that overlap in the time intervals can lead to certain dependency effects of data points with their previous values. To rectify this problem, we performed our MSD calculations without an overlap in the time intervals. These plots clearly show a linear behavior. A representative plot of the MSD for three different clusters at 500 K are presented in Fig. 26.8.

For the cases of the magic cluster sizes, which here contain 19, 21, and 38 atoms, we have already seen from the plots of their traces that they hardly move at 300 K in our simulations of 10^8 MC steps. Thus, we cannot extract the MSDs at this temperature. This is not the case at higher temperatures for which the diffusion coefficient has been extracted from the MSD plot shown in Fig. 26.8b. The plot of the MSD for a 100 atom cluster at 500 K, again illustrate the mobility of these clusters. The diffusion coefficients of clusters of various sizes calculated at three temperatures is summarized in Table 26.1. Here, we have also included the case of the 100 atom compact cluster, to get some preliminary ideas of scaling with size. As compared to a 10 atom cluster,

Table 26.1. Diffusion coefficient of clusters as a function of temperatures

Cluster size (atoms)	Diffusion Coefficient D ($\text{\AA}^2/\text{s}$)		
	Temperature of simulation		
	300 K	500 K	1000 K
10	40	1.12×106	1.62×109
14	6.0	2.0×105	4.35×108
18	3.8	1.09×105	2.32×108
19	—	1.29×104	3.08×108
21	—	9.49×103	1.59×108
26	0.48	2.24×104	1.54×108
38	—	2.22×104	7.61×107
100	—	3.2×103	4.7×106

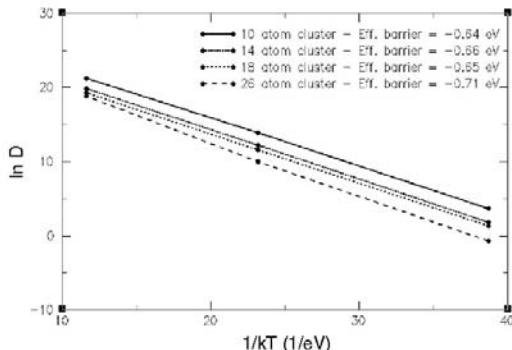


Fig. 26.9. Arrhenius plot with the effective barriers for the 10 atom, 14 atom, 18 atom and 26 atom clusters

we find that the diffusion coefficient of a 100 atom cluster to be three orders of magnitude slower at temperatures of 500 and 1000 K.

The above Arrhenius plot (Fig. 26.9) drawn for clusters in the general category is interesting since it yields values of their effective energy barriers which may also be extracted from experiments. From our simulations the 10, 14, 18 and 26 atom clusters have effective barriers of 0.64, 0.66, 0.65 and 0.71 eV, respectively. As can be observed, there is a slight increase in the effective barrier for the 26 atom cluster as compared to the 10, 14 and 18 atom clusters which have almost the same value. We need to carry out further investigations of the microscopic details of the diffusion processes to draw definite conclusion in this regard.

Some insights into the microscopic details of the cluster size and temperature dependencies of the diffusion processes may be obtained from the frequencies of the various events as recorded in our KMC simulations. Table 26.2 shows a comparison of the frequencies of events for specific examples from each of the four categories of clusters at 300 K. A common feature for clusters in all four categories is that diffusion along step edge A dominates that along step edge B. Of course, this also implies larger occurrence of step A as compared to step B, the rationale for which is in turn related to the corner rounding processes. Several subtle differences can be found for the frequencies of processes involving kink incorporation, corner rounding, etc., for the four types of clusters considered. At 300 K, these small differences manifest themselves in quite striking differences in the traces of the center of mass motion of the clusters. Note that of the 49 processes, only a handful actually participates in the diffusion.

In Tables 26.3–26.6 the frequencies of events executed by a representative cluster from each of the four types, as a function of temperature, are presented. It is observed that the corner rounding processes show a higher count in the more diffusive clusters and their count rises for the magic clus-

Table 26.2. Comparison of frequency of processes in the 4 cluster categories – 300 K

Processes	Cluster Type	Energy Barrier (eV)	10 atom (General)	19 atom (Hexagon)	21 atom (6n + 3)	38 atom (Hexagon+1)
Step Edge A		0.252	0.63	0.62	0.71	0.72
Step Edge B		0.295	0.13	0.17	0.14	0.20
Kink Detach along Step A		0.519	0.0	0.0	0.0	0.0
Kink Detach along Step B		0.556	0.0	0.0	0.0	0.0
Kink Detach along Step (small) A		0.608	0.0086	0.026	0.017	0.0
Kink Detach along Step (small) B		0.680	0.0	0.0016	0.0010	0.0
Kink Incorp. A		0.220	0.0	0.0	0.0	0.0
Kink Incorp. B		0.265	0.0	0.0	0.0	0.0
Kink Incorp. (small) A		0.0075	0.0081	0.025	0.017	0.0
Kink Incorp. (small) B		0.0810	0.0	0.0	0.0	0.0
Kink Detach out of Step A		0.658	0.0	0.0	0.0	0.0
Kink Detach out of Step B		0.590	0.0	0.0	0.0	0.0
Kink Fall into Step A		0.074	0.0	0.0	0.0	0.0
Kink Fall into Step B		0.0069	0.0	0.0	0.0	0.0
Kink Rounding A		0.656	0.0	0.0	0.0	0.0
Kink Rounding B		0.678	0.0	0.0	0.0	0.0
KESE A		0.374	0.0	0.0	0.0	0.0
KESE B		0.402	0.0	0.0	0.0	0.0
Corner Rounding at AA stage 1		0.313	0.033	0.0	0.0031	0.0
Corner Rounding at AA stage 2		0.143	0.0	0.0	0.0	0.0
Corner Rounding at AA stage 3		0.0096	0.031	0.0	0.0031	0.0
Corner Rounding at BB stage 1		0.374	0.0033	0.0	0.0	0.0
Corner Rounding at BB stage 2		0.038	0.0024	0.0	0.0	0.0
Corner Rounding at BB stage 3		0.052	0.0012	0.0	0.0	0.0
Corner Rounding at AB stage 1		0.317	0.066	0.066	0.045	0.029
Corner Rounding at AB stage 2		0.0839	0.0042	0.0053	0.0029	0.0021
Corner Rounding at BA stage 1		0.396	0.0043	0.0047	0.0027	0.0021
Corner Rounding at BA stage 2		0.0148	0.066	0.067	0.045	0.029
Rounding a chain A		0.0626	0.0030	0.0	0.0	0.0
Rounding a chain B		0.0191	0.0023	0.0	0.0	0.0

ters as the temperature is increased. Also note that the frequency of the motion against step A is between 60–70% at 300 K for all clusters. Although at higher temperatures this frequency decreases, it still remains the dominant one compared to the other processes. This can be explained by looking at the barriers for the corner rounding at the AB and BA junctions, where we can see that the barrier to go from the tip of the corner towards step A (BA stage 2) is much lower than that from tip of corner towards step B (AB stage 2). As a result the atom prefers to diffuse along the step edge A.

Table 26.3. Frequency of processes for the 10 atom cluster (general) at different temperatures

Processes	Temperature	Energy Barrier (eV)	300 K	500 K	1000 K
Step Edge A		0.252	0.63	0.38	0.19
Step Edge B		0.295	0.13	0.14	0.11
Kink Detach along Step A		0.519	0.0	0.0	0.0072
Kink Detach along Step B		0.556	0.0	0.0	0.0044
Kink Detach along Step (small) A		0.608	0.0086	0.019	0.031
Kink Detach along Step (small) B		0.680	0.0	0.0035	0.013
Kink Incorp. A		0.220	0.0	0.0	0.0059
Kink Incorp. B		0.265	0.0	0.0	0.0035
Kink Incorp. (small) A		0.0075	0.0081	0.017	0.027
Kink Incorp. (small) B		0.0810	0.0	0.0026	0.010
Kink Detach out of Step A		0.658	0.0	0.0	0.0
Kink Detach out of Step B		0.590	0.0	0.0	0.0
Kink Fall into Step A		0.074	0.0	0.0	0.0
Kink Fall into Step B		0.0069	0.0	0.0	0.0
Kink Rounding A		0.656	0.0	0.0	0.0
Kink Rounding B		0.678	0.0	0.0	0.0
KESE A		0.374	0.0	0.0020	0.0051
KESE B		0.402	0.0	0.0	0.0039
Corner Rounding at AA stage 1		0.313	0.033	0.049	0.054
Corner Rounding at AA stage 2		0.143	0.0	0.0025	0.011
Corner Rounding at AA stage 3		0.0096	0.031	0.042	0.041
Corner Rounding at BB stage 1		0.374	0.0033	0.013	0.028
Corner Rounding at BB stage 2		0.038	0.0024	0.011	0.025
Corner Rounding at BB stage 3		0.052	0.0012	0.0059	0.014
Corner Rounding at AB stage 1		0.317	0.066	0.11	0.12
Corner Rounding at AB stage 2		0.0839	0.0042	0.023	0.055
Corner Rounding at BA stage 1		0.396	0.0043	0.023	0.055
Corner Rounding at BA stage 2		0.0148	0.066	0.11	0.12
Rounding a chain A		0.0626	0.0030	0.010	0.018
Rounding a chain B		0.0191	0.0023	0.0086	0.017

We have extended the simulations to a set of larger clusters containing 223, 409, 612 and 1003 atoms, to gain further information on the scaling of the diffusion coefficient with size. Table 26.4 summarizes our results for the diffusion coefficients at 500 K. As expected, there is a decrease in the value of D with the size of the cluster. Figure 26.10 below shows a plot of Log D vs. Log N , where D is the diffusion coefficient and N is the size of the cluster. The scaling exponent that we obtain for the larger cluster sized ranging from 55 atoms to 1003 atoms is close to -0.6 . Values of the exponent

Table 26.4. Frequency of processes for the 19 atom cluster (Hexagon) at different temperatures

Processes	Temperature Barrier (eV)	Energy	300 K	500 K	1000 K
Step Edge A	0.252	0.62	0.42	0.23	
Step Edge B	0.295	0.17	0.24	0.15	
Kink Detach along Step A	0.519	0.0	0.0020	0.019	
Kink Detach along Step B	0.556	0.0	0.0	0.012	
Kink Detach along Step (small) A	0.608	0.026	0.012	0.043	
Kink Detach along Step (small) B	0.680	0.0016	0.0023	0.019	
Kink Incorp. A	0.220	0.0	0.0020	0.018	
Kink Incorp. B	0.265	0.0	0.0	0.011	
Kink Incorp. (small) A	0.0075	0.025	0.011	0.037	
Kink Incorp. (small) B	0.0810	0.0	0.0012	0.012	
Kink Detach out of Step A	0.658	0.0	0.0	0.0013	
Kink Detach out of Step B	0.590	0.0	0.0	0.0027	
Kink Fall into Step A	0.074	0.0	0.0	0.0016	
Kink Fall into Step B	0.0069	0.0	0.0	0.0023	
Kink Rounding A	0.656	0.0	0.0	0.0011	
Kink Rounding B	0.678	0.0	0.0	0.0	
KESE A	0.374	0.0	0.0011	0.0078	
KESE B	0.402	0.0	0.0	0.0066	
Corner Rounding at AA stage 1	0.313	0.0	0.0	0.014	
Corner Rounding at AA stage 2	0.143	0.0	0.0	0.0050	
Corner Rounding at AA stage 3	0.0096	0.0	0.0	0.012	
Corner Rounding at BB stage 1	0.374	0.0	0.0	0.0079	
Corner Rounding at BB stage 2	0.038	0.0	0.0	0.012	
Corner Rounding at BB stage 3	0.052	0.0	0.0	0.0049	
Corner Rounding at AB stage 1	0.317	0.066	0.11	0.11	
Corner Rounding at AB stage 2	0.0839	0.0053	0.024	0.055	
Corner Rounding at BA stage 1	0.396	0.0047	0.023	0.051	
Corner Rounding at BA stage 2	0.0148	0.067	0.12	0.11	
Rounding a chain A	0.0626	0.0	0.0	0.0053	
Rounding a chain B	0.0191	0.0	0.0	0.0054	

from earlier work are found to be around -1.5 [10, 17, 18]. Our results thus imply that the calculated decrease in D with cluster size is slower than that found in earlier works. This could be attributed to the inclusion of a larger number of processes in our database, some of which have very low energy barriers allowing the atoms of the cluster to fluctuate in a smaller region of phase space and hence leading to lower diffusivity of the center of mass. It could also be that some of the activation energy barriers calculated from EAM potentials are not accurate. It would thus be worthwhile to carry out

Table 26.5. Frequency of Processes for the 21 atom cluster ($6n + 3$) at different temperatures

Processes	Temperature	Energy	300 K	500 K	1000 K
		Barrier (eV)			
Step Edge A		0.252	0.71	0.48	0.24
Step Edge B		0.295	0.14	0.19	0.15
Kink Detach along Step A		0.519	0.0	0.0018	0.021
Kink Detach along Step B		0.556	0.0	0.0	0.012
Kink Detach along Step (small) A		0.608	0.017	0.029	0.046
Kink Detach along Step (small) B		0.680	0.0010	0.0054	0.020
Kink Incorp. A		0.220	0.0	0.0018	0.019
Kink Incorp. B		0.265	0.0	0.0	0.011
Kink Incorp. (small) A		0.0075	0.017	0.026	0.039
Kink Incorp. (small) B		0.0810	0.0	0.0029	0.013
Kink Detach out of Step A		0.658	0.0	0.0	0.0016
Kink Detach out of Step B		0.590	0.0	0.0	0.0033
Kink Fall into Step A		0.074	0.0	0.0	0.0019
Kink Fall into Step B		0.0069	0.0	0.0	0.0032
Kink Rounding A		0.656	0.0	0.0	0.0017
Kink Rounding B		0.678	0.0	0.0	0.0012
KESE A		0.374	0.0	0.0027	0.0086
KESE B		0.402	0.0	0.0020	0.0071
Corner Rounding at AA stage 1		0.313	0.0031	0.0063	0.014
Corner Rounding at AA stage 2		0.143	0.0	0.0	0.0050
Corner Rounding at AA stage 3		0.0096	0.0031	0.0062	0.012
Corner Rounding at BB stage 1		0.374	0.0	0.0016	0.0081
Corner Rounding at BB stage 2		0.038	0.0	0.0019	0.011
Corner Rounding at BB stage 3		0.052	0.0	0.0010	0.0052
Corner Rounding at AB stage 1		0.317	0.045	0.090	0.10
Corner Rounding at AB stage 2		0.0839	0.0029	0.019	0.050
Corner Rounding at BA stage 1		0.396	0.0027	0.018	0.046
Corner Rounding at BA stage 2		0.0148	0.045	0.09	0.10
Rounding a chain A		0.0626	0.0	0.0011	0.0052
Rounding a chain B		0.0191	0.0	0.0013	0.0054

ab initio calculations for some of the key processes using electronic structure calculations. However, since there is still ambiguity about the scaling of D with size from experiments, we are not in a position to comment on the validity of these results.

Table 26.6. Frequency of processes for the 38 atom cluster (Hexagon+1) at different temperatures

Processes	Temperature Barrier (eV)	Energy	300 K	500 K	1000 K
Step Edge A	0.252	0.72	0.53	0.33	
Step Edge B	0.295	0.20	0.25	0.42	
Kink Detach along Step A	0.519	0.0	0.012	0.021	
Kink Detach along Step B	0.556	0.0	0.0046	0.014	
Kink Detach along Step (small) A	0.608	0.0	0.0098	0.032	
Kink Detach along Step (small) B	0.680	0.0	0.0017	0.013	
Kink Incorp. A	0.220	0.0	0.012	0.021	
Kink Incorp. B	0.265	0.0	0.0044	0.014	
Kink Incorp. (small) A	0.0075	0.0	0.0094	0.030	
Kink Incorp. (small) B	0.0810	0.0	0.0010	0.014	
Kink Detach out of Step A	0.658	0.0	0.0	0.0038	
Kink Detach out of Step B	0.590	0.0	0.0013	0.0077	
Kink Fall into Step A	0.074	0.0	0.0	0.0	
Kink Fall into Step B	0.0069	0.0	0.0011	0.0072	
Kink Rounding A	0.656	0.0	0.0	0.0	
Kink Rounding B	0.678	0.0	0.0	0.0	
KESE A	0.374	0.0	0.0012	0.0020	
KESE B	0.402	0.0	0.0	0.0020	
Corner Rounding at AA stage 1	0.313	0.0	0.0	0.0026	
Corner Rounding at AA stage 2	0.143	0.0	0.0	0.0	
Corner Rounding at AA stage 3	0.0096	0.0	0.0	0.0022	
Corner Rounding at BB stage 1	0.374	0.0	0.0	0.0014	
Corner Rounding at BB stage 2	0.038	0.0	0.0	0.0023	
Corner Rounding at BB stage 3	0.052	0.0	0.0	0.0	
Corner Rounding at AB stage 1	0.317	0.029	0.062	0.011	
Corner Rounding at AB stage 2	0.0839	0.0021	0.013	0.0080	
Corner Rounding at BA stage 1	0.396	0.0021	0.012	0.0064	
Corner Rounding at BA stage 2	0.0148	0.029	0.063	0.013	
Rounding a chain A	0.0626	0.0	0.0	0.0	
Rounding a chain B	0.0191	0.0	0.0	0.0	

26.6 Conclusions

In summary, we have performed a kinetic Monte Carlo study of the diffusion of small Cu clusters on Cu(111) surface, using a unique pattern recognition scheme and allowing only single atom peripheral motion. A database of 49 key processes is used in the simulations and the range of MC steps is from 100–500 million. The temperatures used in the simulation are 300, 500 and 1000 K. At 300 K, three types of magic clusters with sizes forming a perfect

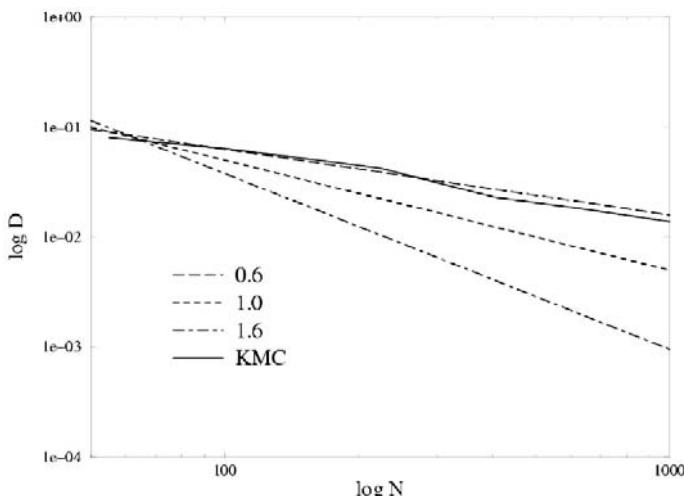


Fig. 26.10. Scaling of diffusion coefficient D with size of cluster N . The plot of the KMC simulation is for the larger cluster sizes ranging from 55–1003 atoms. The slope gives the scaling exponent as -0.6

Table 26.7. Diffusion coefficient for large clusters at 500 K

Cluster size (atoms)	Diffusion coefficient D ($\text{\AA}^2/\text{s}$)
55	4.0×10^{-3}
100	3.2×10^{-3}
223	2.12×10^{-3}
409	1.14×10^{-3}
612	0.92×10^{-3}
1003	0.69×10^{-3}

hexagon, $6n + 3$ and hexagon+1, are found to exhibit non random motion with striking paths traced by the center of mass. At higher temperatures, all four types of clusters display random motion of their center of mass. This is attributed to the fact that many other processes are executed by the atoms at higher temperatures, thus enhancing their motion. The frequency list of the processes executed show that diffusion against step edge A is the most frequently executed process accounting for 60–70% of the total count at 300 K, about 45% at 500 K and about 25% at 1000 K. This can be explained by looking at the barriers for the corner rounding at the AB and BA junctions, where we see that the barrier to go from the tip of the corner towards step A (BA stage 2) is much lower than that from tip of corner towards step B (AB stage 2). Atoms thus prefer to diffuse along the step edge A. The corner rounding processes have a high count in the cases of general clusters and also for all clusters at elevated temperatures, especially the count at the AA

and BB junctions increases with temperature and this could be a key factor for the enhanced diffusivities at these temperatures. We also find that the diffusion coefficient D shows a difference of 3 orders of magnitude in the diffusivities between a 10 atom and a 100 atom cluster. The shapes of the clusters during growth and post deposition are strongly dependent on the number of atoms in the cluster, whether they fall under any of the above mentioned magic categories or the general category. The clusters evolve with time and try to achieve a compact shape which is close to hexagonal, or an irregular hexagon. We have also found that at higher temperatures of 1000 K, vacancies begin to form within the cluster interior. This could lead to the locking of the CM of the cluster to within a few angstroms, since peripheral atoms in such a configuration have fewer options for motion. The disagreement between our scaling behavior and previous results suggests that more accurate studies of the activation energy barriers is in order. A point missing for this and previous similar calculations in the literature is the role of the diffusion pre-exponential factor which we, like others, have taken to have the same generic value for all processes. This assumption is bound to have its limitations. Future work should address this particular issue.

Acknowledgement

We would like to express our thanks to Oleg Trushin and Altaf Karim for many illuminating discussion. This work was supported by the grant NSF-EEC0085604.

References

1. G.L. Kellogg, A.F. Voter: Phys. Rev. Lett. **67**, 622 (1991)
2. S.C. Wang, U. Kürpick, G. Ehrlich: Phys. Rev. Lett. **81**, 4923 (1998)
3. W.W. Pai, A.K. Swan, Z. Zhang, J.F. Wendelken: Phys. Rev. Lett. **79**, 3210 (1997)
4. K. Morgenstern, G. Rosenfeld, B. Poelsema, G. Comsa: Phys. Rev. Lett. **74**, 2058, (1995)
5. K. Morgenstern, G. Rosenfeld, G. Comsa: Phys. Rev. Lett. **76**, 2113, (1996)
6. J.M. Soler: Phys. Rev. B **50**, 5578 (1994)
7. J.M. Wen, S.L. Chang, J.W. Burnett, J.W. Evans, P.A. Thiel: Phys. Rev. Lett. **73**, 2591 (1994)
8. J.M. Soler: Phys. Rev. B **53**, R10 540 (1996)
9. S.V. Khare, N.C. Bartelt, T.L. Einstein: Phys. Rev. Lett. **75**, 2148 (1995); J.C. Hamilton, M.S. Daw, S.M. Foiles: Phys. Rev. Lett. **74**, 2760 (1995); C. DeW. Van Siclen: Phys. Rev. Lett. **75**, 1574 (1995)
10. A. Bogicevic, S. Liu, J. Jacobsen, B. Lundqvist, H. Metiu: Phys. Rev. B **57**, R9459 (1998)
11. S.M. Foiles, M.I. Baskes, M.S. Daw: Phys Rev. B **33**, 7983 (1986)

12. L. Yang, T.S. Rahman: Phys. Rev. Lett. **67**, 2327 (1991); L. Yang, T.S. Rahman, M.B. Daw: Phys. Rev. B **44**, 13725 (1991); T.S. Rahman, in: Cond. Matt. Theory 9 (Nova, New York 1994) p. 299
13. S. Durukanoglu, A. Kara, T.S. Rahman: Phys. Rev. B **67**, 235405 (2003)
14. T.S. Rahman, A. Kara, S. Durukanoglu: J. Phys. Condens. Matter **47**, 15 (2003)
15. U. Kürpick, A. Kara, T.S. Rahman: Phys. Rev. Lett. **78**, 1086 (1997)
16. H. Jonsson, G. Mills, K.W. Jacobsen, in: Classical and Quantum Dynamics in Condensed Phase Simulations, ed. by B.J. Berne et al. (World Scientific, Singapore 1998)
17. D.S. Sholl, R.T. Skodje: Phys. Rev. Lett. **75**, 3158 (1995)
18. S.V. Khare, T.L. Einstein: Phys. Rev. B **54**, 11752 (1996)
19. M. Giesen, G. Schulze Icking-Konert, H. Ibach: Phys. Rev. Lett. **80**, 552 (1998)
20. P.J. Feibelman: Surf. Sci. Lett. **478**, L349 (2001)

27 Including Long-Range Interactions in Atomistic Modelling of Diffusional Phase Changes

D.R. Mason¹, R.E. Rudd², and A.P. Sutton^{1,3}

¹ Department of Materials, Oxford University, OX1 3PH, UK

² Lawrence Livermore National Laboratory, 7000 East Ave., L-045, Livermore, CA 94550, USA

³ Helsinki University of Technology, Laboratory of Computational Engineering, PO.Box 9203, FIN 02015 HUT, Finland

Abstract. Phase transformations in 2xxx series aluminium alloys (Al-Cu-Mg) are investigated with an off-lattice atomistic kinetic Monte Carlo simulation incorporating the effects of strain around misfitting atoms and vacancies. Vacancy diffusion is modelled by comparing the energies of trial states, where the system is partially relaxed for each trial state. Only a limited precision is required for the energy of each trial state, determined by the value of $k_B T$. Since the change in the relaxation displacement field caused by a vacancy hop decays as $1/r^3$, it is sufficient to determine the next move by relaxing only those atoms in a sphere of finite radius centred on the moving vacancy. However, once the next move has been selected, the entire system is relaxed. Simulations of the early stages of phase separation in Al-Cu with elastic relaxation show an enhanced rate of clustering compared to those performed on the same system with a rigid lattice. However on a flexible lattice vacancy trapping by Mg atoms in the ternary Al-Cu-Mg system makes clustering slower than the corresponding rigid lattice calculation.

27.1 Introduction

The mechanical properties of commercial aluminium alloys are manipulated by the process of ageing the supersaturated solid solution. In the Al-4 wt.% Cu system, copper atoms cluster together on $\{200\}_{\text{Al}}$ host planes to form metastable particles known as Guinier-Preston (GP) zones. On further ageing the GP zones undergo a change of structure before transforming eventually into the thermodynamically stable θ phase CuAl_2 , which is incoherent with the host. In the 2xxx series of alloys, magnesium is added and the precipitation sequence changes. First coherent GPB zones form as $\langle 100 \rangle$ rods, which ripen into laths of the orthorhombic S phase Al_2CuMg [1–5].

While the observed structures of GP zones are well established [6–8], and the effects of the addition of certain trace elements have been studied experimentally [9–11], modelling the effect of trace elements on kinetics is a difficult problem. Their influence may stem from the formation of complexes with vacancies, which will reduce the rate of diffusion and of phase separation. It is also possible that large trace element atoms may be attracted to small

solute atoms by the concomitant reduction in strain energy of the system, and in this way promote the formation of a co-cluster. Apart from reductions in the total strain energy of the system through such an association there may also be much shorter-range gains in bond energies. It follows that the influence of trace elements is inherently atomistic in nature, in that it stems from discrete atomic interactions with point defects such as vacancies and other solute atoms. However, misfitting atoms and vacancies interact over large distances compared with a bond length through their elastic fields. It is the requirement to treat both discrete atomic interactions and long-range elastic fields on an equal footing in dilute alloys that makes the modelling of the influence of trace elements so challenging. Our aim has been to develop a methodology to meet this challenge. Since our methodology treats atomic vacancies explicitly we are able to include the possibility that vacancies may become trapped at interfaces surrounding second-phase particles, which may lead to Brownian motion of particles as described in [12, 13].

A vacancy is a centre of tensile dilation, and so it will be attracted to regions of compressive stress. The pattern of compressive stress will follow the underlying local composition of the system, enhancing the time spent by a vacancy in some areas, while reducing it in others. If the pattern of compressive stress is enhanced along particular crystallographic directions, owing to elastic anisotropy, the increased probability of finding a vacancy in these regions may be expected to influence the mobility of solute atoms there, which in turn may influence the morphology of the second-phase particle.

There are three principal steps we have taken to enable us to model diffusional processes atomistically with elastic interactions. Firstly we have improved the kinetic Monte Carlo algorithm used to select each vacancy hop, by developing an efficient combination of the stochastic first and second order residence algorithms which outperforms the n -fold way for this problem. This is described in detail elsewhere [14]. Secondly we have developed an iterative scheme based on the Lanczos method to tridiagonalise a matrix, which appears particularly suited to relaxing the system following a vacancy–atom exchange. Finally, although the relaxation propagates to the boundaries of the system, we have found that the majority of the *change* of the elastic energy due to a vacancy–atom exchange can be recovered by relaxing only those atoms in a near field region.

We demonstrate the importance of considering elastic stress effects when performing kinetic Monte Carlo simulations. We show that the harmonic lattice approximation is insufficient for accurately predicting the difference in energy between trial Monte Carlo states. However, an approximation based on relaxing those atoms within a small sphere centred on the vacancy is enough to determine the next move, provided that eventually the whole system is relaxed after each move. Simulations of diffusion performed in the Al–Cu binary alloy system show that for a fixed number of vacancy–atom exchanges, larger clusters grow on a flexible lattice owing to the reduction in elastic

strain energy this can provide. However, it is also found that since vacancies are preferentially attracted to compressive regions, they can become trapped within emerging magnesium-rich regions in an Al-Cu-Mg ternary alloy, and this can have the opposite effect of slowing the growth rate of second-phase particles.

27.2 Computational Method

27.2.1 Kinetic Monte Carlo

In our example of vacancy diffusion, adjacent atomic configurations are linked by a single vacancy-atom exchange event. Kinetic Monte Carlo [14–17] enables us to associate a meaningful time with simulated stochastic events. The probability that any one of these events will be the next to occur is proportional to the rate at which the event occurs. We have used Flynn's approximation [18] derived from dynamical theory to determine a rate for the transition based on the difference in internal energy between states. The rate of migration from state i to state j is given by

$$r_{i \rightarrow j} = \left(\frac{3}{5} \right)^{1/2} \nu_D \exp \left(- \frac{\langle c \rangle \Omega \Delta^2}{k_B T} \right) \exp \left(- \frac{E_j - E_i}{2k_B T} \right), \quad (27.1)$$

where E_i is the energy of the system in state i when the system is fully relaxed, ν_D is the Debye frequency and Ω the volume of the migrating atom. We have assumed a constant Debye frequency equal to that of pure aluminium, taken to be 1.19×10^{13} Hz. The parameter Δ is the displacement, expressed as a fraction of the bond length, at which the force on the migrating atom is a maximum, and is taken to be the constant value 0.316. $1/\langle c \rangle$ is an effective elastic compliance. For fcc crystals it is given by $1/\langle c \rangle \approx 2/15(3/c_{11} + 2/c_{11-c_{12}} + 1/c_{44})$. Only $1/2\langle 110 \rangle$ hops are considered. Saddle points for exchanges with the next-nearest neighbour shell of atoms are around 2 eV higher in energy and so are assumed too infrequent to affect the kinetics.

We have constructed a set of relatively short-ranged Finnis–Sinclair potentials [19] for this study, described fully in [14]. They have been parameterised to reproduce elastic constants, the cohesive energy and lattice parameters of the pure metals, and we use a simple interpolation scheme for the parameters of the potentials for interactions between dissimilar atoms.

27.2.2 Lanczos Method for Relaxation

An efficient algorithm for finding relaxed states is required because there may be many millions of configurations generated during a kinetic Monte Carlo simulation of the early stages of diffusional phase separation by a vacancy

mechanism. Consider first the harmonic lattice approximation. The energy may be expanded as a Taylor series about the current configuration which may be represented by a vector \mathbf{x}_0 describing the positions of each of the N atoms in the system:

$$E(\mathbf{x}_0 + \mathbf{u}) = E(\mathbf{x}_0) + \left. \frac{\partial E}{\partial \mathbf{x}} \right|_{\mathbf{x}_0} \cdot \mathbf{u} + \frac{1}{2} \mathbf{u} \cdot \left. \frac{\partial^2 E}{\partial \mathbf{x} \partial \mathbf{x}} \right|_{\mathbf{x}_0} \mathbf{u}. \quad (27.2)$$

The vector \mathbf{u} is a displacement from the current configuration \mathbf{x}_0 , which is assumed to be small in magnitude. We identify the first derivative of the energy as (minus) the $3N$ dimensional force $\mathbf{f} = -\partial E / \partial \mathbf{x}|_{\mathbf{x}_0}$ and the second as the matrix of force constants $\mathbf{D} = \partial^2 E / \partial \mathbf{x} \partial \mathbf{x}|_{\mathbf{x}_0}$. These two derivatives are evaluated analytically at the current configuration \mathbf{x}_0 using the interatomic potentials.

Equation (27.2) has a minimum with respect to \mathbf{u} given by

$$\mathbf{u} = \mathbf{D}^{-1} \mathbf{f}. \quad (27.3)$$

The change in energy brought about by making this displacement from the current configuration is obtained by substituting (27.3) into (27.2), giving $\Delta E = -1/2 \mathbf{f} \cdot \mathbf{D}^{-1} \mathbf{f}$.

For the aluminium-copper alloys of interest to this study the displacements of atoms near a vacancy are relatively large and sensitive to the local atomic structure. We have found the difference in energy between harmonically relaxed states to be in error, compared to fully relaxed states, by the order of 0.1 eV. Since this is the same order as the total energy difference between states, counting both chemical bond energy changes and elastic energy changes, the harmonic relaxation approximation is seen to be unsuitable for this study.

To implement an anharmonic relaxation we have found that significant computational gains may be made if an algorithm derived from the Lanczos method of tridiagonalizing a matrix [20] is employed. The strategy is to find an approximate solution to (27.3) and displace atoms through this vector. The force and force constant matrix elements are then recalculated and further displacements constructed until the system energy has converged to an acceptable tolerance. Restarting the calculation in this way not only prevents error accumulation, allowing a machine precision solution if desired, but also side-steps the known problems associated with keeping the Lanczos basis vectors orthonormal.

The Lanczos method is perhaps most familiar in solid state physics in the context of calculating local densities of electronic or vibrational states [21]. Our use of the method here is quite different, namely we are generating an approximate solution for \mathbf{u} in (27.3) for which ΔE has converged. The method constructs a new orthonormal basis set of vectors. After m such basis vectors have been generated, the subspace they span is known as the Krylov subspace: $\mathcal{K}_m(\mathbf{D}, \mathbf{f}) = \{\mathbf{f}, \mathbf{D}\mathbf{f}, \dots, \mathbf{D}^{m-1}\mathbf{f}\}$.

The representation of \mathbf{D} in this basis is the tridiagonal $m \times m$ matrix \mathbf{T}_m .

$$\text{define } \mathbf{T}_m : \mathbf{T}_m \equiv \Phi_m^T \mathbf{D} \Phi_m = \begin{pmatrix} \alpha_0 & \beta_1 & & & \\ \beta_1 & \alpha_1 & \beta_2 & & \\ & & & \ddots & \\ & & & \ddots & \beta_{m-1} \\ & & & \beta_{m-1} & \alpha_{m-1} \end{pmatrix}, \quad (27.4)$$

where the columns of the $3N \times m$ matrix Φ_m are the m generated basis vectors $\{\phi_0, \phi_1, \dots, \phi_{m-1}\}$ each of dimension $3N$. This matrix spans the same subspace as $\mathcal{K}_m(\mathbf{D}, \mathbf{f})$.

The coefficients α and β and the basis vectors ϕ are found by the Lanczos algorithm, which is an iterative process (see, e.g, [20]) requiring only matrix-vector multiplies:

$$\phi_{-1} = \mathbf{0}, \quad \phi_0 = \frac{1}{|\mathbf{f}|} \mathbf{f}, \quad \beta_{m+1} \phi_{m+1} = \mathbf{D} \phi_m - \alpha_m \phi_m - \beta_m \phi_{m-1},$$

$$\text{where } \alpha_m = \phi_m \cdot \mathbf{D} \phi_m \quad \text{and} \quad \beta_m = \begin{cases} \phi_{m-1} \cdot \mathbf{D} \phi_m, & m > 0 \\ |\mathbf{f}|, & m = 0 \end{cases}. \quad (27.5)$$

We may obtain approximate solutions for ΔE and \mathbf{u} in the subspace $\mathcal{K}_m(\mathbf{D}, \mathbf{f})$. The predicted relaxation energy at this stage is

$$\Delta E_m = -\frac{1}{2} |\mathbf{f}|^2 [\mathbf{T}_m^{-1}]_{00}, \quad (27.6)$$

$[\mathbf{T}_m^{-1}]_{00}$ is the element in the top left hand-corner of the inverse of the matrix \mathbf{T}_m . The predicted displacement at this level is $\mathbf{u}_m = |\mathbf{f}| \Phi_m \mathbf{T}_m^{-1} \mathbf{e}_0$, where \mathbf{e}_0 is the m -dimensional unit vector $[1, 0, 0, \dots, 0]^T$. The iterative procedure of (27.5) can be halted when ΔE_m has converged, and then the displacement vector may be found.

We note that the displacement vector may be constructed directly from (27.5) as the basis vectors are constructed, in a manner analogous to that used by the method of Conjugate Gradients. However, we use an alternative method for calculating the displacement vector which is an iterative scheme using the change in energy at level m . Our method does require storing the basis vector set Φ_m , but we believe this extra cost to be offset by the improvement in convergence available by the method of Sect. 27.2.3. We expand the displacement in the m generated basis vectors: $\mathbf{u}_m = \sum_{j=0}^{m-1} \gamma_j \phi_j$. The orthonormality condition of the basis vectors ϕ_m and (27.5) gives

$$\gamma_0 = -\frac{2\Delta E_m}{\beta_0}, \quad \gamma_1 = \frac{\beta_0 - \gamma_0 \alpha_0}{\beta_1}, \quad \gamma_k = -\frac{\gamma_{k-2} \beta_{k-1} + \gamma_{k-1} \alpha_{k-1}}{\beta_k}. \quad (27.7)$$

As seen in (27.5), the initial vector ϕ_0 is a unit vector in $3N$ dimensional space in the direction of the force \mathbf{f} . Each successive multiplication by \mathbf{D} in

(27.5) enables the relaxation to spread further from the centres where the forces in \mathbf{f} are located. Since the change in the relaxation energy associated with a vacancy hop decays rapidly with distance from the vacancy, a rapidly convergent estimate of the relaxation energy may be obtained, together with a good approximation to \mathbf{u} . Thus there is a sound physical basis for our choice of the Lanczos method to relax the system for each trial vacancy-atom exchange.

27.2.3 Improving the Solution

We have found that the efficiency of the convergence of the Lanczos procedure can be improved by storing and reusing information generated during previous relaxations. The values of α and β found during the tridiagonalisation process can improve the estimate of the energy change after only m levels have been performed. By substituting this improved energy into (27.7), the displacement vector is also improved.

The change in energy after m levels can be written as a continued fraction (see, e.g., Heine in [21]). From (27.6) we write

$$\Delta E_m = \cfrac{-1/2\beta_0^2}{\alpha_0 - \cfrac{\beta_1^2}{\alpha_1 - \cfrac{\beta_2^2}{\alpha_2 - \cfrac{\vdots}{\alpha_{m-1} - \beta_m^2}}}} \quad (27.8)$$

Going to another level $m+1$ does not affect the previously found coefficients α and β , but can improve the energy estimate by adding to the end of the continued fraction. It would be possible to improve the estimate of the energy change if we could terminate the continued fraction at level m with something which better represents the tail of the fraction. It is found numerically that averages a and b constructed from the coefficients α and β converge to constant values, but that even after thousands of levels the values for α and β rapidly oscillate about these averages. As a first choice, a quadratic terminator constructed from the averages could be used to replace the tail of the continued fraction at level m , by substituting β_m^2 in (27.8) with β_m^2/t , where

$$a = \frac{1}{m} \sum_{j=0}^{m-1} \alpha_j, \quad b = \frac{1}{m} \sum_{j=1}^m \beta_j, \quad t = a - \frac{b^2}{t} = \frac{a}{2} \left(1 + \sqrt{1 - \left(\frac{2b}{a} \right)^2} \right). \quad (27.9)$$

The form for the quadratic terminator in (27.9) is valid because the discriminator is always positive.

While it is possible to ensure convergence in these averages by going to enough levels, we are looking for a displacement vector estimated from only

the first few basis vectors. Too many levels of the procedure gives a well-converged solution only to the harmonic problem of (27.3). We find ΔE_m has typically converged to 0.1% when \mathbf{T}_m is constructed to level $m = 6$ to 12. Note that the numerical errors at this level remain small. This is why we have no need for renormalisation of the basis set Φ_m . We are trying to find a displacement vector which will minimise the system energy on a weakly anharmonic potential energy surface. To converge to the true minimum we calculate the force and matrix elements at the current atom displacement, call the Lanczos procedure of (27.5) to generate a better solution and move the atoms. Starting from an initial position \mathbf{x}_0 , this produces a succession of estimates for the local energy minimum $\mathbf{x}_0, \mathbf{x}_0 + \tilde{\mathbf{u}}^{(0)}, \mathbf{x}_0 + \tilde{\mathbf{u}}^{(0)} + \tilde{\mathbf{u}}^{(1)}, \dots$, where $(\)$ denotes an approximated value. The displacement vectors are generated by

$$\begin{aligned} \mathbf{D}|_{\mathbf{x}_0} \tilde{\mathbf{u}}^{(0)} &\approx \mathbf{f}|_{\mathbf{x}_0} \\ \mathbf{D}|_{\mathbf{x}_0 + \tilde{\mathbf{u}}^{(0)}} \tilde{\mathbf{u}}^{(1)} &\approx \mathbf{f}|_{\mathbf{x}_0 + \tilde{\mathbf{u}}^{(0)}} \\ &\dots \end{aligned} \quad (27.10)$$

We are using the Newton-Raphson method to approach the solution, so the input forces $\mathbf{f}|_{\mathbf{x}_0}, \mathbf{f}|_{\mathbf{x}_0 + \tilde{\mathbf{u}}^{(0)}}, \dots$ diminish in magnitude quadratically. However, the input matrices $\mathbf{D}|_{\mathbf{x}_0}, \mathbf{D}|_{\mathbf{x}_0 + \tilde{\mathbf{u}}^{(0)}}, \dots$ change only slightly as the the position of the atoms is updated, reflecting the weakly anharmonic nature of the potentials near the minimum position.

As a function of the number of levels m , information about the extreme eigenvalues converges faster than information about the eigenvalues in the middle of the spectrum [21]. Therefore, it is possible to estimate the limits of the eigenvalue spectrum from a finite number of values of α and β . Long-range elastic interactions are determined by the properties of the lower limit of the spectrum. The short-range displacement field around a vacancy is determined primarily by the first few levels of the continued fraction. By attaching a terminator to the continued fraction not only are we able to describe long-range elastic interactions but we also embed a finite cluster centred on the vacancy in an infinite medium characterised by average elastic properties. This embedding also influences the short-range displacement field around the vacancy.

The values of the coefficients α and β to level m' are stored, and they are over-written by new values as they are generated until the criterion for convergence of the relaxation energy is satisfied. Subsequently we wish to calculate the change in the relaxation energy when a vacancy is exchanged with a neighbour. Convergence of the change in the elastic relaxation energy is attained with a smaller number, m , of exact levels. To calculate the change in the elastic relaxation we use the following continued fraction, where we note that the first m pairs of coefficients α and β are those computed for the displaced vacancy, the remaining $m' - m$ pairs of coefficients α and β are

those stored and not over-written, and t is the terminator of (27.9), averaged over all m' coefficients:

$$\Delta\tilde{E}_m = \frac{-1/2\beta_0^2}{\alpha_0 - \frac{\beta_1^2}{\vdots}} \quad (27.11)$$

$$\alpha_{m-1} - \frac{\beta_m^2}{\alpha_m - \frac{\beta_{m+1}^2}{\vdots}}$$

$$\alpha_{m'-1} - \frac{\beta_{m'}^2}{t}$$

The improved estimate for the energy change calculated using (27.11) is substituted in (27.7) to get an improved estimate for the displacement vector. Note that since the values of α and β are exact to level m , the projection of this new estimate on the true solution to (27.3) is greater than that available previously, even though it is within the same Krylov subspace $\mathcal{K}_m(\mathbf{D}, \mathbf{f})$.

This algorithm can be efficiently coded in parallel. A domain decomposition is used to construct the elements of \mathbf{D} . The matrix elements do not then need to be broadcast to perform the matrix-vector multiply in (27.5).

27.2.4 Short-Range Relaxation Region

The displacement field around a point defect scales in magnitude as $1/r^2$, where r is the distance from the defect. Starting from a fully relaxed system of atoms, a single atom-vacancy exchange is performed to construct a new trial configuration. The additional displacement field introduced by this exchange will be dipolar, and so scale as $1/r^3$. This displacement is short-ranged and so the greatest contribution to the elastic energy change due to an exchange will come from the minor corrections of the positions of the atoms close to the exchange event.

As we are undertaking a Monte Carlo simulation, it is necessary that the energies be accurate only to within a fraction of the thermal energy. At 300 K, this energy is 0.025 eV. We need an approximation method which can reliably make estimates of the energy difference between states to within about 1 meV.

A good approximation to the elastic energy change is found by pinning all atoms more than a fixed radius R distant from the original vacancy location, and only relaxing those within. If this is done, after relaxation only those atoms immediately exterior to the sphere will still have a finite net force acting on them. The error in the elastic energy will be due to the failure of the embedding external medium to respond to the change in shape of the interior and vice-versa. This energy error will be positive, and scale as $1/R^3$.

However, if a second virtual state is considered, it too will have an error of roughly the same size. All virtual states will therefore have a systematic component to the error due to the non-propagation of the force dipole to infinite range. An additional error component will be due to the coupling of the inhomogeneities in the far field and the force dipole. As larger near field regions are assumed, this error component will tend to zero. As this error must be kept small compared to the thermal energy, we have determined our cut-off range R to be three times the range of the potential, covering a sphere of 754 lattice sites plus the original vacancy position.

The systematic error component may be removed without affecting the statistical likelihood of selecting a given state. This error will appear in the free energy as a constant additional term, and so in the rate as a constant multiplier. If after the move the system is allowed to relax fully, then a comparison may be made between exact and approximate energies. This difference can be ascribed to the systematic error, and so the rates may be corrected. If the system is relaxed after every move, the energy of the previously visited configuration will always be stored exactly, so that the rates may be corrected even before the exchange is made.

The difference in energy between elastically relaxed states made using the short-range approximation is shown in Fig. 27.1. It can be seen that if a sufficiently large relaxation region is chosen the magnitude of the energy error can indeed be made sufficiently small. Note that our relaxation region is larger than that used previously [22–24], and that we can achieve very small errors in the energy difference by ensuring that the whole system is relaxed before each move is performed. We have chosen to use a relaxation region of 754 atoms about the original vacancy location, which produces errors in the energy of the order of 1 meV.

27.3 Clustering in the Al-Cu and Al-Cu-Mg Systems

The alloy compositions we have studied are listed in Table 27.1. We have chosen a supercell of 10976 atoms ($14 \times 14 \times 14$ fcc unit cells) run for at least one hundred thousand vacancy exchanges. The traces for the flexible lattice calculations each took about one month of computation on four Sun Ultra-SPARC III 900-MHz Processors. For this reason these simulations were done once only. We have modelled natural ageing, that is to say ageing at room temperature (300 K) as opposed to artificial ageing at elevated temperatures. This is to provide a large driving force for clustering. We have also chosen to use a single vacancy in these diffusion simulations.

27.3.1 Effect of Elastic Strain on Vacancy Diffusion

Only one type of transition, a nearest neighbour vacancy-atom exchange, is permitted in this simulation. The time taken for the same number of vacancy

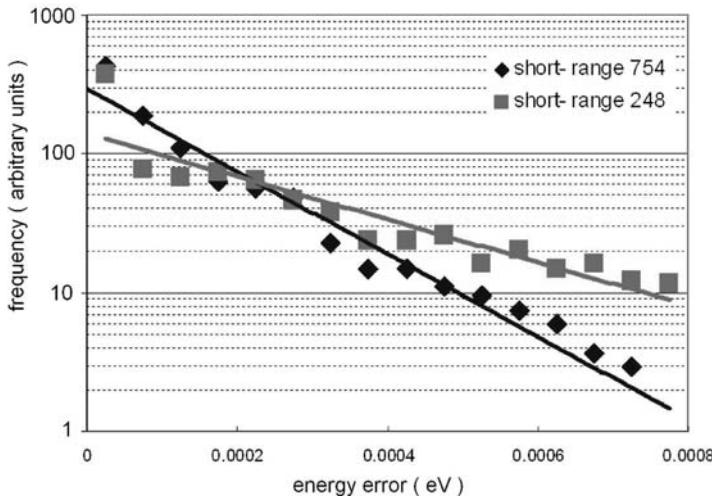


Fig. 27.1. Errors introduced using the short-range relaxation region, with cut-off regions encapsulating 248 and 754 atom sites centred on the original vacancy position. The vertical axis shows the frequency of errors. It is seen that the probability of the difference in energy between relaxed states being in error by more than 0.001 eV becomes very small with the larger cut-off region

Table 27.1. Aluminium alloy compositions

alloy	aluminium		copper		magnesium				
	at.%	(wt.%)	n_{Al}	at.%	(wt.%)	n_{Cu}	at.%	(wt.%)	n_{Mg}
1	98	(95.4)	10755	2	(4.6)	220			
2	97.8	(95.2)	10733	2	(4.6)	220	0.2	(0.2)	22
3	97.5	(95.0)	10700	2	(4.6)	220	0.5	(0.4)	55
4	96	(91.1)	10536	4	(8.9)	439			
5	95.8	(90.9)	10514	4	(8.9)	439	0.2	(0.2)	22
6	95.5	(90.6)	10481	4	(8.9)	439	0.5	(0.4)	55
7	94.0	(86.9)	10316	6	(13.1)	659			

moves with and without atomic relaxation is shown in Fig. 27.2. In the case of the Al-Cu alloys, we find that running the simulation on a flexible lattice increases the real time taken for a given number of exchanges. On a flexible lattice, the real time simulated is roughly 10–20% greater over 100,000 exchanges, reflecting a commensurate decrease in the mean vacancy hopping rate. In more familiar terms the *mobility* of the vacancy is reduced on a flexible lattice. The origin of this reduction must be that the relaxations in a flexible lattice produce a greater variability in the configurational energy of the system as the vacancy migrates. In a flexible lattice the vacancy is attracted to certain sites, where it spends longer periods of time.

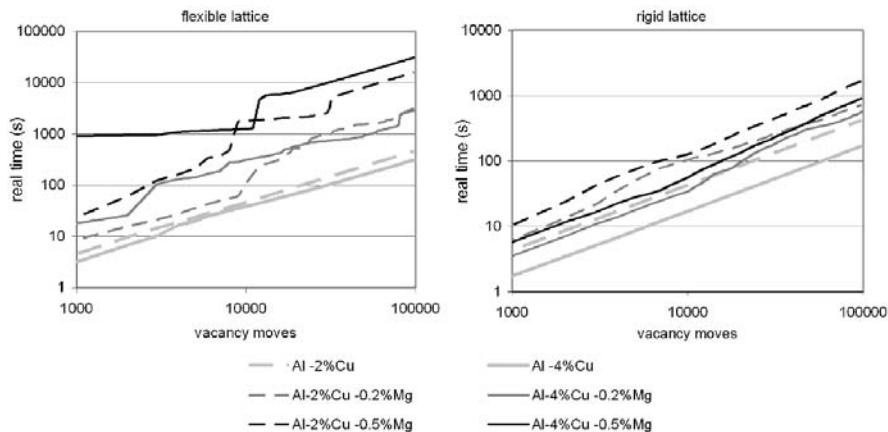


Fig. 27.2. Effect of performing the simulation on a flexible lattice on the real time simulated. Times shown are the simulated real time (in s) for a sequence of 100,000 vacancy moves

In the case of Al-Cu-Mg alloys, the time differences in the flexible and rigid lattices are even more pronounced, with the time simulated being 5 times greater with 0.2% Mg (alloys 2,5), and around 10 times greater with 0.5% Mg (alloys 3,6). Note that steep steps in the plot are seen in the flexible lattice calculation when Mg is present, indicating that the vacancy is trapped temporarily. In the rigid lattice calculation these steps are absent in these early stages, although steps do appear later when there is significant clustering [14].

27.3.2 Effect of Elastic Strain on Clustering

Al-Cu

We now consider the short-range ordering of copper atoms. We know from separate Monte Carlo simulations using the same potentials that $L1_0$ and $L1_2$ type intermetallic structures, rather than the θ phase, are preferentially generated. There are n_{Cu} copper atoms in the system. At any time, there will be a certain number of pairs of copper atoms into pairs separated by a $\langle 100 \rangle$ vector, $n_{\langle 100 \rangle}$. Our short-range order parameter, sro_{Cu} , measures how close we are getting to this low energy intermetallic phase Al_3Cu , with structure $L1_2$: $sro_{\text{Cu}} = n_{\langle 100 \rangle} / ((1/2) \times 6 \times n_{\text{Cu}})$.

We can investigate how the clustering is affected by the lattice being flexible by plotting this short-range order parameter as a function of the real time simulated. In Fig. 27.3 we plot the evolution of the short-range ordering of copper atoms in Al-Cu on a flexible and on a rigid lattice. The traces corresponding to the rigid lattice calculations are much less computationally

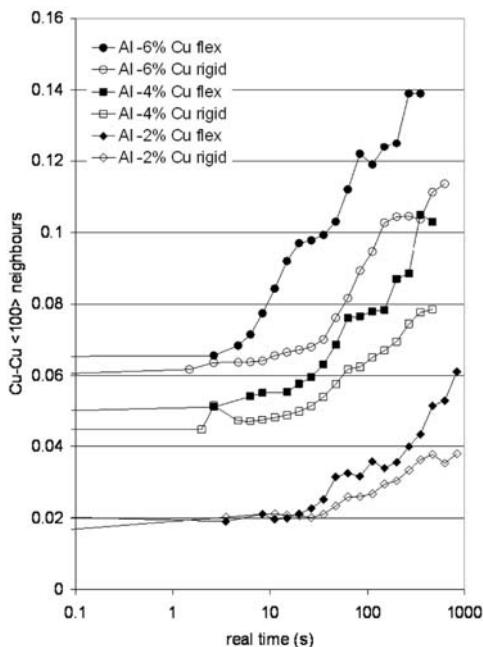


Fig. 27.3. Ordering in the Al-Cu system. These simulations were performed with a single vacancy in a periodic supercell of 10976 lattice points, with zero external stress, at a simulated temperature of 300. The simulations comprised approximately 200,000 vacancy moves for alloys 1 and 7 and 250,000 moves for alloy 4. The order parameter is defined in the text. The lines are to guide the eye

expensive than those on a flexible lattice, and so have been generated by averaging over five independent simulation runs. The order parameter is found as the average over fixed time intervals, evenly spaced on the logarithmic real time scale.

It is clear that at all three alloy compositions simulated there is a marked *increase* in the rate of ordering of copper atoms on a flexible lattice, despite the reduction of the vacancy mobility seen in Fig. 27.2. The speed of diffusing copper atoms is determined by the product of the vacancy mobility and the force driving their motion. Therefore there must be an additional driving force for clustering on a flexible lattice as compared with a rigid lattice. Once clustering has initiated somewhere in the system it will bias the motion of other copper atoms towards the cluster along lobes of compressive stress surrounding the cluster. The additional driving force must more than compensate the reduction of the vacancy mobility to produce the increased rate of ordering of copper atoms on a flexible lattice.

Al-Cu + 0.2 at.% Mg

Copper atoms are expected to be centres of tension, and magnesium atoms centres of compression in an aluminium matrix. Therefore, there will be an attractive elastic interaction between Cu and Mg atoms on a flexible Al lattice. We might therefore expect to see Cu atoms segregating to Mg rich regions. However, the rate at which such a process occurs may be reduced by vacancies becoming trapped by Mg atoms. This trapping may be expected to occur more effectively on a flexible lattice where the motion of vacancies is biased by the elastic fields of Mg atoms, and more stable vacancy-Mg atom complexes may form through elastic relaxation.

In Fig. 27.4 we examine the effect of adding microalloying quantities of Mg atoms to Al-Cu alloys on the clustering of Cu and Mg atoms. The degree of clustering of magnesium atoms may be followed by defining an order parameter for the evolution of close packed clusters of magnesium atoms: $s_{\text{rOMg}} = n_{\langle(1/2)(1/2)0\rangle} / ((1/2) \times 12 \times n_{\text{Mg}})$ where $n_{\langle(1/2)(1/2)0\rangle}$ is the number of $(\text{Mg} - \text{Mg})_{\langle(1/2)(1/2)0\rangle}$ pairs, and n_{Mg} is the number of magnesium atoms present in the system. It can be seen that the rate of ordering of copper atoms is still enhanced on a flexible lattice in the presence of 0.2 at.% Mg. But, the rate of clustering of magnesium atoms is reduced on a flexible lattice until later times.

Al-Cu + 0.5 at.% Mg

In Fig. 27.5 the effect of adding a higher concentration of Mg is simulated. With the higher Mg content it is no longer the case that the rate of ordering of copper atoms on a flexible lattice is enhanced.

It is seen that again the Mg atoms are clustering at a later time when the simulation is run with a flexible lattice. However, when it occurs the degree of clustering of Mg is enhanced at this higher Mg concentration.

27.3.3 Association of Vacancies with Mg Atoms

The number of magnesium atoms surrounding each vacancy is plotted as a function of time for rigid and flexible lattices in Fig. 27.6. On a rigid lattice we see that the number of magnesium atoms around a vacancy increases with time, but that this is not the case on a flexible lattice, at least for real times of up to 1000 s. On the rigid lattice, magnesium clusters are forming, and the vacancies are associated with these clusters.

27.4 Conclusions

We have presented a method to rapidly evaluate transition rates for possible vacancy moves incorporating atomistic elastic relaxation, by approximating

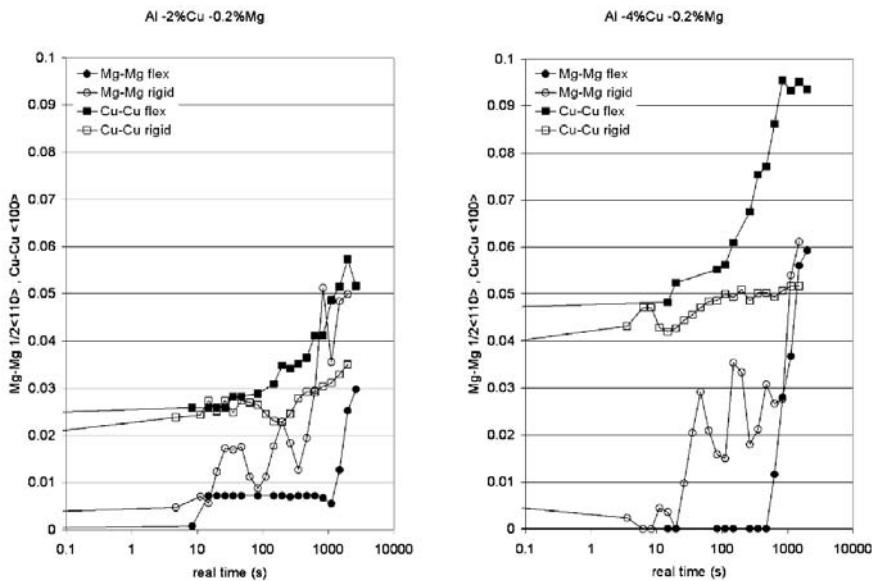


Fig. 27.4. Ordering in the Al-Cu-0.2% Mg system. These simulations used the same parameters as those in Fig. 27.3, and the traces were generated in the same way. The degree of clustering of Mg atoms is measured by the short range order parameter defined in the text. We see that there is some clustering of Mg atoms, and ordering of Cu atoms, present on both rigid and flexible lattices. The rate of ordering of copper atoms is significantly increased on the flexible lattices. The clustering of magnesium atoms seems to occur at later times on a flexible lattice as compared with a rigid lattice

the saddle points as a function of the internal energy of the initial and final states. We found that the harmonic lattice approximation was insufficient for measuring the tiny difference in the elastic relaxation energy between possible vacancy moves. However a scheme based on relaxing only the local region around the vacancy was capable of reducing errors in the system energy to an acceptable fraction of the thermal energy. This short-range scheme was found to be successful for evaluating the transition energies for possible vacancy moves, although a full atomistic relaxation was required after the move was decided.

The Lanczos method is particularly suitable for atomic relaxation after a single vacancy move, as each iteration effectively spreads the relaxation region further from the atoms experiencing a force around the vacancy. Although the relaxations remain time-consuming, our use of this relaxation scheme has enabled us to model hundreds of thousands of vacancy moves while keeping the system atomistically relaxed.

The rate at which phase separation can occur in alloys is proportional to the mobility of vacancies and to the driving force for diffusion. Both of these

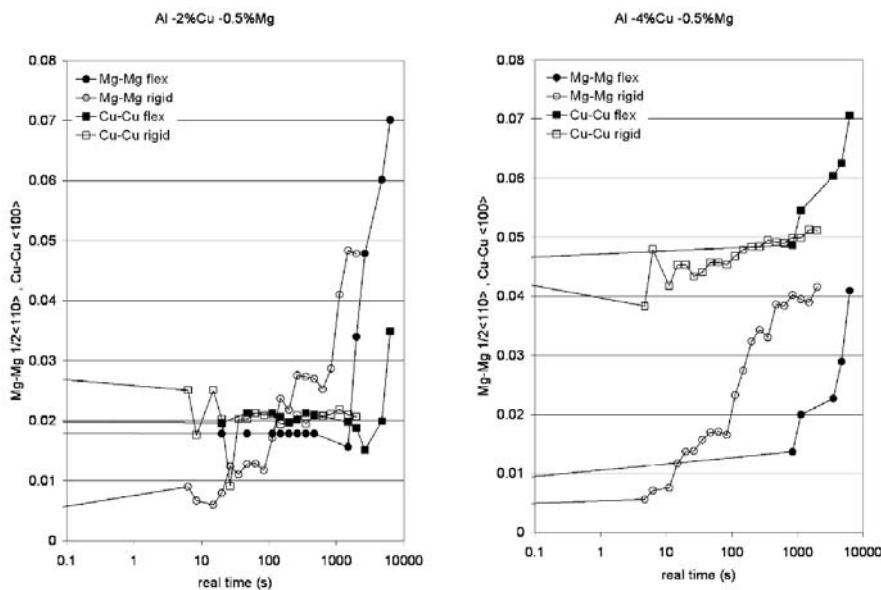


Fig. 27.5. Ordering in the Al-Cu-0.5%Mg system. These simulations used the same parameters as those in Fig. 27.4, and the traces were generated in the same way. Again the magnesium atoms are seen to cluster at later times on a flexible lattice than were seen on the rigid lattice. However, here the clustering of Mg atoms is more extensive, and the clustering of copper atoms reduced

may be affected by performing the simulation on a rigid lattice rather than on a flexible lattice. We have seen in Fig. 27.2 that the vacancy mobility is quite different on flexible and rigid lattices. Vacancies may be attracted to impurities through long-range elastic fields, and may form complexes that are more stable on a flexible lattice than on a rigid lattice. As a result the vacancy mobility may be reduced significantly on a flexible lattice. The exclusion of elastic relaxation in a rigid lattice has the effect of smoothing the configurational potential energy surface, especially in those regions of configuration space where vacancies would have become trapped if elastic interactions had been included.

However, we also saw in Fig. 27.3 that the rate of ordering of copper atoms is lower in a rigid lattice than in a flexible lattice. The apparent paradox is resolved by noting that there is a reduced driving force for ordering of copper atoms on a rigid lattice, where bonds are constrained to remain longer than they would be in a relaxed lattice, and this more than outweighs the greater vacancy mobility in a rigid lattice.

A small concentration (0.2–0.5 at.%) of magnesium atoms added to the system are rapidly precipitated out of solution in a rigid lattice, owing to the very high energy compressed bonds around each Mg atom. In a flexible lattice

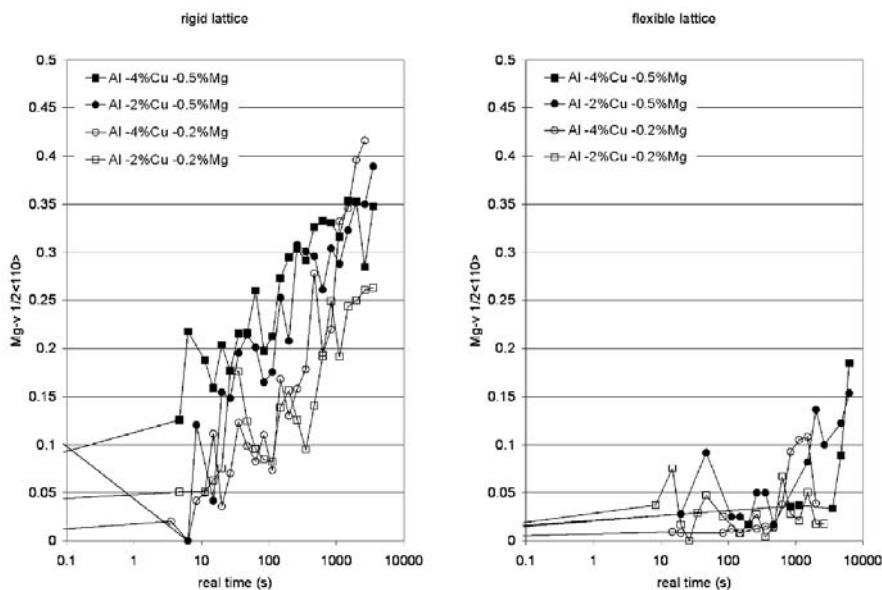


Fig. 27.6. The proportion of atoms surrounding the vacancy which are magnesium. On a rigid lattice (*left*) the number of magnesium atoms surrounding a vacancy increases with time. This, together with the clustering of magnesium atoms seen in Figs. 27.4 and 27.5 shows that on a rigid lattice magnesium clusters are forming, and the vacancies are associated with them. On a flexible lattice (*right*) the number of magnesium atoms around a vacancy is not increasing

the solubility of Mg is enhanced somewhat, and instead their principal action is to reduce severely the vacancy mobility. Mg atoms are especially trapping in a flexible lattice as the compression in the lattice around each atom is alleviated by the formation of a Mg–vacancy pair.

Acknowledgements

D.R.M was supported by an EPSRC grant. The computer simulations were performed on Oswell, the Oxford University Supercomputer. This work was performed in part under the auspices of the U.S. Department of Energy by the University of California, Lawrence Livermore National Laboratory, under Contract No. W-7405-Eng-48.

References

1. J.T. Veitz, I.J. Polmear: *J. Inst. Met.* **94**, 410 (1966)
2. B.C. Muddle, I.J. Polmear: *Acta Metall.* **37**, 777 (1989)
3. S.P. Ringer, T. Sakurai, I.J. Polmear: *Acta. Metall.* **45**, 3731 (1997)

4. H.K. Hardy: *J. Inst. Metall.* **84**, 17 (1954)
5. A.K. Mukhopadhyay: *Met. Mat. Trans.* **33A**, 3635 (2002)
6. K. Hono: *Progress Mat. Sci.* **47**, 621 (2002)
7. V. Gerold: *Scripta Met.* **22**, 927 (1989)
8. J.M. Silcock, T.J. Heal, H.K. Hardy: *J. Inst. Met.* **82**, 239 (1954)
9. S.P. Ringer, K. Hono, T. Sakurai: *Met. Mat. Trans. A* **26**, 2207 (1995)
10. J.M. Silcock, T.J. Heal, H.K. Hardy: *J. Inst. Met.* **84**, 23 (1955)
11. J.M. Silcock, H.M. Flower: *Scripta Mater.* **46**, 389 (2002)
12. T.T. Rautiainen, A.P. Sutton: *Phys. Rev. B* **59**, 13681 (1999)
13. R. Weinkamer, P. Fratzl: *Europhys. Lett.* **61**, 261 (2003)
14. D.R. Mason, R.E. Rudd, A.P. Sutton: Stochastic Kinetic Monte Carlo Algorithms for long-range Hamiltonians. *Comp. Phys. Comm.* **160**, 140 (2004)
15. K.A. Fichthorn, W.H. Weinberg: *J. Comp. Phys.* **95**, 1090 (1991)
16. A.B. Bortz, M.H. Kalos, J.L. Lebowitz: *J. Comp. Phys.* **17**, 10 (1975)
17. M.A. Novotny: *Phys. Ref. Lett.* **74**, 1 (1995)
18. C.P. Flynn: Point Defects and Diffusion (Clarendon Press 1972)
19. M.W. Finnis, J.E. Sinclair: *Phil. Mag. A* **50**, 45 (1984)
20. Y. Saad: Iterative Methods for Sparse Linear Systems (PWS 1996) Available online at <http://www-users.cs.umn.edu/saad/books.html>
21. V. Heine, D.W. Bullett, R. Haydock, M.J. Kelly: Solid State Physics (Academic Press 1980)
22. H. Ikeda, H. Matsuda: *Mater. Trans. JIM* **33**, 466 (1992)
23. H. Ikeda, H. Matsuda: *Mater. Trans. JIM* **32**, 651 (1993)
24. H. Ikeda, H. Matsuda: *Mater. Trans. JIM* **37**, 1413 (1996)

28 Br Electrodeposition on Au(100): From DFT to Experiment

S.J. Mitchell and M.T.M. Koper

Schuit Institute for Catalysis and Department of Chemical Engineering,
Eindhoven University of Technology, 5600 MB Eindhoven, The Netherlands

Abstract. In recent years, the accuracy and technical sophistication of electrochemical surface models has dramatically increased. We report on the most recent advances in which DFT calculations are used to obtain the necessary parameters for the construction of an off-lattice Monte Carlo model for Br electrodeposition on Au(100). Previous Monte Carlo methods have been applied to the study of Br/Ag(100), which displays only commensurate ordered phases, thus suggesting that a lattice-gas treatment of the adlayer is sufficient. However, since Br/Au(100) displays an incommensurate ordered phase, the use of a lattice-gas model is not justified. Since an off-lattice model requires many additional parameters to that of a lattice-gas model, DFT calculations have been used. Much effort has been to extract the best possible parameters from the DFT calculations, and the model is justified by the remarkable similarities to experiment.

Previously, electrochemical applications of Monte Carlo simulations have focused on systems for which a lattice-gas model is reasonable [1]. Among these studies, halide adsorption stands out as the most easily understood. For a review of experiments, see [2]. Bromine adsorption on Ag(100) stands out as one of the most extensively studied examples, both experimentally and theoretically [3–14]. However, no development of computational techniques would be complete without an investigation of a system for which a lattice-gas model is not sufficient. We therefore extend the study of halide adsorption to include Br/Au(100), which is known to display an incommensurate phase transition [15–17]. A more complete theoretical study of Br/Au(100) can be found in [18].

The details of the grand-canonical off-lattice Monte Carlo simulation method are far too complex to present here. We refer the reader to [11] which has all of the necessary details, corrections, and relevant historical references for the off-lattice model presented here. It is sufficient to understand that the surface model includes a two dimensional square surface region with periodic boundary conditions. Adsorbates enter and leave this surface via grand-canonical Metropolis moves, and adsorbates move laterally with the 2D surface via continuous displacement Metropolis moves. The interactions include adsorbates interacting with each other, adsorbates interacting with the surface corrugation potential, and an effective interaction with the chemical or electrochemical potential, $\bar{\mu}$.

All DFT calculations were performed using the Vienna Ab Initio Simulation Package (VASP) [19–21] using a plane-wave basis set, a periodic slab geometry, the generalized gradient-corrected exchange-correlation functional (GGA, PW91) [22], Vanderbilt ultra-soft pseudopotentials [23], and a cut-off energy of 400 eV. All slabs contained three metal layers separated by six vacuum layers, each vacuum layer being the equivalent thickness of a metal layer. To examine possible coverage dependence, supercells of two different sizes were examined, a $2(2.97) \times 2(2.97) \times 12.6 \text{ \AA}$ cell (denoted as 2×2), containing 4 surface Au atoms, and a $3(2.97) \times 3(2.97) \times 12.6 \text{ \AA}$ cell (denoted as 3×3), containing 9 surface Au atoms. The nearest-neighbor separation for the Au(100) surface atoms was $a = 2.97 \text{ \AA}$, as determined by our own calculations. The k -point mesh was generated using the Monkhorst method with a $7 \times 7 \times 1$ grid for the 2×2 cells and a $5 \times 5 \times 1$ grid for the 3×3 cells.

The off-lattice model requires us to determine essentially three different behaviors, the corrugation potential or surface binding energy as a function of the lateral position, $U(x, y)$, the long-range dipole-dipole interactions, and the short-range interactions. Each of these was determined by repeated analysis of a series of DFT configurations, and although configurational relaxations were performed for both the substrate and the adlayer, in general, these relaxations had little effect, unless surface buckling was observed. When buckling of the Au(100) was seen after relaxation, anomalously large surface dipole moments were observed, which would seem to suggest a very different surface model, which would not be expected to reproduce the experimentally observed coverage isotherm. For this reason, no relaxations of the substrate were performed for the results reported here.

Figure 28.1 shows an analytic sinusoidal approximation for the corrugation potential,

$$U(x, y) = \Delta_1 \frac{[\cos\left(\frac{2\pi x}{a}\right) \cos\left(\frac{2\pi y}{a}\right) + 1]}{2} + \Delta_2 \frac{[\cos\left(\frac{2\pi x}{a}\right) + \cos\left(\frac{2\pi y}{a}\right)]}{2}, \quad (28.1)$$

where $\Delta_1 = 274 \text{ meV}$ and $\Delta_2 = 234 \text{ meV}$. This approximation was created by examining the DFT total energy results as the binding site of a single Br adsorbate was varied across the 3×3 surface in steps of $a/4$ in both the x and y directions. For each lateral position, the surface normal, z , position of the adsorbate was relaxed. Equation (28.1) reproduces the DFT results in the vicinity of the preferred binding region, the bridge and hollow site areas, but the analytic approximation is less accurate near the top site. However, since the top site is energetically unfavorable, this difference is not expected to significantly effect the Monte Carlo results. Note that (28.1) reproduces the experimentally observed binding site and has a symmetry consistent with that expected for the experimentally observed commensurate to incommensurate phase transition.

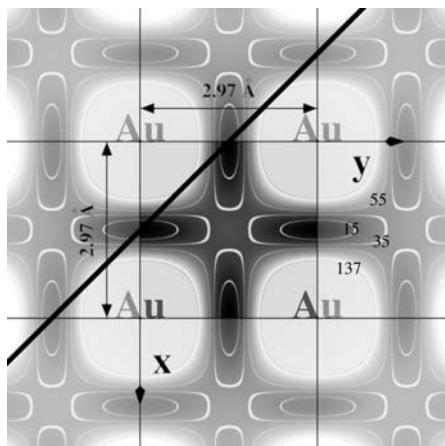


Fig. 28.1. Analytic sinusoidal corrugation potential, $U(x, y)$. The grid lines indicate the boundaries between surface unit cells, and the underlying substrate atoms are labeled. The grayscale image in the central square shows $U(x, y)$, where black indicates the lowest value, 0 meV, and white indicates $U(x, y) = 137$ meV. The gray shading is truncated above $U(x, y) = 137$ meV, and the contour lines as labeled in the plot with units of meV. The diagonal line indicates the compression direction for the experimentally observed incommensurate phase

Next, we must determine the lateral interactions between adsorbed Br. To do this, we consider six configurations, each with four-fold symmetry, where the Br are constrained to bridge site adsorption. Four of the six configurations are shown in Fig. 28.2, each of which corresponds to a physically realistic coverage. However, to determine the short-range excluded volume interaction, two additional configurations were examined, with coverage $\Theta = 1$ and $\Theta = 2$, corresponding to every other bridge site occupied and every bridge site occupied, respectively. When the DFT total energy is corrected for binding energy, different numbers of adsorbates in the cell, and the residual energies due to the substrate and adsorbate itself, one can determine the interaction energy per adsorbate, Φ/N , where N is the number of adsorbates in the DFT cell (see Fig. 28.3). From this short-range behavior, we can extract the short-range interactions necessary for the Monte Carlo model. Note that the DFT total energy calculations do not seem to be sensitive to the long-range dipole-dipole interactions, although an analysis of the electron density clearly shows a surface dipole moment associated with the Br/Au(100) bond.

Finally, we must calculate the dipole moment of the Br/Au(100) system, which can easily be calculated by integrating total charge density, including both electrons and nuclear ionic cores. Such an analysis indicates a surface dipole moment, $s = 0.3 \text{ e}\AA$, where the full dipole moment is assumed to be $s = s\hat{z}$. For all configurations shown in Fig. 28.2, this is true.

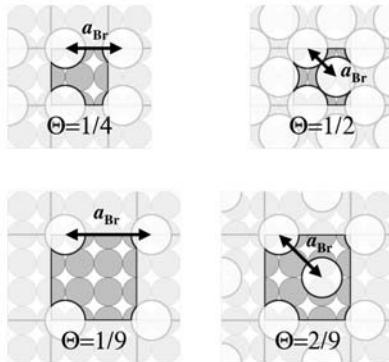


Fig. 28.2. Four of the six adlayer configurations used in the DFT calculations. The small gray circles indicate surface Au(100) atoms while the larger white circles indicate adsorbed Br. The DFT cell configurations have been repeated in each lateral direction. The arrows indicate the Br-Br nearest neighbor distance, a_{Br} , for each cell

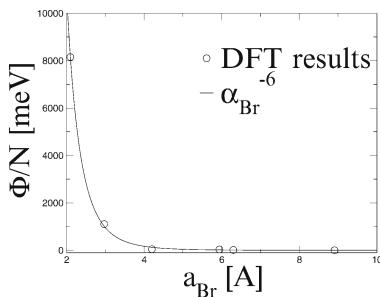


Fig. 28.3. The lateral interaction energy derived from DFT calculations for the four configurations shown in Fig. 28.2 plus two additional configurations with $\Theta = 1$ and $\Theta = 2$. The approximate analytic form, found by a non-linear fit to $1/r^6$, suggests the functional form of the short-range interactions

To see the surface polarization more clearly, we define the charge transfer function as

$$\begin{aligned} \Delta\rho(\mathbf{x}) &= \frac{\left\{ \rho(\mathbf{x})_{N\text{Br-Au}(100)} - \sum_i \rho_i(\mathbf{x})_{\text{Br}} - \rho(\mathbf{x})_{\text{Au}(100)} \right\}}{N} \\ &= -\frac{\left\{ \rho_e(\mathbf{x})_{N\text{Br-Au}(100)} - \sum_i \rho_{ei}(\mathbf{x})_{\text{Br}} - \rho_e(\mathbf{x})_{\text{Au}(100)} \right\}}{N}, \quad (28.2) \end{aligned}$$

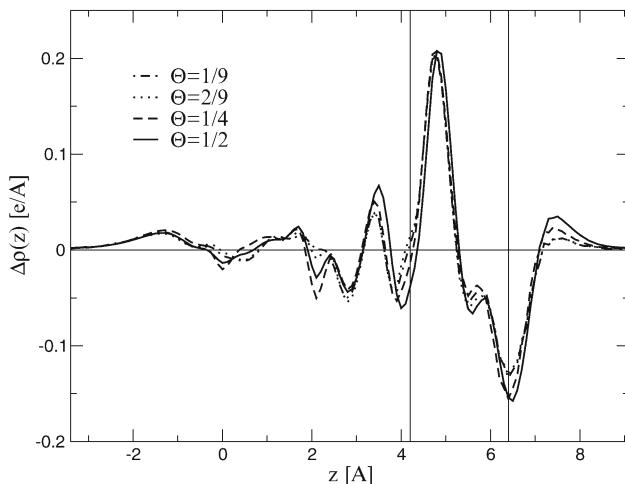


Fig. 28.4. The charge transfer function, $\Delta\rho(z)$, for the four adlayer configurations shown in Fig. 28.2. The surface polarization is clearly seen. The straight horizontal line indicates $\Delta\rho(z) = 0$, and the two vertical lines indicate the surface Au(100) location and the Br adlayer location, $z_{\text{top}} = 4.2 \text{ \AA}$ and $z_{\text{Br}} = 6.4 \text{ \AA}$ respectively

where $\rho(\mathbf{x})_{N\text{Br}-\text{Au}(100)}$ is the full charge density of the adlayer system with N adsorbed Br in the cell, $\rho_i(\mathbf{x})_{\text{Br}}$ is the full charge density of a single Br atom at the same position as that in the Br-Au bonded system, i indexes the N adsorbed Br, $\rho(\mathbf{x})_{\text{Au}(100)}$ is the charge density of the Au(100) slab with all atoms at the same positions as in the Br-Au bonded system, and the subscripts “e” denote electron only densities, having the sign convention that positive indicates greater electron density. The charge transfer function is calculated using only charge densities for the same size cells. The charge transfer function is shown in Fig. 28.4 for each of the configurations of Fig. 28.2.

Having derived all of the necessary parameters for the off-lattice model from DFT calculations, a grand-canonical off-lattice Monte Carlo simulation can now be performed using the corrugation potential of (28.1), the short-range $1/r^6$ repulsions shown in Fig. 28.3, and the long-range $1/r^3$ dipole-dipole repulsions due to the surface polarization shown in Fig. 28.4. The results of the Monte Carlo simulation for average coverage vs electrochemical potential are shown in Fig. 28.5. The simulations are quite slow, requiring approximately 6 weeks of computing time on a single GHz processor PC. The simulation results agree quite well with the observed experimental results, except the experiments see only one incommensurate phase, whereas

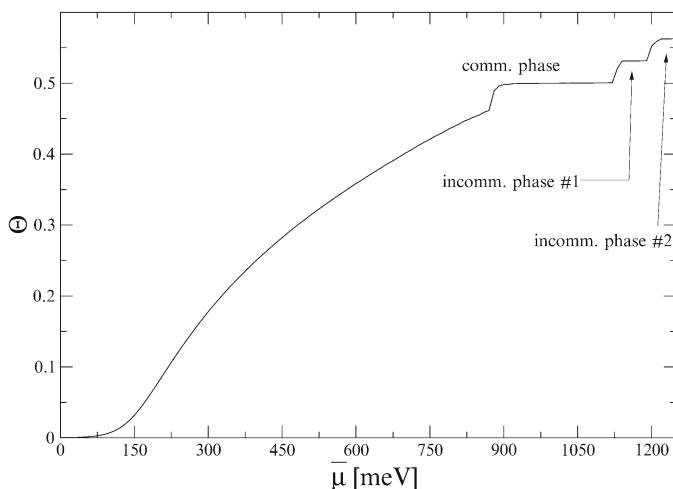


Fig. 28.5. The results of the grand-canonical off-lattice Monte Carlo simulation for a surface of 32×32 Au surface atoms. Three different ordered phases of the Br adlayer are seen in the simulations as coverage plateaus. From left to right, a disordered commensurate phase, an ordered commensurate phase, and two apparently different incommensurate ordered phases

the simulations seem to suggest two incommensurate phases. The observation of two incommensurate phases in the simulation results is most likely an artifact of the small system size and the strain placed on the adlayer from trying to fit an essentially hexagonal adlayer into a square simulation volume.

Future work should consist of larger simulations to try and eliminate the “extra” observed incommensurate phase. However, since much of the simulation time involves a very slow relaxation to equilibrium, some clever procedures will be necessary to overcome the slow equilibration times. Additionally, it is known that the electrochemical experiments proceed in the presence of water. However, neither the DFT calculations nor the Monte Carlo simulations include water. Future work should concentrate on including water into the simulations.

Acknowledgments

This work was funded by the Netherlands Organization for Scientific Research (NWO). The authors would like to thank T. Wandlowski for useful discussions, encouragement, and the use of his experimental data.

References

1. G. Brown, P.A. Rikvold, S.J. Mitchell, M.A. Novotny: Monte Carlo Methods for Equilibrium and Nonequilibrium Problems in Interfacial Electrochemistry. In: *Interfacial Electrochemistry* (Marcel Dekker, New York 1999) p. 47
2. O.M. Magnussen: *Chem. Rev.* **102**, 679 (2002)
3. O. Endo, M. Kiguchi, T. Yokoyama, M. Ito, T. Ohta: *J. Electroanal. Chem.* **473**, 19 (1999)
4. A. Ignaczak, J.A.N.F. Gnomes: *J. Electroanal. Chem.* **420**, 71 (1997)
5. M.T.M. Koper: *J. Electroanal. Chem.* **450**, 189 (1998)
6. S.J. Mitchell, G. Brown, P.A. Rikvold: *J. Electroanal. Chem.* **493**, 68 (2000)
7. S.J. Mitchell, G. Brown, P.A. Rikvold: *Surf. Sci.* **471**, 125 (2001)
8. S.J. Mitchell, P.A. Rikvold, G. Brown, in: *Computer Simulation Studies in Condensed-Matter Physics XIII*. Ed. by D.P. Landau, S.P. Lewis, H.-B. Schüttler (Springer, Berlin Heidelberg New York 2001)
9. B.M. Ocko, J.X. Wang, T. Wandlowski: *Phys. Rev. Lett.* **79**, 1511 (1997)
10. S.J. Mitchell, S. Wang, P.A. Rikvold, G. Brown, in: *Computer Simulation Studies in Condensed Matter Physics XIV*. Ed. by D.P. Landau, S.P. Lewis, H.B. Schuttler (Springer, Berlin Heidelberg New York 2001)
11. S.J. Mitchell, S. Wang, P.A. Rikvold: *Faraday Discussions* **121**, 53 (2002)
12. B.M. Ocko, J.X. Wang, T. Wandlowski: *Phys. Rev. Lett.* **79**(8), 1511 (1997)
13. G. Valette, A. Hamelin, R. Parsons: *Z. Phys. Chem. Neue Folge* **113**, 71 (1978)
14. J.X. Wang, T. Wandlowski, in: *Proceedings of the Symposium on The Electrochemical Double Layer*, Vol. 97–17. (The Electrochemical Society, New Jersey 1997) p. 293
15. T. Wandlowski, J.X. Wang, O.M. Magnussen, B.M. Ocko: *J. Phys. Chem.* **100**, 10277 (1996)
16. A. Cuesta, D.M. Kolb: *Surf. Sci.* **465**, 310 (2000)
17. F. Reniers, D.H. Fairbrother, S. Wu, J. Lipkowski: *Surf. Sci.* **433–435**, 12 (1999)
18. S.J. Mitchell, M.T.M. Koper: Submitted to *Surf. Sci.*
19. G. Kresse, J. Hafner: *Phys. Rev. B* **47**, 558 (1993)
20. G. Kresse, J. Hafner: *Phys. Rev. B* **49**, 14251 (1994)
21. G. Kresse, J. Furthmüller: *Comput. Mater. Sci.* **6**, 15 (1996)
22. J.P. Perdew, J.A. Chevary, S.H. Vosko, K.A. Jackson, D.J. Pederson, M.R. Singh, C. Fiolhais: *Phys. Rev. B* **46**, 6671 (1992)
23. D. Vanderbilt: *Phys. Rev. B* **41**, 7892 (1990)

29 Simulation of ZnSe, ZnS Coating on CdSe Substrate: The Electronic Structure and Absorption Spectra

J.S. Kim¹, S.J. Yun¹, and G.S. Lee¹

Department of Physics, Pohang University of Science and Technology
Republic of Korea

Abstract. The electronic structure and absorption spectra are computed for ZnSe, ZnS coating on CdSe substrate-CdSe/ZnSe/ZnS. In CdSe(0001) coated with ZnSe, the energy of zincblende(ZB) ZnSe(111) epilayers on CdSe(0001) is about 80 meV lower and more stable than hexagonal wurtzite (WZ) ZnSe(0001) epilayers on CdSe(0001). Coating CdSe(0001) layers with ZnSe(111) and ZnS(111) introduced large red-shift and enhanced peak amplitude of the imaginary part of dielectric function. This peak is due to the topmost ZnSe or ZnS 3 epilayers contribution.

29.1 Introduction

Since the low dimensional structures of semiconductors have unusual properties compared to those of bulk materials, they have recently attracted considerable interest [1]. These differences are due to the quantum confinement, surface effects and so on. As the size of the crystal of the semiconductor decreases from the bulk crystal to atomic level, many kinds of fundamental properties are altered significantly. Therefore, one can improve and tailor desired properties by controlling the size of a material for one's purposes [2].

Owing to these properties of low dimensional structures of semiconductors, they have been considered as promising materials for many applications. Their nonlinear optical properties and wide band gap can be used to develop photonic applications such as ultrafast optoelectronic switches, short-wavelength light-emitting diodes and laser diodes. In addition, they also can be used to get information for understanding of quantum confinement and surface effect of low dimensional structures of the semiconductors. Semiconductor nanocrystals embedded in many different media have been investigated for this reason in these days.

Other outstanding applications are biological labeling and nanoscale topographic localization of probes in biological systems [3–8]. Luminescence, second harmonic generation and linear fluorescence optical properties connected with a near-field scanning optical microscopy have showed the interesting possibility of creating multicolor labeling and dynamic three-dimensional nanoscale optical imaging. The nanostructures, surface-passivated quantum

dots used to labeling DNA or living cells, are highly stable against photo-bleaching and have a bright, narrow and symmetrical emission [5–8]. For the perfect realization of these potential applications, a better understanding of nonlinear or linear optical properties of low dimensional structures of semiconductors are still necessary. In this paper, the electronic structure and absorption spectra (imaginary part of dielectric function) are investigated for ZnSe, ZnS coating on CdSe-CdSe/ZnSe/ZnS.

29.2 Calculation Details

DFT calculation, as implemented in the Vienna Ab-initio Simulation Package (VASP) [9, 10], is performed using the plane-wave-basis within the local density approximation (LDA) as parameterized by Ceperly and Alder. The ionic potential was treated with the projector augmented wave (PAW) method [11], and the cut-off energy of plane wave basis was set to 280.0 eV. Real space mesh was used for non-local potential. For calculation of the imaginary part of dielectric function, oscillator strength method was used as described in [12].

We consider the substrate CdSe(0001) with eight-atomic layers and 10 vacuum layers. In the bottom layers Cd and Se are fixed at the bulk position and their dangling bonds are terminated by H atoms. Six epilayers (6 ELs) of ZnSe are coated on CdSe (0001) in various ways; CdSe (0001)/ZnSe (0001)-6 ELs, CdSe (0001)/ZnSe (0001)-3 ELs+ZnSe((111)-3 ELs, CdSe (0001)/ZnSe (111)-3 ELs+ZnSe (0001)-3 ELs and CdSe (0001)/ZnSe (111)-6 ELs. In addition we considered ZnS (111)-3 ELs coating on CdSe (0001)/ZnSe(111)-6 ELs. All epilayers are fully relaxed with the exception of the bottommost five layers of CdSe(0001). ($15 \times 15 \times 1$) meshes of k points in the first Brillouin zone generated in the scheme of Monkhorst-Pack are used for the calculation of all structures and energy, and ($23 \times 23 \times 1$) meshes are used in absorption spectra calculation. The error in energy was limited to 1 meV ($\delta E \sim 1$ meV).

29.3 Results and Discussion

CdSe(0001)/ZnSe(0001) and CdSe(0001)/ZnSe(111)

Bulk crystals of CdSe, ZnSe and ZnS have two structures, hexagonal wurtzite (WZ) and zincblende (ZB) structures. Experimentally all ZnSe epilayers

Table 29.1. Total energy of CdSe/ZnSe: (a) CdSe(0001)/ZnSe(0001)-6 ELs (b) CdSe(0001)/ZnSe(0001)-3 ELs+ZnSe((111)-3 ELs (c) CdSe(0001)/ZnSe(111)-3 ELs+ZnSe(0001)-3 (d) CdSe(0001)/ZnSe(111)-6 ELs

(a)	(b)	(c)	(d)
-98.368	-98.399	-98.428	-98.452 (eV)

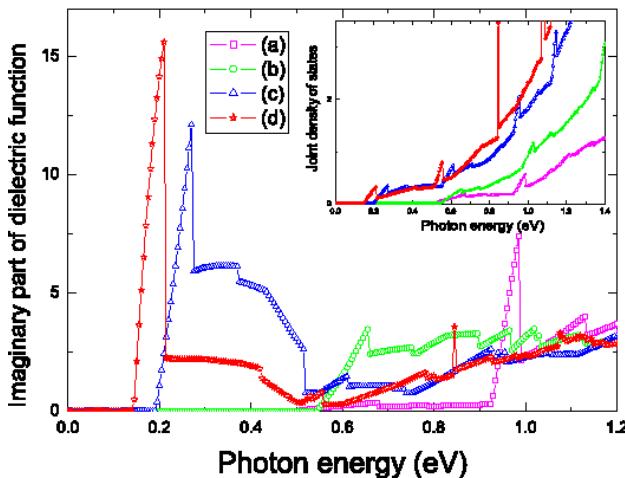


Fig. 29.1. Imaginary part of dielectric function: (a) CdSe(0001)/ZnSe(0001)-6 ELs. (b) CdSe(0001)/ZnSe(0001)-3 ELs+ZnSe(111)-3 ELs. (c) CdSe(0001)/ZnSe(111)-3 ELs+ZnSe(0001)-3. (d) CdSe(0001)/ZnSe(111)-6 ELs. The inserted plot is joint density of states

grown on the wurtzite CdSe substrate do not have the wurtzite structure but only the zincblende structure. This indicates that wurtzite ZnSe epilayers could not be grown even on the wurtzite CdSe epilayers [?, 14]. Therefore, it is theoretically important to investigate which of two structures, WZ or ZB of ZnSe on CdSe(0001), is preferred and more stable. The calculated total energies of each case are summarized in Table 29.1. It shows that CdSe(0001)/ZnSe(111)-6 ELs is 80 meV lower than CdSe(0001)/ZnSe(0001)-6 ELs. Therefore, the ZB ZnSe epilayers on the WZ CdSe epilayers are more stable than the WZ ZnSe epilayers on the WZ CdSe epilayers. This is consistent with experimental results [13, 14]. In Fig. 29.1 linear optical absorption spectra (imaginary part of dielectric function) and joint densities of states (JDOS's) are plotted for the four structures of ZnSe listed in Table 29.1. Imaginary part of dielectric function from ≈ 0.2 eV to ≈ 0.75 eV in all cases is not due to the bottommost CdSe 8 layers but the topmost ZnSe 3 epilayers. CdSe contribution appears over ≈ 0.75 eV. The amplitude of the first peak of CdSe(0001)/ZnSe(111)-6 ELs is remarkably increased than CdSe(0001)/ZnSe(0001)-6 ELs. Although CdSe(0001)/ZnSe(111)-3 ELs+ZnSe(0001)-3 ELs is more stable than CdSe(0001)/ZnSe(0001)-3 ELs+ZnSe(111)-3 ELs, the amplitude of the first peak of former is lower. Therefore, the intensity of optical absorption is modified by the type of ZnSe epilayers. JDOS's show compatible results with imaginary part of dielectric

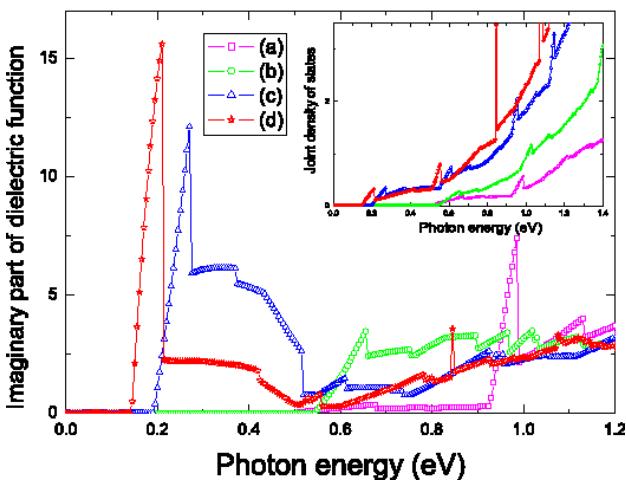


Fig. 29.2. Imaginary part of dielectric function: (a) CdSe(0001). (b) CdSe(0001)/ZnSe(111)-3 ELs. (c) CdSe(0001)/ZnSe(111)-6 ELs. (d) CdSe(0001)/ZnSe(111)-6 ELs/ZnS(111)-3 ELs. The inserted plot is joint density of states

function. Finding the variation of optical property according to the thickness of ZnSe and ZnS on CdSe, we calculated the imaginary part of dielectric function for structure of CdSe(0001), CdSe(0001)/ZnSe(111)-3 or 6 ELs and CdSe(0001)/ZnSe(111)-6 ELs/ZnS(111)-3 ELs.

Figure 29.2 shows the imaginary part of dielectric function and JDOS's for four cases mentioned above. As CdSe(0001) is coated with ZnSe(0001) and ZnS (0001) epilayers, the red-shift and amplitude of the first peak are greatly increased. In bare CdSe(0001) the photon energy range from ≈ 0.5 eV to ≈ 1.2 eV is due to the topmost CdSe(0001) 3 layers. In CdSe(0001)/ZnSe(111)-3 ELs, the photon energy range from ≈ 0.5 eV to ≈ 0.8 eV is caused by the topmost ZnSe 3 layers and the CdSe contribution appears over ≈ 0.8 eV. In CdSe(0001)/ZnSe(111)-6 ELs, the contribution of the topmost ZnSe 3 layers is dominant in the range from ≈ 0.2 eV to ≈ 0.8 eV and the contribution of CdSe make an appearance gradually over ≈ 0.8 eV. In the last case, CdSe(0001)/ZnSe(111)-6 ELs/ZnS(111)-3 ELs, the contribution of the topmost ZnS 3 layers is superior to the contribution of CdSe in the range from ≈ 0.1 eV to ≈ 0.6 eV. The first peak in all cases is mainly due to the topmost ZnSe 3 layers and ZnS 3 layers, that is, unusual peaks are caused by the surface of ZnSe and ZnS. Although our models are not quantum dots but thin films and we calculated not photoluminescence spectra but absorption spectra, the trend is similar to the experimental result for photoluminescence of CdSe/ZnSe/ZnS-core/shell quantum dots(QDs). The CdSe/ZnSe/ZnS QDs exhibit higher fluorescence efficiency than CdSe QDs and CdSe/ZnSe QDs [8].

29.4 Conclusion

We investigated the absorption spectra of multi-layer structures CdSe/ZnSe/ZnS. On the CdSe(0001) substrate, the most stable structure of ZnSe epilayers is not hexagonal wurtzite but zincblende. As CdSe(0001) layers are coated with ZnSe(111) and ZnS(111), large red-shifts and enhanced peak amplitude of the imaginary part of dielectric function are found to occur. This peak is due to the topmost ZnSe or ZnS 3 layers contribution.

Acknowledgement

We thank Dr. S.K. Shin of Postech Chemistry Department for calling our attention to nanostructures with CdSe type materials. We appreciate Dr. S.H. Lee's help for us to learn the VASP package. We benefited from discussions with Dr. Y-G Yoon of Chung-Ang University. This work was supported by Postech Research Fund for Basic Sciences.

References

1. A.P. Alivisatos: Science **271**, 933 (1996)
2. L. Brus: Appl. Phys. A **53**, 465 (1991)
3. M. Bruchez Jr., M. Moronne et al.: Science **281**, 2013 (1998)
4. D. Jakubczyk, Y. Shen et al.: Opt. Lett. **24**, 1151 (1999)
5. W.C.W. Chan, S. Nie: Science **281**, 2016 (1998)
6. J.K. Jaiswal, H. Mattossi et al.: Nature Biotechnology **21**, 47 (2003)
7. M. Han, X. Gao et al.: Nature Biotechnology **19**, 631 (2001)
8. P. Reiss, S. Carayon et al.: Synthetic Metals **139**, 649 (2003)
9. G. Kresse, J. Hafner: Phys. Rev. B **48**, 13115 (1993)
10. G. Kresse, J. Hafner: Phys. Rev. B **49**, 14251 (1994)
11. G. Kresse, D. Joubert: Phys. Rev. B **59**, 1758 (1999)
12. B. Adolph et al.: Phys. Rev. B **63**, 125108 (2001)
13. N. Matsumura, J. Ueda et al.: Jpn. J. Appl. Phys. **39**, L1026 (2000)
14. N. Matsumura, K. Yasui et al.: J. Crystal Growth **237-239**, 1536 (2002)
15. M. Danek et al.: J. Crystal Growth **145**, 714 (1994)

30 Simulation of Islands and Vacancy Structures for Si/Ge-covered Si(001) Using a Hybrid MC-MD Algorithm

L. Nurminen¹, F. Tavazza², D.P. Landau^{1,2}, A. Kuronen¹, and K. Kaski¹

¹ Helsinki University of Technology, Laboratory of Computational Engineering
P.O.Box 9203, 02015 HUT, Finland

² The University of Georgia, Center for Simulational Physics
Athens, GA 30602-2451, USA

Abstract. A classical hybrid MC-MD algorithm is applied to study physically interesting phenomena, such as island stability and formation of dimer-vacancy structures, on Si/Ge-covered Si(001). The method introduces collective moves into the standard MC algorithm to allow the system to escape from metastable states. Simulation results are found to be in good qualitative agreement with experiments.

30.1 Introduction

Understanding the properties of thin layers on semiconductor surfaces is crucial for the development of new fabrication methods for a wide variety of nano-structure and optoelectronics applications. The Ge/Si(001) system is an ideal model for studies of heteroepitaxial growth where structural modifications, such as step-edge evolution or formation of defect structures, play a central role in determining the large-scale morphology of the surface.

Structural changes often involve propagation of long-range elastic effects. Studying such phenomena requires not only sufficiently large system sizes but also efficient simulation methods which can produce large configurational changes in the system. Semiconductor surfaces are often characterized by extremely complicated energy landscapes. In simulations of such systems, conventional algorithms become inefficient in dealing with long time-scale processes associated with escape from metastable states.

In order to overcome these problems, we have developed a new hybrid Monte-Carlo Molecular-Dynamics algorithm which uses collective moves of several atoms to produce large-scale changes in the system. The version presented here is intended for the study of Si(001) systems, but the idea itself is quite general and could potentially be applied to other semiconductor surfaces. In the current version, the method is based on introducing long-range dimer displacements into the standard MC algorithm which significantly speed up the evolution of the system by allowing the system to overcome the large energy barriers associated with dimer diffusion. In this paper, we

apply this method to study two physically interesting phenomena, namely the relaxation of islands and the formation of vacancy structures.

30.2 Simulation Method

In this work, off-lattice constant-pressure Monte Carlo simulations [1] are used to model temperature-dependent structural properties of the Si- or Ge-covered Si(001) surface. The Si-Si, Si-Ge and Ge-Ge interactions are modeled using the classical Stillinger–Weber (SW) [2, 3] potential which we have recently tested for use in finite-temperature simulations involving the Si(001) surface [4]. The system is modeled using slab geometry such that the unit cell is constructed along the [110], [1 $\bar{1}$ 0] and [001] directions and periodic boundaries are applied in the x and y directions. The simulation slab consists of 20 atomic layers with the surface size varying from 20×20 to 40×40 atoms. The method is not limited to these system sizes but can be used to treat much larger systems (over 10^5 atoms).

In the standard constant-pressure Metropolis MC algorithm, two kinds of trial moves are attempted: small random displacements of individual particles (*single-atom moves*) and random variations of the sides of the simulation cell (*volume variations*) in order to keep the pressure constant at $P = 0$. In both cases, the acceptance probability is given by the Metropolis form [5]. More details can be found from [3, 4].

The Si(001) surface consists of rows of dimerized atoms (2×1 reconstruction) which induces significant displacements of the atoms in the underlying layers. The potential barrier for a jump of a single dimer from one binding site to a neighboring site is extremely high, and therefore such events never occur if only small displacements of individual atoms are attempted in the simulation. Moreover, breaking a dimer bond would require a large amount of energy, which means that the two atoms must move simultaneously without breaking the bond. For this reason, traditional algorithms become inefficient in relaxing the system.

In order to overcome these problems we have developed a *hybrid MC-MD algorithm* [6, 7] which introduces large dimer displacements to circumvent the high potential barrier between two binding sites. In this so-called *dimer-jump algorithm*, two atoms comprising a dimer undergo the same displacement (translation and a small rotation) which typically allows the particles to move a distance comparable to the Si lattice constant ($a_{0,\text{Si}} = 5.43 \text{ \AA}$).

Let us now consider the arrangement of atoms in the underlying layers. Figure 30.1 shows a snapshot of a typical step edge on Si(001). We see that the four nearest atoms under a dimer have relaxed to near-tetrahedral positions, whereas those atoms which are not directly under the step are themselves dimerized. If we now displace one of the upper-layer dimers by an amount corresponding to the average distance between binding sites, the dimer lands in a position where its nearest neighbors are either too close or too distant.

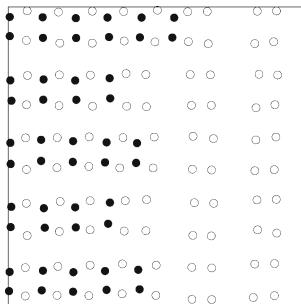


Fig. 30.1. Snapshot of a typical S_B step edge on the Si(001) surface after relaxation at 700 K. The open and solid circles are lower and upper terrace atoms, respectively

Moreover, the configuration at the initial site has become unfavorable to those atoms which were previously under the dimer. Consequently, the dimer jump will result in a very high-energy configuration which will almost certainly be rejected in the Metropolis trial.

In order to achieve a much better acceptance rate, we have introduced a scheme in which the local environments around the initial and the landing site are relaxed prior to the acceptance test. The dimer itself is also included in this group of atoms. The relaxation is performed using standard Molecular Dynamics with Velocity Verlet for integrating the equations of motion. Initial velocities are assigned from the Maxwell-Boltzmann distribution at the simulation temperature and constant-temperature conditions are maintained using velocity rescaling. In most cases, a relaxation of 7–10 MD steps using a time step of 10^{-15} s is enough for achieving a good acceptance rate.

30.3 Relaxation of Islands and Step Edges

As the first application example of the metjod described above, we investigate the relaxation of islands and step edges on Si(001). We begin the simulations with a square Si island on the Si(001) surface, which enables us to study both the S_A and S_B steps at the same time (see Fig. 30.2 for an illustration of the different step edges). The system is allowed to evolve using the dimer-jump algorithm until equilibrium is reached, and no further qualitative changes in the configuration are then observed.

Figure 30.2 shows typical snapshots of an island at two different temperatures. Looking first at the island on the left, we notice that the shape of the island has changed to a slightly anisotropic form. The S_A and S_B steps have evolved in a very different manner: the S_A steps are longer and have a smooth appearance, whereas the S_B steps are clearly more rough and irregular. This behavior is consistent with STM images of Si islands and step edges which all show very similar features as observed in our simulations (see e.g. [8, 9]).

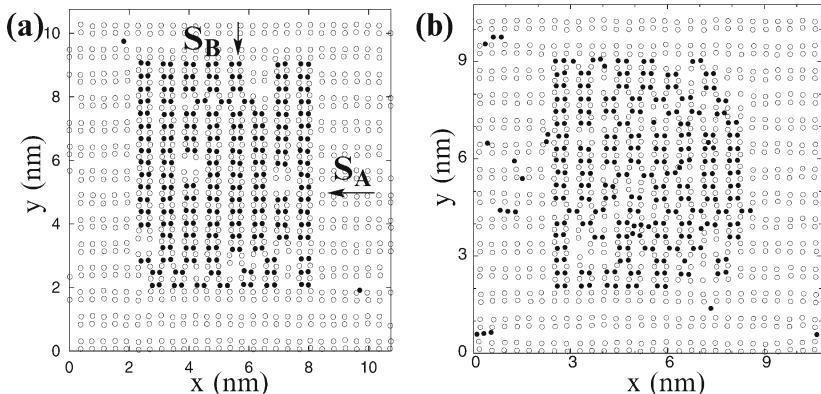


Fig. 30.2. Snapshots of typical Si islands on Si(001). (a) $T = 700\text{ K}$, after 7×10^4 MCS. (b) $T = 930\text{ K}$, after 4×10^3 MCS. The solid circles are adatoms and the open circles are atoms in the first surface layer

Comparison of the two figures at 700 K and 930 K shows that island stability is clearly dependent on temperature. At 700 K, the island changes its shape during equilibration but stays compact and well defined for very long simulation times. In contrast, when $T = 930\text{ K}$, the island dissociates very quickly. If we reduce the initial island size, we observe that the dissociation takes place at much lower temperatures (see e.g. [6, 7]). These observations indicate that there exists a temperature-dependent critical size beyond which islands are stable against dissociation and smaller islands will decay. It is not the aim of this work to address the subject in detail, but from these results it is evident that the algorithm is suitable for such an investigation.

30.4 Formation of Vacancy Structures

Experimental studies have produced consistent evidence that dimer vacancies (DV) are intrinsic defects of the Si(001) surface [10]. They are known to cluster into distinctive complexes (e.g. 2, 2 + 1, 3 DVs etc.) [11], and under certain conditions, elongated vacancy islands are also observed to form. In the case of Ge growth on Si(001), an ordered vacancy-line pattern forms spontaneously on the surface and the periodicity of the pattern depends on the thickness of the Ge layer [12]. This so-called $2 \times n$ reconstruction has a profound influence on the large-scale morphology of the surface, leading to e.g. a reversal of the step-edge roughness and surface stress anisotropy.

We began to investigate the vacancy-line formation by comparing the relative energy of different $2 \times n$ reconstructions. The details of the calculations are given in [13]. The results show that the $2 \times n$ pattern is stabilized by the Ge layer and increasing Ge coverage leads to a decrease in the optimal spac-

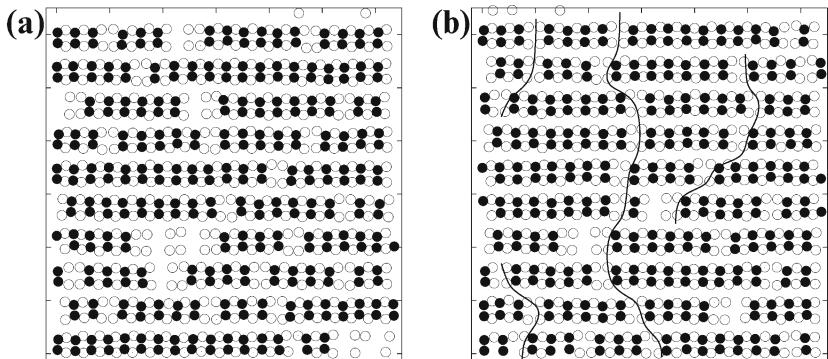


Fig. 30.3. Formation of dimer vacancy structures at $T = 930\text{ K}$. (a) Clean Si(001). (b) 2 ML of Ge on Si(001). In both cases, the initial configuration was a random distribution of vacancies

ing between the vacancy lines. Our results are in excellent agreement with experimental observations, which confirms that our classical model seems to capture the essential features of dimer vacancies in Si(001) surfaces. As the next step, it is interesting to ask whether the ordering process can be simulated using the dimer-jump algorithm.

Figure 30.3 shows two snapshots from simulations using the dimer-jump algorithm for a pure Si(001) surface and for a surface covered with 2 ML of Ge. In both cases, we used the same initial configuration which consists of a random distribution of dimer vacancies. All dimers are allowed to move on the surface during the simulation, which in effect means that the vacancies migrate. We notice that on the clean Si(001) surface, the vacancies show a tendency to cluster together but the overall arrangement is random. On the Ge-covered surface, on the other hand, we observe that clear segments of vacancy lines are forming (one line passes through the whole system and two shorter segments have also developed). Once formed, the vacancy lines are very stable and do not dissociate.

The drawback in these particular simulations is that we were forced to use a rather high temperature (930 K) in order to have sufficiently many dimer jumps accepted within a reasonable simulation time. However, the equilibrium shape of the vacancy lines is expected to be quite kinked and disordered at this temperature. Therefore, an approach such as parallel tempering (see e.g. [14]) could be useful in speeding up the equilibration at lower temperatures. Nevertheless, even these first results are promising and suggest that the dimer-jump algorithm gives the possibility to perform atomic-scale simulations of large systems such that experimentally accessible length scales are within reach.

30.5 Conclusion

We have used a newly developed hybrid MC-MD algorithm to study the relaxation of 2D Si islands on Si(001) and the formation of vacancy structures on clean and Ge-covered Si(001). The method is designed to overcome problems related to long time and length scales in physical phenomena occurring on semiconductor surfaces. This approach has the advantages of being suitable to off-lattice simulations and utilizing bulk-fitted potentials without the introduction of any additional parameters. The simulation results are in good qualitative agreement with experiments, and the investigation suggests that the algorithm is also suitable for quantitative studies.

Acknowledgments

This work was supported by the Academy of Finland, Project No. 1169043, and by NSF Grant No. DMR-0094422.

References

1. D.P. Landau, K. Binder: *A Guide to Monte Carlo Simulations in Statistical Physics*, Cambridge U. Press, (2000)
2. F.H. Stillinger, T.A. Weber: *Phys. Rev. B* **31**, 5262 (1985)
3. M. Laradji, D.P. Landau: *Phys. Rev. B* **51**, 4894 (1995)
4. L. Nurminen, F. Tavazza, D.P. Landau, A. Kuronen, K. Kaski: *Phys. Rev. B* **67**, 035405 (2003)
5. N. Metropolis, A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller, E. Teller: *J. Chem. Phys.* **21**, 1087 (1953)
6. F. Tavazza, Ph.D. Thesis, The University of Georgia, (2003)
7. F. Tavazza, L. Nurminen, D.P. Landau, A. Kuronen, K. Kaski: *Phys. Rev. E*, in press
8. Y.W. Mo, B.S. Swartzentruber, R. Kariotis, M.B. Webb, M.G. Lagally: *Phys. Rev. Lett.* **63**, 2393 (1989)
9. N.C. Bartelt, R.M. Tromp, E.D. Williams: *Phys. Rev. Lett.* **73**, 1656 (1994)
10. R.M. Tromp, R.J. Hamers, J.E. Demuth: *Phys. Rev. Lett.* **55**, 1303 (1985)
11. J. Wang, T.A. Arias, J.D. Joannopoulos: *Phys. Rev. B* **47**, 10497 (1993)
12. F. Liu, F. Wu, M.G. Lagally: *Chem. Rev.* **97**, 1045 (1997)
13. L. Nurminen, F. Tavazza, D.P. Landau, A. Kuronen, K. Kaski: *Phys. Rev. B* **68**, 085326 (2003)
14. D. Frenkel, B. Smit: *Understanding Molecular Simulation. From Algorithms to Applications*. (Academic Press, 1996)

31 Spin-Polarons in the FM Kondo Model

M. Dagofer¹, W. Koller², A. Prüll¹, H.G. Evertz¹, and W. von der Linden¹

¹ Institut für Theoretische Physik, TU Graz, Austria

² Department of Mathematics, Imperial College

Abstract. We show that holes doped into the FM Kondo model form independent FM spin-polarons with single holes inside, instead of showing phase separation. As a consequence there is a pseudogap in the DOS, which has been observed experimentally. Phase separation does appear at vanishing coupling J_{AF} of the corespins, where we also observe stripe-like configurations. At large J_{AF} and large doping, a flux phase appears. We present the phase diagram of this model in 2d.

Manganite compounds like $\text{La}_{1-x}\text{Sr}_x\text{MnO}_3$ and $\text{La}_{1-x}\text{Ca}_x\text{MnO}_3$ exhibit a very complex phase diagram involving, e.g., ferromagnetic (FM) metallic, FM insulating and different combinations of FM and AF behaviour in different lattice directions. They have been at the forefront of both experimental and theoretical research for a number of years [1]. One of the most striking effects is the so-called “Colossal Magneto-Resistance” (CMR), a fast change of resistivity as a function of magnetic field which occurs in part of the phase diagram and may lead to technical applications. The manganites have a perovskite lattice structure, similar to cuprates. The rich phase diagram is due to a complicated interplay of kinetic, spin, orbital, Coulomb, and Jahn-Teller lattice degrees of freedom, and has been a tremendous challenge for theoretical descriptions.

In Mn^{3+} the 3d orbitals are relevant. Three of them, called t_{2g} , are lower in energy and effectively form a spin- $\frac{3}{2}$ “corespin”. Only one of the two higher e_g -orbitals is occupied. The simplest model which captures an essential part of the relevant physics is the FM Kondo model (double exchange model)

$$\hat{H} = -t \sum_{\langle ij \rangle, \sigma} c_{i\sigma}^\dagger c_{j\sigma} - J_H \sum_i \hat{s}_i \cdot \hat{S}_i + J_{AF} \sum_i \hat{S}_i \cdot \hat{S}_i \quad (31.1)$$

in which lattice couplings and the second e_g -orbital have been neglected. The first term describes hopping of e_g electrons, the second is the Hund coupling to corespins, and the third the AF coupling among corespins. Because of strong Hund coupling J_H , the e_g -spins like to be almost parallel to the corespins which therefore tend to align ferromagnetically in the doped case. In the undoped case ($n = 1$) the corespins are AF. This model has been extensively investigated in the past, and has been claimed to show phase separation into a conducting hole-rich FM region and an AF insulator [1,2]. After consideration

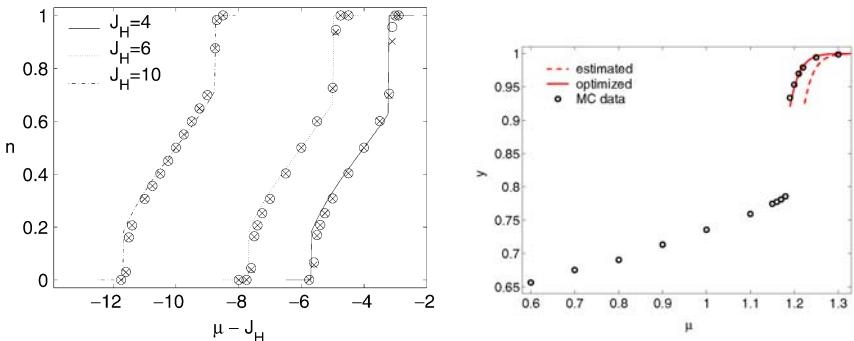


Fig. 31.1. Electron filling $n(\mu)$ for 1d (left) and 2d (right). The symbols denote MC-results

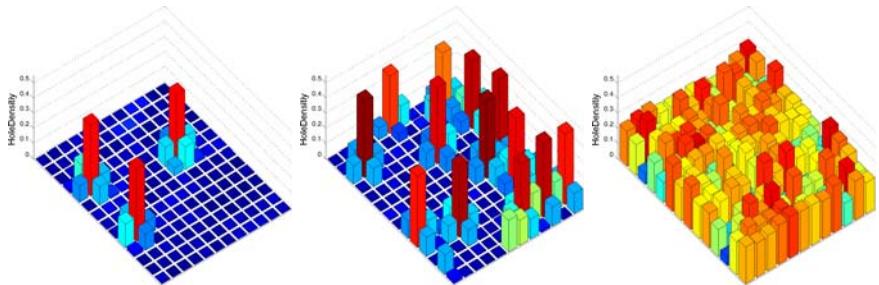


Fig. 31.2. MC snapshots of hole-density in 2d. AF-region (left), gap-region (middle), and the weakly FM-region (right)

of long range Coulomb repulsion, this gives rise to a model of “nanoscale phase separation” at the root of CMR [1,3]. We have, however, shown recently [4,5], that in 1d and 2d within physically relevant parameter ranges, there is instead a separation of individual holes into tiny FM regions, so-called spin-polarons, even without long-range Coulomb forces.

In our calculations, we first integrate out the high energy state in which the e_g electron has spin opposite to the corespin [6–8]. This is similar to the derivation of the tJ model from the Hubbard model. It results in an effective coupling $J_{\text{eff}} = J_{AF} + \frac{1}{2J_H}$. We can thus include effects of finite J_H without extra effort, instead of setting it to infinity as is often done. We treat the corespins as classical, which is justified because of their size [9,10]. \hat{H} is then bilinear in e_g electrons and can be integrated. What remains is a classical Monte Carlo simulation over corespin configurations [2], for which we have used suitable updates and taken care to treat autocorrelations correctly [4,5,8]. In the following, we show results for physically relevant parameter values $\beta = 50$, $J_{AF} = 0.02$, and $J_H = 6$, with t as our unit of energy. The calculations were performed on a lattices of size 50 in 1d and 12×14 in 2d.

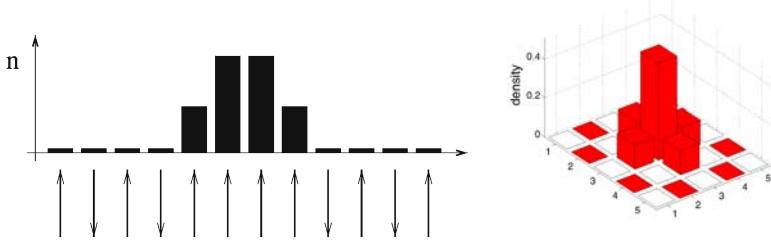


Fig. 31.3. Simple model for single-hole polarons in 1d (left) and 2d (right)

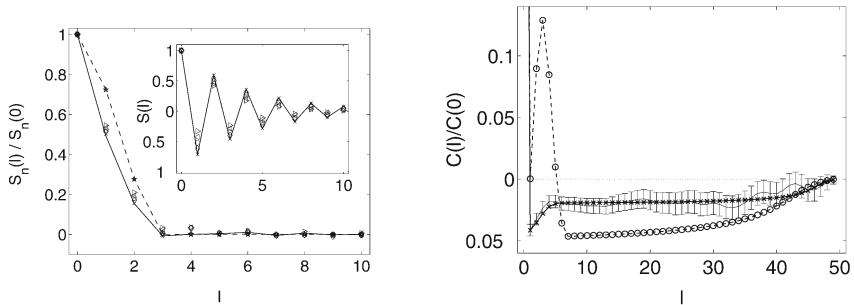


Fig. 31.4. *Left:* Dressed corespin-correlations for 1...5 holes in 1d. *Lines:* model polarons with perfect order (dashed) and in UHA (solid). *Inset:* spin-spin-correlation. *Right:* density correlations in 1d

In grand canonical simulations a jump in the electron filling $n(\mu)$ appears close to $n = 1$ (Fig. 31.1). This is usually interpreted as the above mentioned phase separation [2]. There are, however, strong indications against this interpretation, as we will now demonstrate.

One piece of evidence is the distribution of fermion occupation numbers in the simulations, which is smooth and broad in the FM region below the gap, but which is strongly peaked at integer electron occupations within the gap-region. Indeed, MC snapshots within the gap region (and above) do not show phase separation, but instead separated individual holes in a small FM corespin region, in which (in 2d) a single corespin is flipped with respect to the AF configuration (Fig. 31.2). These FM spin-polarons do not attract each other, but appear to occur at random positions.

We developed a simple model to describe such objects, as depicted in Fig. 31.3. Each polaron here contains a single hole and 3–5 FM corespin sites, embedded in an AF corespin background. In the simplest case, FM and AF order are taken to be perfect. This can be improved by the so-called Uniform Hopping Approximation (UHA) [8,11,12], which uses constant angles between corespins in the FM resp. AF phases. Within this model, single-hole polarons

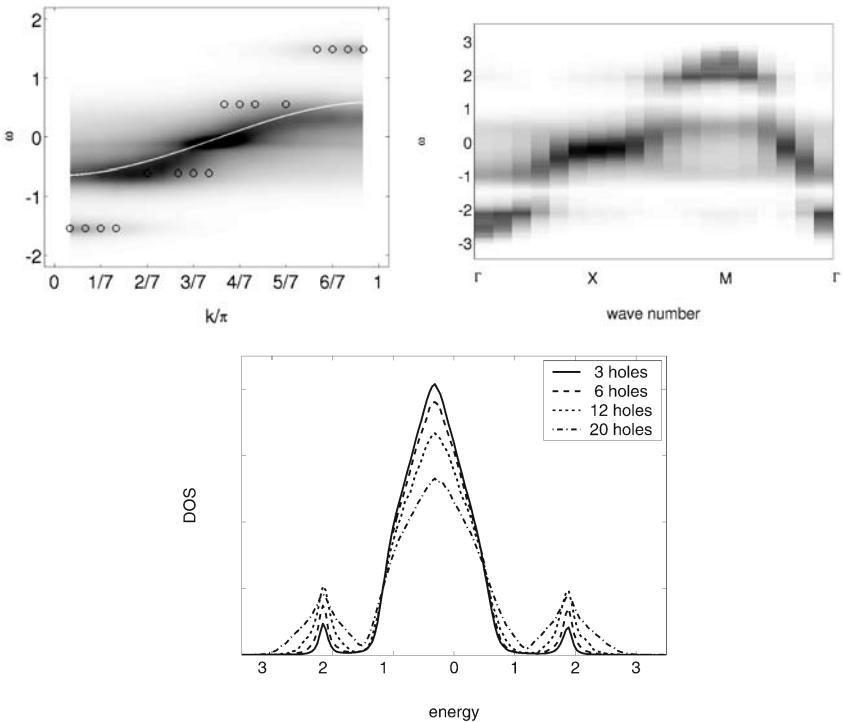


Fig. 31.5. Spectral function in 1d with 1 hole (*left*) and in 2d with 20 holes in the gap region (*middle*). *Right:* DOS in 2d for different dopings

turn out to be energetically favored over larger objects, except at very small J_{AF} (see below). At the critical chemical potential μ_c , the addition of another polaron costs no energy, which results in large fluctuations of the number of holes and can explain the jump in $n(\mu)$. We generated ensembles of such model polarons in random locations. Comparison with MC snapshots shows strikingly good agreement [5].

More quantitative evidence for *single-hole* polarons is shown in Fig. 31.4. On the left, the corespin-correlations $n_i^{\text{hole}} S_i \cdot S_j$ around a hole are seen to be almost independent of the number of holes. This is similar in 2d. On the right, the MC-data (bars) are perfectly described by single-hole polarons (solid line), but do not match at all the behavior of a bipolaron with two holes (dashed line).

The single-hole polaron of typically 4 sites in 1d is like a quantum mechanical well with just 4 states. The outermost of these states lie outside the band of the AF background (Fig. 31.5 left). Thus as a consequence of the small polaron, a symmetric *pseudogap* appears in the density of states, which has also been observed experimentally [13–16] and in previous calcula-

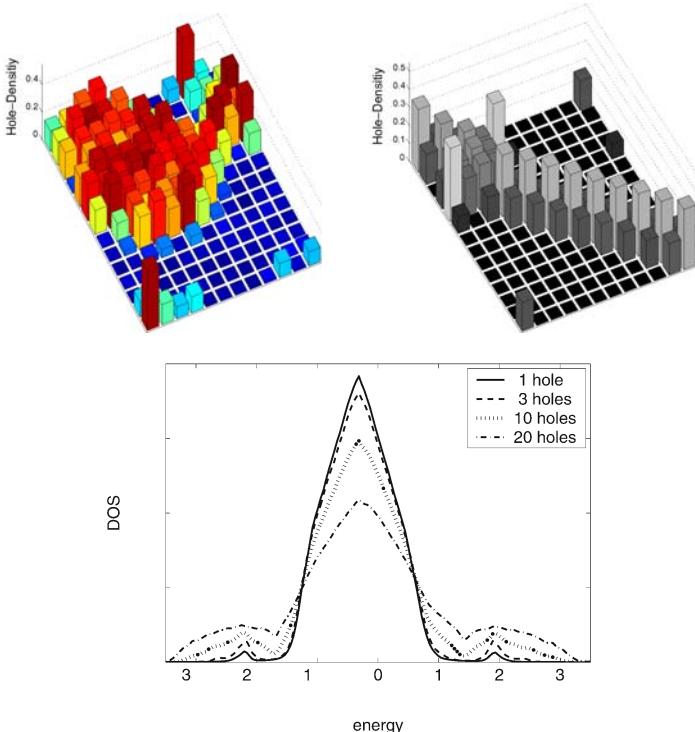


Fig. 31.6. $J_{AF} = 0$. Snapshots with phase separation (20 holes, *left*) and a stripe (11 holes, *middle*) at $\beta = 80$, both in the gap region. Closing of the pseudogap in the DOS (*right*, $\beta = 50$)

tions [1]. In 2d, the situation is similar. Additional holes go into more small polarons, and the pseudogap remains, independent of the number of holes, until the FM region is reached (Fig. 31.5 right). In fact, the transition from polaronic to FM behavior occurs when there is no more space for individual polarons.

The situation is different for $J_{AF} = 0$ (unphysical for the manganites), when there is less resistance to phase separated FM domains of corespins. We still see individual polarons, but at larger filling phase separation can occur now. As a consequence, the pseudogap fills up when the number of holes increases (Fig. 31.6). In addition, other exotic configurations occur at low temperature, like the stripe depicted in Fig. 31.6.

We have determined the phase diagram of the FM Kondo model in 2d (Fig. 31.7) at $\beta = 50$ and $\beta = 80$. For large J_{AF} , a so-called flux phase appears around $x = 1 - n = 0.5$ [17–19]. In the interesting doping range below about 20%, but with small J_{AF} , there is coexistence of polarons and

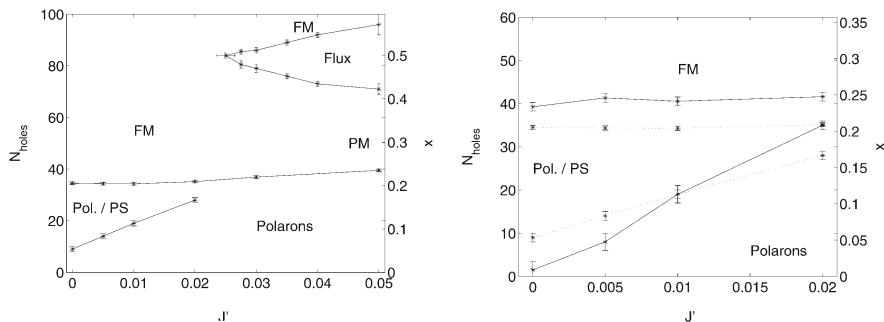


Fig. 31.7. Phase diagram in 2d, as a function of doping x (right scale) and AF coupling $J' \equiv J_{AF}$. Left: $\beta = 50$. Right: $\beta = 50$ (dashed) and $\beta = 80$ (solid) in a smaller range of parameters

phase separation, with strengthened phase separation for low temperature. For the likely physically relevant range $J_{AF} = 0.02 \dots 0.05$, however, polarons become even stronger when the temperature is reduced. At the same time, the ferromagnetism of corespins in the FM/PM phase becomes weaker as J_{AF} rises.

In summary, we have shown strong evidence, from MC simulations of the FM Kondo model with classical corespins, that holes doped into the model do not phase separate, but instead form small individual FM polarons. This effect becomes even stronger at lower temperatures in the relevant range of couplings.

References

1. E. Dagotto, T. Hotta, A. Moreo: Phys. Rep. **344**, 1 (2001)
2. S. Yunoki *et al.*: Phys. Rev. Lett. **80**, 845 (1998)
3. T. Wu *et al.*: Phys. Rev. Lett. **86**, 5998 (2001)
4. W. Koller, A. Prüll, H.G. Evertz, W. von der Linden: Phys. Rev. B **67**, 174418 (2003)
5. M. Daghofer, W. Koller, H.G. Evertz, W. von der Linden: cond-mat/0312394, and in preparation
6. S. Yarlagadda, C.S. Ting: Int. J. Mod. Phys. B **15**, 2719 (2001)
7. S.-Q. Shen, Z.D. Wang: Phys. Rev. B **61**, 9532 (2000)
8. W. Koller, A. Prüll, H.G. Evertz, W. von der Linden: Phys. Rev. B **66**, 144425 (2002)
9. N. Furukawa, in: *Physics of manganites*, 1. edn. (Kluwer Academic Publishing, New York 1998)
10. E. Dagotto *et al.*: Phys. Rev. B **58**, 6414 (1998)
11. J. van den Brink, D. Khomskii: Phys. Rev. Lett. **82**, 1016 (1999)
12. W. Koller, A. Prüll, H.G. Evertz, W. von der Linden: Phys. Rev. B **67**, 104432 (2003)

13. D.S. Dessau *et al.*: Phys. Rev. Lett. **81**, 192 (1998)
14. T. Saitoh *et al.*: Phys. Rev. B **62**, 1039 (2000)
15. Y.-D. Chuang *et al.*: Science **292**, 1509 (2001)
16. J.H. Park *et al.*: Phys. Rev. Lett. **76**, 4215 (1996)
17. H. Aliaga *et al.*: Phys. Rev. B **64**, 024422 (2001)
18. D.F. Agterberg, S. Yunoki: Phys. Rev. B **62**, 13816 (2000)
19. M.Y.W. Koshibae, S. Maekawa: Phys. Rev. Lett. **81**, 5604 (1998)

List of Contributors

- Adler, J., 169
Bertino, E., 61
Bhatt, R.N., 136
Boettcher, S., 74
Brown, G., 157, 196
Bruce, K.B., 61
Chambers, C., 61
Chen, D., 174
Daghofer, M., 276
Dean, J., 61
El-bayyari, Z., 205
Endres, R.G., 100
Evertz, H.G., 276
Frank, S., 80
Ghosh, C., 215
Grove. D., 61
Hansmann, U.H.E., 87
Hihinashvilli, R., 169
Ito, K., 95
Kara, A., 215
Kaski, K., 270
Kato, K., 95
Kim, J.S., 265
Koller, W., 276
Koper, M.T.M., 258
Körner, M., 142
Krawczyk, J., 122
Kuronen, A., 270
Landau, D.P., 1, 169, 270
Lee, G.S., 265
Lee, H.K., 157
Lewis, S.P., 1
Liu, J., 109
Luijten, E., 109
Mason, D.R., 241
Mitchell, S.J., 258
Müller, M., 151
Novotny, M.A., 201
Nurminen, L., 270
Okamoto, Y., 146
Okumura, H., 146
Percus, A.G., 74
Prellberg, T., 122
Prüll, A., 276
Rahman, T.S., 215
Rapaport, D.C., 7
Rechnitzer, A., 122
Rikvold, P.A., 90, 196
Rudd, R.E., 241
Schmickler, W., 80
Schulthess, T.C., 100, 157
Schüttler, H.-B., 1
Sevim, V., 90
Shimada, T., 95
Shumway, J., 181
Sutton, A.P., 241
Tavazza, F., 270
Temam,R., 61
Thompson, S.H., 196
Troyer, M., 142
Tsai, S.-H., 169

- Vink, R., 45
Virnau, P., 151
von der Linden, W., 276
Wingreen, N.C., 100
- Yamamoto, R., 61
Yukawa, S., 19
Yun, S.J., 265
- Zheng, B., 25
Zhou, C., 136